Diese Arbeit wurde vorgelegt am Lehrstuhl A für Mathematik

Machine Learning: automatische Bilderkennung mit Mathematik?!

Ein Lehr-Lern-Modul im Rahmen eines mathematischen Modellierungstages für Schülerinnen und Schüler der Sekundarstufe II

Machine Learning: automatic image recognition with mathematical methods?!

A teaching and learning module in the context of a mathematical modeling day for pupils at German secondary level II

Masterarbeit Mathematik

März 2019

Vorgelegt von Presented by	Lars Schmidt [redacted]
Erstprüfer First examiner	Prof. Dr. Sebastian Walcher Lehrstuhl A für Mathematik RWTH Aachen University
Zweitprüfer Second examiner	Prof. Dr. Martin Frank Steinbuch Centre for Computing (SCC) Karlsruhe Institute of Technology
Betreuerin Supervisor	Sarah Schönbrodt Lehrstuhl II für Mathematik RWTH Aachen University

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

1. Einleitung

Amazon, Netflix, Zalando, Allianz, Apple, Google, Facebook. Die Auflistung soll keinesfalls Schleichwerbung für einen der genannten Konzerne sein. Alle sieben sind Firmen, die in verschiedenen Branchen aktiv sind. Doch von sozialen Netzwerken über Versicherungen und Technikproduktion, bis zu Online Shopping und Video-Streaming vereint sie etwas. Die Gemeinsamkeit ist verantwortlich für den Bekanntheitsgrad und den Einfluss der Konzerne. Sie ist ein aktuelles Forschungsgebiet, das unzählige Anwendungen bietet und in den unterschiedlichsten Bereichen genutzt wird (vgl. Guido & Müller, 2017, S. 1). Es handelt sich dabei um das sogenannte *Machine Learning*.
Dieses höchst relevante Thema soll in einem Lehr-Lern-Modul für das Schülerlabor CAMMP aufgegriffen werden. Dazu wurde ein aktueller Anwendungsbereich ausgewählt, in dem Machine Learning genutzt wird: die Bilderkennung. Die Bilderkennung wird im Modul zunächst auf Verkehrsampeln und anschließend auf Gesichter angewendet. Ziel ist es den Schülerinnen und Schülern[1] aus der Sekundarstufe II, die am Workshops teilnehmen, bewusst zu machen, wie präsent das Thema ist. Gleichzeitig soll ihnen verdeutlicht werden, dass es keine riesigen Forschungsabteilungen wie bei den oben genannten Firmen braucht, um Machine Learning zu nutzen (vgl. Guido & Müller, 2017, S. IX). Neben diesen spezifischen Zielen des entwickelten Lernmoduls, sollen natürlich auch die allgemeinen Ziele des CAMMP Programms umgesetzt werden. Allen voran, soll die mathematische Modellierungskompetenz der Schüler an lebensrelevanten Problemstellungen gefördert werden. Die weiteren Ziele, die CAMMP verfolgt, werden in Kapitel 2 erläutert.
Die Motivation für die Entwicklung eines Lernmoduls im Bereich Machine Learning als Thema meiner Abschlussarbeit, hat im wesentlichen zwei verschiedene Gründe. Zum einen fasziniert mich das Thema. Die Tatsache, dass es in zahlreichen, aktuellen und alltäglichen Situationen genutzt wird, macht eine intensive Auseinandersetzung damit sehr interessant. Des Weiteren bieten die Anwendungsbereiche aus dem Lebensumfeld der Schüler eine ideale Möglichkeit für die Entwicklung eines Lehr-Lern-Moduls. Dies führt direkt zum zweiten Grund, warum ich mich für das Thema entschieden habe. Als angehende Lehrkraft mit den Fächern Mathematik und Physik, ist mir die Notwendigkeit mathematische Modellierung verstärkt in die Schulen zu bringen, durchaus bewusst (vgl. Ferri & Greefrath, 2013, S.2f). Die Möglichkeit, einen Workshop für ein Schülerlabor zu entwickeln kann einen wichtigen Beitrag dazu leisten. Gleichzeitig wird das erarbeitete Modul im Rahmen des CAMMP Programms ständig weiter entwickelt. Indem der Workshop wiederholt durchgeführt wird und das Modul dabei stetig verbessert und angepasst wird, bleibt die Arbeit aktuell.
Kernpunkt dieser Arbeit ist es, das erarbeitete Modul vorzustellen und in seinen Einzelheiten zu beschrieben. Dazu wird in Kapitel 2 zunächst ein Überblick über das CAMMP Schülerlabor und dessen Veranstaltungsformate, sowie die Theorie mathematischer Modellieren gegeben.
Anschließend wird in Kapitel 3 der für den Workshop grundlegende mathematische

[1] Im Folgenden nur noch mit Schüler abgekürzt. Gleiches gilt für Lehrer, Betreuer, usw.

Hintergrund beschrieben. Zu Beginn wird eine Einführung in den Begriff Machine Learning vorgenommen und verschiedene Lernarten werden vorgestellt. Anschließend wird auf die *Support Vector Machine*, kurz SVM, eingegangen. Dabei handelt es sich um einen Algorithmus, welcher auf Grundlage von Beispielen Klassifizierungsprobleme löst. Die SVM stellt den Kernpunkt des mathematischen Hintergrundes dar und wird deshalb detailliert beschrieben. Da das Lernmodul für die Bearbeitung durch Schüler entworfen wurde, sollen hier auch Bezüge zum Kernlehrplan des Landes Nordrhein-Westfalen hergestellt werden. Der Kernlehrplan wurde letztlich nach den Beschlüssen der Kultusministerkonferenz (KMK) formuliert. Diese gehen wiederum auf den Fachdidaktiker Winter zurück (vgl. Kultusministerkonferenz, 2012, S. 11); (vgl. Reit, 2016, S. 1). Daher lassen sich die Inhalte des Moduls auch für andere Lerngruppen anwenden, die nicht aus Nordrhein-Westfalen kommen.
In Kapitel 4 wird das entworfene Modul in seinen einzelnen Bestandteilen erklärt. Zu Beginn wird der typische Ablauf eines Modellierungstages aus dem Schülerlabor vorgestellt. So soll auch der CAMMP-unerfahrene Leser einen guten Eindruck vom Ablauf erhalten. Anschließend werden die erstellten Materialien vorgestellt und dabei Bezug auf den Modellierungsbegriff aus Kapitel 2 genommen. Die Entscheidungen für den Aufbau des Moduls und die Formulierung einzelner Aufgabenteile wird didaktisch begründet und am Verlauf der sogenannten Modellierungsspirale (siehe Abschnitt 2.2.2) erklärt.
Im letzten Kapitel der vorliegenden Arbeit wird das Modul nach seiner Durchführung evaluiert und ein Ausblick auf mögliche Erweiterungen gegeben.

Zum Schluss der Einleitung soll noch beschrieben werden, worin das Ziel der Arbeit besteht und an welche Leser sie sich wendet. Im Vordergrund steht die Entwicklung des Lernmoduls. In diesem Zusammenhang wurden alle Arbeitsmaterialien und Vorträge erstellt, die im Anhang aufgeführt sind. Diese Materialien wurden im Rahmen zweier Durchführungen des Moduls getestet, evaluiert und verbessert. Die Ziele des Lernmoduls werden in den Kapiteln 2 und 4 detailliert beschrieben. Neben den Zielen des Lernmoduls liegen dieser Arbeit jedoch noch andere Intentionen zu Grunde: dem Leser soll vor allem einen Überblick über den didaktischen Hintergrund mathematischer Modellierung, sowie über den fachlichen Hintergrund der SVM geben werden. Dabei wird dem Leser ein gewisses mathematisches Verständnis unterstellt. Die Arbeit richtet sich unter anderem an Lehrkräfte, weshalb auch Vergleiche zur Schulmathematik gezogen werden sollen. Leser der Arbeit sollen im Anschluss nicht nur das Prinzip der SVM verstanden haben. Sie sollen auch eine Möglichkeit kennen lernen, wie das Thema mit Schulmathematik behandelt werden kann. Ohne, dass sie bei einer Durchführung dabei gewesen seien müssen, sollen sie einen Eindruck davon haben, wie der Workshop aufgebaut ist.

2. Didaktischer Hintergrund

Ziel dieses Kapitels ist es, dem Leser in erster Linie eine Einführung in den Begriff der mathematischen Modellierung zu geben. Zu Beginn wird das Angebot und die Philosophie des Schülerlabors CAMMP beschrieben. Anschließend soll ausgehend vom Modellierungsbegriff der *Modellierungskreislauf* erklärt werden. Dabei wird zunächst die Variante beschreiben, welche in den Lernmodulen verwendet wird. Während der Lernmodule soll der Modellierungskreislauf den Schülern als Orientierung dienen. Er wird auch als *didaktischer* Modellierungskreislauf bezeichnet (Ferri & Greefrath, 2013, S. 17). An dieser Stelle wird er nicht verwendet, um Diagnosen bezüglich der Modellierungskompetenz der Schüler treffen zu können. Kreisläufe, die dafür vorgesehen sind, bezeichnet man als *diagnostische* Modellierungskreisläufe (vgl. Brand, 2014, S. 10, 14). Sie werden am Ende dieses Kapitels beschrieben.

2.1. Das Schülerlabor CAMMP

Das Schülerlabor CAMMP ist eines von insgesamt acht Schülerlaboren der RWTH Aachen. Das Projekt wurde 2011 von Prof. Martin Frank, Prof. Ahmed Ismail, Dr. Christina Roeckerath und Dr. Nicole Faber gegründet und entwickelt sich seitdem immer weiter. Dabei steht die Abkürzung CAMMP für **C**omputational and **M**athematical **M**odeling **P**rogram. Wie dieser Name schon verrät, steht mathematische Modellierung im Mittelpunkt des Programms. Im Wesentlichen geht es darum Schüler und Wissenschaftler zusammen zu führen, um anhand von praktischen Beispielen die Grundlagen mathematischer Modellierung kennenzulernen. Für diese Zusammenarbeit gibt es zwei verschiedene Veranstaltungsformate: den CAMMP day und die CAMMP week.
CAMMP days sind eintägige Workshops für einen gesamten Mathematik Schulkurs zu verschiedenen lebensnahen Fragestellungen. Diese Fragestellungen sind bewusst aus dem Kontext der teilnehmenden Schüler gewählt. Im Rahmen des Workshops lösen die Teilnehmer mit der Unterstützung zweier studentischer Hilfskräfte die jeweilige Fragestellung. Dabei arbeiten sie in Zweier- bis Dreiergruppen, unter der Verwendung der Software MATLAB[2] an einem Laptop. Bereits entwickelte Module für die Workshops sind zum Beispiel *Wie funktioniert eigentlich Google und was hat das mit Mathe zu tun?*, *Von der Schallplatte bis zum mp3 Player: wie funktioniert das Komprimierungsverfahren?* oder aber *Wie sicher ist meine Privatsphäre in sozialen Netzwerken und was hat das mit Mathe zu tun?*[3]. Diese Auswahl wird durch das Modul *Künstliche Intelligenz: Gesichtserkennung und autonomes Fahren mit Mathematik?!*, welches in dieser Arbeit entwickelt wurde, ergänzt. Wie an den Beispielen zu sehen ist, stammen die Fragestellungen zwar aus verschiedensten Bereichen, können jedoch alle mittels mathematischer Modellierung gelöst werden. Die Module richten sich dabei nach den aktuellen Kernlehrplänen des Landes Nordrhein-Westfalen, sodass die Lehrkräfte immer das passende Modul zum im Unterricht behandelten Stoff auswählen können. An dieser Stelle soll jedoch betont werden, dass die Fragestellungen bei CAMMP noch

[2]https://de.mathworks.com/products/matlab.html
[3]https://blog.rwth-aachen.de/cammp/

über den Kernlehrplan hinaus gehen. So lauten die Leitlinien für Aufgaben, dass sie möglichst offen, flexibel und tiefer mathematisch seien sollen.

Neben den CAMMP days gibt es noch die CAMMP week. Einmal im Jahr fahren 35 bis 40 Schüler verschiedener Schulen mit begleitenden Lehrkräften für eine Woche nach Belgien in eine Jugendherberge. Dort arbeiten sie in Neunergruppen bestehend aus sechs Schülern, zwei Lehrern und einem wissenschaftlichen Mitarbeiter an teilweise ungelösten Fragestellungen. Diese realen Probleme stammen meist aus der Forschung von Firmen oder Universitätsinstituten. Beispiele sind *Wie kann man optimale Knochenimplantate für Unfallverletzte entwickeln? (Uniklinik RWTH Aachen)* oder *Wie muss man Solarkraftwerke konstruieren, damit sie wettbewerbsfähig werden? (TSK Flagsol)*. Auch hier werden die Schüler von Wissenschaftlern unterstützt. Parallel zur Gruppenarbeit, wird eine Fortbildung für die Lehrkräfte zum Thema „Mathematische Modellierung mit Schüler/innen“ angeboten. Am Ende der Woche präsentieren die Schülergruppen ihre Ergebnisse vor den Firmenvertretern, ihren Familien und anderen Interessenten auf einer repräsentativen Abschlussveranstaltung in der Universität.
Neben der oben genannten Idee, Schülern mathematische Modellierung näher zu bringen, verfolgt das Programm noch weitere Ziele. Angelehnt an den Modellierungskreislauf, welcher im Abschnitt 2.2.2 detaillierter beschrieben wird, sollen die Schüler zunächst ein umgangssprachlich formuliertes Problem mathematisch ausdrücken. Dieses Problem soll dann mittels vorhandener mathematischer Kenntnisse gelöst werden. Bevor die Ergebnisse interpretiert werden, findet eine Bewertung und Überprüfung der Lösung statt. An diesem Ablauf lässt sich erkennen, dass die Lernziele der jeweiligen Module an die prozessbezogenen Kompetenzen des Kernlehrplans angelehnt sind. So werden neben dem Modellieren, besonders das Problemlösen und, durch die Arbeit im Team, das Kommunizieren trainiert (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014, S. 19ff). Neben diesen mathematischen und prozessbezogenen Zielen, ist auch die Berufs- und Studieninformation, sowie die Begeisterung für ein Mathematik-, oder CES (Computaional Engeneering Science) Studium Teil der CAMMP Philosophie. Daher wird bei jedem CAMMP day ein Vortrag zum Thema Berufs- und Studienorientierung gehalten, bei dem der Studiengang CES vorgestellt wird.
Außerdem sollen auch die begleitenden Lehrkräfte vom Angebot profitieren. Sei es im Rahmen der Fortbildung auf der CAMMP week, oder aber bei den Tagesworkshops. Sie lernen, wie komplexe aber anwendungsorientierte Probleme die Schüler motivieren und in den Schulunterricht eingebunden werden können. Teil der CAMMP Philosophie ist es, Lehrkräfte zu befähigen und zu ermutigen ihren Schülern forschungsbasierte Modellierung zu vermitteln. Neben den Teilnehmern profitieren auch die betreuenden Hilfswissenschaftler. Gerade die Lehramtsstudenten im CAMMP Team können während der Module wichtige Unterrichtserfahrungen machen.

2.2. Mathematische Modellierung

In der Einleitung dieser Arbeit wurde bereits beschrieben, welch wichtige Rolle der Anwendungsbezug bei Schulaufgaben spielt. Nach Borromeo Ferri sind gute Modellierungsaufgaben ein Kernmerkmal für qualitativ hochwertigen Unterricht. Außerdem sind sie zentral für die Rekonstruktion kognitiver Prozesse der Lernenden (vgl. Ferri & Greefrath, 2013, S. 4). Für die CAMMP Workshops sollen Modellierungsaufgaben jedoch nicht zu diagnostischen Zwecken, wie der Rekonstruktion kognitiver Prozesse, sondern zum stärken der Modellierungskompetenz eingesetzt werden. Auch die Kultusministerkonferenz (KMK) hat die Bedeutung mathematischer Modellierung erkannt und sie als eine der sechs allgemeinen mathematischen Kompetenzen festgelegt (vgl. Kultusministerkonferenz, 2012, S. 11). Durch diesen Beschluss wird mathematisches Modellieren im Kernlehrplan des Landes Nordrhein-Westfalen als eine der vier prozessbezogenen Kompetenzen genannt (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014, S. 16). Im Folgenden soll der Modellierungsbegriff unter Einbezug der didaktischen Forschung näher erläutert werden.

2.2.1. Der Modellierungsbegriff

Als einer der ersten Vertreter für Realitätsbezüge im Mathematikunterricht in den 70er Jahren, unterscheidet der US-amerikanische Mathematiker Henry Pollak vier verschiedene Aspekte zum Modellierungsbegriff (vgl. Ferri & Greefrath, 2013, S. 11) nach (Pollak, 1977, S. 255f). Zum ersten nennt er die klassische *Angewandte Mathematik*, welche zum Beispiel die physikalische Anwendung der Analysis beinhaltet. Zusammen mit dem zweiten Begriff, der *Anwendbaren Mathematik*, betonen diese beiden Aspekte die fachmathematische Seite. Unter Anwendbare Mathematik fallen die verschiedenen mathematischen Disziplinen wie Statistik, Lineare Algebra, Analysis oder Stochastik. Die beiden anderen Begriffe, welche explizit das prozesshafte Lösen eines Problems betreffen, sind das *Vereinfachte Modellieren* und das *Modellieren*. Während bei ersterem der Modellierungskreislauf nur ein Mal durchlaufen wird, wird er beim Modellieren mehrfach durchlaufen (vgl. Ferri & Greefrath, 2013, S. 11). Auf die CAMMP days lässt sich letzterer Begriff übertragen. Während der Module durchlaufen die Teilnehmer mehrere Male den Modellierungskreislauf. Wirft man einen Blick in den Duden, so findet man unter dem Verb modellieren (wissenschaftlich): „von etwas ein Modell herstellen, bilden“ (Duden, 2018). Im Fall der Mathematik müsste man ein mathematisches Modell der Realität aufstellen. Bevor also ein Blick auf den oben erwähnten Modellierungskreislauf geworfen wird, muss zunächst der Begriff mathematisches Modell betrachtet werden.

Greefrath schreibt in Anlehnung an Leuders und Maaß, dass ein mathematisches Modell eine isolierte Darstellung der Welt sei, die vereinfacht wurde und zur Anwendung von Mathematik geeignet sei (vgl. Padberg & Greefrath, 2010, S. 43). Durch die Vereinfachung und die isolierte Betrachtung hat ein mathematisches Modell in seiner Anwendung auf ein reales Problem jedoch Grenzen. Wie im unteren Abschnitt 2.2.2 beschrieben wird, werden nie alle Eigenschaften eines realen Problems im mathematischen

Modell ausgedrückt. Es werden zunächst Annahmen und Vereinfachungen getroffen. Im Regelfall ist es aber gar nicht Zweck des Modells die komplexe Realität vollständig abzubilden (vgl. Padberg & Greefrath, 2010, S. 43); (Roddeck, 2008, S. 9). Aus den hier genannten Eigenschaften folgert Greefrath weiter, dass Modelle nicht eindeutig seien könnten. Sie seien subjektiv und je nachdem welche Vereinfachungen getroffen wurden, können unterschiedliche Ergebnisse resultieren (vgl. Padberg & Greefrath, 2010, S. 44). Für das Lehr-Lern-Modul bedeutet das, dass die Diskussion verschiedener Schülerlösungen eine wichtige Rolle spielt. Zuletzt sollen noch zwei Klassen von Modellen unterschieden werden. Abhängig davon, welchen Zweck sie erfüllen, kann man sie in *normative* und *deskriptive* Modelle unterteilen. Normative Modelle dienen als Vorlage für die reale Situation. Padberg (2010) definiert normative Modelle als solche, die mathematische Vorschriften entwickeln, die in bestimmten Situationen für Entscheidungen verwendet werden können (vgl. Padberg & Greefrath, 2010, S. 94). Ein Beispiel für eine normative Modellierung wäre die Verteilung der Heizkosten in einem Haus mit mehreren Wohnungen nach vorgegebenen Regeln. Bei einem deskriptiven Modell soll die reale Situation beschrieben werden. Es steht ein Original zu Verfügung, von dem dann eine Modell gebildet werden soll. Ein Beispiel dafür wäre die Beschreibung der Entwicklung einer Tierpopulation durch eine geeignete Funktion. Deskriptive Modelle können noch weiter unterteilt werden, je nachdem ob sie rein beschreibend, bereits erklärend oder sogar voraussagend sind (vgl. Ferri & Greefrath, 2013, S. 13f). Um bei dem Beispiel der Tierpopulation zu bleiben, könnte man mit der aufgestellten Funktion auch Voraussagen über die zukünftige Entwicklung der Population treffen.
Beim Lesen des obigen Abschnitts mag der Eindruck entstanden sein, dass Modellierung bzw. das Aufstellen eines Modells lediglich den Übersetzungsprozess von der Realität zur Mathematik beinhaltet. Um diesem Eindruck entgegen zu wirken, wird in der Literatur auch von Modellieren und Anwenden gesprochen. Es soll betont werden, dass die mathematische Lösung nach der Berechnung wieder auf die Realität zurück übertragen werden muss (vgl. Reit, 2016, S. 9). Blum schreibt, dass es neben dieser Begriffswahl eine Vielzahl von Auffassungen und Formulierungen über den Modellierungsbegriff gibt. Die Spannweite reiche vom Mathematisieren im engeren Sinn, also dem Aufstellen eines mathematischen Modells als geeignetes Abbild eines Ausschnitts der Welt, bis hin zum angewandten Problemlösen (vgl. Blum, 2006, S. 9). Blum schlägt für seine Sichtweise einen Modellierungskreislauf vor, welcher den gesamten Lösungsprozess eines realen Problems unter der Verwendung von mathematischen Methoden beschreibt. Sein Kreislauf, der im folgenden Abschnitt beschrieben wird, dient in dieser Arbeit als Leitlinie für die Entwicklung des Lernmoduls.

2.2.2. Der Modellierungskreislauf

Um zu beschreiben, was Modellieren ist, wird der sogenannte *Modellierungskreislauf* genutzt. Er ist genau genommen auch ein Modell und beschreibt den idealisierten Lösungsprozess, welcher durchlaufen werden muss, um ein bestimmtes Problem zu lösen. *Idealisiert* meint in diesem Zusammenhang, dass der Kreislauf ein normatives Vorgehen zur Lösung darstellt (vgl. Reit, 2016, S. 11). Der Kreislauf bei CAMMP soll den

Schülern Anhaltspunkte liefern, an welchem Schritt der Gesamtlösung sie sich gerade befinden. Ein solcher „didaktischer Modellierungskreislauf“(Ferri & Greefrath, 2013, S. 17) weist eine geringe Komplexität auf, die die gewünschte Verortung des eigenen Standpunktes im Lösungsprozess erleichtert. Außerdem ist er in einer für Schüler angemessenen Sprache formuliert (vgl. Brand, 2014, S. 15). Einer der ersten und bekanntesten Kreisläufe dieser Art geht auf Blum zurück und stammt aus den 1980er Jahren. Er wird in der folgenden Abbildung 1 dargestellt.

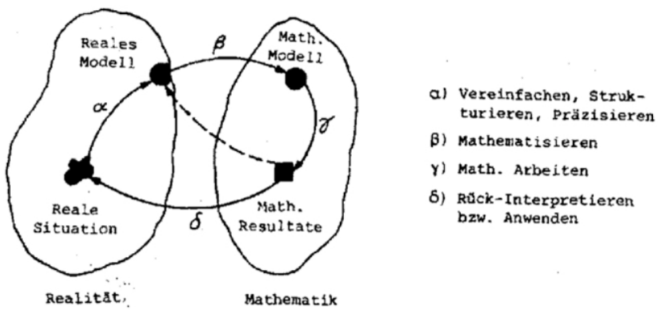


Abb. 1: Reduzierter, didaktischer Kreislauf nach Blum (aus: Reit, 2016, S. 12).

Wie in Abbildung 1 zu sehen ist, besteht der Kreislauf aus vier Punkten: der *realen Situation*, dem *realen Modell*, dem *mathematischen Modell* und den *mathematischen Resultaten*. Diese Punkte sind mit Pfeilen verbunden, welche die einzelnen Arbeitsschritte zur Lösung des Problems darstellen. Die Schritte sind mit α, β, γ und δ gekennzeichnet. Im Folgenden werden die einzelnen Schritte des Kreislaufs anhand des erarbeiteten Lernmoduls erklärt.
Ausgangspunkt eines jeden Modellierungsprozesses ist die *reale Situation*. Bevor mit der Lösung begonnen werden kann, muss das zugrundeliegende Problem verstanden werden. Für Schukajlow bedeutet Verstehen einer Aufgabenstellung, dass der Problemlöser eine mentale Repräsentation dieser bildet (vgl. Schukajlow, 2011, S. 78). Der Text, das Bild der Realsituation und das Vorwissen des Lesers stellen die Grundlage für die Repräsentation dar. Informationen aus der Aufgabenstellung werden geordnet und mit individuellem Wissen ergänzt. Fehlende Informationen werden recherchiert. Im Workshop wird das reale Problem durch die automatische Klassifizierung von Bildern gegeben. Zunächst müssen Informationen über die Bilder gesammelt werden. Welche Größe (Pixelzahl) haben sie? Wie viele verschiedene Bildklassen sollte es geben?
Im ersten Schritt (α), dem **Vereinfachen und Strukturieren**, wird aus der realen Situation das *reale Modell* gebildet. In diesem Schritt werden keine zusätzlichen Informationen hinzugefügt, sondern bestehende Informationen reduziert. Der Problemlöser muss irrelevante Informationen herausfiltern und Vereinfachungen treffen (vgl. Schukajlow, 2011, S. 78). Im Workshop wird zunächst die Vereinfachung getroffen Bilder aus zwei Pixeln zu betrachten. Um einen Anwendungsbezug herzustellen, werden Verkehrsampeln ausgewählt. Diese bestehen, nach dem vereinfachten Modell aus zwei übereinander angebrachten Boxen, vergleichbar mit zwei Pixeln eines Bildes. Es werden zunächst lediglich rote und grüne Ampelphasen betrachtet. Für rote Ampeln werden

verschiedene rote Farbtöne für die obere und verschiedene Schwarztöne für die untere Box angenommen. Die Farbtöne sind zunächst als [r,g,b]-Tripel gegeben, wobei r den Rot-, g den Grün-, und b den Blauanteil des jeweiligen Farbtons angibt. Für grüne Ampeln werden die verschiedenen Schwarztöne für die obere und verschiedene Grüntöne für die untere Box angenommen. Es sollte darauf geachtet werden, das Modell nicht zu weit zu reduzieren, es also zu einfach zu gestalten. Dadurch könnte es passieren, dass das reale Modell die Situation nicht mehr korrekt beschreibt. Um bei dem Beispiel zu bleiben, könnte es sich als zu ungenau erweisen nur einen bestimmten Farbton für die jeweilige Ampelphase zuzulassen. Durch verschiedene Tageszeiten, erscheinen die Ampelfarben nicht immer genau gleich. Alle Informationen, welche dann noch im Realmodell enthalten sind, müssen im nächsten Schritt mathematisiert werden. Um die Informationen auf ihre Mathematisierbarkeit überprüfen zu können, wird die Kenntnis mathematischer Operationen benötigt (vgl. Schukajlow, 2011, S. 79).
Als nächstes erfolgt das **Mathematisieren** (β). Dabei werden die Objekte und ihre Relationen untereinander zu einem neuen Modell zusammengebracht. Das Realmodell wird unter der Verwendung mathematischer Ausdrücke in ein *mathematisches Modell* überführt (vgl. Schukajlow, 2011, S. 78). Der Mathematikdidaktiker Heinrich Winter (1975) weist Mathematikunterricht vier allgemeine Lernziele zu. Eines dieser Ziele ist es, dass die Schüler lernen, Situationen der Wirklichkeit zu mathematisieren (vgl. Winter, 1975, S. 50). Genau dieses Ziel wird im zweiten Modellierungsschritt verfolgt. Die Bilder werden in diesem Schritt als 2×1-Vektoren dargestellt. Dafür müssen die [r,g,b]-Tripel des jeweiligen Pixels aus dem realen Modell zunächst zu einem Wert zusammen gefasst werden. Dies kann durch die Grauwertberechnung aus der Bildverarbeitung realisiert werden. Um aus einem [r,g,b]-Tripel den entsprechenden Grauwert g zu berechnen wird die Formel $g(r, g, b) = 0,299 \cdot r + 0,587 \cdot g + 0,114 \cdot b$ genutzt[4]. Der erste Eintrag der 2×1-Vektoren beinhaltet den Grauwert für den oberen, der zweite für den unteren Pixel. Die Merkamlsvektoren werden dann zu Punkten in einem zweidimensionalen Raum. Um das mathematische Modell zu vervollständigen, muss eine Funktion gefunden werden, welche diese Datenpunkte voneinander trennt. Das mathematische Modell, welches soeben beschrieben wurde, bleibt in seiner Form für den gesamten CAMMP day bestehen. Das mathematische Problem wird jedoch immer komplexer. Die steigende Komplexität äußert sich zum einen durch den Übergang von zwei- auf dreidimensionale Daten und schließlich zum hochdimensionalen Fall der Gesichtserkennung. Zum anderen wird sie durch die Unterscheidung von mehr als zwei Klassen und dem Problem nicht linear-separierbarer Daten realisiert.
Nachdem das mathematische Modell konstruiert wurde, müssen für das **mathematische Arbeiten** (γ) die passenden mathematischen Werkzeuge ausgewählt werden. Die im Modell formulierten Operationen müssen ausgeführt werden, um eine Lösung, also *mathematische Resultate* zu erhalten (vgl. Schukajlow, 2011, S. 80). In diesem Schritt könnten vor allem die *formalen Fertigkeiten* in den Mittelpunkt gestellt werden. Diese Fertigkeiten beinhalten unter anderem das Kennen und Anwenden mathematischer Definitionen und Regeln und das formale Arbeiten mit Variablen, Termen und Funktio-

[4]Diese Formel wird auch von MATLAB verwendet.

nen. Das Erlernen formaler Fertigkeiten stellt ein weiteres der vier allgemeine Ziele von Mathematikunterricht nach Winter dar (vgl. Winter, 1975, S. 54). Für die Workshops ist dieses Ziel eher sekundär. Wie im weiteren Verlauf des Abschnitts erwähnt wird, werden bei CAMMP days Computer verwendet, welche die Schüler zur Bearbeitung des Moduls nutzen. Diese Computer können kalkülhafte Berechnungen und Lösungsverfahren von Standardaufgaben erledigen, sodass der Fokus auf den Modellierungsprozess gelegt werden kann (vgl. Frank et al., 2018, S. 141). Im Workshop werden mittels der *Support Vector Machine* (vgl. Kapitel 3) Parameter für eine Funktion bestimmt, welche die Merkmalsvektoren einordnen kann. Die Schüler können dann einen Merkmalsvektor in diese Funktion eingeben und erhalten einen bestimmten Wert, welcher die mathematische Lösung darstellt.
Im vierten Schritt (δ) des Modellierungskreislaufs erfolgt das **Interpretieren** der mathematischen Resultate. Die mathematischen Ausdrücke müssen wieder in die Realität übersetzt und ein Bezug zur realen Situation hergestellt werden (vgl. Schukajlow, 2011, S. 80). Der Wert, den die Schüler aus der Funktion erhalten haben, muss interpretiert werden um abschließend eine Einordnung des Bildes vornehmen zu können. Wie diese Einordnung genau funktioniert ist ebenfalls in Kapitel 3 nachzulesen. Durch diese Resultate muss jedoch nicht zwangsweise eine Lösung der Realsituation gegeben sein. Inwiefern die Lösung plausibel und zufriedenstellend ist, muss bei der Interpretation herausgefunden werden (vgl. Schukajlow, 2011, S. 80). Daher wird beim vierte Schritt nicht nur vom Interpretieren, sondern auch vom **Validieren** gesprochen. Eine einfache Plausibilitätsprüfung wäre beispielsweise die Betrachtung der Einheiten und der Größenordnung. Eine andere Möglichkeit wäre es, das Modell anhand von Aufgaben zu überprüfen, deren Lösungen bereits bekannt sind (vgl. Hinrichs, 2008, S. 26). Gelangt man zu dem Schluss, dass die Resultate nicht zufriedenstellende für die Realsituation sind, müssen die Schritte eins bis vier wiederholt werden. Bei den meisten Problemen reicht ein einmaliges Durchlaufen des Kreislaufs nicht aus, da die gefundene Lösung die Realsituation noch nicht gut abbildet. Ein Fehler kann in jedem Schritt des Kreislaufs liegen. Zum einen könnten schon falsche Annahmen oder zu starke Vereinfachungen getroffen worden sein. Zum anderen könnte bei der mathematischen Formulierung oder Berechnung ein Fehler unterlaufen sein (vgl. Schukajlow, 2011, S. 81). Die Schüler werden im Verlauf des Workshops mehrfach die Erfahrung machen, dass das Ergebnis noch nicht zufriedenstellend ist. Wie in Abschnitt 4.1 beschrieben wird, werden sie mindestens fünf Durchläufe durch den Modellierungskreislauf machen. Hinter jedem neuen Durchlauf steht eine Steigerung der Komplexität des mathematischen Problems. Allen Durchläufen liegt jedoch das gleiche mathematische Modell zu Grunde. Die Daten werden durch Merkmalsvektoren beschrieben, die als Punkte in einem n-dimensionalen Raum getrennt werden müssen. Aufgabe der Schüler ist für jeden Durchlauf kognitive, stellenweise auch innermathematische Arbeit zu leisten, um die Trennfunktion für das immer komplexer werdende Problem zu finden.
Der Kreislauf nach Blum dient als direktes Vorbild für den Modellierungskreislauf, welcher in den CAMMP days verwendet wird (vgl. Abb. 3).

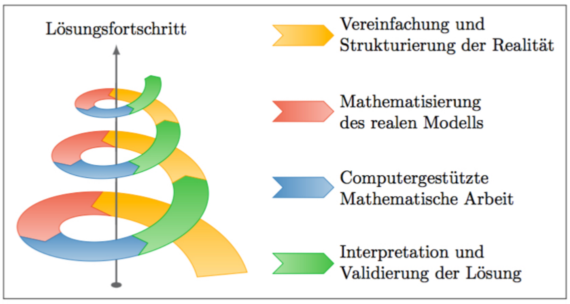


Abb. 2: Solution Helix of Math (aus: Frank et al., 2018, S. 141).

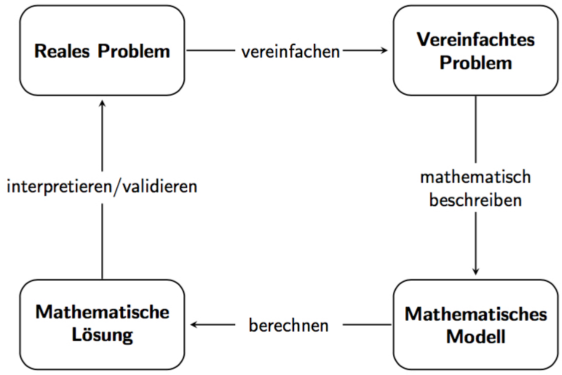


Abb. 3: In den Modulen gezeigter Kreislauf (angelehnt an Blum, 1985).

Wie in Abbildung 3 zu sehen ist, werden die gleichen Punkte und Schritte wie in Abbildung 1 verwendet. Lediglich die Darstellungsform unterscheidet sich geringfügig.
Neben dem Kreislauf in Abbildung 3, welcher in den CAMMP days verwendet wird, liegt bei der Entwicklung der Lernmodule ein leicht abgewandeltes Modell zu Grunde. Dieses Modell berücksichtigt den Einsatz digitaler Werkzeuge. Anstelle eines Kreises, wird der Modellierungsprozess durch eine aufwärts drehende Spirale verdeutlicht. Das Modell stammt ursprünglich von der Initiative *Computer-Based Math*[5], welche sie als *Solution Helix of Math* betitelt hat. In Abbildung 2 ist die von CAMMP verwendete Form dargestellt (vgl. Frank et al., 2018, S. 140). Im Unterschied zu Büchter und Leuders, welche ebenfalls die Form einer Spirale vorschlagen (vgl. Padberg & Greefrath, 2010, S. 48) nach (Büchter & Leuders, 2005, S. 76), die sich jedoch immer weiter öffnet, wird die von CAMMP verwendete Spirale immer enger. Durch eine solche Darstellung soll der Lösungsfortschritt während des Modellierungsprozesses betont werden. Mit jedem Umlauf nähern sich die Modellierenden immer weiter der Lösung (vgl. Padberg & Greefrath, 2010, S. 48f); (Frank et al., 2018, S. 139ff). In diesem Zusammenhang wird auch klar, warum von einer Spirale, anstatt einer Helix gesprochen wird.
Eine weitere Besonderheit der von CAMMP verwendeten Spirale im Vergleich zu Büchter und Leuders, besteht im Einbezug des Computers für den Modellierungsprozess. Nach Greefrath & Weitendorf (2013) stellt der Computer besonders bei realitätsbezogenen Problemen ein sinnvolles Werkzeug dar, das sowohl Lernende als auch Lehrende unterstützen kann (vgl. Greefrath & Weitendorf, 2013, S. 181). Der Computer kann an jedem Schritt, des in Abbildung 1 und 3 beschriebenen Kreislaufs eingesetzt werden. So ermöglicht der Computereinsatz im ersten Schritt bereits die Recherche wichtiger Informationen, welche für das Realmodell benötigt werden. Weiter können das aufgestellte Modell und die gefundenen Ergebnisse am Computer visualisiert werden. Für die teilnehmenden Schüler ist es wichtig, das Modell nicht nur durch die mathematischen Symbole, sondern auch auf der bildliche Repräsentationsebene zu sehen. Der Wechsel zwischen diesen Ebenen wird als wesentlich für das Verstehen mathematischer Zusammenhänge gesehen (vgl. Schmitz, 2017, S. 2). Es können leicht Parameter variiert und

[5]https://www.computerbasedmath.org/maths-process-poster/

deren Auswirkungen direkt am Modell beobachtet werden. Durch den Computereinsatz können die Schüler in kurzer Zeit viel mehr experimentieren. Beim mathematischen Arbeiten in den CAMMP Workshops ist der Computer essenziell. Meist werden numerische Berechnungsverfahren benötigt, welche die Schüler nicht in überschaubarer Zeit erledigen können (vgl. Greefrath & Weitendorf, 2013, S. 182). Aus diesem Grund wird in den CAMMP Angeboten die Software MATLAB verwendet, welche größtenteils zur Datenerfassung, -auswertung und -analyse, sowie zur numerischen Simulation genutzt wird. Sie hat sich gegen andere Numerikprogramme im professionellen Bereich von der Hochschule bis in die Industrie durchgesetzt (vgl. Thuselt & Gennrich, 2014, S. 2). Neben der konkreten Berechnung, wird MATLAB in dem hier erarbeiteten Lernmodul auch für das Algebraisieren und Visualisieren verwendet. Bei der Bilderkennung werden die realen Farbwerte der n Pixel eines Bildes mit dem Computer zunächst in Grauwerte umgewandelt. Die Grauwerte aller n Pixel werden anschließend in einem n-dimensionalen Vektor gespeichert. Im Falle der Verkehrsampeln, die als Bilder aus zwei Pixeln zu verstehen sind, werden die 2×1-Vektoren mittels der Software als Datenpunkte in einem zweidimensionalen Koordinatensystem visualisiert. Dabei wird auf der x-Achse der x_1-Wert und auf der y-Achse der x_2-Wert des entsprechenden Vektors aufgetragen wird. Für diese Datenpunkte wird mit MATLAB eine Trennfunktion zur Klassifizierung der Ampeln gesucht.
Da die Schüler während der Module möglichst selbstständig arbeiten sollen, ist eine selbständige Ergebniskontrolle von großer Wichtigkeit. Mittels MATLAB können direkte Rückmeldungen erstellt werden, mit denen die Schüler ihre Ergebnisse überprüfen können. In Abschnitt 4.3 wird noch genauer auf diese Rückmeldefunktion eingegangen. Haben die Schüler einen Fehler gemacht, so erhalten sie nicht das korrekte Ergebnis, sondern einen Hinweis, an welcher Stelle sie ihren Lösungsweg überdenken müssen. Die Möglichkeit der Selbstkontrolle, sei es durch direkte Rückmeldungen, oder auch durch Grafiken, ist eine weiterer Vorteil des Einsatzes digitaler Werkzeuge (vgl. Greefrath & Siller, 2018, S. 10).
Es sollte an dieser Stelle betont werden, dass digitale Werkzeuge keinesfalls das Verstehen der mathematischen Modelle und der außermathematischen Probleme ersetzen. Durch die oben beschriebenen Funktionen wie Visualisieren, Experimentieren, Algebraisieren und Berechnen kann der Verstehensprozess jedoch unterstützt werden (vgl. Greefrath & Weitendorf, 2013, S. 183). Wird der Computer für den Modellierungsprozess hinzugezogen, so muss dem Kreislauf ein zusätzlicher Schritt hinzugefügt werden. Dieser Schritt wird in Abbildung 4 verdeutlicht. Wie beim ursprünglichen Kreislauf, muss die reale Situation vereinfacht und in das Realmodell umgewandelt werden. In diesem Schritt ist das Problem immer noch in der *Alltagssprache* formuliert. Anschließend wird es bei beiden Kreisläufen mit *Mathematischer Sprache* in das mathematische Modell überführt. Nach diesem Schritt müssen die mathematischen Ausdrücke beim Computereinsatz jedoch zusätzlich in die *Befehlssprache* der verwendeten Software übersetzt werden. Nachdem mit ihrer Hilfe eine Lösung bestimmt wurde, muss diese wieder zweifach zurückübersetzt werden (vgl. Greefrath & Weitendorf, 2013, S. 184f). Da die wenigsten Schüler im Vorfeld schon Erfahrungen mit MATLAB gemacht haben, könnte ihnen dieser Schritt Schwierigkeiten bereiten. Um dem entgegenzuwirken, wer-

den an vielen Stellen, an denen die Schüler Eingaben im Programm machen müssen, schon vorgefertigte Befehlsbausteine oder Anleitungen gegeben.

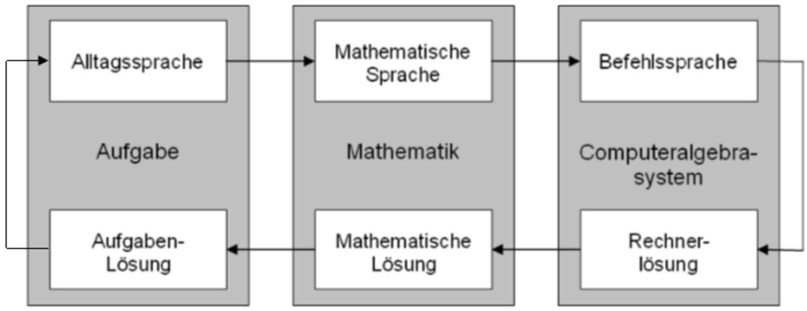


Abb. 4: Übersetzungsschritte beim Modellieren mit Computern (aus: Greefrath & Weitendorf, 2013, S. 184).

Neben dem didaktischen Modellierungskreislauf, welcher im vorherigen Abschnitt ausführlich vorgestellt wurde, gibt es noch den sogenannten *diagnostischen* Modellierungskreislauf, der die Untersuchung kognitiver Prozesse der Problemlöser ermöglichen soll (vgl. Brand, 2014, S. 10, 14). Dieser Kreislauf geht ebenfalls auf Blum zurück (vgl. Blum & Leiß, 2005, S. 18f). Er besteht anstatt aus vier, aus sieben Teilaktivitäten, die idealtypisch jeweils durchlaufen werden müssen, um die Problemstellung zu lösen (vgl. Blum, 2006, S. 9). In folgender Abbildung 5 wird der diagnostische Kreislauf dargestellt:

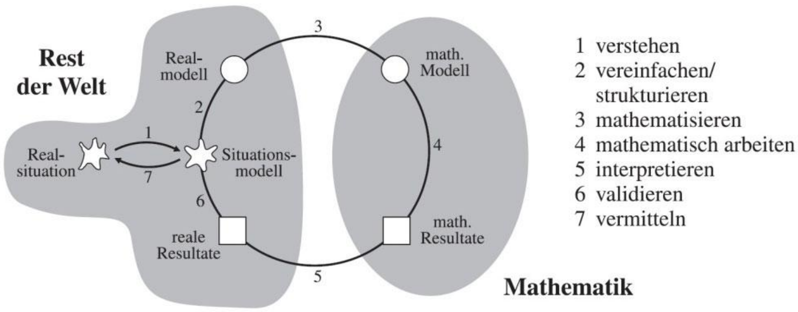


Abb. 5: Siebenschrittiger Modellierungskreislauf (aus: Blum, 2006, S. 9).

Er unterscheidet sich vom didaktischen Kreislauf (vgl. Abb. 1) im Wesentlichen durch den Einbezug des Situationsmodells. Dies soll die Untersuchung kognitiver Prozesse der Problemlöser ermöglichen (vgl. Brand, 2014, S. 10, 14). Borrommeo Ferri und Greefrath (2013) ordnen solche Kreisläufe dem *dreischrittigen Mathematisieren* unter, da sie neben dem realen und dem mathematischen das Situationsmodell enthalten. Auch sie betonen die Eigenschaft „möglichst genau beschreiben zu können, wie Lernende mit Modellierungsaufgaben umgehen" (Ferri & Greefrath, 2013, S. 18).
Ein weiterer Unterschied ist noch die Hinzunahme des siebten Schrittes, dem *Vermitteln.* Die Problemlöser sollen ihre Resultate in angemessener Weise präsentieren, wobei

vor allem die Kompetenzen Argumentieren und Kommunizieren im Vordergrund stehen. Nicht nur im Unterricht ist die Präsentation der gefundenen Lösung wichtig. Immer wenn in der Wirtschaft oder der Wissenschaft eine Problem modelliert und gelöst wurde, müssen die Resultate dem Auftraggeber in einer passenden Form vorgestellt werden (vgl. Hinrichs, 2008, S. 28). In den Workshops erfolgt dieser Schritt durch die gemeinsame Besprechung der Arbeitsblätter. Des weiteren wurde den Punkte Interpretieren und Validieren im Vergleich zum didaktischen Kreislauf jeweils ein eigener Schritt zugewiesen.
Durch die zusätzlichen Schritte weist der diagnostische Kreislauf eine, für den CAMMP Einsatz unnötige Komplexität auf. Da der Modellierungskreislauf nicht zu diagnostischen Zwecken eingesetzt werden soll, sondern eher als Werkzeug zur Metakognition, wird die vereinfachte Version benutzt (vgl. Abb. 3).

2.2.3. Modellieren als Kompetenz

Eine der Grundideen von CAMMP ist es die Modellierungskompetenz der Teilnehmer zu stärken. Diese Idee geht auf die drei Grunderfahrung für den Mathematikunterricht nach Winter zurück. Ihm zufolge, soll der Mathematikunterricht allgemein bildend sein, indem er drei Grunderfahrungen ermöglicht:

- „Erscheinungen der Welt um uns, die uns alle angehen oder angehen sollten, aus Natur, Gesellschaft und Kultur, in einer spezifischen Art wahrzunehmen und zu verstehen, (G1)

- mathematische Gegenstände und Sachverhalte, repräsentiert in Sprache, Symbolen, Bildern und Formeln, als geistige Schöpfungen, als eine deduktiv geordnete Welt eigener Art kennen zu lernen und zu begreifen, (G2)

- in der Auseinandersetzung mit Aufgaben Problemlösefähigkeiten, die über die Mathematik hinaus gehen, (heuristische Fähigkeiten) zu erwerben.“ (G3) (vgl. Winter, 1996, S. 37).

Durch den Einsatz von Modellierungsaufgaben und digitalen Werkzeugen, können alle drei Erfahrung bedient werden. Die Anwendungsorientiertheit und der realer Kontext solcher Aufgaben bieten den Schülern die Möglichkeit, ihre Umwelt mit mathematischen Mitteln zu erschließen (vgl. Ferri & Greefrath, 2013, S. 20). So wird Grunderfahrung (G1) erfüllt. Durch den Einsatz des Computers bei Modellierungsaufgaben, können Grafiken und dynamische Visualisierungen genutzt werden, wie bereits oben beschrieben wurde. Auf diese Weise können die Repräsentationsformen Sprache, Symbole, Bilder und Formeln aus Grunderfahrung (G2) miteinander verknüpft werden (vgl. Greefrath & Siller, 2018, S. 8). Des Weiteren können die in (G3) genannten Problemlösefähigkeiten durch den Einsatz von Modellierungsaufgaben gefördert werden, da die

Schüler neue Strategien entwickeln müssen um die realen Probleme zu lösen (vgl. Ferri & Greefrath, 2013, S. 20). Die Problemlösefähigkeit zu fördern ist, wie oben beschrieben, ein Ziel von CAMMP. Es sollen eher diese, als die formalen Fertigkeiten geschult werden, was durch den Computereinsatz ermöglicht wird. Durch den Computereinsatz kann der Fokus anstatt auf Lösungsverfahren von Standardaufgaben, auf die einzelnen Modellierungsschritte gelegt werden (vgl. Frank et al., 2018, S. 141).

3. Mathematischer Hintergrund

In diesem Kapitel wird der mathematische Hintergrund des entwickelten Lernmoduls beschrieben. Im Kern geht es um die *Support Vector Machine* (SVM). Dieser Algorithmus wird in erster Linie zur Lösung von Klassifikatiosproblemen eingesetzt. In der Masterarbeit von Sarah Schönbrodt (vgl. Schönbrodt, 2018) wurde er bereits beschrieben und zur Klassifikation von Bildern genutzt. Im Unterschied dazu, soll der mathematische Hintergrund der SVM in dieser Arbeit kleinschrittiger beschrieben werden. Das Kapitel soll den Interessenten des Lernmoduls die Möglichkeit geben, den fachlichen Hintergrund zu verstehen. Bei diesen Interessenten wird vor allem an die begleitenden Lehrkräfte gedacht, welche zwar ein gewisses mathematischen Vorwissen, jedoch noch keine genaueren Kenntnisse des Algorithmus haben. Das Kapitel teilt sich in drei Abschnitte auf. Im ersten Abschnitt 3.1 wird er Begriff *Machine Learning* erklärt und in drei Kategorien unterteilt. Der zweite Abschnitt 3.2 stellt den mathematischen Teil dieser Arbeit dar. Die SVM wird detailliert beschrieben. Der Ablauf orientiert sich dabei an dem Aufbau des Lernmoduls, in dem die Schüler den Algorithmus zum Teil herleiten. Im dritten Abschnitt 3.3 wird der Bezug zum Kernlehrplan gesucht und erklärt, welche Teile der SVM an den Schulstoff anknüpfen.

3.1. Was ist *Machine Learning?*

In diesem Abschnitt soll zunächst der Begriff *Machine Learning* erläutert werden. Es werden drei verschiedene Unterkategorien des Begriffs beschrieben und anhand von Beispielen verdeutlicht, welche Anwendungen in die jeweilige Unterkategorien zählen.

Der Begriff Machine Learning, oder übersetzt *Maschinelles Lernen*, setzt sich aus zwei Bestandteilen zusammen. Zunächst soll der hintere Teil, das *Lernen*, betrachtet werden. Schlägt man den Begriff in psychologischen Lexika nach, so findet man etliche Definitionen. Zimbardo definiert „Lernen als einen Prozess, [...] der zu relativ stabilen Veränderungen im Verhalten oder im Verhaltenspotential führt und auf Erfahrung aufbaut.“ (Zimbardo, 2013, S. 227).
Schmitt versteht unter Lernen „den Erwerb, die Veränderung oder den Abbau von Erlebens- und Verhaltensweisen durch bestimmte Umwelterfahrungen.“ (Schmitt, 1999, S. 1).
An diesen zwei Formulierungen wird ersichtlich, was auch für einen Großteil der anderen Definitionen gilt. Die meisten Definitionen für den Lernbegriff enthalten die beiden Punkte, dass Lernen erstens ein Prozess ist, welcher auf Erfahrungen aufbaut und zweitens, dass es zu einer Verhaltensänderung kommt. Genau diese beiden Aspekte können auch auf das Lernen von *Maschinen* übertragen werden. Diese Maschinen sind meist Computerprogramme, welche ähnlich wie ihre lebendigen Vorbilder einen Lernprozess durchlaufen sollen (vgl. Frochte, 2018, S. 13). Die Erfahrungen, welche in den obigen Definitionen genannt wurden, stellen Datensätze dar. Mit diesen Datensätzen soll das Programm arbeiten und davon ausgehend bestimmte Aufgaben erledigen. Die Aufgabe eines Computerprogramms der Firma *Amazon* aus der Einleitung dieser Arbeit könnte

zum Beispiel die Einordnung von Kunden sein. Jeder *Amazon* Nutzer kennt wohl Sätze wie „Das könnte Sie auch interessieren“ oder „Kunden, die dies gekauft haben, kauften auch...“. Grundlage für diese Vorschläge sind Lernprogramme, welche mit Nutzerdaten arbeiten und sie nach ihrem Verhalten in verschiedene Klassen einteilen (vgl. Frochte, 2018, S. 25).

Die Anwendungsbeispiele, die für das Lehr-Lern-Modul genutzt werden, weisen einen ähnlichen Aufgabentyp auf. Die automatische Bilderkennung wird im ersten Teil des Workshops auf Bilder mit zwei beziehungsweise drei Pixeln vereinfacht. Als Kontext wurde die Klassifizierung von Verkehrsampeln für autonomes Fahren gewählt. Der Computer im Auto soll Verkehrsampeln anhand ihrer Farbe erkennen und entscheiden, ob das Auto anhalten muss oder weiterfahren kann. Bei der automatischen Gesichtserkennung, welche im zweiten Teil des Workshops thematisiert wird, setzen sich die Bilder aus mehr als drei Pixeln zusammen. Es soll erkannt werden, welche Person auf einem Bild zu sehen ist, um beispielsweise ein Handy zu entsperren. Bei den beschriebenen Problemen geht es darum Daten auf Grund ihrer verschiedenen Merkmale einer Klasse zuzuordnen. Merkmale eines *Amazon* Kunden könnten bereits bestellte Artikel, oder die aufgerufenen Seiten sein. Merkmale bei der Bildklassifizierung sind die Farbwerte der einzelnen Pixel.

Mathematisch können die Daten als sogenannte *Merkmalsvektoren* geschrieben werden (vgl. Géron, 2018, S. 39). Ein Datensatz X mit m Datenpunkten enthält die reellen Merkmalsvektoren $\vec{x}_i$, $i = 1, ..., m$. Besteht ein Datenpunkt aus n Merkmalen, so haben die Vektoren $\vec{x}_i$ die Dimension n. Es gilt:

$$X = \{\vec{x}_1, ..., \vec{x}_m\} \subset \mathbb{R}^n \ . \tag{3.1}$$

Für 200 Datenpunkte eines Gesichter-Datensatzes, welcher aus Fotos der Größe 20×30 Pixeln bestehen würden X also wie folgt aussehen:

$$X = \{\vec{x}_1, ..., \vec{x}_{200}\} \subset \mathbb{R}^{600} \ .$$

Die Einträge von $\vec{x}_i$ wären dann durch die Farbwerte der 600 Pixel gegeben. Aufgabe des Machine Learning-Programms ist es nun, eine Funktion f zu suchen, welche den Merkmalsvektoren $\vec{x}_i$ einen bestimmten Ausgabewert in der abgeschlossen Menge Y zuordnet. Gesucht ist also

$$f : X \to Y \ . \tag{3.2}$$

Bezieht man die Terminologie aus Abschnitt 2.2 mit ein, so handelt es sich bei der Darstellung eines Datensatzes durch Merkmalsvektoren bereits um ein mathematisches Modell. Die Daten sind zunächst durch verschiedene Bilder von Gesichtern gegeben. Diese wurden vereinfacht und lediglich die Farbwerte der einzelnen Pixel als Merkmale ausgewählt. Die 600 Merkmale jedes Gesichtes wurden dann in die Einträge eines 600×1-Vektors übertragen.

Im Folgenden soll zwischen drei verschiedenen Lernarten unterschieden werden. Bei allen drei Arten soll eine Funktion f wie in (3.2) beschrieben gefunden werden. Der Unterschied ist jedoch jeweils die Gestalt der Mengen X und Y (vgl. Frochte, 2018, S. 20).

3.1.1. Drei verschiedene Lernarten

Die Lernalgorithmen, mit denen eine Funktion f gesucht werden soll, lassen sich in drei verschiedene Kategorien einteilen. Es wird zwischen *überwachtem, unüberwachtem* und *bestärkendem Lernen* unterschieden. Der beim CAMMP day angewendete Lernalgortihmus, die Support Vector Machine, zählt zur Kategorie des überwachten Lernens. Daher wird diese Kategorie im Folgenden ausführlicher beschrieben als die anderen beiden.

Überwachtes Lernen

Typische Aufgaben, welche von überwachten Lernalgorithmen gelöst werden, sind *Klassifizierungsprobleme* (vgl. Frochte, 2018, S. 20). Bei Klassifizierungsproblemen sollen Daten in verschiedene Klassen eingeordnet werden. Dazu liegt ein Datensatz X von *Beispieldaten* vor. Auf Grundlage dieses Datensatzes soll eine handhabbare Funktion f gesucht werden, die neue Daten, die nicht im Beispieldatensatz liegen, einordnen kann. Die Menge X der Beispieldaten besteht aus Merkmalsvektoren $\vec{x}_i \in \mathbb{R}^n$, wie sie bereits in Gleichung (3.1) beschrieben wurden. Zu diesen Merkamlsvektoren liegt beim *überwachten Lernen* (engl. *supervised learning*) noch eine Zusatzinformation vor. Diese Zusatzinformation beinhaltet die Klasse, zu der der jeweilige Merkmalsvektor gehört. Um bei dem Ampelbeispiel aus dem Workshop zu bleiben: von den Verkehrsampeln, welche durch die Merkmalsvektoren repräsentiert werden, ist bekannt, welche grün und welche rot sind. Im Gegensatz zu dem in Gleichung 3.1 beschrieben Datensatz, wird ein solcher Datensatz X aus m Merkmalsvektoren folgendermaßen beschrieben:

$$X = \{(\vec{x}_1, t_1), (\vec{x}_2, t_2), ..., (\vec{x}_m, t_m)\} \subseteq \mathbb{R}^n \times C \tag{3.3}$$

Die Merkmalsvektoren $\vec{x}_i$ werden in Tupeln zusammen mit der Zusatzinformation $t_i \in C$ abgespeichert. Die abgeschlossene und diskrete Menge $C \subseteq \mathbb{R}$ steht für die Menge der Klassen. Die $t_i,\ i = 1, .., m$ werden als *Labels* bezeichnet und geben an, in welcher Klasse der jeweilige Merkmalsvektor liegt. Man spricht auch von einem markierten oder *gelabelten* Datensatz. Bei dem Beispiel der roten und grünen Ampeln, würde man der einen Ampelklasse das Label -1 und der anderen das Label $+1$ zuweisen. Man spricht auch von einem *binären Klassifizierungsproblem* (vgl. Raschka & Mirjalili, 2017, S. 27f), da die Merkmalsvektoren in zwei Klassen aufgeteilt sind. Dementsprechend wäre $C = \{-1, 1\}$. Der Computer im Auto soll dann anhand dieser Daten eine handhabbare Funktion $f: \ X \to C$ erlernen, die neue, ungelabelte Daten korrekt in die Klassen zuordnet (vgl. Wroble et al., 2013, S. 405). Das Prinzip des überwachten Lernens wird in der folgenden Abbildung 6 verdeutlicht.

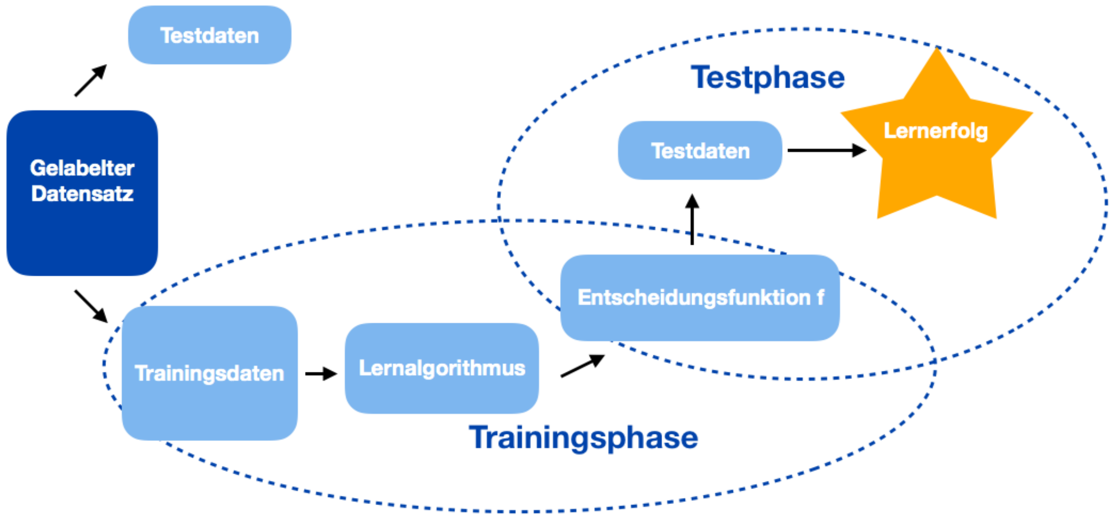


Abb. 6: Schema des überwachten Lernens. Einteilung eines gelabelten Beispieldatensatzes in Test- und Trainingsdaten. Aufstellen der Entscheidungsfunktion f mit Trainingsdaten durch Lernalgorithmus. Anschließendes Testen der Funktion f mit Testdaten. Bewertung des Lernerfolgs.

Beim überwachten Lernen wird der gesamte gelabelte Beispieldatensatz zunächst in *Trainings-* und *Testdaten* aufgeteilt. Die Angaben zum Verhältnis von Trainings- zu Testdaten variieren. Gängige Einteilung sind $70\% - 80\%$ Trainings- und $20\% - 30\%$ Testdaten (vgl. Whitenack, 2017, S. 77). Mit Hilfe des Lernalgorithmus wird dann eine *Entscheidungsfunktion* f gesucht, welche die Testdaten möglichst gut einteilt. Was „möglichst gut" heißt, wird im Abschnitt 3.2 genauer beschrieben. Für die Entscheidungsfunktion müssen (fast) alle Trainingsdaten korrekt eingeordnet werden. Die Phase, in der die Funktion f gesucht wird, nennt man *Trainingsphase*. Wurde die Funktion f gefunden, so wird sie auf die Testdaten angewendet. Da die Testdaten ebenfalls gelabelt sind, also bekannt ist, wie sie korrekt eingeordnet werden müssten, kann einfach überprüft werden, ob die Funktion f den richtigen Wert ausgibt. Auf diese Weise kann der Lernerfolg eingeschätzt werden. Diese Phase wird *Testphase* genannt.
Ein oft genanntes Anwendungsbeispiel für ein binäres Klassifizierungsproblem Problem wäre die Erkennung von Spam E-Mails (vgl. Guido & Müller, 2017, S. 27); (Géron, 2018, S. 8), wie es in Abbildung 7 dargestellt wird.

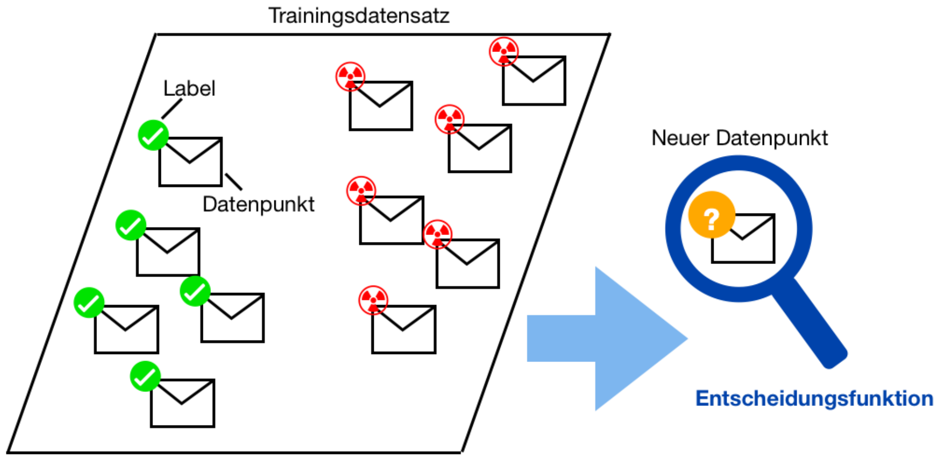


Abb. 7: Beispiel für ein binäres Klassifizierungsproblem überwachten Lernens. Zuordnung von E-Mails in die Klassen Spam (rot) und nicht Spam (Ham genannt) (grün). Datenpunkte durch Merkmalsvektoren der E-Mails gegeben. Aufstellen der Entscheidungsfunktion f mit der unbekannte E-Mails eingeordnet werden können (angelehnt an: Géron, 2018, S. 8).

Die beiden Klassen sind durch *Spam* (rot) oder nicht Spam (grün), auch *Ham* genannt, gegeben. Die Merkmalsvektoren werden durch die Briefe symbolisiert und bestehen aus verschiedenen Eigenschaften von E-Mails. Es könnten zum Beispiel gewisse Wortkombinationen wie „Für Sie“, „Herzlichen Glückwunsch“, „Gewinner“, „kostenlos“ oder „Kreditkarte“ darin codiert sein (vgl. Géron, 2018, S. 4f). Zu jedem Merkmalsvektor aus den Beispieldaten gehört ein Label. Die Labels sind in Abbildung 7 durch die grünen und roten Symbole gegeben. Anhand der Beispieldaten wird dann, wie in Abbildung 6 beschrieben, eine Entscheidungsfunktion f gesucht, welche die Merkmalsvektoren den beiden Klassen zuordnet. Anschließend können dann mit f neue, unbekannte E-Mails eingeordnet werden. Dieser Schritt wird durch die Lupe und die Mail mit dem gelben Symbol in Abbildung 7 verdeutlicht.
Natürlich könnte man einen Spamfilter auch ohne Machine Learning programmieren. Die übliche Vorgehensweise wäre dann, den Datensatz händisch nach Gemeinsamkeiten und Mustern von Spam E-Mails zu durchsuchen. Wenn man Muster und Unterschiede zu den Ham E-Mails gefunden hat, so könnte man für jede Auffälligkeit eine Regel formulieren. Das Ganze würde in einer langen Liste von Regeln resultieren, die für jede einzelne E-Mail abgearbeitet werden müsste, um sie einzuordnen. Wenn die Versender von Spam E-Mails dann bemerken, warum ihre Mails geblockt werden, so können sie einfach kleine Änderungen vornehmen. So könnten sie zum Beispiel einfach andere Wörter benutzen und der Spam Filter würde nicht mehr funktionieren (vgl. Géron, 2018, S. 4f).
Ein auf Machine Learning basierender Filter, würde im Idealfall erkennen, dass in den geänderten Spam E-Mails nun andere Wörter gehäuft auftreten und automatisch seine Zuordnung anpassen. Der Vorteil liegt also sowohl in der höheren Flexibilität, als auch in der Rechengeschwindigkeit von Lernalgorithmen gegenüber langen, händisch

erstellten Regellisten (vgl. Géron, 2018, S. 4f).
Kennzeichnend für Klassifizierungsprobleme ist, dass die Werte der Menge C der Klassen nominal skaliert sind (vgl. Frochte, 2018, S. 22). Vereinfacht gesagt, lassen die Klassen sich nicht ordnen. Im E-Mail Beispiel kann man nicht sagen, dass eine Ham Mail doppelt so gut ist wie eine Spam Mail oder ähnliches. Im Fall der Ampelerkennung, was ebenfalls ein Klassifizierungsproblem darstellt, kann nicht gesagt werden, dass gelbe Ampeln besser als rote oder grüne sind. Die Tatsache, dass man am liebsten grüne Ampeln beim Auto fahren hat, ändert daran nichts. Im Unterschied dazu gibt es so genannte *Regressionsprobleme*. Sie unterscheiden sich von den Klassifikationsproblemen lediglich in der Menge C. Sie besteht hier meist aus stetigen Größen, welche metrisch skaliert sind. Ein Anwendungsbeispiel aus einer der sieben großen Firmen *Amazon, Netflix, Zalando, Allianz, Apple, Google, Facebook*, die am Anfang genannt wurden, wäre die Festlegung eines Versicherungstarifs auf Grundlage verschiedener Kundenmerkmale. Der Zielwert in C wäre dann ein bestimmter Geldbetrag. Bei den Zielwerten kann es sich jedoch auch um diskrete Werte handeln. Wenn man beispielsweise vorhersagen möchte, wie viele Unfälle abhängig von Witterung, Wochentag etc. an einem Tag passieren, so wären die Zielwerte in C diskret. Trotzdem kann man Häufigkeit, Reihenfolge und Abstand messen. 3 Unfälle pro Tag sind halb so viele wie 6, die Zielwerte sind also metrisch skaliert (vgl. Raschka & Mirjalili, 2017, S. 28); (Frochte, 2018, S. 22).

Unüberwachtes Lernen

Die zweite Kategorie von Lernalgorithmen gehören zum *unüberwachten Lernen* (engl. *unsupervised Learning*). Im Unterschied zu überwachten, bekommen unüberwachte Lernalgorithmen keine gelabelten Beispieldatensätze (vgl. Wroble et al., 2013, S. 405). Lernalgorithmen aus diesem Bereich haben das Ziel interessante Strukturen und Muster in einem Datensatz zu finden. Daher werden sie meist in der Wissensentdeckung in Datenbanken, dem sogenannten *Data Mining* eingesetzt (vgl. Wroble et al., 2013, S. 406). Von den eingangs erwähnten Firmen finden sich hier vor allem *Amazon, Netflix* und *Zalando* wieder. Nutzer, beziehungsweise Kunden, sollen anhand ihres Kaufverhaltens in Gruppen eingeteilt werden. Solche Gruppen, auch *Cluster* genannt, benötigen kein Label wie beim überwachten Lernen. Wichtig ist das ähnliche Verhalten (vgl. Frochte, 2018, S. 25). Wie in Abbildung 8 zu sehen ist, ist die Einteilung in die verschiedenen Gruppen keinesfalls eindeutig. Im Vergleich zum überwachten Lernen gibt es keine richtig oder falsch klassifizierten Daten.

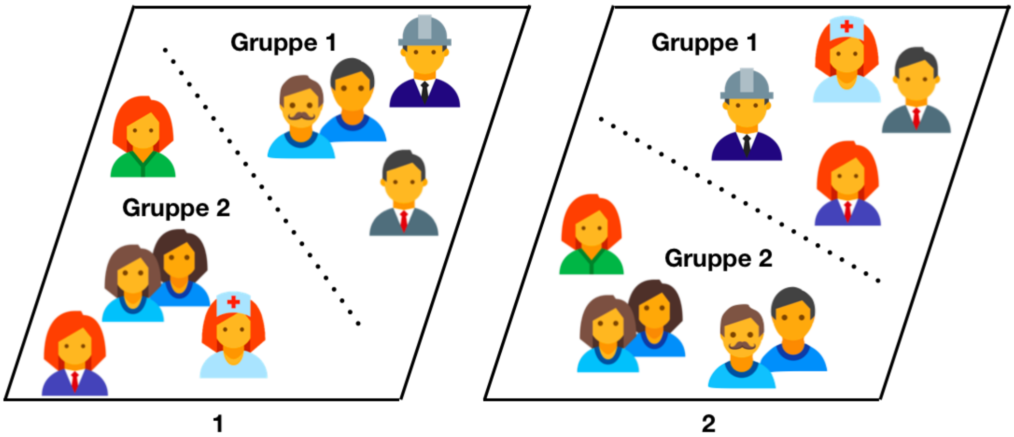


Abb. 8: Beispiel für eine Cluster-Bildung mit unüberwachtem Lernen. Zwei verschiedene Zuordnungen von einer Personengruppe: Zuordnung 1 (links) nach Geschlecht. Zuordnung 2 (rechts) nach Art der Kleidung (Arbeit/Freizeit).

Links, in Teil eins der Abbildung 8, wird der Datensatz aus neun Personen nach Geschlecht in zwei Gruppen geteilt. Rechts, in Teil zwei der Abbildung, werden ebenfalls 2 Gruppen gebildet, wobei hier ein anderes Merkmal, nämlich die Art der Kleidung ausgewählt wurde. Beide Einteilungen erscheinen sinnvoll. Je nachdem, nach welchen Merkmalen die Cluster erstellt wurden, ist beispielsweise die Krankenschwester in Gruppe eins oder in Gruppe zwei. Bei einem Klassifizierungsproblem des überwachten Lernens könnte man vorgeben, dass nach Geschlecht klassifiziert werden soll. Man würde alle Daten vorher labeln und könnte am Ende sagen, dass die richtige Einteilung der Krankenschwester in die Klasse weiblich ist. Das ist beim unüberwachten Lernen nicht der Fall (vgl. Frochte, 2018, S. 25). Dadurch entsteht das Problem, dass man zur Evaluation eines solchen Lernalgorithmus das Ergebnis per Hand prüfen muss. Ein unüberwachter Algorithmus zur Gesichtserkennung könnte beispielsweise alle Bilder, welche ein Lachen aufweisen zusammenfassen. So wurde zwar eine Gruppierung erreicht, diese ist jedoch nicht sinnvoll, wenn das Ziel ist einzelne Personen auf Bildern zu erkennen (vgl. Guido & Müller, 2017, S. 124).

Bestärkendes Lernen

Der letzte Unterbereich von Machine Learning, welcher im Rahmen dieser Arbeit vorgestellt werden soll ist das *bestärkende Lernen* (engl. *reinforcement learning*). Genau wie beim unüberwachten Lernen liegt auch hier ein nicht-gelabelter Datensatz vor. Was bekannt ist, sind wünschenswerte Ergebnisse (vgl. Frochte, 2018, S. 23). Zum besseren Verständnis, soll direkt mit einem Beispiel für bestärkendes Lernen begonnen werden.

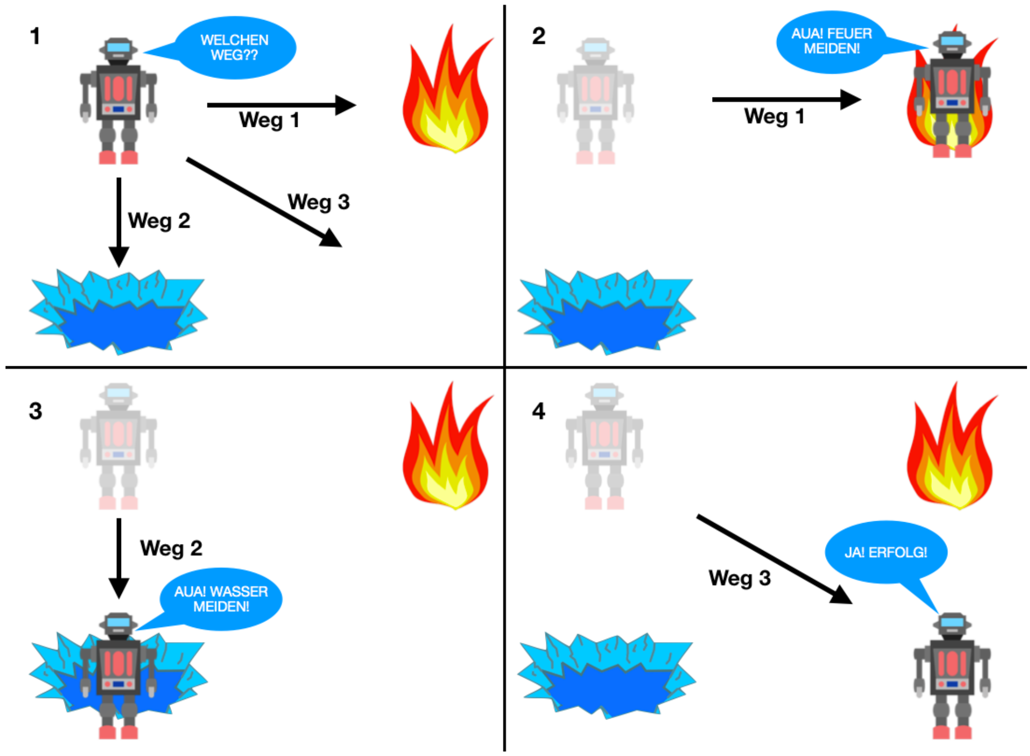


Abb. 9: Beispiel für bestärkendes Lernen. Bild 1: Agent (Roboter) beobachtet die Umgebung: drei mögliche Wege. Bild 2: erster Weg führt zu Bestrafung → Feuer wird gemieden. Bild 3: zweiter Weg auch bestraft → Wasser wird gemieden. Bild 4: dritter Weg hat keine Strafen → wird beibehalten.

Die wünschenswerte Lösung wäre, dass der Roboter unbeschadet und in kürzester Zeit durch den Hindernisparcours in Abbildung 9 kommt. Zunächst beobachtet der Roboter die Umgebung (1). Anschließend probiert er verschiedene Lösungswege aus und erhält dabei kontinuierlich eine Rückmeldung in Form von Belohnung oder Bestrafung (2-3). Die Leistung des Roboters verbessert sich durch Interaktion mit der Umwelt. Mit der Zeit sucht er den optimalen Lösungsweg der Aufgabe, bei der er die meisten Belohnungen erhält, beziehungsweise Strafen vermeidet (4) (vgl. Géron, 2018, S. 13). Auch ein Schachcomputer arbeitet nach diesem Prinzip. Für eine bestimmte Folge von Zügen erhält er eine positive oder negative Rückmeldung. Positive Rückmeldungen erhält er beispielsweise, wenn er Figuren schlägt - negative, wenn seine eigenen geschlagen werden (vgl. Raschka & Mirjalili, 2017, S. 29).
Der Roboter aus den beiden Beispielen wird auch als *Agent* bezeichnet (vgl. Frochte, 2018, S. 23). Für Algorithmen des bestärkende Lernens wird immer ein Agent benötigt. Dieser Agent soll eine bestimmte Aufgabe erfüllen. Der Weg zum Erfüllen der Aufgabe besteht jedoch aus mehreren Schritten, was diese Lernmethode von den anderen beiden unterscheidet (vgl. Raschka & Mirjalili, 2017, S. 30).
Ein aktuelles Beispiel für den Einsatz von Bestärkendem Lernen ist das Projekt *AlphaGo* der Google-Firma *DeepMind*. Das Programm schlug 2017 den Weltmeister Ke Jie im Brettspiel *Go*, indem es Millionen von Spielen analysierte und dann gegen sich

selbst spielte. Durch diese immense Menge an Spielen konnte das Programm einen Erfahrungsschatz aufbauen, welcher die menschlichen Möglichkeiten übersteigt (vgl. Zhou, 2018, S. 5).

Nachdem nun drei verschiedenen Kategorien von Machine Learning vorgestellt und voneinander abgegrenzt wurden, soll im folgenden Abschnitt ein spezieller Lernalgorithmus betrachtet werden. Wie oben bereits erwähnt, wird für den Workshop die *Support Vector Machine*, kurz SVM genutzt. Im Folgenden soll der mathematische Hintergrund dieses Algorithmus beschrieben und seine Funktionsweise erläutert werden. In Abschnitt 3.3 werden Verweise auf den Kernlehrplan im Fach Mathematik für das Land Nordrhein-Westfalen gegeben. Inhaltlich sind die Anknüpfungspunkte im Bereich der Analytischen Geometrie und Linearen Algebra des Kernlehrplans einzuordnen.

3.2. Mathematischer Hintergrund der Support Vector Machine

Die SVM ist ein Lernalgorithmus aus dem Bereich des überwachten Lernens. Sie eignet sich besonders gut zum Klassifizieren kleinerer, bis mittel großer Datensätze (vgl. Géron, 2018, S. 145). SVMs behandeln die Aufgabe eine Entscheidungsfunktion f auf Grundlage von Beispielen zu erlernen, mit der Daten klassifiziert werden können, die nicht im Beispieldatensatz liegen (vgl. Wroble et al., 2013, S. 427). Um die Entscheidungsfunktion f zu finden, sucht die SVM zunächst eine Funktion, welche die verschiedenen Datenklassen voneinander trennt. In Abschnitt 3.1 wurde unter dem Punkt überwachtes Lernen bereits beschrieben wie ein solcher Beispieldatensatz für Klassifizierungsprobleme aussieht. Der Übersicht halber, soll er hier jedoch noch kurz angeführt werden.

Datensatz als Teil des mathematischen Modells
Sei B eine Menge aus m gelabelten Beispieldaten mit

$$B = \{(\vec{x}_1, t_1), (\vec{x}_2, t_2), ..., (\vec{x}_m, t_m)\} \subseteq \mathbb{R}^n \times C \; . \tag{3.4}$$

Die Vektoren $\vec{x}_i \;, i = 1, ...m \in \mathbb{R}^n$ sind die Merkmalsvektoren mit n Merkmalen. Die Werte $t_i \;, i = 1, ...m \in C$ sind die zugehörigen Labels der Merkamlsvektoren. Es gilt $|C| =$ Anzahl der Klassen. Bei binären Klassifikationsproblemen schreibt man $C = \{-1, 1\}$. SVMs können sowohl auf binäre Klassifizierungsprobleme, also solche mit 2 Klassen, als auch auf auf Probleme mit mehreren Klassen angewendet werden (vgl. Abschnitt 3.2.3). Der Datensatz B wird in die Menge D der *Trainingsdaten* und die Menge E der *Testdaten* eingeteilt:

$$D = \{(\vec{x}_1, t_1), (\vec{x}_2, t_2), ..., (\vec{x}_k, t_k)\} \subset B \; , \tag{3.5}$$
$$E = \{(\vec{x}_{k+1}, t_{k+1}), (\vec{x}_{k+2}, t_{k+2}), ..., (\vec{x}_m, t_m)\} \subset B \; . \tag{3.6}$$

Im Workshop sind durch die Darstellung der Verkehrsampeln als Merkmalsvektoren die ersten beiden Schritte gegeben, die die Schüler im Modellierungskreislauf machen.

Zunächst müssen die Schüler Annahmen treffen, welche Farbtöne die Ampeln haben. Die Situation wird vereinfacht, indem nur rote und grüne Ampeln mit zwei Boxen betrachtet werden. Für jede Box wird der entsprechende Farbton als RGB-Wert angegeben. Der RGB-Wert eines Farbtons liegt als Tripel $[r, g, b]$ vor, bei dem r den roten Anteil, g den grünen Anteil und b den blauen Anteil des entsprechenden Farbtons angibt. Ausgehend davon recherchieren die Schüler eine Funktionsgleichung, die aus den RGB-Tripeln einen Grauwerte erzeugt. Mit Hilfe dieser Gleichung kann ein Teil des mathematischen Modells aufgestellt werden: Für jede Box liegt der Grauwert des Farbtons vor. Die beiden Werte können als Einträge in einen 2×1-Vektor geschrieben und somit grafisch dargestellt werden. Der nächste Schritt zur Vervollständigung des mathematischen Modells wäre nun die Funktion zu suche, welche die Vektoren der beiden Klassen mittels SVM voneinander trennt.

3.2.1. Lineare SVM für binäres Klassifizierungsproblem

Um den mathematischen Hintergrund der SVM zu erklären, soll zunächst ein binäres Klassifizierungsproblem betrachtet werden. Des Weiteren werden für den Workshop zunächst nur Daten verwendet, welche *linear separierbar* sind. Das bedeutet, dass die Daten durch eine lineare Funktion getrennt werden können. Im Fall von zweidimensionalen Merkmalsvektoren $\vec{x}_i \in \mathbb{R}^2$ könnten sie durch eine Gerade (siehe Abbildung 10) und im dreidimensionalen Fall durch eine Ebene getrennt werden. Haben die Merkmalsvektoren $n > 3$ Dimensionen, so werden die Daten durch sogenannte *Hyperebenen* getrennt (vgl. Hornsteiner, 2010, S. 19).
Der Übersicht halber soll kurz dargestellt werden, wie der Ablauf des SVM Verfahrens zum Lösen von Klassifizierungsproblemen aussieht:

1. Zunächst wird der gelabelte Beispieldatensatz in Test- und Trainingsdaten aufgeteilt (vgl. Gleichungen 3.4-3.6).

2. In der Trainingsphase wir die optimale Trennfunktion anhand der Trainingsdaten bestimmt.

3. Um zu beurteilen, wie gut die Trennfunktion ist, wird sie in der Testphase genutzt um die Testdaten zu klassifizieren. Da es sich um einen Beispieldatensatz handelt, bei dem die Zuordnung schon bekannt war, kann die Aussage getroffen werden, wie gut die Trennfunktion arbeitet.

An diesem Ablauf orientiert sich die folgende mathematische Beschreibung des SVM Verfahrens. Der zu Grunde liegende Datensatz wurde bereits in den Gleichungen (vgl. 3.4-3.6) beschrieben und getrennt. Es folgt also der zweite Schritt im oben beschriebenen Ablauf.
Die Trennfunktion wird immer anhand der Trainingsdaten aufgestellt. Mathematisch lässt sich das ganze wie folgt ausdrücken (vgl. Bishop, 2006, S. 326):
Für einen linear separierbaren Trainingsdatensatz D gibt es mindestens eine Trennfunktion g mit

$$g: \ \vec{n} * (\vec{x} - \vec{p}) = 0 \ , \tag{3.7}$$

die die Trainingsdatenpunkte fehlerfrei trennt. Dabei bezeichnet $*$ das Skalarpodukt zweier Vektoren, was mit kartesischen Koordinaten wie folgt definiert ist:
Gegeben seien die Vektoren $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2$.
Dann wird das Skalarprodukt $\vec{a} * \vec{b}$ der beiden gegeben durch:

$$\vec{a} \cdot \vec{b} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} * \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1 b_1 + a_2 b_2.$$

Die Vektoren im Workshop sind zwei- beziehungsweise dreidimensional, weshalb eine geometrische Definition sehr anschaulich ist. Geometrisch kann das Skalarprodukt über

$$\vec{a} * \vec{b} = \cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}| \tag{3.8}$$

definiert werden. Dabei meint $\cos(\alpha)$ den Winkel zwischen den beiden Vektoren[6] und $|\cdot|$ den Betrag eines Vektors. Der Betrag eines Vektors $\vec{a} \in \mathbb{R}^n$ ist gegeben durch

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2 + ... + a_n^2}\ . \tag{3.9}$$

Für Vektoren, die im rechten Winkel zueinander stehen, wird der $\cos(\alpha) = 0$ und somit ergibt sich für das Skalarprodukt der beiden ebenfalls 0. Für die mathematische Herleitung der SVM werden im Folgenden vor allem geometrische Begründungen gegeben. Dies soll das Verständnis erleichtern und orientiert sich direkt an der Vorgehensweise im Workshop (vgl. Abschnitt 4.3). Gleichung (3.7) ist die sogenannte *Normalenform* einer Hyperebene. Im zweidimensionalen Fall kann die Gleichung leicht nachvollzogen werden (vgl. Abb. 10).

[6] In höheren Dimensionen $n \geq 4$ ist die geometrische Anschauung des Winkelbegriffs schwierig. Die Definition des Skalarproduktes über den Kosinus lässt sich jedoch ohne weiteres auf die höheren Definitionen übertragen (vgl. Gleichung 3.10).

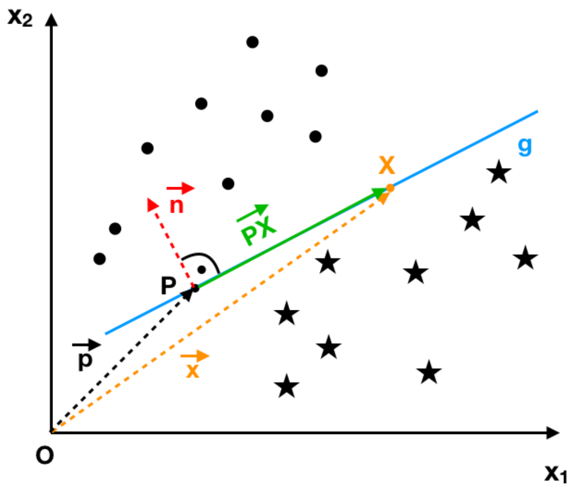


Abb. 10: Skizze zur Normalenform einer Gerade in zwei Dimensionen. Abgebildet sind zwei verschiedene Klassen (Kreis & Stern), die Gerade g (blau) Normalenvektor $\vec{n}$ (rot), Stützvektor der Geraden $\vec{p}$ (schwarz) und Verbindungsvektor $\overrightarrow{PX}$ (grün).

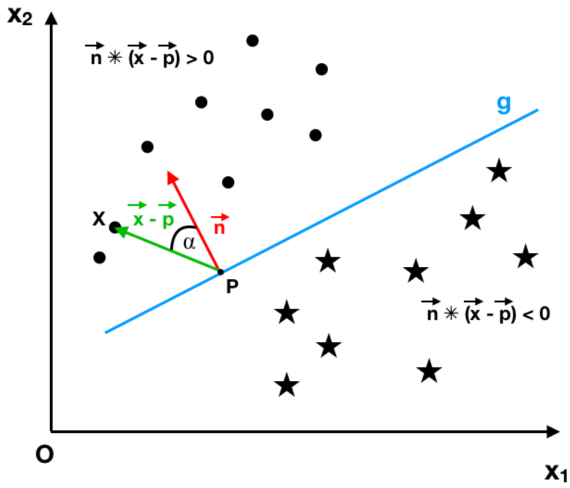


Abb. 11: Unterschiedliche Vorzeichen für $\vec{n} * (\vec{x} - \vec{p})$ für die Trainingsdaten oberhalb und unterhalb der Geraden.

In Abbildung 10 sind zwei verschiedene Klassen von Trainingspunkten dargestellt. Zum einen gibt es die Klasse der Punkte und zum anderen die der Sterne. Die Daten sind offensichtlich linear durch die Gerade g separierbar. Die Gerade g kann durch die Normalenform in Gleichung (3.7) beschrieben werden. Da der Normalenvektor $\vec{n}$ senkrecht auf der Geraden g steht, steht er auch senkrecht auf allen Verbindungsvektoren $\overrightarrow{PX}$ von Punkt P mit Ortsvektor $\vec{p}$ zu Punkten X mit Ortsvektor $\vec{x}$, die auf der Geraden liegen. Wie oben bereits beschrieben, ist das Skalarprodukt für senkrecht stehende Vektoren stets 0. Alle Punkte, die auf der Geraden g liegen erfüllen also die Bedingung $\vec{n} * (\vec{x} - \vec{p}) = 0$.

In der Schulmathematik wird die Normalenform im Zusammenhang mit Ebenen eingeführt (vgl. Griesel et al., 2011, S. 291f). Dort wird die Normalenform genutzt, um die Lagebeziehungen von Geraden und Ebenen zu bestimmen. Das ganze lässt sich für den Einsatz im Workshop nutzen und zunächst um eine Dimension auf die Lage von Punkten und Geraden reduzieren. Für die Schüler, die die Normalenform noch nicht kennen, wird im Workshop eine kurze Einführung gegeben.

Entscheidungsfunktion

In Abbildung 11 wird gezeigt, was mit Punkten passiert, die nicht auf der Geraden g liegen. Dies sind die Trainingsdatenpunkte, für die gilt

$$\vec{n} * (\vec{x} - \vec{p}) \neq 0 \, .$$

Das Skalarprodukt ist wie in Gleichung (3.8) zu sehen über den Winkel α zwischen den zwei Vektoren $\vec{n}$ und $\overrightarrow{PX}$ definiert. Im $\mathbb{R}^n$ verliert diese geometrische Betrachtung über den Winkel an Anschaulichkeit. Da es sich jedoch um euklidische Vektorräume handelt, lässt sich auch für höhere Dimension der Winkel $\angle$ zwischen zwei Vektoren $\vec{a}$ und $\vec{b}$, sowie die Länge $||\cdot||$ eines Vektors $\vec{a}$ definieren:

$$\angle(\vec{a},\vec{b}) = \arccos\frac{\vec{a}*\vec{b}}{|\vec{a}|\cdot|\vec{b}|} \tag{3.10}$$

$$||\vec{a}|| = \sqrt{\vec{a}*\vec{a}}\ . \tag{3.11}$$

Auf diese Weise lässt sich das Verfahren, welches hier detailliert für zwei Dimensionen beschrieben wird unmittelbar auf höhere Dimensionen anwenden. Dieser Schritt wird auch im Workshop praktiziert. Die Schüler starten mit der Bilderkennung im $\mathbb{R}^2$ und wenden das erarbeitete Modell am Ende des Workshops im $\mathbb{R}^{1782}$ an. Der Anschaulichkeit halber, nun jedoch wieder zurück zum zweidimensionalen Fall. Für den Kosinus gilt:

$$\cos(\alpha)\begin{cases} > 0 & \text{für } 0^\circ \leqslant \alpha < 90^\circ \text{ und } 270^\circ < \alpha < 360^\circ \\ = 0 & \text{für } \alpha = 90^\circ \text{ und } \alpha = 270^\circ \\ < 0 & \text{für } 90^\circ < \alpha < 270^\circ\ . \end{cases}$$

Zeigt der Normalenvektor $\vec{n}$, wie in Abbildung 11, nach oben aus der Geraden[7], so gilt für Trainingsdatenpunkte oberhalb der Geraden g

$$\vec{n}*(\vec{x}-\vec{p}) > 0\ , \tag{3.12}$$

da der Winkel α hier immer im Intervall $[0^\circ, 90^\circ)$ oder $(270^\circ, 360^\circ)$ liegt.
Für Trainingsdatenpunkte unterhalb der Geraden g gilt hingegen

$$\vec{n}*(\vec{x}-\vec{p}) < 0\ , \tag{3.13}$$

da der Winkel α hier immer im Intervall $(90^\circ, 270^\circ]$ liegt.
Anhand des Vorzeichens von $\vec{n}*(\vec{x}-\vec{p})$, kann für Merkmalsvektoren $\vec{x}$ eines Datensatzes bestimmt werden, in welcher Klasse sie liegen. Diese Erkenntnis liefert bereits die Funktion f, welche bei einem Klassifizierungsproblem gesucht wird (siehe Abschnitt 3.1). Nach Auflösen der Klammer und Umschreibung von $-\vec{p}*\vec{n} = b$ können wir festhalten:

[7] Sollte er nach unten zeigen, so werden die Ungleichungszeichen in (3.12) und (3.13) umgedreht. In dem Fall werden einfach die Labels für (3.15) anders herum gewählt.

Sei $B \subset \mathbb{R}^n$ ein Datensatz aus zwei Klassen mit den Merkmalsvektoren $\vec{x}$ und $C = \{-1, 1\}$ die Menge der Klassen.
Die Funktion f, mit

$$\begin{aligned} f : B &\to C \\ \vec{x} &\mapsto f(\vec{x}) = \operatorname{sgn}(\vec{n} * \vec{x} + b) \ . \end{aligned} \tag{3.14}$$

wird **Entscheidungsfunktion** genannt. Sie ordnet den Merkmalsvektoren $\vec{x}$ aus B Klasse $f(\vec{x})$ aus C zu.

Dabei ist $\operatorname{sgn}(x)$ die sogenannte Vorzeichenfunktion:

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{für } x < 0 \\ 0 & \text{für } x = 0 \\ +1 & \text{für } x > 0 \end{cases} \ .$$

Ohne Beschränkung der Allgemeinheit können die Labels t_i der Merkmalsvektoren so gewählt werden, dass die Datenpunkte oberhalb der Geraden die Labels $t_i = +1$, und die unterhalb der Geraden die Labels $t_i = -1$ erhalten. Damit lassen sich die beiden Gleichungen (3.12) und (3.13) zusammenfassen zu

$$t_i\,(\vec{n} * \vec{x}_i + b) > 0 \ \forall\ \vec{x}_i \in D \ , \ i = 1, .., k \ . \tag{3.15}$$

Auch hier wurde wieder $-\vec{p} * \vec{n}$ durch b ersetzt. Dies muss für alle Trainingsdatenpunkte des in Gleichung (3.6) beschriebenen Trainingsdatensatzes D gelten, die richtig klassifiziert werden. Wird ein Punkt falsch klassifiziert, so würde sich in Gleichung (3.15) für diesen Punkt ein negatives Vorzeichen ergeben. Fürs Erste stellen wir die Bedingung an die SVM, dass sie eine Trennfunktion sucht, welche Gleichung (3.15) erfüllt:

Bedingung 1:
Sei $D \subset \mathbb{R}^n \times C$ ein Trainingsdatensatz mit k Merkmalsvektoren $\vec{x}_i \in \mathbb{R}^n$ und zugehörigen Labels $t_i \in C = \{-1, 1\}$, $i = 1, ..., k$.
Finde eine affine Trennfunktionen g mit den Parametern $\vec{n}$ und b, welche die Gleichung

$$t_i\,(\vec{n} * \vec{x}_i + b) > 0 \ \forall\ \vec{x}_i \in D \ , \ i = 1, .., k$$

erfüllt, sprich alle Trainingsdatenpunkte richtig eingeordnet werden.

Im Fall der geforderten linear separierbaren Daten, gibt es immer eine solche Trennfunktion g. Solche Datensätze sind jedoch keine Selbstverständlichkeit. Es gibt Datensätze, welche nicht durch eine Gerade getrennt werden können. Ein solcher Fall ist in Abbildung 12 dargestellt.

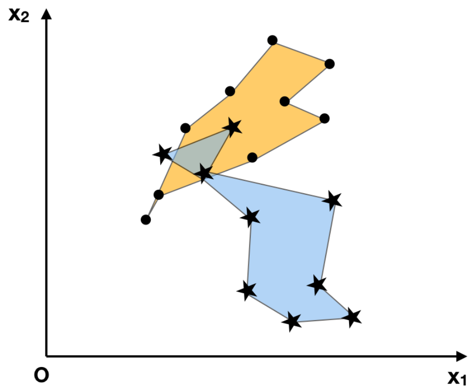


Abb. 12: Zwei verschiedene Klassen (Kreis & Stern). Die in orange eingefärbte Fläche ist der Bereich aus Datenpunkten der Klasse Kreis und die in blau eingefärbte Fläche der Bereich der Klasse Stern. Daten sind nicht linear separierbar.

Wie in Abbildung 12 zu sehen ist, überlappen sich die eingefärbten Flächen der beiden Klassen Stern und Punkt. Dies macht eine Trennung durch eine Gerade unmöglich. Auf einen solchen Fall wird im Abschnitt 3.2.2 eingegangen.
Obwohl die Daten beim Workshop linear separierbar sind, gibt es ein Problem. Es kann viele Trennfunktionen geben, welche Bedingung 1 erfüllen. Es gibt mehrere Kandidaten für die Trennfunktion, welche die SVM sucht.
Dies wird in Abbildung 13 dargestellt. Der Datensatz besteht wieder aus den beiden Klassen der Punkte und der Sterne. Eingezeichnet sind drei lineare Funktionen g_1 (blau), g_2 (grün) und g_3 (orange). Die drei Funktionen trennen den Trainingsdatensatz alle fehlerfrei. Würde man mit den drei Geraden jedoch einen neuen Datensatz klassifizieren, so liegt der Verdacht nah, dass Geraden g_2 und g_3 fehleranfälliger sind.

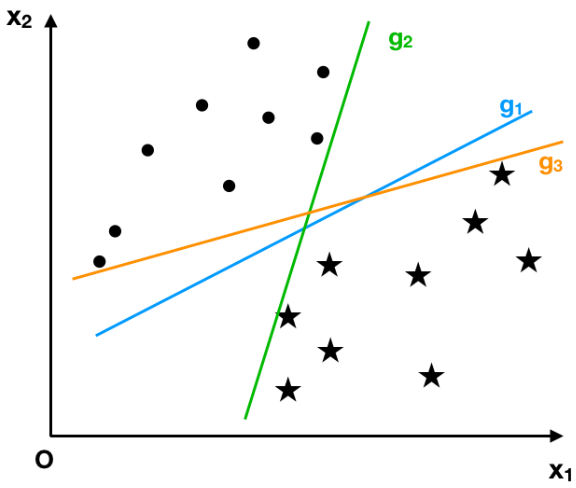


Abb. 13: Trainingsdatensatz aus zwei Klassen: Punkte und Sterne. Mehrere mögliche Trennfunktion g_1 (blau), g_2 (grün) und g_3 (orange) für den Trainingsdatensatz.

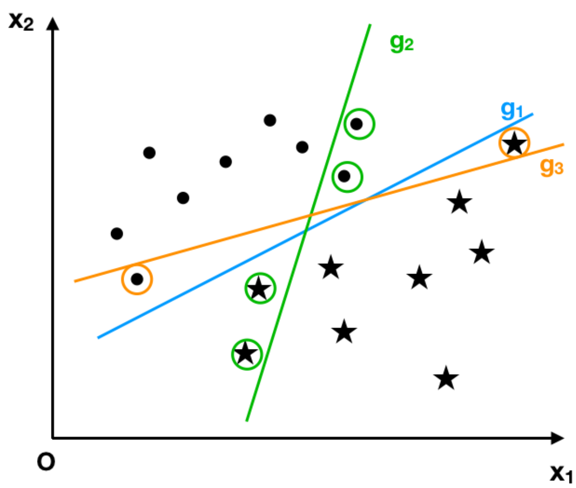


Abb. 14: Testdatensatz: umkreist sind falsch eingeordnete Punkte. Gerade g_2 (grün) ordnet 4 Datenpunkte falsch ein. Gerade g_3 (orange) ordnet 2 Datenpunkte falsch ein. Gerade g_1 (blau) ordnet alle Punkte richtig ein.

Eine solche Situation wird in Abbildung 14 gezeigt. Die Geraden g_2 und g_3 liegen so nah an den Trainingsdaten in Abbildung 13, dass bei leichten Veränderungen im Test Datensatz in Abbildung 14 falsche Zuordnungen auftreten können. Lässt sich eine gefundene Trennfunktion nicht gut auf neue Datensätze übertragen, spricht man auch von einer schlechten *Generalisierbarkeit* (vgl. Whitenack, 2017, S. 75). Die Gerade g_1 hingegen weist einen größeren *Abstand* zu den Trainingspunkten auf und scheint somit besser zum Einordnen neuer Daten geeignet.
Hieraus ergibt sich unmittelbar die zweite Bedingung für die SVM. Sie soll die Trenngerade finden, welche den Datensatz *am besten* trennt. Wie in den Abbildung 13 und 14 verdeutlicht wird, ist die Trenngerade am besten, zu der die *nächstgelegenen* Datenpunkten den *maximalen* Abstand haben. (vgl. Wroble et al., 2013, S. 424). Der Bereich um die Trennfunktion, der frei von Datenpunkten ist, wird auch *Margin* genannt. Die zweite Bedingung an die SVM lautet also:

Bedingung 2:
Finde die Trennfunktion g mit den Parametern $\vec{n}$ und b, die einen maximalen Margin hat.

Diese Bedingung sollen die Schüler während des Workshops selbständig formulieren. Die wörtliche Formulierung stellt den Ausgangspunkt für das Aufstellen der Abstandsformel im nächsten Schritt dar.

Abstandsbegriff

Für Bedingung 1 haben wir mit Gleichung (3.15) bereits einen mathematischen Ausdruck gefunden. Nun soll auch Bedingung 2 für die SVM mathematisiert werden. Dazu ist es notwendig, zunächst eine Formel für den Abstand von Datenpunkten zu einer Trennfunktion aufzustellen.

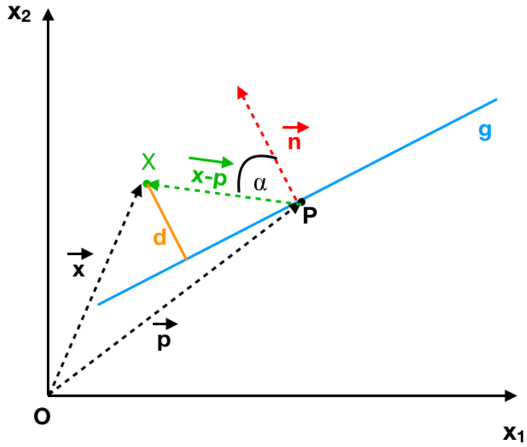


Abb. 15: Skizze zur Berechnung des Abstandes d (orange) eines Punktes X von einer Geraden g (blau). Außerdem dargestellt sind der Normalenvektor $\vec{n}$ (rot), der Stützvektor $\vec{p}$ der Geraden und der Verbindungsvektor $\overrightarrow{PX}$.

Gesucht ist eine Formel für den Abstand d eines Trainingsdatenpunktes $\vec{x}$ von einer Hyperebene. Der Anschaulichkeit halber, wird das Problem in Abbildung 15 in zwei Dimensionen dargestellt. Die trennende Hyperebene ist dann wie schon in Abbildung 10 und 14 durch eine Gerade gegeben.
Es wird deutlich, warum das Skalarprodukt in Gleichung (3.8) über die geometrische Definition eingeführt wurde. Der Winkel α zwischen den Vektoren $\overrightarrow{PX}$ und $\vec{n}$ wird zur geometrischen Herleitung der Abstandsformel benötigt.

<u>1. Fall:</u>
Es soll zunächst der Fall betrachtet werden, dass der Datenpunkt X, wie in Abbildung 15 dargestellt, oberhalb der Trennfunktion g liegt. Außerdem zeigt der Normalenvektor $\vec{n}$ nach oben. In diesem Fall gilt:

$$\cos(\alpha) = \frac{d}{|\overrightarrow{PX}|} \tag{3.16}$$

$$\cos(\alpha) = \frac{\vec{n} * \overrightarrow{PX}}{|\vec{n}| \cdot \left|\overrightarrow{PX}\right|} \tag{3.17}$$

Für die erste Gleichung kann in Abbildung 15 der Wechselwinkel in das rechtwinklige Dreieck mit Hypotenuse $|\overrightarrow{PX}|$ eingezeichnet werden.
Den Kosinus für rechtwinklige Dreiecke kennen die Schüler bereits. Außerdem wird

daran angeknüpft, dass die Schüler den kürzesten Abstand eines Punktes von einer Geraden bereits als Lot kennen (vgl. Ernst Klett Verlag, 2017, S. 80f). Gleichsetzen von (3.16) und (3.17) liefert:

$$\begin{aligned}
\frac{d}{|\overrightarrow{PX}|} &= \frac{\overrightarrow{PX} * \vec{n}}{|\overrightarrow{PX}| \cdot |\vec{n}|} && | \cdot |\overrightarrow{PX}| \\
d &= \frac{\overrightarrow{PX} * \vec{n}}{|\vec{n}|} && | \text{ ersetze } \overrightarrow{PX} = \vec{x} - \vec{p} \\
d &= \frac{\vec{n} * (\vec{x} - \vec{p})}{|\vec{n}|} && (3.18)
\end{aligned}$$

<u>2. Fall:</u>
In diesem Fall werden Punkte X betrachtet, welche unterhalb der Geraden g in Abbildung 15 liegen. Der Normalenvektor $\vec{n}$ zeigt wie in Fall 1 nach oben. Für solche Punkte kann man die gleiche Herleitung verwenden. Dadurch, dass die Winkel α in den cos-Gleichungen (3.16) und (3.17) jedoch nicht mehr identisch sind, muss ein weiterer Schritt mehr gemacht werden.
Für den $\cos(\alpha)$ gilt:

$$\cos(180° - \alpha) = -\cos(-\alpha) = -\cos(\alpha)\ .$$

Damit erhält man in Gleichung (3.18) ein negatives Vorzeichen, weshalb sich für Punkte unterhalb der Trennfunktion der Abstand d zu

$$d = -\frac{\vec{n} * (\vec{x} - \vec{p})}{|\vec{n}|} \tag{3.19}$$

ergeben würde. An dieser Stelle liegen zwei Probleme vor: Erstens ist der Abstandsbegriff positiv definiert und zweitens unterscheiden sich die Formeln (3.18) und (3.19). Es sollte jedoch für alle Datenpunkte eine einheitliche Abstandsformel aufgestellt werden. Um beide Probleme zu lösen, liegt es nahe den Betrag zu verwenden und

$$d = \left| \frac{\vec{n} * (\vec{x} - \vec{p})}{|\vec{n}|} \right| \tag{3.20}$$

zu definieren. Mit Gleichung (3.20) wird der elementargeometrische Abstand definiert. Die Schüler kennen diesen schon als Lotbegriff aus der Schule (vgl. Ernst Klett Verlag, 2017, S. 80f).
Da die Zuordnung der Daten bereits bekannt ist, kann auch diese Information mit einbezogen werden. Wie schon in Ungleichung (3.15) nutzen wir die Labels um eine einheitliche, positive Abstandsformel für alle Trainingsdaten zu erhalten. Trainingsdatenpunkte unterhalb der Trennfunktion haben laut Gleichung (3.19) einen negativen Wert für d und außerdem die Labels $t = -1$ zugewiesen bekommen. Trainingsdatenpunkte oberhalb haben einen positiven Wert für d und die Labels $t = +1$. Anstelle des Betrags in (3.20) können die t_i verwendet werden. Für den Trainingsdatensatz D aus

Gleichung (3.6) mit den Merkmalsvektoren $\vec{x}_i$ und zugehörigen Labels t_i, $i = 1, ..., k$ ergibt sich dann:

$$d = t_i \cdot \frac{\vec{n} * (\vec{x} - \vec{p})}{|\vec{n}|} > 0 \quad \forall \, \vec{x}_i \; \in D \; , \; i = 1, ..., k. \tag{3.21}$$

Wie schon in Gleichung (3.14) können in (3.21) die Klammer im Zähler aufgelöst und $-\vec{p} * \vec{n}$ durch b ersetzt werden. Dann ergibt sich die folgende Ungleichungsbedingung:

$$d = t_i \cdot \frac{\vec{n} * \vec{x}_i + b}{|\vec{n}|} > 0 \quad \forall \, \vec{x}_i \; \in D \; , \; i = 1, ..., k. \tag{3.22}$$

Mit der Aufstellung des Abstandsbegriffs, welche im Workshop analog zur Herleitung hier erfolgt, machen die Schüler einen wichtigen Schritt, um ihr mathematisches Modell zu vervollständigen. Wie in Abschnitt 4.3 erklärt werden wird, haben die Schüler Ampeln mit einer Art Black-Box (vgl. Greefrath & Siller, 2018, S. 7) eingeordnet, bevor sie eine Formel für den Abstand hergeleitet haben. Indem sie dem Computer jetzt eine Formel vorgeben wurde ein Schritt gemacht, um die Black-Box ein Stück weit aufzulösen.

Margin maximieren

In Bedingung 2 haben wir formuliert, dass der Margin maximiert werden soll. Die Breite des Margins ist durch den doppelten Abstand d zu den nächst gelegenen Datenpunkten gegeben, wie in Abbildung 16 verdeutlicht werden soll.

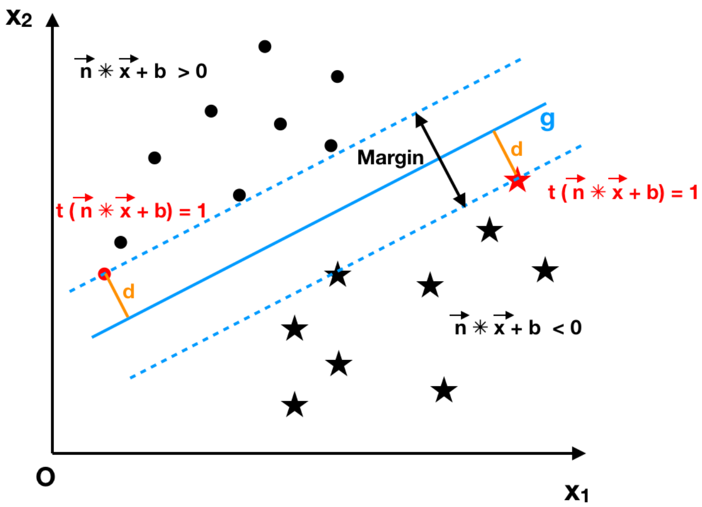


Abb. 16: Trainingsdatensatz aus zwei Klassen: Punkte und Sterne. Trennfunktion g (blau) mit *Margin*. Breite der Margin gegeben durch $2 \cdot d$ (orange).

In Abbildung 16 ist zu erkennen, dass die Breite des Margins gegeben ist durch

$$2 \cdot d = 2 \cdot \min_i t_i \cdot \frac{\vec{n} * \vec{x}_i + b}{|\vec{n}|} \; . \tag{3.23}$$

Für die beste Trennfunktion, soll diese Breite nun maximiert werden. Die Parameter, welche dazu variiert werden können sind $\vec{n}$ und b. Damit ergibt sich

$$\max_{\vec{n},b}\left[2\cdot\min_i t_i\cdot\frac{\vec{n}*\vec{x}_i+b}{|\vec{n}|}\right]$$
$$\text{unter den Nebenbedingunen } t_i\left(\vec{n}*\vec{x}_i+b\right)>0 \qquad \forall\ \vec{x}_i\in D\ ,\ i=1,..,k\ . \tag{3.24}$$

Für die Berechnung des Minimums spielt $|\vec{n}|$ im Nenner keine Rolle, weshalb (3.24) umgeschrieben werden kann zu

$$\max_{\vec{n},b}\left[2\cdot\frac{1}{|\vec{n}|}\cdot\min_i t_i\cdot\left(\vec{n}*\vec{x}_i+b\right)\right]$$
$$\text{unter den Nebenbedingunen } t_i\left(\vec{n}*\vec{x}_i+b\right)>0 \qquad \forall\ \vec{x}_i\in D\ ,\ i=1,..,k\ . \tag{3.25}$$

Die Nebenbedingungen stammen aus Bedingung 1 (3.15) und geben an, dass alle Trainingsdatenpunkte richtig eingeordnet werden müssen. Der Ausdruck in Gleichung (3.25) lässt sich durch Ausnutzen der Nebenbedingung vereinfachen. Es gilt

$$\begin{aligned}
& & & t_i\left(\vec{n}*\vec{x}_i+b\right)>0 & & \\
\Leftrightarrow & & \exists\,\beta>0\ : \quad & t_i\left(\vec{n}*\vec{x}_i+b\right)\geq\beta & & |\div\beta \\
\Leftrightarrow & & & \frac{t_i\left(\vec{n}*\vec{x}_i+b\right)}{\beta}\geq 1 & & \text{mit } \vec{n}/\beta=\vec{n}',\ b/\beta=b' \\
\Leftrightarrow & & & t_i\left(\vec{n}'*\vec{x}_i+b'\right)\geq 1. & &
\end{aligned} \tag{3.26}$$

$\vec{n}$ und b können also einfach beliebig mit einem $\beta>0$ skaliert werden, ohne dass die dadurch beschriebene Hyperebene geändert wird (vgl. Süße & Rodner, 2014, S. 475); (Hamel, 2011, S. 80). Es gilt

$$t_i\left(\vec{n}*\vec{x}_i+b\right)\geq 1\ \forall\ \vec{x}_i\in D\ ,\ i=1,..,k\ . \tag{3.27}$$

Für die nächst gelegenen Datenpunkte kann dann

$$t_i\cdot\left(\vec{n}*\vec{x}_i+b\right)=1 \tag{3.28}$$

geschrieben werden. Dies wird in Abbildung 16 verdeutlicht. Diese nächst gelegenen Datenpunkte werden im hinteren Teil von Gleichung (3.25) gesucht. Damit vereinfacht sich der Ausdruck in (3.25)

$$\max_{\vec{n},b}\left[2\cdot\frac{1}{|\vec{n}|}\cdot\underbrace{\min_i t_i\cdot\vec{n}*\vec{x}_i+b}_{=1}\right]$$
$$\text{unter den Nebenbedingunen } t_i\left(\vec{n}*\vec{x}_i+b\right)\geq 1 \qquad \forall\ \vec{x}_i\in D\ ,\ i=1,..,k$$

zu

$$\max_{\vec{n}} \left[2 \cdot \frac{1}{|\vec{n}|} \right]$$

$$\text{unter den Nebenbedingungen } t_i \left(\vec{n} * \vec{x}_i + b \right) \geq 1 \qquad \forall \; \vec{x}_i \in D \;,\; i = 1,..,k \;. \tag{3.29}$$

Den mathematischen Teil der Margin-Maximierung, werden die Schüler im Workshop nicht bearbeiten. Sie sollen lediglich in Worten formulieren, dass der Computer die Gerade mit maximalem Margin bestimmen soll.

Das primale Optimierungsproblem

Anstatt in (3.29) das Maximum zu suchen, kann auch das Minimum des Kehrwertes bestimmt werden (vgl. Hornsteiner, 2010, S. 22). Es ergibt sich

$$\min_{\vec{n}} \frac{1}{2} \cdot |\vec{n}|$$

$$\text{unter den Nebenbedingungen } t_i \left(\vec{n} * \vec{x}_i + b \right) \geq 1 \qquad \forall \; \vec{x}_i \in D \;,\; i = 1,..,k \;. \tag{3.30}$$

Da die Werte für $\vec{n}$, die $|\vec{n}|$ minimieren, auch das Quadrat minimieren, kann (3.30) umgeschrieben werden zu

$$\min_{\vec{n}} \frac{1}{2} \cdot |\vec{n}|^2$$

$$\text{unter den Nebenbedingungen } t_i \left(\vec{n} * \vec{x}_i + b \right) \geq 1 \qquad \forall \; \vec{x}_i \in D \;,\; i = 1,..,k \;. \tag{3.31}$$

Der Vorteil ist, dass $\frac{1}{2} \cdot |\vec{n}|^2$ überall differenzierbar ist, während $\frac{1}{2} \cdot |\vec{n}|$ in 0 nicht differenzierbar ist. Optimierungsalgorithmen, wie die SVM, arbeiten jedoch besser mit differenzierbaren Funktionen (vgl. Géron, 2018, S. 158); (Hamel, 2011, S. 81). Außerdem liegt mit Gleichung (3.31) ein sogenanntes *konvexes* quadratisches Optimierungsprogramm mit Nebenbedingungen in primaler Form vor. Genau genommen handelt es sich um eine vereinfachte Version, welche wie folgt beschrieben werden kann:

Konvexes Quadratisches Optimierungsprogramm (vereinfacht)
Sei $\vec{p} \in \mathbb{R}^{n_p}$ ein n_p-dimensionaler Vektor, welcher die zu optimierenden Parameter beinhaltet. Die Anzahl der Nebenbedingungen sei durch n_c gegeben. Ferner sei $\vec{b} \in R^{n_c}$ ein n_c-dimensionaler Vektor und $A \in \mathbb{R}^{n_c \times n_p}$ eine $n_c \times n_p$-Matrix.
Dann wird durch

$$\min_{\vec{p}} \frac{1}{2} \vec{p}^T \cdot \vec{p} \tag{3.32}$$

$$\text{mit Nebenbedingungen} \qquad A \cdot \vec{p} \leq \vec{b} \tag{3.33}$$

ein konvexes *quadratisches Optimierungsprogramm* (QP) mit Nebenbedingungen definiert.

Für die Nebenbedingungen $A \cdot \vec{p} \leq b$ in der obigen Definition kann auch

$$\vec{p} * \vec{a}_i \leq b_i \tag{3.34}$$

geschrieben werden, wobei $\vec{a}_i$ die i-te Zeile der Matrix A und b_i das i-te Element von $\vec{b}$ repräsentiert. In der Hochschulmathematik würde man nun sogenannte *Lagrange Multiplikatoren* einführen um das Programm zu lösen. Dies wird in diesem Abschnitt ab Gleichung (3.41) näher erläutert.
Das in den Formeln (3.32) und (3.33) beschriebene konvexe QP ist ein Spezialfall. Im Allgemeinen sind quadratische Optimierungsprobleme komplexer definiert. Dies wird in den Formulierungen (3.38) und (3.39) beschrieben. An der Stelle wird auch erklärt, warum das vereinfachte Problem bereits ein Konvexes ist.
Die obige vereinfachte Beschreibung wird nun auf unseren Fall übertragen. Dieser Schritt soll zeigen, dass es sich in 3.31 um ein QP handelt.

Tab. 1: Übertragen der Definition (linke Spalte) auf unsere Situation (rechte Spalte).

Definition	Situation
n_p	n Dimensionen der Merkamlsvektoren
n_c	k Anzahl der Trainingsdatenpunkte
$\vec{p}$	$\vec{n}$
b	Vektor der Dimension k mit Einträgen $-(1 - t_i \cdot b)$
$\vec{a}_i$	$-t_i \cdot \vec{x}_i$

Damit ergibt sich Formel (3.32) zu

$$\begin{aligned} & \min_{\vec{n}} \frac{1}{2} \vec{n}^T \cdot \vec{n} \\ \Leftrightarrow \qquad & \min_{\vec{n}} \frac{1}{2} \left| \vec{n} \right|^2 \end{aligned} \tag{3.35}$$

Nun muss nur noch die Gleichung (3.33) für die Nebenbedingungen übertragen werden:

$$\begin{aligned} & t_i \left(\vec{n} * \vec{x}_i + b \right) \geq 1 && \forall \ \vec{x}_i \in D \ , \ i = 1, .., k \\ \Leftrightarrow \qquad & t_i \left(\vec{n} * \vec{x}_i \right) \geq 1 - t_i \cdot b && \forall \ \vec{x}_i \in D \ , \ i = 1, .., k \\ \Leftrightarrow \qquad & \underbrace{-t_i \left(\vec{n} * \vec{x}_i \right)}_{\vec{p} \ * \ \vec{a}_i} \leq \underbrace{-(1 - t_i \cdot b)}_{b_i} && \forall \ \vec{x}_i \in D \ , \ i = 1, .., k \end{aligned} \tag{3.36}$$

Insgesamt ergibt sich damit, dass die beiden aufgestellten Bedingungen 1 & 2 mathematisch als primales Optimierungsproblem ausgedrückt werden können:

Gegeben sei $D \subset \mathbb{R}^n \times C$ ein Trainingsdatensatz aus k Merkmalsvektoren $\vec{x}_i \in \mathbb{R}^n$, $i = 1, ..., k$ und den zugehörigen Labels $t_i \in C = \{-1, 1\}$. Dann werden die Parameter $b \in \mathbb{R}$ und $\vec{n} \in \mathbb{R}^n$ gesucht, sodass

$$\begin{aligned} & \min_{\vec{n} \in \mathbb{R}^n} \frac{1}{2} |\vec{n}|^2 \\ \text{unter den Nebenbedingungen} \quad & t_i (\vec{n} * \vec{x}_i + b) - 1 \geq 0, \quad i = 1, .., k \end{aligned} \tag{3.37}$$

In den Nebenbedingungen wurde die 1 auf die linke Seite der Ungleichung gebracht. Dieser Schritt wird später für Gleichung (3.41) benötigt. Ein solches Optimierungsproblem erinnert an eine Extremwertaufgabe mit Nebenbedingung aus der Schulmathematik. Ein entscheidender Unterschied besteht jedoch darin, dass in (3.37) die Nebenbedingung nicht durch eine Gleichung, sondern durch eine Ungleichung gegeben ist. Außerdem werden in der Schule lediglich Funktionen einer Veränderlichen betrachtet. Das Lösen des Problems ist hier komplizierter und geht über die in der Schule behandelten Aspekte hinaus. Im Workshop erhalten die leistungsstärkeren Schüler die Möglichkeit, das primale Problem im Rahmen einer Bonusaufgabe zu formulieren. Dabei wird jedoch auf die mathematische Darstellung der Nebenbedingungen verzichtet.

Nun soll noch kurz betrachtet werden, warum durch die Formulierungen (3.32) und (3.33) ein vereinfachtes QP gegeben ist. Ein QP wird normalerweise durch

$$\min_{\vec{p}} \frac{1}{2} \vec{p}^T \cdot H \cdot \vec{p} + \vec{f}^{\,T} \cdot \vec{p} \tag{3.38}$$

$$\text{mit Nebenbedingungen} \quad A \cdot \vec{p} \leq b \tag{3.39}$$

beschrieben. Im Vergleich zum vereinfachten Problem beinhaltet es noch die Matrix $H \in R^{n_p \times n_p}$ und den Vektor $\vec{f} \in \mathbb{R}^{n_p}$. Im vereinfachten Problem wurde $\vec{f}$ durch den Nullvektor und H durch die Identitätsmatrix gegeben. An dieser Stelle wird ersichtlich, warum es sich bei dem vereinfachten QP um ein konvexes Optimierungsproblem handelt. Für solche Probleme muss die Matrix H positiv semidefinit sein (vgl. Alt & Seydenschwanz, 2013, S. 23). Das bedeutet, dass

$$\vec{p}^T \cdot H \cdot \vec{p} \geq 0 \qquad \forall \vec{p} \in \mathbb{R}^{n_p} \tag{3.40}$$

gelten muss. In unserem Fall ist H durch die Identitätsmatrix, und $\vec{p}$ durch $\vec{n}$ gegeben, also

$$\begin{aligned} & \vec{n}^T \cdot \mathbf{Id} \cdot \vec{n} \geq 0 \\ \Leftrightarrow \quad & \vec{n}^T \vec{n} \geq 0 \end{aligned}$$

was eine wahre Aussage darstellt. Damit wurde gezeigt, dass es sich bei den beiden Bedingungen 1 und 2 an die SVM um ein konvexes QP handelt (vgl. Géron, 2018, S. 159). Eine wichtige Eigenschaft eines konvexen QP ist, dass jede lokale Lösung auch eine globale Lösung ist (vgl. Alt & Seydenschwanz, 2013, S. 15f); (Bishop, 2006, S.

325).
Für einen kleinen Datensatz, wie er zu Beginn des Workshops betrachtet wird, ist das Problem (3.37) direkt lösbar. Die Anwendung der Gesichtserkennung stellt für das primale Problem jedoch einen zu hohen Rechenaufwand dar. Bei dem Gesichter-Datensatz beinhalten die Merkmalsvektoren mehr Einträge als bei dem Datensatz aus Verkehrsampeln. Genau genommen wird das Bild eines Gesichtes aus $n \times n$ Pixeln in einen $n^2 \times 1$-Merkmalsvektor überführt. Die Einträge des Vektors werden, analog zu den Verkehrsampeln, durch den Grauwert des entsprechenden Pixels gegeben. Ein Bild in der Auflösung 32×32 hätte dementsprechend schon einen 1024×1-Merkmalsvektor. Damit übersteigt die Dimension der Merkmalsvektoren die Anzahl der Trainingsdatenpunkte. In einem solchen Fall ist es von Vorteil, anstatt des *primalen* Optimierungsproblems, das zugehörige *duale* Optimierungsproblem zu lösen (vgl. Géron, 2018, S. 160). Außerdem ermöglichen duale Optimierungsprobleme den sogenannten *Kernel-Trick*. Dieser ermöglicht es Trainingsdaten, welche in ihrem ursprünglichen Raum nicht linear separierbar sind, in eine höhere Dimension zu transformieren, in der sie dann linear separierbar sind. Dabei ist es jedoch weder nötig die Formel für die Transformation anzugeben, noch die einzelnen Transformationen explizit zu berechnen (vgl. Süße & Rodner, 2014, S. 480). Da der Kernel-Trick für das Lernmodul jedoch keine Rolle spielt wird er hier nicht weiter erläutert. Der interessierte Leser wird an dieser Stelle auf die Masterarbeit von Schönbrodt verwiesen (vgl. Schönbrodt, 2018).

Das duale Problem und die Lagrange-Schreibweise

Das duale Problem wird während des Workshops zwar nicht von den Schülern formuliert, es steckt jedoch in den Algorithmen, mit denen sie die Trennfunktionen für $n \geq 3$-dimensionale Daten suchen. Aus diesem Grund soll das duale Problem hier kurz vorgestellt und in seiner Schreibweise erläutert werden.
Das duale Problem wird über sogenannte *Lagrange Multiplikatoren* $a_i \in \mathbb{R}^{\geq 0}$ eingeführt und das primale Problem (3.37) mit den Lagrange Multiplikatoren ausgerückt.

$$\mathcal{L}(\vec{n}, b, \vec{a}) = \frac{1}{2}|\vec{n}|^2 - \sum_{i=1}^{k} a_i \left[t_i \left(\vec{n} * \vec{x}_i + b\right) - 1\right] \ . \tag{3.41}$$

Dabei steht $\vec{a}$ für den $k \times 1$-Vektor aus den Lagrange Multiplikatoren a_i. Jede der k Nebenbedingungen aus dem QP in (3.37) hat ein zugehöriges a_i. Die durch $\mathcal{L}$ definierte Funktion wird als *Lagrange Funktion* bezeichnet. Für das Optimierungsproblem muss nun $\mathcal{L}(\vec{n}, b, \vec{a})$ bezüglich $\vec{a}$ maximiert und bezüglich $\vec{n}$ und b minimiert werden. Anschaulich wird also ein Maximum oder ein Sattelpunkt der Lagrange Funktion bestimmt, bei dem die Ableitungen nach $\vec{n}$ und b verschwinden (vgl. Schölkopf & Smola, 2001, S. 12); (Frochte, 2018, S. 290f). Der hintere Teil der Lagrange Funktion gibt die Nebenbedingung aus (3.37) wieder. Die Nebenbedingungen liegen bei Optimierungsproblemen stets als Nullgleichungen bzw. Ungleichungen vor.
Würde es sich bei dem Optimierungsproblem um eine Extremwertaufgabe aus der Schulmathematik handeln, so würde man jetzt die Stelle der Funktion suchen, bei der die erste Ableitung verschwindet. Bei dem Schulproblem liegt die Nebenbedingung jedoch als Gleichung und nicht wie in unserem Fall, als Ungleichung vor. In einem solchen

Fall müssen die sogenannten *Karush-Kuhn-Tucker*-Bedingungen (KKT-Bedingungen) einbezogen werden (Frochte, 2018, S. 290f):

$$a_i \geq 0 \qquad i = 1, .., k \tag{3.42}$$

$$t_i\left(\vec{n} * \vec{x}_i + b\right) - 1 \geq 0, \qquad i = 1, .., k \tag{3.43}$$

$$\text{und} \quad a_i \cdot \left[t_i\left(\vec{n} * \vec{x}_i + b\right) - 1\right] = 0, \qquad i = 1, .., k \tag{3.44}$$

Bedingung (3.42) lieferte schon die Definition der Lagrange Multiplikatoren. Ungleichung (3.43) stimmt mit der Nebenbedingung (3.37) des primalen Problems überein. Gleichung (3.44) wird auch als *Komplementäritätsbedingung* bezeichnet. Aus ihr geht hervor, dass für alle Trainingsdatenpunkte $\vec{x}_i$ entweder $a_i = 0$, oder $t_i \cdot (\vec{n} * \vec{x}_i + b) = 1$ gelten muss. In Gleichung (3.28) und zugehöriger Abbildung 16, wurde begründet, dass für die Datenpunkte, welche den Margin begrenzen gerade $t_i \cdot (\vec{n} * \vec{x}_i + b) = 1$ gilt. Diese Punkte werden auch *Stützvektoren* (engl. *support vectors*) genannt und geben dem Lernalgorithmus seinen Namen. Die Menge der Stützvektoren wird mit S bezeichnet. Für alle anderen Punkte $\vec{x}_i \notin S$ ist $t_i\left(\vec{n} * \vec{x}_i + b\right) > 1$, sodass nach Bedingung (3.44) für sie $a_i = 0$ gelten muss (vgl. Frochte, 2018, S. 290f).

Wie beim Extremwertproblem aus der Schulmathematik, müssen die Stellen gesucht werden, bei denen die Ableitung der Lagrangefunktion unter den KKT- Nebenbedingungen verschwindet. Da die Funktion $\mathcal{L}\left(\vec{n}, b, \vec{a}\right)$ jedoch von mehreren Parametern abhängt, müssen partielle Ableitungen verwendet werden (vgl. Schölkopf & Smola, 2001, S. 12):

$$\frac{\partial}{\partial \vec{n}} \mathcal{L}\left(\vec{n}, b, \vec{a}\right) = 0 \qquad \Rightarrow \vec{n} = \sum_{i=1}^{k} a_i t_i \vec{x}_i \tag{3.45}$$

$$\frac{\partial}{\partial b} \mathcal{L}\left(\vec{n}, b, \vec{a}\right) = 0 \qquad \Rightarrow 0 = \sum_{i=1}^{k} a_i t_i \tag{3.46}$$

Unter der Verwendung von (3.45) und (3.46) kann die Lagrangefunktion (3.41) nach diversen Umformungsschritten, die in (Schönbrodt, 2018) aufgeführt sind, in die duale Form $\mathcal{L}_d$ gebracht werden:

$$\mathcal{L}_d\left(\vec{a}\right) = \sum_{i=1}^{k} a_i - \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} a_i a_j t_i t_j \left(\vec{x}_j * \vec{x}_i\right). \tag{3.47}$$

Also kann das primale Problem (3.37) umgeschrieben werden zu

Gegeben sei $D \subset \mathbb{R}^n \times C$ ein Trainingsdatensatz aus k Merkmalsvektoren $\vec{x}_i \in \mathbb{R}^n$, $i = 1, ..., k$ und den zugehörigen Labels $t_i \in C = \{-1, 1\}$. Dann werden die Lagrange Multiplikatoren $a_i \in \mathbb{R}^{\geq 0}$ gesucht, sodass

$$\begin{aligned} & \max_{a_i \in \mathbb{R}^{\geq 0}} \mathcal{L}_d(\vec{a}) = \sum_{i=1}^{k} a_i - \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} a_i a_j t_i t_j (\vec{x}_j * \vec{x}_i) \\ \text{unter den Nebenbedingungen} \quad & a_i \geq 0, \quad i = 1, .., k \\ \text{und} \quad & \sum_{i=1}^{k} a_i t_i = 0 \ . \end{aligned} \tag{3.48}$$

Berechnung von $\vec{n}$

Im Gegensatz zum primalen Problem (3.37) werden nicht direkt die optimalen Parameter für den Normalenvektor $\vec{n}$, sondern die optimalen Lagrange Multiplikatoren a_i gesucht. Über Gleichung (3.45) kann $\vec{n}$ jedoch direkt berechnet werden:

$$\vec{n} = \sum_{i=1}^{k} a_i t_i \vec{x}_i \ .$$

Die Komplimentäritätsbedingung (3.44) gibt außerdem vor, dass nicht alle Trainingsdatenpunkte $\vec{x}_i$, sondern lediglich die Stützvektoren $\vec{x}_i \in S$ in diese Summe eingehen. Somit können wir die Gleichung umschreiben zu

$$\vec{n} = \sum_{\vec{x}_i \in S} a_i t_i \vec{x}_i \ . \tag{3.49}$$

An dieser Stelle wird deutlich, warum das duale Problem von Vorteil ist, wenn große Datensätze oder mehr Merkmale als Trainingsdaten gibt. Für $\vec{n}$ müssen, wie in Gleichung (3.49) beschrieben, Linearkombinationen aus den Stützvektoren gebildet werden. Ist die Dimension von $\vec{n}$ größer als die Anzahl der Trainingspunkte k, so erspart man sich mindestens[8] $\dim(\vec{n}) - k$ Schritte. Außerdem müssen nicht alle Trainingspunkte, sondern lediglich die Stützvektoren betrachtet werden.

Berechnung von b

Um die Lagrangefunktion im dualen Problem (3.48) zu maximieren kann beispielsweise die Lagrange-Newton Iteration genutzt werden (vgl. Wroble et al., 2013, S. 430). Wenn die Funktion maximiert wurde, kann der optimale Normalenvektor über Gleichung (3.49) berechnet werden. Es fehlt lediglich noch eine Formel zur Berechnung von b. Dafür kann die Eigenschaft genutzt werden, dass für alle Stützvektoren

$$t_i \cdot (\vec{n} * \vec{x}_i + b) = 1 \tag{3.50}$$

gilt. Da $\vec{n}$ bereits berechnet wurde, könnte in (3.50) ein beliebiger Stützvektor für $\vec{x}_i$ eingesetzt und die Gleichung nach b aufgelöst werden. Ein potentieller Fehler, den man

[8] Mindestens tritt nur dann ein, wenn alle Trainingsdatenpunkte in der Menge S lägen, was bei realen Daten sehr unrealistisch ist.

bei der Berechnung von $\vec{n}$ gemacht hat, würde sich bei diesem Vorgehen jedoch direkt auf b übertragen (vgl. Frochte, 2018, S. 291). Ein besserer Weg zur Berechnung von b ist es, über alle Stützvektoren zu mitteln. Dafür wird Gleichung (3.50) noch umgestellt zu

$$\begin{aligned}
& t_i \cdot (\vec{n} * \vec{x}_i + b) = 1 && \cdot t_i \\
\Leftrightarrow\quad & t_i^2 \cdot (\vec{n} * \vec{x}_i + b) = t_i && \mid t_i^2 = 1\ ;\ -t_i \\
\Leftrightarrow\quad & 1 \cdot (\vec{n} * \vec{x}_i + b) - t_i = 0 && -b\ ;\ \cdot(-1) \\
\Leftrightarrow\quad & -(\vec{n} * \vec{x}_i) + t_i = b && .
\end{aligned} \tag{3.51}$$

Setzt man für $\vec{n}$ die Linearkombination der Stützvektoren aus Gleichung (3.49) ein und mittelt über alle s Stützvektoren in S, so erhält man

$$b = \frac{1}{s} \cdot \sum_{\vec{x}_i \in S} \left(t_i - \sum_{\vec{x}_j \in S} a_j t_j \vec{x}_j * \vec{x}_i \right) \ . \tag{3.52}$$

Resultat

Nach den umfangreichen mathematischen Schritten der letzten Seiten gerät schnell das eigentliche Ziel in Vergessenheit. Es sollten optimale Werte für die Parameter $\vec{n}$ und b gefunden werden. Die optimierten Parameter beschreiben die Funktion, welche einen zu Grunde liegenden Trainingsdatensatz am besten trennt und mit der auch unbekannte Datenpunkte klassifiziert werden können. Mit den Ausdrücken (3.49) und (3.52) haben wir die optimalen Lösungen für die Parameter $\vec{n}$ und b gefunden, sodass die Entscheidungsfunktion aus (3.14) angepasst werden kann:

Sei $B \subset \mathbb{R}^n \times C$ ein Datensatz mit den Merkmalsvektoren $\vec{x}_i \in \mathbb{R}^n$ und $C = \{-1, 1\}$ die Menge der Klassen. Außerdem sei $D \subset B$ der Trainingsdatensatz aus k Merkmalsvektoren $\vec{x}_i \in \mathbb{R}^n\ ,\ i = 1, ..., k$ und den zugehörigen Labels $t_i \in C = \{-1, 1\}$. Sei S die Menge der s Stützvektoren. Ferner seien $a_i \in \mathbb{R}^{\geq 0}$ die optimierten Lagrange Multiplikatoren.
Die Funktion f, mit

$$\begin{aligned}
& f : B \to C \\
& \vec{x} \mapsto f(\vec{x}) = \operatorname{sgn}\left(\underbrace{\sum_{\vec{x}_i \in S} a_i t_i \vec{x}_i}_{\vec{n}} * \vec{x} + \underbrace{\frac{1}{s} \cdot \sum_{\vec{x}_i \in S} t_i - \sum_{\vec{x}_j \in S} a_j t_j \vec{x}_j * \vec{x}_i}_{b} \right) \ .
\end{aligned} \tag{3.53}$$

wird **Entscheidungsfunktion** genannt. Sie ordnet jedem Merkmalsvektor $\vec{x}$ aus B eine Klasse $f(\vec{x})$ aus C zu.

Im CAMMP day ist der Datensatz zunächst durch $B \subset \mathbb{R}^2 \times \{-1, 1\}$ und den Merkmalsvektoren $\vec{x} \in R^2$ gegeben. Die Merkmalsvektoren sind Teil des mathematische Modells und symbolisieren rote und grüne Verkehrsampeln. Es wird zunächst angenommen, dass es nur diese beiden Ampelphasen gibt. Für rote Ampeln ist der x_1-Wert des Merkmalsvektors der Grauwert für einen roten und der x_2-Wert der Grauwert für einen schwarzen Farbton. Bei grünen Ampeln hingegen ist der x_1-Wert des Merkmalsvektors der Grauwert für einen schwarzen und der x_2-Wert der für einen grünen Farbton. Den roten Ampeln werden dann die Label $t = -1$ und den grünen die Labels $t = +1$ zugewiesen. Die Menge C ist somit gegeben durch $C = \{-1, 1\}$. Der gesamte Datensatz $B \subset \mathbb{R}^2 \times C$ wird in Test- und Trainingsdaten unterteilt. Mit der Menge $D \subset B$ der Trainingsdatenpunkte wird die Entscheidungsfunktion (3.53) mittels SVM bestimmt. Die Entscheidungsfunktion wird durch den Normalenvektor $\vec{n}$ und den Wert $b = -\vec{n} * \vec{p}$ beschrieben. Mit der Entscheidungsfunktion können die Testdatenpunkte eingeordnet und die Güte bestimmt werden. Außerdem kann die Entscheidungsfunktion zum Klassifizieren unbekannter, neuer Daten genutzt werden.

3.2.2. Schlupfvariablen für nicht-linear separierbare Daten

Für linear separierbare Daten, wie sie in Abschnitt 3.2.1 zu Grunde gelegt wurden, haben wir damit eine Lösung gefunden. Ein Problem stellen die durch die Ungleichungen (3.15) beziehungsweise (3.27) aufgestellten Nebenbedingungen dar. Sie sagen aus, dass alle Trainingsdatenpunkte korrekt klassifiziert werden müssen und nicht innerhalb des Margins liegen dürfen. Liegt ein Trainingsdatensatz mit nur einem einzigen Ausreißer vor, der es nicht ermöglicht die Daten durch eine lineare Funktion zu trennen, versagt der bisherige Algorithmus. Dieser Fall ist in Abbildung 17 dargestellt. Ein anderes Szenario wäre, dass der Ausreißer zwar noch eine lineare Separation ermöglicht, das Ergebnis jedoch nicht zufriedenstellend ist. Der Algorithmus würde so lange versuchen ein geeigneten Lagrangeparamter a_i für den Ausreißer zu finden, bis er richtig klassifiziert wird. Dabei wächst der Wert für a_i immer weiter an und der Ausreißer erhält ein immer größeres Gewicht (vgl. Hornsteiner, 2010, S. 23f); (Géron, 2018, S. 146f). Dieser Fall ist in Abbildung 18 dargestellt.

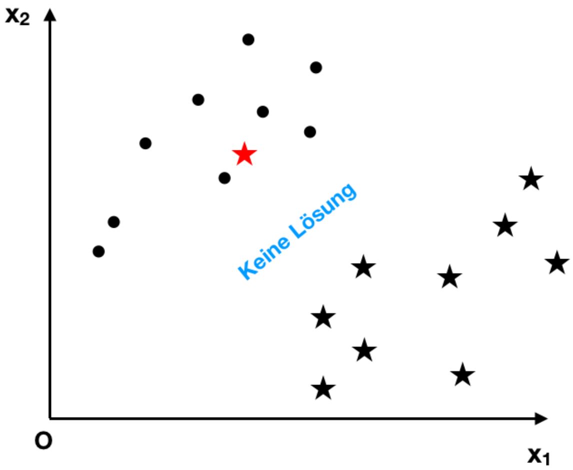


Abb. 17: Abgebildet sind zwei verschiedene Klassen (Kreis & Stern). Der Ausreißer (rot) aus der Klasse Sterne macht eine lineare Separation unmöglich.

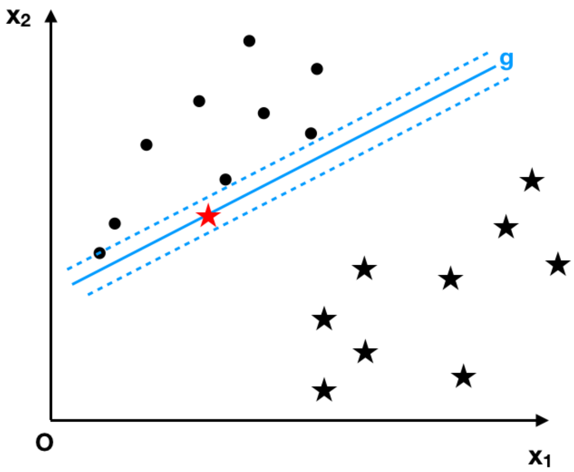


Abb. 18: Abgebildet sind zwei verschiedene Klassen (Kreis & Stern). Der Ausreißer (rot) aus der Klasse Sterne beeinflusst das Ergebnis zu stark. Die Trennfunktion g lässt sich schlecht auf neue Daten anwenden.

Für den Datensatz in Abbildung 17 kann keine lineare Trennfunktion gefunden werden. Mathematisch ausgedrückt ist der Schnitt der konvexen Hüllen der beiden Klassen Stern und Kreis nicht leer. Einfach gesagt: die Mengen der beiden Datenpunktklassen überlappen sich. Eine solche Situation wurde bereits anfangs des Abschnitts 3.2 in Abbildung 12 dargestellt. Um die beiden in Abbildung 17 und 18 dargestellten Szenarien zu vermeiden, gibt es die Möglichkeit die Bedingung aus Ungleichung (3.27) abzuschwächen. Der dazu verwendete Ansatz wird auch als *Soft Margin* Klassifizierung bezeichnet und unterscheidet sich nur gering von dem vorher genutzten sogenannten *Hard Margin* Ansatz (vgl. Géron, 2018, S. 147). Die Bezeichnung *Soft* meint, dass die Nebenbedingungen für die Optimierung abgeschwächt werden. Es wird erlaubt, dass Trainingsdatenpunkte innerhalb des Margins und sogar auf der falschen Seite der Trennfunktion liegen. Mathematisch wird die Abschwächung der Nebenbedingung durch *Schlupfvariablen* (engl. *slack variables*) $\xi_i \in \mathbb{R}^{\geq 0}$ formuliert:

$$t_i \left(\vec{n} * \vec{x}_i + b\right) \geq 1 - \xi_i \; \forall \; \vec{x}_i \in D \; , \; i = 1, .., k \; . \tag{3.54}$$

Jeder Trainingsdatenpunkt erhält eine eigene Schlupfvariable. Dabei liegen Trainingspunkte $\vec{x}_i$ mit $\xi_i = 0$ außerhalb des Margins und werden korrekt klassifiziert. Für sie ist (3.54) identisch mit der Hard Margin Bedingung aus Ungleichung (3.27). Trainingspunkte mit $0 < \xi_i < 1$ liegen innerhalb des Margins, aber noch auf der richtigen Seite der Trennfunktion. Sie werden also auch richtig klassifiziert. Für Trainingspunkte $\vec{x}_i$ mit $\xi_i \geq 1$ gilt, dass sie auf der falschen Seite der Trennfunktion liegen. Summiert man über alle Schlupfvariablen ξ_i erhält man ein Maß dafür, wie viele Trainingsdaten maximal falsch klassifiziert werden (vgl. Frochte, 2018, S. 292). Der Fehler durch die Schlupfvariablen kann durch einen zusätzlichen Parameter $C \in \mathbb{R}^{>0}$ reguliert werden:

$$C \cdot \sum_{i=1}^{k} \xi_i \ . \tag{3.55}$$

Das primale Problem (3.37) ohne Lagrange Multiplikatoren lautet dann wie folgt.

Sei $D \subset \mathbb{R}^n \times C$ ein Trainingsdatensatz aus k Merkmalsvektoren $\vec{x}_i \in \mathbb{R}^n$, $i = 1, ..., k$ und den zugehörigen Labels $t_i \in C = \{-1, 1\}$. Sei außerdem $C \in R^{>0}$. Dann werden die Parameter $b \in \mathbb{R}$, $\vec{n} \in \mathbb{R}^n$ und $\xi_i \in \mathbb{R}^{\geq 0}$ gesucht, sodass

$$\begin{aligned} &\min_{\vec{n}, \xi_i} \frac{1}{2} |\vec{n}|^2 + C \cdot \sum_{i=1}^{k} \xi_i \\ \text{unter den Nebenbedingungen} \qquad &t_i (\vec{n} * \vec{x}_i + b) - 1 + \xi_i \geq 0, \quad i = 1, .., k \ . \end{aligned} \tag{3.56}$$

An (3.56) werden die Regulierungsmöglichkeiten durch den Kostenfaktor C deutlich. Wählt man große Werte für C, so muss die Summe der Schlupfvariablen möglichst klein werden. Wenige Datenpunkten dürfen innerhalb des Margins, oder auf der falschen Seite liegen. Ausreißer erhalten ein höheres Gewicht und im Grenzfall für $C \to \infty$ ist man wieder im Fall der Hard Margin Klassifikation. Wählt man für C jedoch kleine Werte, so wird es mehreren Datenpunkten erlaubt, innerhalb oder auf der falschen Seite des Margins zu liegen. Der Margin wird durch verkleinern von C immer größer und der Fehler durch Falschklassifikation steigt. Ziel ist es C so zu wählen, dass Ausreißern kein zu hohes Gewicht verliehen wird, da die resultierende Entscheidungsfunktion dann nicht gut auf neue Daten angewendet werden könnte (vgl. Abb. 18). Gleichzeitig sollte C nicht zu klein gewählt werden, sodass der Fehler durch Falschklassifikation nicht zu groß wird. Die beiden Extrema werden auch als *Over-* beziehungsweise *Underfitting* bezeichnet. Lässt sich die gefundene Entscheidungsfunktion schlecht generalisieren, das heißt auf neue Datensätze anwenden (zu große C), spricht man von ersterem. Teilt sie den Trainingsdatensatz jedoch nicht mehr sinnvoll ein und macht zu viele Fehler (kleine C), ist von letzterem die Rede (vgl. Whitenack, 2017, S. 76); (Wroble et al., 2013, S. 433).
Das zugehörige duale Problem mit Lagrangemultiplikatoren gleicht dem der Hard Margin Klassifikation (3.48) bis auf die veränderte Nebenbedingung an die Lagrange Multiplikatoren in Form der oberen Schranke C. Interessierte Leser können dies in (Bishop, 2006, S. 332f) nachlesen.

3.2.3. Lineare SVM bei mehr als zwei Klassen

Als letztes soll in diesem Abschnitt erklärt werden, wie die SVM auch auf ein Klassifizierungsproblem angewendet werden kann, das aus mehr als nur zwei Klassen besteht. Die Klassifizierung von Verkehrsampeln, welche im Workshop genutzt wird, stellt ein gutes Beispiel für eine solche Situation dar. Offensichtlich unterscheidet man zwischen roten, gelben und grünen Ampeln. Im Lernmodul gehen die Schüler, wie bereits beschrieben, zunächst nur von grünen und roten Ampeln aus und verbessern dann ihr Modell, sodass auch gelbe Ampeln betrachtet werden. Das mathematische Modell bleibt dabei gleich: Es wird eine Trennfunktion für Punkte in einem n-dimensionalen Raum gesucht. Das mathematische Problem wird hingegen komplexer, da nun drei Klassen voneinander getrennt werden müssen. Bei solchen nicht binären Klassifizierungsproblemen gibt es zwei mögliche Vorgehensweisen, wie man mittels SVM eine Einteilung vornehmen kann. Es wird zwischen der sogenannten *One-versus-One* (OVO) und der *One-Versus-All* Methode (OVA) unterschieden. Grundidee beider Verfahren ist es mehrere SVMs zu trainieren und immer zwei Klassen mit einander zu vergleichen. Die Methoden unterscheiden sich lediglich in der Zusammensetzung der jeweiligen einzelnen Klassen (vgl. Géron, 2018, S. 96f). Beide sollen im Folgenden kurz vorgestellt werden.

One-Versus-All SVM

Die erste Methode, ein Klassifizierungsproblem mit n verschiedenen Klassen zu lösen, ist die *One-Versus-All* Methode (OVA). Wie der Name schon sagt, wird immer eine Klasse (one) mit dem Rest der anderen Klassen (all) verglichen. Bei n Klassen werden also n einzelne SVMs trainiert und zugehörige Trennfunktionen $f_1, ..., f_n$ gesucht. Jede Trennfunktion hat die Gestalt $f(\vec{x}) = \vec{n} * \vec{x} + b$. Für jede einzelne SVM wird die einzelne Klasse positiv $(t = +1)$ und der Zusammenschluss aus den übrigen Klassen negativ $(t = -1)$ gelabelt. Um einen unbekannten Datenpunkt zu klassifizieren, wird er in jede der Trennfunktionen eingesetzt. Dann wird das Maximum der n Werte aus den einzelnen Trennfunktionen gesucht (vgl. Lampert & Lampert, 2009, S. 46):

$$F(\vec{x}) = \max_{i=1,...,n} f_i(\vec{x}). \tag{3.57}$$

Ein Testdatenpunkt wird der Klasse zugeordnet, bei der das Maximum angenommen wird. In einigen Situationen, in denen ein unbekannter Datenpunkt $\vec{x}$ eingeordnet werden soll, liefert nur eine der $f_i(\vec{x}, i = 1, ..., n$ einen positiven Wert. In solchen Fällen kann der unbekannte Datenpunkt direkt in die entsprechende Klasse eingeordnet werden. Das Vorgehen wird in Abbildung 19 an einem Beispiel verdeutlicht.

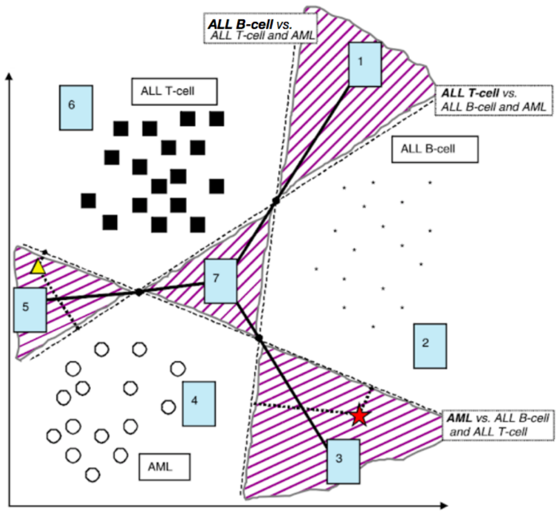


Abb. 19: Beispiel für Mehrklassenproblem. Trainingsdatensatz aus 3 Klassen: ALL T (Quadrate), ALL B (Punkte) und AML (Kreise). Gestrichelte Linien: Trennfunktion für die jeweilige Aufteilung. Schraffierte Bereiche: entweder kein oder zwei positive Werte für Trennfunktion. 7 verschiedene Zonen in Tabelle 2 erklärt. Schwarze dicke Linie: resultierende Entscheidungsfunktion. Gelbes Dreieck und roter Stern: zwei neu zu klassifizierende Datenpunkte (aus: Statnikov et al., 2011, S. 92).

In Abbildung 19 ist ein Trainingsdatensatz aus drei verschiedenen Klassen dargestellt. Klasse eins ALL T sind die Quadrate, Klasse zwei ALL B sind die Punkte und Klasse drei AML sind die Kreise. Die gestrichelten Geraden sind die Trennfunktionen f_1, f_2 und f_3 der jeweiligen SVMs. Für Datenpunkte in den violett schraffierten Bereichen 3, 5 und 7 liefern entweder zwei oder keine der Einzelklassifikationen einen positiven Wert. In solchen Fällen muss nach Gleichung (3.57) der maximale Wert für $\vec{n} * \vec{x} + b$ genommen werden. Dies wird für das gelbe Dreieck angedeutet. Für das gelbe Dreieck liefern die Klassifikationen *ALL T vs. Rest*, sowie *AML vs. Rest* einen positiven Wert. Folglich kann nicht direkt entschieden werden, ob das Dreieck zur Klasse der Kreise oder zu der Klasse der Quadrate gehört. Für die Trennfunktion von *ALL T vs. Rest* ist der Wert $\vec{n} * \vec{x} + b$ jedoch größer, als für *AML vs. Rest*. Das Dreieck hat einen größeren Abstand zu der erst genannten Trennfunktion. Daher wird das Dreieck nach Gleichung (3.57) in die Klasse der ALL T der Quadrate eingeordnet. Für die übrigen Bereiche 2,4 und 6 liefert jeweils nur eine Eizelklassifikation einen positiven Wert, sodass Punkte hier direkt zugewiesen werden können. In Tabelle 2 werden die Zuordnungen für alle Bereiche angegeben.

Tab. 2: Resultate für die 7 Bereiche. Einträge in den Spalten 2-4 bedeuten positive Werte für die jeweilige Trennfunktion. Bei Resultat '?' muss maximaler Wert betrachtet werden.

Bereich	AML vs. Rest	ALL T vs. Rest	ALL B vs. Rest	Resultat
1	–	ALL T	ALL B	?
2	–	–	ALL B	ALL B
3	AML	–	ALL B	?
4	AML	–	–	AML
5	AML	ALL T	–	?
6	–	ALL T	–	ALL T
7	–	–	–	?

Das OVA Vorgehen erfordert viel Rechenzeit, da bei n verschiedenen Klassen n quadratische Optimierungsprobleme gelöst werden müssen. Außerdem ist das Training der einzelnen SVMs sehr unbalanciert, da die positiv gelabelte Datenmenge meistens viel kleiner ist als der Zusammenschluss aus den übrigen Klassen (vgl. Benesty et al., 2007, S. 832); (Schölkopf & Smola, 2001, S. 211). Um dieses Ungleichgewicht zu umgehen, gibt es das *One-Versus-One* Verfahren.

One-Versus-One SVM

Auch bei diesem Verfahren werden wieder mehrere binäre SVMs ausgewertet. Im Unterschied zur OVA Methode, wird jedoch nur jeweils eine Klasse (one) mit einer anderen (one) verglichen und nicht mit dem Rest aus allen anderen Klassen. Dadurch wird ein ungleichmäßiges Training, wie bei der OVA Methode verhindert. Für jede mögliche Kombination der Klassen eines Datensatzes wird eine SVM trainiert, sodass man bei n Klassen $\binom{n}{2} = n(n-1)/2$ verschiedene SVMs trainieren muss. Das bedeutet zwar, dass mehr quadratische Optimierungsprobleme gelöst werden müssen als bei der OVA Methode, jedoch müssen für die einzelnen Probleme deutlich weniger Datenpunkte betrachtet werden. Dadurch erfordern die beiden Verfahren ungefähr die gleiche Rechenzeit (vgl. Lampert & Lampert, 2009, S. 46); (Benesty et al., 2007, S. 832). Bei der OVO Methode wird die Zuteilung eines unbekannten Datenpunktes über die Formel[9]

$$F(\vec{x}) = \max_{i=1,\dots,n} \#\{j \in \{1,...,n\} : f_{i,j}(\vec{x}) > 0\} \tag{3.58}$$

ermittelt. Dabei steht $f_{i,j}$ für die Trennfunktion der SVM von Klasse i und j. Für einen zu klassifizierenden Datenpunkt $\vec{x}$ wird also gezählt wie oft er in eine der n Klassen zugeordnet wurde. Er wird in die Klasse eingeteilt, der er am häufigsten aus den $n(n-1)/2$ Zuordnungen zugeteilt worden ist (vgl. Lampert & Lampert, 2009, S. 46).

[9] # steht hier für Anzahl

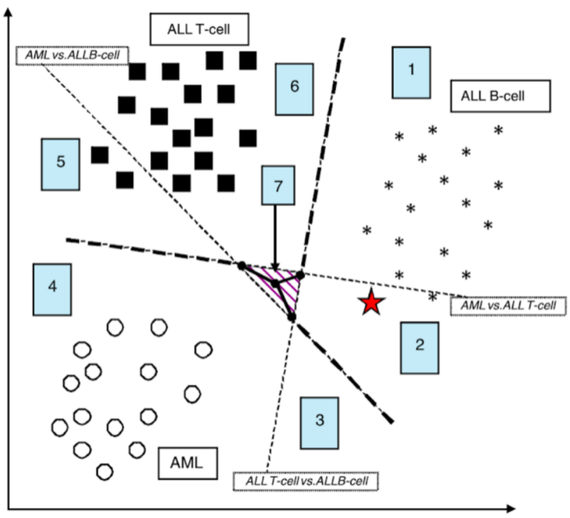


Abb. 20: Beispiel für Mehrklassenproblem. Trainingsdatensatz aus 3 Klassen: ALL T (Quadrate), ALL B (Punkte) und AML (Kreise). Gestrichelte Linien: Trennfunktion für die jeweilige Aufteilung. Schraffierter Bereich: mehrere Klassen haben die maximale Anzahl an Einzelzuteilungen. 7 verschiedene Zonen in Tabelle 3 erklärt. Schwarze dicke Linie: resultierende Entscheidungsfunktion. Roter Stern: neu zu klassifizierende Datenpunkt (aus: Statnikov et al., 2011, S. 94).

In Abbildung 20 wird wieder der Trainingsdatensatz aus den drei Klassen ALL T (Quadrate), ALL B (Punkte) und AML (Kreise) dargestellt. Dieses Mal wurde das OVO Verfahren genutzt, um die drei gestrichelten Trennfunktionen $f_1 - f_3$ der jeweiligen SVMs zu finden. Die dicke, schwarze, teils-gestrichelte Linie ist die resultierende Entscheidungsfunktion. Vergleicht man Abbildung 19 mit Abbildung 20, so fällt auf, dass der schraffierte Bereich hier viel kleiner ist. Nur in diesem Bereich kann es zu einem 'Unentschieden' kommen, bei dem dann analog zur OVA Methode der maximale Wert für $\vec{n} * \vec{x} + b$ genommen werden muss. Der unbekannte Datenpunkt, der durch den roten Stern symbolisiert wird, wird zur Klasse ALL B der Punkte zugeteilt. Er wird durch die SVM von *AML vs. ALL B* und durch die SVM von *ALL T vs. ALL B* beides mal zu ALL B zugeteilt.

Tab. 3: Resultate für die 7 Bereiche. Einträge in den Spalten 2-4 bedeuten Zuordnung für die jeweilige Trennfunktion. Bei Resultat '?' muss maximaler Wert betrachtet werden.

Bereich	AML vs. ALL B	ALL T vs. ALL B	AML vs. ALLT	Resultat
1	ALL B	ALL B	ALL T	ALL B
2	ALL B	ALL B	AML	ALL B
3	AML	ALL B	AML	AML
4	AML	ALL T	AML	AML
5	AML	ALL T	ALL T	ALL T
6	ALL B	ALL T	ALL T	ALL T
7	ALL B	ALL T	AML	?

Die beiden Verfahren wurden so ausführlich vorgestellt, da die Teilnehmer des Workshops eines dieser Verfahren umsetzen sollen. Sie sind somit unmittelbare Bestandteile des CAMMP days. Im Modul trennen die Schüler die gelben, die grünen und die roten Verkehrsampeln. Dazu können sie eines der beiden vorgestellte Verfahren auswählen (vgl. 4.3.4).

3.3. Anknüpfungspunkte zur Schulmathematik

In diesem Abschnitt sollen die Anknüpfungspunkte des in Teil 3.2 erläuterten mathematischen Hintergrunds der SVM zur Schulmathematik beschrieben werden. Zu Grunde liegt der Kernlehrplan im Fach Mathematik des Landes Nordrhein-Westfalen (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014). Der inhaltliche Schwerpunkt des entwickelten Lernmoduls liegt im Teilgebiet der analytischen Geometrie und linearen Algebra. Im Kernlehrplan wird dieses Inhaltsfeld wie folgt umschrieben:

> „Die Geometrie umfasst den quantitativen und den qualitativen Umgang mit ebenen und räumlichen Strukturen. Die Idee der Koordinatisierung ermöglicht deren vertiefte Untersuchung mit algebraischen Mitteln im Rahmen der analytischen Geometrie. Die Beschreibung mittels Vektoren erlaubt dabei den Rückgriff auf das universelle Handwerkszeug der linearen Algebra. Aus der Idee der Parametrisierung ergeben sich Beschreibungen für geometrische Objekte sowie für geradlinige Bewegungen im Raum. Nach der Metrisierung des Raumes mit dem Skalarprodukt lassen sich nicht nur Winkel-, Längen- und Abstandsmessungen durchführen, sondern auch die strategischen und rechnerischen Bearbeitungsmöglichkeiten für geometrische Fragestellungen erweitern.“ (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014, S. 17).

In dieser Beschreibung lassen sich bereits viele Aspekte aus dem mathematischen Hintergrund des CAMMP days wiederfinden. Ein zentraler Aspekt des Moduls, welcher essenziell für die Aufstellung des mathematischen Modells ist, ist die Darstellung der Daten durch Merkmalsvektoren. Die Merkmalsvektoren stellen die Grundlage für die SVM dar, da sie die Verwendung des Skalarproduktes ermöglichen. Dieser Punkt kann in der obigen Beschreibung unter dem *universellen Handwerkszeug der linearen Algebra* aufgeführt werden. Das Aufstellen der Entscheidungsfunktion beruht wie in Abschnitt 3.2.1 beschrieben auf dem Abstands- und dem Winkelbegriff. Die Schüler erarbeiten sich beide Begriffe auf den ersten beiden Arbeitsblättern. Zunächst müssen sie auf dem ersten Blatt eine mathematische Formel für den Abstand eines Punkte von einer Geraden aufstellen. Im Kernlehrplan wird dazu unter den inhaltsbezogenen Kompetenzen geschrieben, dass die Schüler Abstände zwischen Punkten, Geraden und Ebenen bestimmen sollen (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014, S. 33). Die Formel, welche die Schüler aufstellen, wird dabei wie in den Gleichungen (3.16) bis (3.22) über die geometrische Definition des Skalrproduktes hergeleitet. Auf dem zweiten Blatt stellen die Schüler schließlich über die geometrische Betrachtung und die Definition des Skalarproduktes eine Fallunterscheidung für die Entscheidungsfunktion auf (vgl. Abschnitt 4.3.4). Diese Arbeitsschritte lassen sich ebenfalls in der obigen Beschreibung aus dem Kernlehrplan wiederfinden. Sie werden außerdem unter den inhaltsbezogenen Kompetenzerwartungen aufgeführt: „Sie (die Schüler) untersuchen mithilfe des Skalarprodukts geometrische Objekte und Situationen im Raum (Orthogonalität, Winkel- und Längenberechnung).“ (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014, S. 29). Genau dies wird durch die Fallunterscheidung auf dem zweiten Arbeitsblatt umgesetzt. Da die Schüler zunächst mit zwei- und dreidimensionalen Vektoren arbeiten, wird für die Trennfunktion die Normalenform einer Geraden, beziehungsweise einer Ebenen genutzt. Die Darstellung von geometrischen Objekten durch diese Form wird ebenfalls im Kernlehrplan genannt. Dort heißt es, „die Schülerinnen und Schüler stellen Ebenen in Normalenform dar und nutzen diese zur Orientierung im Raum.“ (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014, S. 33). Die Darstellung von Ebenen mit Hilfe der Normalenform wird erst in der Q2 thematisiert, weshalb EF, bzw. Q1 Kurse noch nicht über dieses Wissen verfügen. Im Modul wurde daher eine kurze Einführung in den Begriff vorgenommen. Bei der Durchführung des Moduls wird evaluiert werden, inwiefern der noch unbekannte Stoff verstanden wurde (vgl. Abschnitt 5.4.2).

Ein weiterer Anknüpfungspunkt ist durch die Optimierung des Abstandsbegriffs gegeben im Kernlehrplan wird unter dem Punkt *Anregungen für Aufgaben zur Leistungsüberprüfung und Unterrichtsgestaltung*, die Optimierung von Abständen als vernetzende Aufgabe genannt (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014, S. 40). Da die Nebenbedingungen für die Optimierung im Lernmodul durch Ungleichungen gegeben sind (vgl. z.B. Gleichungen (3.42) bis (3.44)), geht die Mathematik an dieser Stelle über den Schulstoff hinaus. Daher wird der Optimierungsalgorithmus im Modul als Black-Box betrachtet (vgl. Abschnitt 4.3.5).

4. Didaktisch-methodisches Konzept

Im vierten Kapitel der vorliegenden Arbeit sollen das entwickelte Lernmodul vorgestellt werden. Indem die Umsetzung des didaktischen und des mathematischen Hintergrundes für den geplanten CAMMP day beschrieben wird, stellt dieses Kapitel ein Kernelement der Arbeit dar. Zunächst sollen die Struktur eines CAMMP days im Allgemeinen beschrieben werden. Anschließend werden spezielle Bestandteile des Machine Learning Moduls beschrieben. Dazu gehören gestuften Hilfen, Live-Scripts, Antwortzettel, Hilfekarten und Besprechungsfolien. Im letzten Abschnitt werden die digitalen Aufgabenblätter und die Präsentationen vorgestellt, welche für das Modul entwickelt wurden. Dabei wird die Version beschrieben, welche in der ersten Durchführung des CAMMP days verwendet wurde. Wie in den Abschnitten 5.3.3 und 5.4.3 ersichtlich wird, mussten nach den beiden Durchführungen nur geringe Verbesserungen und Änderungen des Materials gemacht werden. Hauptsächlich wurde der Umfang reduziert. Ziel des Kapitels ist, dass der Leser einen Eindruck des entwickelten Moduls gewinnt und ersichtlich wird inwiefern die in Kapitel 2 vorgestellten didaktischen mit den in Kapitel 3 vorgestellten mathematischen Aspekten verknüpft werden.

4.1. Allgemeine Struktur eines CAMMP days

Bevor die einzelnen Bestandteile des entwickelten Workshops vorgestellt werden soll zunächst die allgemeine Durchführung eines CAMMP days beschrieben werden. Der Workshop, dessen Thema vom teilnehmenden Mathematik Kurs im Vorfeld ausgewählt wird, wird unter der Anleitung von zwei Mitarbeitern des CAMMP Teams durchgeführt. Bei den Mitarbeitern handelt es sich um studentische Hilfskräfte. Die Betreuer führen die Lerngruppe mit Hilfe verschiedener Präsentationen durch den Tag. Einige Elemente sind dabei für jeden CAMMP day gleich, unabhängig davon welches Modul ausgewählt worden ist. So startet jeder CAMMP day mit einer Begrüßung, in der das CAMMP Team und die Idee hinter dem Projekt kurz vorgestellt wird. Anschließend hält ein Doktorand einen Vortrag über mathematische Modellierung. Darauf wird im Abschnitt 4.1.1 dieser Arbeit näher eingegangen. Auch der Abschluss jedes Workshops ist identisch. Die Betreuer halten am Ende eine Abschlusspräsentation über weitere Angebote von CAMMP, welche bereits in Abschnitt 2 vorgestellt wurden. Außerdem beinhaltet die Abschlusspräsentation einen Link zur Online-Evaluation des durchgeführten Lernmoduls. Die Teilnehmer sollen hier Rückmeldung geben und ihre Erfahrungen schildern, sodass die Module evaluiert und stetig verbessert werden können. Außerdem ist die Berufs- und Studienorientierung ein Bestandteil jedes CAMMP days. Den teilnehmenden Schülern wird im Rahmen einer Präsentation der Studiengang *Computational Engineering Science* (CES) vorgestellt. Dies hat vor allem für Kurse der Qualifikationsphase, die Zielgruppe des entwickelten Moduls sind, eine hohe Relevanz.

4.1.1. Modellierungsvortrag

Bestandteil jedes CAMMP days ist der Modellierungsvortrag. Nachdem die Schüler begrüßt wurden, werden sie im Rahmen dieses Vortrags zum ersten Mal mit dem Thema der mathematischen Modellierung konfrontiert. Wie bereits oben erwähnt, werden diese Vorträge von Doktoranden aus dem Bereich CES, Mathematik oder Mathematikdidaktik gehalten. Zu Beginn jedes Vortrags stellt der Doktorand sich kurz vor und erklärt den Schülern seinen Werdegang von der Schule, über das Grundstudium, bis hin zu seinem aktuellen Promotionsstudium. Die meisten teilnehmenden Schüler haben nur ungefähre Vorstellungen wie der Bildungsweg nach der Schule weitergehen kann. Da sich das Modul vor allem an Schüler der Qualifikationsphase I und II richtet, ist die Studieninformation zu diesem Zeitpunkt durchaus sinnvoll. Der Anfang des Vortrags soll den Schülern deshalb eine Orientierungsmöglichkeit geben und ihnen einen Eindruck vermitteln, was „Doktorand“ bedeutet und wie ein exemplarischer Werdegang aussehen kann. Im Hauptteil des Vortrags werden die Schüler zum ersten Mal mit dem Modellierungsbegriff konfrontiert. Dabei wird ihnen der Modellierungskreislauf aus Abbildung 3 in Abschnitt 2 schrittweise erklärt. Für die Erklärungen nutzen die Doktoranden Beispiele aus ihrem eigenen Forschungsgebiet. Ein Modellierungsvortrag thematisiert zum Beispiel die Frage, wie ein Flugzeug gebaut werden muss, damit es möglichst weit fliegen kann. Anhand dieses Optimierungsproblems werden die 4 Stationen des Modellierungskreislaufs erläutert. Für das reale Problem werden bestimmte Annahmen über das Gewicht des Flugzeugs getroffen. Außerdem werden Vereinfachungen gemacht. So könnte die Form des Rumpfes beispielsweise vereinfacht durch einen Zylinder beschrieben werden. Eine weitere Vereinfachung wäre es, das Flugzeug als symmetrisch anzunehmen. Auf diese Weise müsste nur eine Seite des Flugzeugs berechnet werden. Ziel ist es, dass die Schüler verstehen warum Vereinfachungen und Annahmen getroffen werden müssen. Um ein mathematisches Modell zu entwickeln stellt der Computer ein hilfreiches und zum Teil notwendiges Werkzeug dar (vgl. 2.2.2). An dieser Stelle wird betont, dass nur eine kleine Anzahl von Problemen exakt analytisch lösbar ist. Solche Probleme werden meistens in der Schulmathematik behandelt. Für die Probleme, die im Rahmen des Forschungsgebiets des jeweiligen Doktoranden oder in den CAMMP days thematisiert werden, funktioniert das nicht mehr. Hier berechnet der Computer eine numerische Lösung. Der Kompetenzbereich *Werkzeuge nutzen*, der im Kernlehrplan (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014, S. 16f) benannt wird, soll durch den Computereinsatz gefördert werden. Den Schülern wird bewusst gemacht, dass die Menge der für sie lösbaren Probleme durch die Verwendung von Computersoftware deutlich größer wird. Im letzten Schritt, dem Interpretieren, macht der Doktorand deutlich, wie wichtig es ist die mathematischen Ergebnisse am Ende zu diskutieren. Erhält man für das Flugzeug beispielsweise die Lösung, dass die Tragflächen nur 5 cm dünn, aber dafür 30 m lang seinen sollen kann dies eine mathematische Lösung des Optimierungsproblems sein. In der Realität ist diese Lösung jedoch nicht umsetzbar. Die Fähigkeit Ergebnisse mit Bezug zur Realität einzuordnen und zu interpretieren ist auch für die Workshops sehr wichtig. Erhalten die Schüler eine mathematische Lösung, deren Interpretation zu

einem nicht zufriedenzustellenden Ergebnis führt, so werden sie motiviert eine Modellverbesserung durchzuführen.
Nachdem der Modellierungsvortrag gehalten wurde, findet die Einführung in die jeweilige Problemstellung des Workshops statt (vgl. Abschnitt 4.3.1).

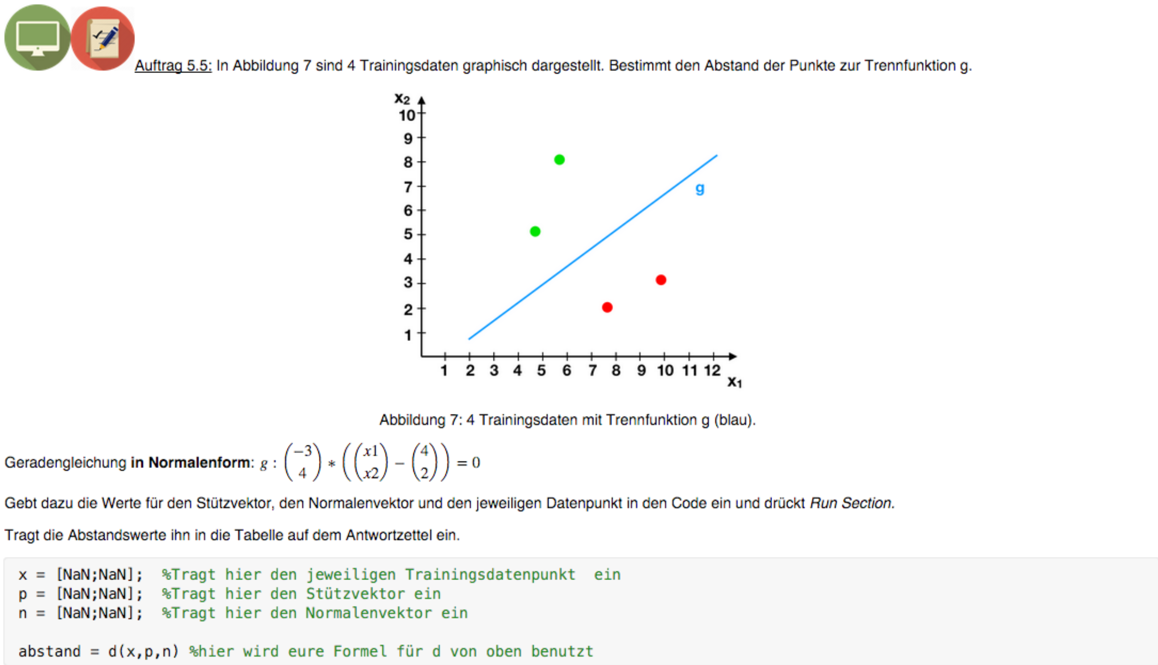
4.2. Methodische Besonderheiten

In diesem Unterpunkt sollen fünf methodische Gesichtspunkte vorgestellt werden, durch die sich das entwickelte Lernmodul auszeichnet und teilweise von anderen Workshops unterscheidet.

4.2.1. Live-Script

Die Arbeitsblätter für den Workshop wurden in digitaler Form als *Live-Script* in MATLAB umgesetzt. Dies sind Programmdateien, die es ermöglichen einen Code zusammen mit einem vorformatierten Text und Grafiken anzeigen zu lassen. Sie ersetzen das Ausdrucken von Arbeitsblättern für die Schüler, wie es bei vorherigen CAMMP Modulen noch nötig war. Da die Aufgabenstellung und der Code in den die Schüler ihre Ergebnisse eintragen gleichzeitig angezeigt werden, wird ein strukturiertes Arbeiten ermöglicht. Der Code auf den Live-Scripts ist ähnlich wie ein Lückentext gestaltet. Stellen, an denen die Schüler eine Eingabe tätigen müssen, werden durch *NaN*[10] gekennzeichnet. Ein Beispiel dafür wird in Abbildung 21 dargestellt.

[10]Steht für Not a Number und kennzeichnet eine Eingabemöglichkeit für die Schüler.



Auftrag 5.5: In Abbildung 7 sind 4 Trainingsdaten graphisch dargestellt. Bestimmt den Abstand der Punkte zur Trennfunktion g.

Abbildung 7: 4 Trainingsdaten mit Trennfunktion g (blau).

Geradengleichung **in Normalenform**: $g : \begin{pmatrix} -3 \\ 4 \end{pmatrix} * \left(\begin{pmatrix} x1 \\ x2 \end{pmatrix} - \begin{pmatrix} 4 \\ 2 \end{pmatrix} \right) = 0$

Gebt dazu die Werte für den Stützvektor, den Normalenvektor und den jeweiligen Datenpunkt in den Code ein und drückt *Run Section*.

Tragt die Abstandswerte ihn in die Tabelle auf dem Antwortzettel ein.

```
x = [NaN;NaN];  %Tragt hier den jeweiligen Trainingsdatenpunkt  ein
p = [NaN;NaN];  %Tragt hier den Stützvektor ein
n = [NaN;NaN];  %Tragt hier den Normalenvektor ein

abstand = d(x,p,n) %hier wird eure Formel für d von oben benutzt
```

Abb. 21: Auszug aus dem Live-Script für Aufgabenblatt 1. In der oberen Hälfte: Arbeitsauftrag mit zugehöriger Abbildung. Darunter im grauen Kasten: Code mit Eingabemöglichkeiten für die Schüler.

Wie in Abbildung 21 zu sehen ist, bietet die Live-Script Oberfläche viele Möglichkeiten, um die Arbeitsblätter ansprechend zu gestalten. In der oberen Hälfte der Abbildung ist der Arbeitsauftrag mit zugehöriger Abbildung zu sehen. Darunter im grau hinterlegten Bereich, ist der Code abgebildet, welcher die Gestalt eines Lückentextes hat. Mit Hilfe der grünen Kommentare in den jeweilige Code Zeilen, können die Schüler beim Ausfüllen unterstützt werden. Die Live-Scripts zu den einzelnen Arbeitsblättern stellen die zentralen Elemente für die Schüler dar. Von dort führen Verweise zu Hilfekarten oder zu den Antwortzetteln, welche im nächsten Unterpunkt betrachtet werden sollen.

4.2.2. Antwortzettel

Wie in Abschnitt 4.3 beschrieben wird, gibt es einige Aufgaben, die die Schüler nicht am Computer lösen. Beispielsweise müssen sie auf dem ersten Aufgabenblatt verschiedene Datensätze zeichnerisch trennen. Deshalb wurden Antwortzettel eingeführt, auf denen die Schüler Teile ihrer Ergebnisse festhalten. Es werden vor allem solche Ergebnisse festgehalten, für die es keine direkte Rückmeldung durch den Code gibt oder die einer Diskussion im Plenum bedürfen. Aufgaben, die bereits eine Rückmeldung im Code beinhalten, werden nur bei Fragen der Schüler besprochen. Außerdem sollen auf diesen Blättern Fragen zum Verständnis beantwortet werden. Im Falle von längeren Rechnungen, wie zum Beispiel bei der Herleitung der Abstandsformel, ist ausreichend Platz auf den Antwortzetteln vorhanden. Die Antwortzettel ermöglichen, dass die Laptops der Gruppen bei der Besprechung der einzelnen Aufgabenblätter zugeklappt werden können. Auf diese Weise kann gewährleistet werden, dass die Schüler ihren Fokus auf die Besprechung legen. Die Antwortzettel sind im Anhang unter den Abschnitten A.2.1 bis A.2.4 angefügt.

4.2.3. Besprechungsfolien

An jedes Arbeitsblatt schließt sich eine Besprechungsphase an. Diese dient als Überleitung zum jeweils nächsten Blatt und zur Sicherung der bisherigen Ergebnisse. Die Schüler sollen während der Besprechungsphasen die Gelegenheit bekommen ihre eigene Lösung vorzustellen und über Lösungswege anderer Gruppen zu diskutieren. Um eine strukturierte Diskussion der Arbeitsergebnisse zu ermöglichen, wurden Besprechungsfolien entwickelt. Mit diesen Folien können einige Teilaufgaben des jeweiligen Arbeitsblatts besprochen werden. Teilaufgaben, die bereits eine direkte Rückmeldung vom Code beinhalten werden hier nicht besprochen. Des weiteren wurden Folien vorbereitet, die während der Arbeitsphasen an schnelle Schülergruppen verteilt werden. Die entsprechende Gruppe hält ihre Ergebnisse dann auf der Folie, anstatt auf dem Antwortzettel fest. In der Besprechungsphase können sie dann ihre Lösung präsentieren. Besonders bei Teilaufgaben, in denen mehrere Lösungswege möglich oder Wahlmöglichkeiten gegeben sind, kann auf diese Weise die Sicherung werden. Die Besprechungsfolien sind im Anhang unter den Abschnitten A.5.1 bis A.5.5 angefügt.

4.2.4. Gestufte Hilfekarten

Ein wichtiger Bestandteil des Lernmoduls sind die *gestuften Hilfekarten*. In diesem Abschnitt soll ihr zu Grunde liegendes didaktisches Konzept kurz erläutert werden. Gestufte Hilfekarten sind ein Mittel zur *inneren Differenzierung* in Lernsituationen und gehen auf den Physik-Fachdidaktiker Prof. Josef Leisen zurück. In einer für alle Schüler geplanten Lerngelegenheit besteht die Gefahr einen Teil der Schüler zu überfordern, während man den Anderen unterfordert. Diese Gefahr kann sowohl durch verschiedene Lerntempi, als auch durch individuelle Lernvoraussetzungen entstehen (vgl. Bruder & Reibold, 2010, S. 2). Um ihr entgegen zu wirken, gibt es verschiedene Konzepte der inneren und äußeren Differenzierung. Während sich äußere Differenzierung auf die Aufteilung der Schüler in verschiedene Lerngruppen oder Kurse bezieht, meint innere Differenzierung die Förderung des Einzelnen innerhalb der Klasse oder des Kurses (vgl. Bruder & Reibold, 2010, S. 2). Bei der äußeren Differenzierung wird versucht durch die richtige Verteilung eine möglichst homogene Lerngruppe zu schaffen (vgl. Schneuwly, 2014, S. 23). Eine Maßnahme der inneren Differenzierung kann zum Beispiel das Anbieten von Aufgaben verschiedener Niveaus während einer Arbeitsphase sein. Dabei suchen die Schüler eigenständig das für sie passende Niveau aus. Eine Maßnahme zur äußeren Differenzierung wäre beispielsweise das Einrichten von Mathematik-Förderkursen für lernschwache, oder das Schaffen von Kursen für begabte Schüler. Bei einem CAMMP day werden ausschließlich Methoden der inneren Differenzierung verwendet.

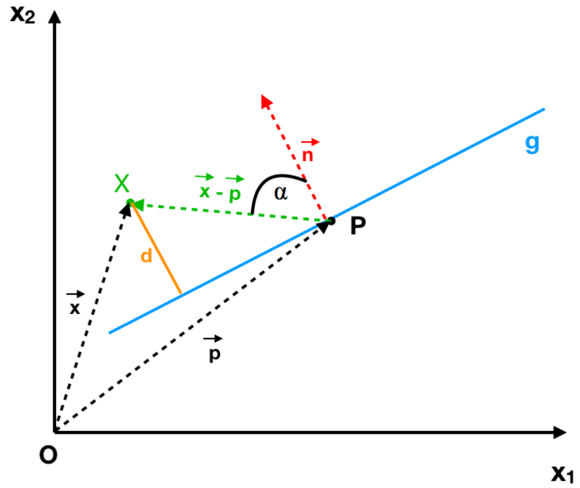


Abb. 22: Skizze zur Berechnung des Abstandes d (orange) eines Punktes X (grün) von der Gerade g (blau) mit Winkel α, Normalenvektor $\vec{n}$ (rot) und Punkt P (schwarz) aus dem Workshop.

Aufbau gestufter Hilfen

Gestufte Hilfekarten sind stets so gestaltet, dass die Schüler nicht direkt die vollstämdige Lösung, sondern Handlungs- und Denkimpulse für einen möglichen Lösungsweg erhalten (vgl. Hänze et al., 2010, S. 64). Sie bestehen aus mehreren einzelnen Hilfeschritten, welche in ihrem Unterstützungsrad aufeinander aufbauen. Der letzte Hilfeschritt ist die Musterlösung für einen möglichen Lösungsweg. Anhand der Lösung können die Schüler die Richtigkeit und die Vollständigkeit ihrer eigenen Lösung überprüfen. Die Hilfekarten dienen somit als Kontrollmöglichkeit und können bei richtiger Lösung das Erfolgserleben der Schüler stärken (vgl. Franke-Braun et al., 2008, S. 29); (Hänze et al., 2010, S. 68).
Ein Beispiel für den Einsatz einer gestuften Hilfekarte aus dem Workshop ist das Folgende:
Die Schüler sollen eine Formel für den Abstand eines Punktes von einer Geraden aufstellen. Dazu erhalten sie die in Abbildung 22 gezeigte Skizze. Die vollständige gestufte Hilfekarte dazu befindet sich im Anhang unter A.4.1. Hier sollen jedoch kurz die einzelnen Stufen erläutert werden, um den Aufbau der entwickelten Lernhilfen zu verdeutlichen. Im ersten Schritt wird der Hinweis gegeben, zwei verschiedene Formeln für $\cos(\alpha)$ in Abbildung 22 aufzustellen und dafür den Wechselwinkel zu nutzen. In der zweiten Stufe werden die allgemeinen Formeln angegeben, welche benutzt werden können. Eine Gleichung für $\cos(\alpha)$ erhält man über die Definition des Skalarproduktes: $\vec{a} * \vec{b} = \cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}|$. Die Andere über den Kosinussatz für ein rechtwinkliges Dreieck: $\cos(\alpha) = \mathit{Ankathete}\ /\ \mathit{Hypotenuse}$. Die Schüler müssen die Formeln nun auf die Skizze in Abbildung 22 anpassen. Fällt ihnen das schwierig, so erhalten sie Hilfestufe 3. Diese Karte beinhaltet zu jeder der beiden Formeln eine Skizze, die fast der in Abbildung 22 gleicht. Der einzige Unterschied ist, dass an die Vektoren die entsprechenden Bezeichnungen aus den beiden Formeln geschrieben wurden, um den Transfer

zu erleichtern. Durch den Wechsel zur grafischen Darstellungsform, sollen außerdem verschiedene Lerntypen angesprochen werden. Im vierten Schritt werden die beiden umgeschriebenen Formeln für $\cos(\alpha)$ gezeigt. Der fünfte Schritt beinhaltet einen Hinweis, wie die Formeln zusammengefügt werden können um schließlich ein Ergebnis für den Abstand d zu erhalten. Im sechsten Schritt wird dann lediglich eine mögliche Musterlösung präsentiert, mit der die Schüler ihr Ergebnis vergleichen können.

Vorteile gestufter Hilfen

Gestufte Hilfekarten fordern den Lernenden zu gezielten Überlegungen oder Handlungen auf. Auf diese Weise sollen sie den Lösungsprozess unterstützen und eine Differenzierung nach Leistungsstand ermöglichen. Sie eigenen sich besonders um Lernende vor komplexere Probleme zu stellen, als im gewöhnlichen Schulunterricht (vgl. Franke-Braun et al., 2008, S. 27); (Hänze et al., 2010, S. 64). Gestufte Hilfekarten gewährleisten, dass sowohl schwächere Schüler komplexe Aufgaben selbständig lösen können, indem sie viele Hilfestufen in Anspruch nehmen. Auf der anderen Seite werden auch leistungsstarke Schüler gefördert, da sie zunächst versuchen können die komplexe Aufgabe ohne Hilfe zu lösen (vgl. Wodzinski, 2013, S. 46). Während des Workshops werden die Schüler durchaus mit komplexen Problemen konfrontiert, weshalb sich diese Differenzierungsmethode besonders gut anbietet. Die höhere Komplexität der in CAMMP behandelten Probleme im Vergleich zu denen aus der Schule, wird unter anderem durch den höheren Grad an Offenheit erreicht.

Die Schüler können selbst entscheiden, wie stark sie die Hilfekarten in Anspruch nehmen. Auf diese Weise passen sie den Schwierigkeitsgrad der Aufgabe ihrem individuellen Leistungs- und Motivationsstand an (vgl. Franke-Braun et al., 2008, S. 29). Gestufte Hilfekarten fördern also eigenverantwortliches Lernen, wie es auch die Kultusministerkonferenz fordert (vgl. Kultusministerkonferenz, 2015, S. 5).

Gerade für leistungsschwächere Schüler stellen sie noch einen weiteren Vorteil dar. Die Schüler können die Hilfekarten selbständig in Anspruch nehmen, ohne die Lehrperson oder den Betreuer fragen zu müssen und müssen so nicht direkt ihr Unwissen zeigen (vgl. Franke-Braun et al., 2008, S. 30).

Eine Studie besagt, dass Schüler sich höchstens fünf Minuten mit einer Teilaufgabe auseinandersetzen. Wird in der Zeit kein Lösungsgedanke entwickelt, wird der Lernprozess abgebrochen (vgl. von Aufschnaiter, 2003, S. 110). Gestufte Hilfekarten können diesem Phänomen entgegen wirken. Mit jeder Stufe wird ein Impuls zur Lösung der Aufgabe gegeben, wodurch die Wahrscheinlichkeit für eine erfolgreiche Bearbeitung steigt.

In der selben Studie wurde ebenfalls herausgefunden, „dass sich Schüler auch in sehr engmaschig angelegten Aufgabenserien als autonom erleben, wenn das Anforderungsniveau gut zu ihren Denk- und Handlungsmöglichkeiten passt." (von Aufschnaiter & von Aufschnaiter, 2001, S. 409). Im Modul gibt es teilweise solche engmaschigen Aufgabenserien. Durch die gestuften Hilfen kann das Anforderungsniveau individuell an die Möglichkeiten der Schüler angepasst werden, sodass die Schüler sich nach von Aufschnaiter (2001) als autonom erleben. Eines der Leitkonzepte der CAMMP days ist es zwar Aufgaben möglichst offen zu gestalten, jedoch gibt es auch Phasen in denen die

Schüler stärker angeleitet werden sollten.

4.2.5. Icons

Durch den Einsatz der vielen verschiedenen Materialien wie Hilfekarten, Besprechungsfolien, Antwortblätter und Live-Scripts können die Schüler leicht den Überblick verlieren. Um dem entgegen zu wirken, wurden verschiedene Icons eingeführt, welche den Schülern eine bessere Orientierung ermöglichen sollen. Die Icons sind in der folgenden Abbildung 23 dargestellt.

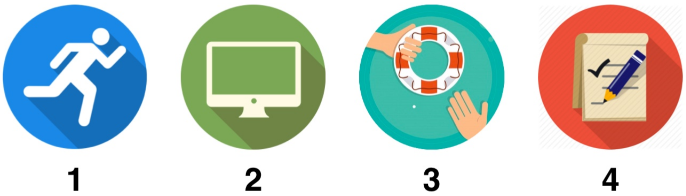


Abb. 23: Im Modul verwendete Icons zur besseren Orientierung. Bedeutung der Icons sind: Icon 1: Bonusaufgaben. Icon 2: Eingabe im Code. Icon 3: Hilfekarte vorhanden. Icon 4: Ausfüllen des Antwortblattes.

Die vier verschiedenen Icons werden alle im Live-Script verwendet. Von dort aus weisen sie den Weg zu den anderen Arbeitsmaterialien. Soll eine Eingabe direkt im Live-Script, also im Code erfolgen, so wird das grüne Icon 2 aus Abbildung 23 verwendet. Gibt es zu der jeweiligen Aufgabe eine gestufte Hilfekarte, so wird dies durch Icon 3 verdeutlicht. Icon 4 wird verwendet, wenn die Schüler etwas auf dem Antwortzettel eintragen müssen. Die Hilfekarten und die Antwortzettel sind in einer Ecke ebenfalls mit dem jeweiligen Icon gekennzeichnet, sodass die Zuordnung erleichtert wird. Icon 1 wird immer am Ende der Arbeitsblätter verwendet. Durch dieses Icon soll angedeutet werden, dass es sich um eine Bonusaufgabe für schnelle Gruppen handelt.

4.3. Aufbau des entwickelten Lernmoduls

In diesem Abschnitt wird der Aufbau des Machine Learning Workshops beschrieben. Dabei sollen die einzelnen Bestandteile inklusive der zugehörigen Materialien vorgestellt werden. Die chronologische Abfolge orientiert sich dabei am geplanten Verlauf des Lernmoduls. Dies soll dem Leser ermöglichen einen guten Eindruck des Ablaufs zu erhalten. Der Modellierungsvortrag (vgl. Abschnitt 4.1.1) ist, abgesehen von der Begrüßung, der erste Bestandteil eines CAMMP days. Im Anschluss startet der eigentliche Workshop, indem die Schüler eine Einführung in die zu Grunde liegende Problemstellung erhalten. Dies wird im erarbeiteten Modul durch eine Einführungspräsentation realisiert.

4.3.1. Einführungspräsentation

Die Einführungspräsentation findet sich im Anhang unter A.6.1 Ziel dieser Präsentation ist es, die Schüler in das Thema Künstliche Intelligenz und Machine Learning einzuführen. Da sie in den ersten zwanzig bis dreißig Minuten weitestgehend zugehört haben, sollen die Schüler durch eine Eröffnungsdiskussion über ihr Vorwissen im Bereich Künstliche Intelligenz zunächst aktiviert werden. Didaktisch liegt hier der Gedanke eines kognitiv aktivierenden Unterrichtseinstiegs zu Grunde (vgl. Rabe, 2014, S. 7). Der Didaktiker Hilbert Meyer (2015) weist guten Unterrichtseinstiegen verschiedene Funktionen zu. Obwohl ein CAMMP day keine Unterrichtsstunde darstellt und die Eröffnungspräsentation somit nicht direkt als Unterrichtseinstieg zu verstehen ist, soll sie sich an einigen dieser Funktionen orientieren. Die Aktivierung von Vorerfahrungen, durch die Eröffnungsdiskussion über den Begriff künstliche Intelligenz ist eine dieser Funktionen (vgl. Meyer, 2015, S. 122). Die Diskussion soll sehr offen gestaltet werden. Die Betreuer sollen die Aussagen der Schüler nicht werten, sondern lediglich als Moderator verschiedene Wortmeldungen annehmen und die Diskussion leiten. Anschließend wird der Begriff Machine Learning über zwei verschiedene Definitionen erklärt. Zunächst wird das Lernen von lebendigen Organsimen nach Zimbardo beschrieben. Anschließend wird diese Definition auf das Lernen von Computern übertragen. Da die meisten Schüler voraussichtlich noch relativ wenig Erfahrung auf dem Gebiet Machine Learning haben, wird als nächstes die Frage geklärt, warum sie überhaupt etwas über das Thema lernen sollten. Mit Hilfe der nächsten Folien soll eine Antwort darauf gefunden und dabei die intrinsische Motivation der Schüler gestärkt werden. Sich nur deshalb mit dem Thema zu beschäftigen, weil der Lehrer das Thema für die Exkursion vorgegeben hat, ist ein schlechter Grund. Auch die Aussagen von bedeutsamen Persönlichkeiten, wie Bill Gates, bezüglich der hohen Bedeutung von Machine Learning sind kaum bessere Argumente. Solche Gründe zielen nur auf die extrinsische Motivation ab. Die Präsentation der verschiedenen und aktuellen Anwendungsbereiche hingegen soll die Schüler von innen heraus motivieren (vgl. Kress & Pappas, 2014, S. 53). Die diskutierten Anwendungsbeispiele wurden bewusst aus dem Umfeld der Schüler gewählt. Auch dies ist didaktisch begründet. Hier lässt sich eine weitere Funktion eines guten Unterrichtseinstiegs nach Meyer wiederfinden. Ein guter Einstieg soll Interesse am Thema erzeugen und neugierig machen (vgl. Meyer, 2015, S. 122). Indem die Anwendungen aus dem unmittelbaren Umfeld der Schüler gewählt wurden, soll das Interesse gefördert werden.

Auf den nächsten Folien werden die für den CAMMP day relevanten Anwendungen, die Gesichtserkennung und das Autonome Fahren genannt. Beide Anwendungsbereiche sind Beispiele für Klassifizierungsprobleme. Den Schülern wird anhand der automatischen Erkennung von Verkehrsampeln erklärt, was man unter einem solchen Problem versteht. Dies soll den Einstieg auf Aufgabenblatt 1 erleichtern, da die Erkennung von Verkehrsampeln hier ebenfalls thematisiert wird (vgl. 4.3.3). Im Folgenden soll kurz das Beispiel beschrieben werden, welches in der Eröffnungspräsentation genutzt wird, um den Begriff zu erklären.

Um eine Software für autonomes Fahren zu entwickeln, sollen rote und grüne Verkehrs-

ampeln voneinander getrennt werden. Die Workshop-Teilnehmer sollen entscheiden, was in dieser Fragestellung die Klasse und was das zu klassifizierende Objekt ist. Außerdem wird diskutiert, wie mögliche Merkmale aussehen könnten. Den Schülern wird die Vereinfachung gezeigt, dass eine Ampel im Modell aus zwei Boxen besteht, die jeweils einen Farbwert annehmen können. Was mathematisch unter einem Klassifizierungsproblem verstanden wird, wurde bereits in Abschnitt 3.2 detaillierter beschrieben. In aller Kürze wird unter einem Klassifizierungsproblem die Zuordnung eines Objektes X in eine Klasse K verstanden. Das Objekt X wird durch einen Merkmalsvektor $\vec{x}$ gegeben. Für die ausgewählte Fragestellung ist eine Verkehrsampel das Objekt X und die Klassen K sind die roten, die gelben und die grünen Ampeln. Die Merkmale, sprich die Einträge des Merkmalsvektors $\vec{x}$, könnten die Farbwerte für die jeweilige Verkehrsampel sein.
An dieser Stelle wird im Vortrag betont, dass für jede Ampel zwei Farbwerte in den Merkamlsvektor eingetragen werden müssen und dass die unterschiedlichen Ampelphasen auch durch die Anordnung der Farben zu unterscheiden sind. Im nächsten Schritt werden die Merkmalsvektoren grafisch als Punkte im zweidimensionalen Raum dargestellt Eine solche Darstellung ist in Abbildung 24 zu sehen.

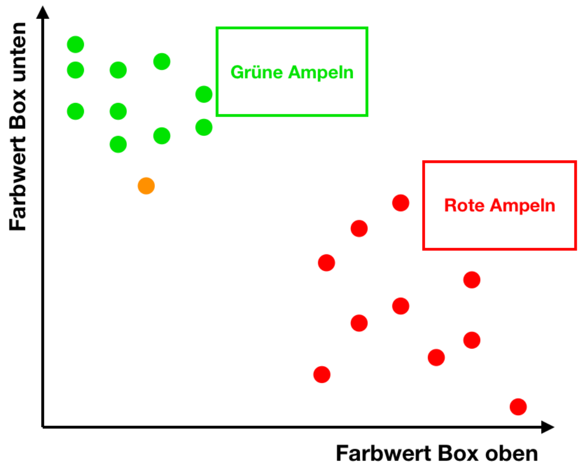


Abb. 24: Dargestellt sind die Merkmalsvektoren repräsentiert durch die roten und grünen Punkte für verschiedene Verkehrsampeln. In orange: Datenpunkt einer zu klassifizierenden Ampel.

Nach Bruner (1974) ist es wichtig Inhalte nicht immer nur auf einer Repräsentationsform darzustellen, damit verschiedene Lerntypen angesprochen werden. Er unterscheidet zwischen der enaktiven, der ikonischen und der symbolischen Ebene (Bruner, 1974, S. 17f). In diesem Fall werden die ikonische und die symbolische Ebene miteinander kombiniert. Auf der ikonischen Ebene werden bildliche Darstellungsformen wie Grafiken oder Tabellen genutzt. Im Workshop wird Abbildung 24 genutzt um die symbolische Darstellung der Merkmalsvektoren mit der ikonischen Ebene zu verknüpfen. Der rote Datenpunkt repräsentiert eine zu klassifizierende Verkehrsampel. Die Teilnehmer des Workshops sollen sie in eine der Klassen einordnen. Dabei ist die Zuordnung, wie in

Abbildung 24 zu sehen, sehr intuitiv. Der Schritt der Einordnung eines Merkmalsvektors in seine Klasse soll zu einem besseren Verständnis der Datenstruktur führen. Der Transfer von einzelnen Merkmalen zu den Vektoren und dann zu den Datenpunkten im Graph ist im Rahmen des Workshops ein wichtiger Prozess. Während des Workshops arbeiten die Schüler oft mit Grafiken, in denen Datenpunkte in dieser Form abgebildet werden.
Oft wird die ikonische auch als Übergang zur symbolischen Ebene eingesetzt (vgl. Jörissen & Schmidt-Thieme, 2015, S. 386). Auf dieser werden schließlich die mathematischen Zeichen benutzt um Formeln aufzustellen. Schülern, die noch wenig Vorwissen im Bereich der Vektoren haben, soll die Darbietung beider Repräsentationsformen ein besseres Verständnis ermöglichen.
In dieser Phase des Einstiegsvortrags sind die Teilnehmer sehr aktiv. Sie sollen anhand des zuletzt behandelten Beispiels den Modellierungskreislauf durchgehen. Dazu wird gemeinsam diskutiert, welche Schritte des Kreislaufs sich an welchen Schritten im Beispiel wiederfinden lassen. Nachdem der Kreislauf bereits im Modellierungsvortrag erklärt worden ist, stellt er hier ein wiederholendes Element dar. Mit dieser Aufgabe wird den Schülern zum einen die Möglichkeit zum aktiven Sich-Erinnern gegeben, zum anderen bekommen sie die Gelegenheit das neue Beispiel in etwas einzuordnen, was sie bereits gehört haben. Nach Heymann sind dies zwei Eigenschaften effizienten Übens und Wiederholens (vgl. Heymann, 2012, S. 8). Da der Workshop nach dem Prinzip des Modellierungskreislaufs aufgebaut ist, ist es hilfreich, wenn die Schüler seinen Ablauf verinnerlichen. Während der Arbeitsphasen sollen sie bei jeder Aufgabe wissen, in welchem Schritt des Kreislaufs sie sich befinden. Dies dient ihnen als Orientierung im umfangreichen Lösungsprozess. Durch die letzte Folie zum Modellierungskreislauf sollen die Schüler in einen Konflikt geführt werden. Dadurch, dass das Objekt im Beispiel nur durch „Hingucken“ einer Klasse zugeordnet wurde, wurde nicht wirklich eine mathematisches Modell gelöst. Für Situationen in denen ein Datenpunkt nicht deutlich einer Klasse zugeordnet werden kann, ist es nicht mehr möglich durch „Hingucken“ eine eindeutige Zuordnung vorzunehmen. Außerdem würde ein solches Vorgehen bei großen Datensätzen zu viel Zeit in Anspruch nehmen, sodass das Verfahren automatisiert werden muss. An dieser Stelle soll den Schülern die Relevanz eines Algorithmus zur Lösung des mathematischen Modells verdeutlicht werden. Mit Hilfe des Algorithmus sollen Datenpunkte unter Zuhilfenahme eines Computers klassifiziert werden. Auch hier werden wieder Grafiken genutzt, um die Problematik zu verdeutlichen und mehrere Repräsentationsformen zu bedienen.
Auf den nächsten Folien wird das Prinzip der Support Vector Machine (SVM) stichpunktartig verdeutlicht. Dabei wird noch nicht auf die mathematische Form, sondern nur auf den Grundgedanken eingegangen. Mit Hilfe der Abbildung aus dem Anfangsbeispiel soll den Schülern verdeutlicht werden, dass eine Trennfunktion gesucht wird, mit der die Datenpunkte zweier Klassen separiert werden können. Hier werden zum ersten Mal die Begriffe *Test-* und *Trainingsdaten* erwähnt. Um sich die neuen Begriffe besser einprägen zu können, werden sie wieder mit einer Grafik kombiniert (vgl. Brzezinska, 2009, S. 47). Am Ende des Vortrags wird die Leitaufgabe des Moduls formuliert. Die Schüler sollen eine SVM zur Erkennung von Verkehrsampeln entwickeln

und diese anschließend auf die Einordnung von einem eigenen Gesichter-Datensatz übertragen. Die Wahl einer solchen Abschlussfolie wurde bewusst getroffen, denn eine weitere Funktion guter Unterrichtseinstiege ist es, die Schüler über das Kommende zu informieren (vgl. Meyer, 2015, S. 122).
Bei der nun folgenden Vorstellung der digitalen Aufgabenblätter in MATLAB, soll wiederholt auf den in Abschnitt 2.2 vorgestellten Modellierungsbegriff verwiesen werden. Die Aufgabenblätter wurden nach dem Schema des Modellierungskreislaufs beziehungsweise der Modellierungsspirale aus Abbildung 2 und 3 in Abschnitt 2.2.2 aufgebaut. Bevor das erste Aufgabenblatt vorgestellt wird, soll dem Leser ein Überblick über die einzelnen Modellierungsschritte während der Arbeitsphasen gegeben werden.

4.3.2. Modellierungsschritte im Überblick

Wie in Kapitel 2 beschrieben, liegt den Lernmodulen in CAMMP ein Modellierungskreislauf, beziehungsweise eine Modellierungsspirale zu Grunde. Die erstellten Materialien lassen sich den einzelnen Stufen der Spirale zuordnen. Dies ist in der folgenden Abbildung 25 dargestellt.

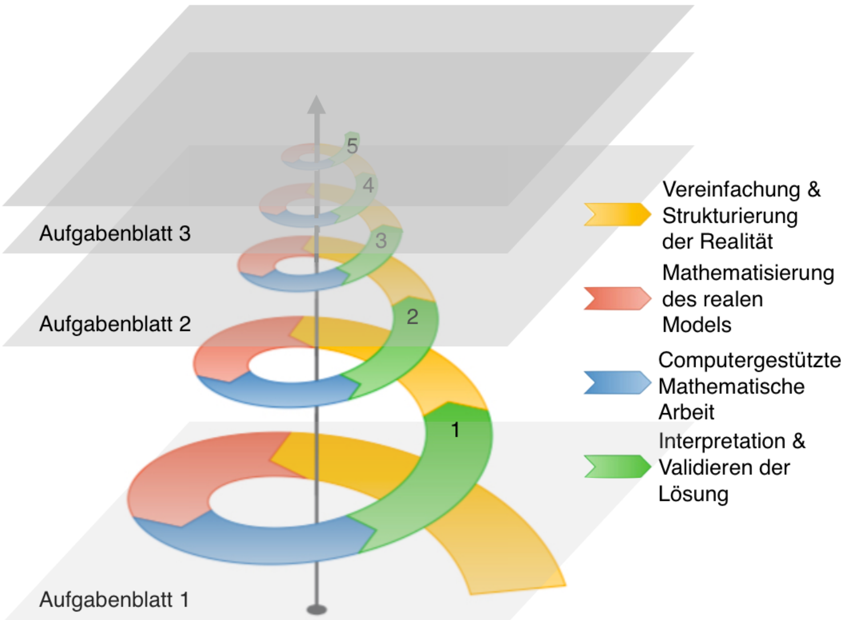


Abb. 25: Modellierungsspirale für das erarbeitete Lernmodul. Ebenen symbolisieren Anfang und Ende eines Arbeitsblattes. Verschiedene Farben für den jeweiligen Modellierungsschritt. Insgesamt werden 5 Durchläufe gemacht.

Die Spirale für das Lernmodul beinhaltet fünf komplette Umdrehungen, welche auf den drei Aufgabenblättern realisiert werden. Eine Umdrehung wird durch die entsprechende Zahl gekennzeichnet. Die Stufen der Spirale werden in vier verschiedenen Farben dargestellt. Diese Stufen sind analog zu den Schritten im Modellierungskreislauf aus Abbildung 3 in Abschnitt 2.2.2, welcher mit den Schülern im Einführungsvortrag

diskutiert wird. Gelbe Stufen stehen für die Vereinfachung und Strukturierung der Realsituation. Bei roten Stufen mathematisieren die Schüler die reale Situation. Befinden die Schüler sich auf einer blauen Stufe, wird die mathematische Lösung mit dem Computer bestimmt. Bei einer grünen Stufe wurde der Kreislauf vollständig durchlaufen. Hier wird das mathematische Ergebnis interpretiert und die Schüler validieren ihre Lösung. An welchen Stellen des jeweiligen Aufgabenblattes die Schüler den entsprechenden Schritt der Spirale durchlaufen, wird in der Vorstellung der Aufgabenblätter in den Abschnitten 4.3.3 bis 4.3.5 erläutert.
Das reale Problem, welches am Anfang der Spirale steht, ist die automatische Klassifizierung von Bildern. Auf dem ersten Arbeitsblatt werden Bilder von Verkehrsampeln gewählt. Während dieses Blattes (vgl. Abschnitt 4.3.3), vollziehen die Schüler die ersten beiden Umdrehungen auf der Modellierungsspirale. Dies wird durch die Ziffern 1 und 2 in Abbildung 25 gekennzeichnet. Im Rahmen des zweiten Aufgabenblattes (vgl. Abschnitt 4.3.4) wird der Kreislauf abermals doppelt durchlaufen, sodass die Schüler zwei weitere Umdrehungen auf der Spirale vollführen. Diese beiden Durchläufe sind durch die Nummer 3 und 4 in Abbildung 25 beschrieben. Schnelle Schüler können das erarbeitete Verfahren noch verbessern, weshalb hier auch drei Durchläufe möglich sind.
Für das dritte Arbeitsblatt (vgl. Abschnitt 4.3.5) wurde ein anderer Kontext gewählt. Das reale Problem ist nach wie vor die automatische Erkennung von Bildern, jedoch sollen nun Gesichter anstatt Ampeln klassifiziert werden. Auf diesem letzten Arbeitsblatt durchlaufen die Schüler ein fünftes Mal den Modellierungskreislauf, was durch die Ziffer 5 in Abbildung 25 angegeben wird.
Es muss betont werden, dass mit einer Umdrehung kein neues mathematisches Modell aufgestellt wird. Das Modell bleibt für den gesamten Workshop die Trennung von Datenpunkten in einem n-dimensionalen Raum. Mit jeder Umdrehung wird jedoch das zu Grunde liegende mathematische Problem komplexer.
Schnelle Schülergruppen erhalten zusätzlich die Möglichkeit im Anschluss an das zweite Arbeitsblatt ein Bonusaufgabenblatt (vgl. Abschnitt 4.3.6) zu bearbeiten. Im Rahmen dieses Blattes wird eine weitere Modellverbesserung eingefügt, sodass solche Schülergruppen sogar sieben Umdrehungen auf der Spirale absolvieren.
In der folgenden Tabelle 4 ist eine Übersicht über die entwickelten Aufgabenblätter zu sehen. Die Übersicht soll einen besseren Überblick über die folgende Vorstellung der Aufgabenblätter ermöglichen. Diese werden in der Tabelle zusammen mit ihren jeweiligen Kerninhalten aufgelistet. Außerdem wird noch angegeben, ob für das jeweilige Aufgabenblatt Hilfekarten zu Verfügung stehen. In der Spalte *Durchläufe Mod. Kreisl.* wird vermerkt, wie viele Durchläufe im Modellierungskreislauf auf dem entsprechenden Aufgabenblatt absolviert werden. Die Angabe '+1' bedeutet, dass im Rahmen der Extraaufgaben noch ein weiterer Durchlauf möglich ist.

Tab. 4: Übersicht über die verwendeten digitalen Aufgabenblätter.

Aufgabenblatt	**Kerninhalte**	**Hilfekarten**	**Durchläufe Mod. Kreisl.**
1	- Klassifizierung roter und grüner Verkehrsampeln - Merkmalsvektoren: 2×1 - Trennfunktion & Abstandsbegriff - primales Optimierungsproblem	ja	2
2	- Klassifizierung roter, gelber und grüner Ampeln - Merkmalsvektoren: 3×1 - Entscheidungsfunktion & Skalarprodukt - Mehrklassen: OVA & OVO	ja	2+1
3	- Klassifizierung von Gesichtern - Merkmalsvektoren: 1782×1 - Training und Test eines eigenen Klassifizierers nach bekanntem Verfahren	nein	1
Bonus	- Klassifizierung roter, gelber und grüner Ampeln - Datensätze nicht linear trennbar - Einführung in Schlupfvariablen & Kostenfaktor	nein	1

4.3.3. Aufgabenblatt 1 | *Farberkennung beim autonomen Fahren*

Der MATLAB Code für Aufgabenblatt 1 findet sich im Anhang unter A.1.1. Das erste Arbeitsblatt beschäftigt sich inhaltlich mit der Einordnung von zweidimensionalen Merkmalsvektoren in zwei verschiedene Klassen. Damit knüpft es direkt an das Beispiel der Verkehrsampeln aus dem Einstiegsvortrag an.
Die erste Aufgabe stellt den ersten Schritt im Modellierungskreislauf dar. Für die reale Situation, das Erkennen von Bildern, wird zunächst ein vereinfachtes Modell entwickelt. Es wird angenommen, dass die Bilder von Ampeln nur aus zwei untereinander angeordneten Pixeln bestehen, die die zwei Boxen einer Ampel repräsentieren. Für rote Ampeln steht die obere Box auf rot und die untere auf schwarz. Für grüne Ampeln steht hingegen die obere auf schwarz und die untere auf grün. Die Schüler sollen dann verschiedene Werte für den Farbcode der einzelnen Boxen ausprobieren und erhalten als Ergebnis entsprechend eine Grafik der daraus resultierenden Ampeln. Ziel ist es, dass sie den Zusammenhang zwischen der Eingabe von Zahlenwerten im Code und der daraus resultierenden Farbe der Ampel erkennen. Die Aufgabe ist sehr frei und spielerisch gestaltet. Die Schüler können ausprobieren, welche Werte sie für realistisch halten und zu verschiedensten Ergebnissen kommen.

In der zweiten Aufgabe wird der zweite Schritt des Modellierungskreislaufs durchlaufen. Die Schüler sollen aus dem vereinfachten realen Modell, ein Mathematisches erstellen. Der Arbeitsauftrag ist zu Beginn sehr offen formuliert. So sollen die Schüler eine Möglichkeit formulieren, wie man die drei RGB-Werte einer Ampelbox in nur einem Wert zusammenfassen kann. Durch diesen Auftrag wird bereits der erste Schritt zur Aufstellung des mathematischen Modells gemacht. Hintergrund ist, dass für jede Box ein Wert abgespeichert werden soll, damit die Ampeln als 2×1-Vektoren dargestellt werde können. Im weiteren Verlauf der zweiten Aufgabe, lernen die Schüler den Grauwert aus der Bildverarbeitung kennen. Er bietet eine Möglichkeit die RGB-Werte zusammenzufassen. Die Schüler sollen sich über den Begriff Grauwert informieren und eine Formel zu seiner Berechnung suchen. Hier werden die Kompetenzen des Problemlösens geschult. Im Kernlerhplan wird unter dem Unterpunkt *Erkunden* gefordert, dass die Schüler selbständig Informationen recherchieren (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2014, S. 20). Die Auswirkung der Formel wird in einer Grafik dargestellt. Sie zeigt die zuletzt eingegebene Ampel aus Aufgabe 1 in Graustufen. Die Merkmalsvektoren, welche die Grauwerte als Einträge beinhalten, stellen den ersten Schritt zum mathematischen Modell des CAMMP days dar. Nach Heymann ist Lernen ein Prozess, sodass nach der Informationsaufnahme ein wiederholtes Bewusstmachen sowie Anwenden des neuen Wissens erfolgen muss (vgl. Heymann, 1998, S. 7). Daher wird die neu erlernte Formel direkt angewendet. Die Schüler sollen verschiedene RGB-Werte von Verkehrsampeln mit Hilfe der Formel in Grauwerte umwandeln. Als Rückmeldung erhalten sie ein Koordinatensystem, in dem die umgewandelten Daten als Punkte dargestellt werden.

In Aufgabe 3 werden die Begriffe Trainings- und Testdaten aus dem Einstiegsvortrag aufgegriffen. Dies wird durch eine Grafik realisiert, in der ein Trainingsdatensatz aus 200 Verkehrsampeln abgebildet wird. Diese Grafik befindet sich jedoch nicht im Live-Script, sondern auf dem entsprechenden Antwortzettel (vgl. Abschnitt 4.2.2).
Die Schüler sollen ohne jegliche Vorgaben eine möglichst einfache Funktion zeichnen, welche die Trainingsdaten der beiden Klassen rot und grün voneinander trennt. Auch hier wird den Schülern zunächst viel Freiheit gelassen, wie die Funktion aussehen soll. Hintergrund ist, dass sie sich bei einer solch offenen Aufgabe schon Gedanken darüber machen, worauf beim Zeichnen einer Trennfunktion geachtet werden muss. Im weiteren Verlauf von Aufgabe 3 werden die Schüler mit dem Problem konfrontiert, dass es mehrere Möglichkeiten gibt die Datenpunkte zu trennen. Die Frage, wie die *beste Trenngerade* ausgewählt wird, ist leitend für den Rest des ersten Arbeitsblattes. Als erste Möglichkeit die Frage zu beantworten, sollen die Schüler verschiedene Testdaten klassifizieren. Dazu wird die eigens erstellte Gerade mit drei anderen Geraden verglichen, die bereits im Code eingespeichert sind. Je nachdem wie gut die selbst erstellte Gerade die Daten einordnet, kann es zu dem Konflikt kommen, dass zwei Geraden alle Testdaten richtig einordnen. Es muss also eine andere Möglichkeit gefunden werden die beste Gerade festzulegen. An dieser Stelle haben die Schüler bereits die erste Umdrehung auf der Modellierungsspirale in Abbildung 25 absolviert. Ausgehend von dem

realen Problem, Bilder zu erkennen, wurde die Vereinfachung getroffen Bilder aus zwei Pixeln zu betrachten. Diese wurden dann als 2×1 Merkmalsvektoren beschrieben. Mit den Vektoren und dem Aufstellen einer Funktionsgleichung für die Trenngerade wurde das mathematischen Modell formuliert. Durch Eingabe der Funktion in den Computer konnten die Schüler dann Testdatenpunkte einordnen. Wie genau die Punkte eingeordnet, werden ist für die Schüler hier noch eine Art Black-Box und wird erst im weiteren Verlauf des ersten Blattes geklärt. Hinter dem Black-Box Ansatz steht die Idee, dass Schüler einen Computeralgorithmus ohne genauere Kenntnis der Funktionsweise nutzen. Im Vordergrund stehen die passende Eingabe und die Diskussion des Resultats, nicht der Algorithmus an sich (vgl. Greefrath & Siller, 2018, S. 7). Nach einmaligem Durchlaufen des Modellierungskreislaufs ergibt sich jedoch das Problem, dass es nicht ausreicht nur die Testdatenpunkte als Kriterium für die beste Trenngerade festzulegen. Mehrere Geraden ordnen alle Testdatenpunkte korrekt ein und es kann nicht entschieden werden, welche die bessere ist. Es muss eine Modellverbesserung erfolgen. Das mathematische Modell, eine Trennfunktion für Datenpunkte im zweidimensionalen Raum zu finden wurde jedoch bereits formuliert. Während dieses Modell für den Rest des CAMMP days gleich bleibt, wird das mathematische Problem komplexer, da im Folgenden nach der optimalen Trennfunktion gesucht wird.

Eine zweite Möglichkeit zu entscheiden, welche Gerade am besten geeignet ist, um die Daten zu trennen wird anhand von Aufgabe 4 und 5 erarbeitet. Hier wird wieder Bezug auf den Verlauf der Geraden genommen. Anhand von vier Beispieldatensätzen, sollen die Schüler zunächst wieder vier Trennfunktionen per Hand einzeichnen. Dabei sollen sie die Punkte umkreisen, welche sie beim Zeichnen als Referenz benutzt haben. Durch wiederholtes Zeichnen sollen die Schüler zu der Erkenntnis gelangen, dass der Bereich um die Trenngerade, in dem keine Datenpunkte liegen (Margin), möglichst groß sein soll. Auch Schüler, die nicht zu dieser Erkenntnis gelangen, werden in Aufgabe 5 darüber informiert. Dadurch können auch sie die nächsten Aufgaben bearbeiten. In dieser Aufgabe wird den Schülern grafisch verdeutlicht, dass der Margin gerade durch den doppelten Abstand zu den nächst gelegenen Datenpunkten gegeben ist. Sie sollen für vier Abbildungen mit jeweils vier verschiedenen Trenngeraden die beste auswählen. Als Orientierung sind für jede Trenngerade jeweils die Abstände zu den nächstgelegenen Datenpunkten aufgelistet. Die Verknüpfung des neu eingeführten Begriffs der Margin mit einer Aufgabe und zugehörigen Grafiken soll zum besseren Verständnis beitragen (vgl. Brzezinska, 2009, S. 47). An dieser Stelle wird zum ersten Mal mit einer direkten Rückmeldung des Codes bezüglich des Ergebnisses gearbeitet. Nur wenn die Schüler für jede Abbildung die richtige Trenngerade auswählen, erhalten sie als Rückgabe `Euer Ergebnis ist korrekt!`. Die kontrollierende Funktion digitaler Werkzeuge im Modellierungsprozess (vgl. Greefrath & Siller, 2018, S. 10) wird noch an weiteren Stellen in den Aufgabenblättern verwendet. Anschließend sollen die Schüler einen Befehl an den Computer formulieren, mit dem er die beste Gerade sucht. Der Suchbefehl soll zunächst nur in Worten formuliert werden. Dieser Schritt dient dazu, die Bedeutung der später aufzustellende mathematische Formel zu verstehen. Gleichzeitig stellt er die Ausgangsbasis für das weitere Vorgehen dar. Es ist notwendig den wörtlichen Befehl

im Folgenden eine mathematische Formel zu übersetzen, damit der Computer ihn verarbeiten kann. In Abschnitt 2.2.2 wurde bereits beschrieben, dass durch Hinzunahme des Computers zusätzliche sprachliche Übersetzungsschritte nötig werden. Genau diese Übersetzungsschritte sollen an dieser Stelle im Lernmodul berücksichtigt werden. Das Aufstellen einer mathematischen Formel für eine reale Situation stellt nach Malle (2013) eine Grundlage des Mathematikunterrichts dar: „Ein Ziel des Algebraunterrichts sollte also darin bestehen, dass Schüler das 'Formelaufstellen' als eine sinnvolle und grundlegende mathematische Tätigkeit erkennen, die nicht weniger sinnvoll bzw. grundlegend ist als 'Rechnen' (oder vielleicht sogar 'Schreiben' und 'Lesen')." (Bürger et al., 2013, S. 56). Indem die Schüler an dieser Stelle eine Formel aufstellen, welche der Algorithmus später zur Klassifizierung nutzen wird, soll dieser didaktische Gedanke nach Malle verfolgt werden. Bezogen auf den Modellierungskreislauf wird hier die Vorarbeit für die Anwendung des bereits aufgestellten mathematischen Modells auf das komplexere Problem der optimalen Trennfunktion geleistet.
In den Aufgabenteilen 5.3 und 5.4 wird zunächst eine mathematische Formel für den Abstandsbegriff aufgestellt. Da der Abstandsbegriff höchst wahrscheinlich im umgangssprachlichen Befehl der Schüler vorkommen wird, ist dies der erste Schritt um den gesamten Befehl zu mathematisieren. Dazu gibt es eine Einführung in die Normalenform für Geraden inklusive Abbildungen. Unter der Nutzung von Hilfekarten soll die Formel für den Abstand eines Punktes von einer Geraden hergeleitet werden. Als Unterstützung erhalten die Schüler eine Übersicht, wie das Skalarprodukt oder Betrag in MATLAB eingegeben werden kann. Auch in Auftrag 5.4 wird wieder mit einer direkten Rückmeldung für das Ergebnis gearbeitet. Mit der neu aufgestellten Formel sollen in Aufgabe 5.5 die Abstände von vier Punkten zu einer vorgegebenen Gerade bestimmt werden. Dadurch soll die Formel an Bedeutung gewinnen. Hier wird wieder der Ansatz nach Heymann verfolgt, nachdem das neue Wissen angewendet werden muss um sich zu festigen (vgl. Heymann, 1998, S. 7). Außerdem sollen die Schüler auf ein Problem aufmerksam gemacht werden. Die aufgestellte Formel liefert für Punkte unterhalb und oberhalb der Geraden jeweils ein anderes Vorzeichen. Für die Entscheidungsfunktion wird diese Eigenschaft auf Aufgabenblatt 2 ausgenutzt. An dieser Stelle soll jedoch der Abstand für alle Datenpunkte optimiert werden. Daher ist es problematisch, dass sich unterschiedliche Vorzeichen ergeben. Dies erfordert eine kleine Änderung der aufgestellten Formel. Didaktisch steckt die Idee der Erzeugung eines kognitiven Konflikts dahinter (vgl. Kircher et al., 2009, S. 181). Ein solcher Konflikt entsteht, wenn das Wahrgenommene nicht mit den bisherigen Erfahrungen übereinstimmt. Im Fall der Aufgabe im Workshop wird der Konflikt durch die unterschiedlichen Vorzeichen erzeugt. Die Schüler kennen für Abstände bisher jedoch nur positive Werte. Das nun auch negative Werte aus ihrer Formel resultieren sollte einen solchen Konflikt erzeugen. Ist die Stärke des Konfliktes nicht zu groß, so werden die Schüler motiviert eine Lösung dafür zu suchen (vgl. Kircher et al., 2009, S. 181).

In Aufgabe 6 sollen die Schüler ihre Formel so ändern, dass nur noch positive Werte resultieren. Da die meisten vermutlich den Betrag nehmen würden, wurde in der Aufgabe noch die Information über die Labels der Datenpunkte hinzugefügt. An dieser

Stelle werden die Schüler etwas stärker gelenkt, da sie sonst vermutlich nicht die Labels als Möglichkeit in Betracht ziehen würden. Der Betrag würde für die Formel zunächst keinen Unterschied machen, auch er liefert nur positive Werte. Es erweist sich für den Optimierungsalgorithmus jedoch als sinnvoll, die Vorinformationen über die Daten in Form der Labels mit einzubeziehen. Würde nur der Betrag verwendet werden, so erhält der Algorithmus keine Information darüber, auf welcher Seite der gesuchten Gerade der Datenpunkt liegen soll. Dadurch kann der Algorithmus eine völlig falsche Gerade aufstellen, welche die Datenpunkte nicht trennt. Bezieht man die Vorinformationen in Form der Labels mit ein, so tritt dieses Problem nicht auf.

Nachdem die Schüler einen Teil ihres wörtlichen Befehls umgeschrieben haben, indem sie den Abstand mathematisch ausgerückt haben, muss das Optimierungsproblem formuliert werden. Ein solches Optimierungsproblem mit Ungleichungen als Nebenbedingung geht über den Schulstoff hinaus (siehe Abschnitt 3.3). Es soll jedoch herausgestellt werden, worin das Problem liegt. Den Schülern wird bewusst gemacht, dass gleichzeitig die Trennfunktion gesucht wird und zu dieser die kürzesten Abstände berechnet werden. Sie sollen das Optimierungsproblem teilweise selbständig formulieren, indem sie die zu optimierende Funktion in MATLAB eingeben. Hierfür erhalten sie wieder gestufte Hilfen und Hinweise auf dem Arbeitsblatt. Die Richtigkeit ihrer Eingabe, können die Schüler wieder grafisch überprüfen. MATLAB liefert die Trenngerade, die sich aus der Optimierung der eingegebenen Funktion ergibt. Die genutzten Optimierungsverfahren werden von den Schülern lediglich als Werkzeuge benutzt. Sie sollen nicht hergeleitet oder selbst programmiert werde. An dieser Stelle wird wieder der oben beschriebene Black-Box Ansatz verfolgt (vgl. Greefrath & Siller, 2018, S. 7). Mit Bezug auf den Modellierungsprozess, leisten die Schüler hier die Übersetzungsschritte nach Greefrath, siehe Abbildung 4 in Abschnitt 2. Sie übersetzen das Problem aus der Alltagssprache in die Mathematische Sprache, indem sie eine Formel für den Abstand formulieren. Diese überführen sie anschließend in die Befehlssprache für die Computereingabe.

Am Ende des ersten Arbeitsblattes gibt es noch Zusatzaufgaben für schnelle Schüler. Dies ist eine Maßnahme, um auf die Heterogenität der Lerngruppe einzugehen. Motivierte, schnelle Schüler können aus zwei verschiedenen Arbeitsaufträgen auswählen. Sie können einerseits ihre Geradengleichung aus Aufgabe 3 so anpassen, dass sie alle Testdaten richtig klassifiziert. Dies können die Schüler mit Hilfe der gewonnen Erkenntnisse erreichen. Indem sie darauf achten, dass die Trenngerade einen möglichst großen Abstand zu den Trainingsdaten hat, werden alle Testampeln richtig klassifiziert. Außerdem gibt es noch eine zweite Wahlaufgabe. Die Schüler sollen die Koordinatenform der Geradengleichung $y = m \cdot x + n$ in die Normalenform umformen. Auch hierfür gibt es Hilfekarten. Während durch die gestuften Hilfen eher die leistungsschwächeren Schüler gefördert werden, können Leistungsstärkere durch schwierigere Extraaufgaben motiviert werden. Dabei ist darauf zu achten, dass die Schüler die Aufgaben selbständig auswählen können (vgl. Hennen, 2008, S. 124); (Storz, 2009, S. 234). Auf diese Weise kann ihnen mehr Selbstverantwortung für das Lernen gegeben werden. Die Schüler „werden als aktive Lernende angesehen, die selbst aus einem Lernangebot auswählen“

(Storz, 2009, S. 234).

Das erste Arbeitsblatt stellt einen doppelten Durchlauf durch den Modellierungskreislaufs dar. Am Ende des Arbeitsblatts haben die Schüler ein Verfahren aufgestellt, welches die optimale Trennfunktion für zweidimensionale Datensätze berechnet. Im Folgenden wird das mathematische Problem komplexer, indem auch dreidimensionale Daten mit mehr als zwei Klassen getrennt werden können.

4.3.4. Aufgabenblatt 2 | *Modellverbesserung: drei Ampelphasen*

Das zweite Aufgabenblatt beginnt in der ersten Aufgabe mit einer Modellverbesserung. Es ist unter A.1.2 im Anhang zu finden. Die Schüler sollen ihr bisheriges Modell von Bildern mit nur zwei Pixeln erweitern, indem sie zusätzlich zur roten und grünen noch die gelbe Ampelphase berücksichtigen. Da die Ampeln nun aus drei Boxen bestehen, erhöht sich auch die Dimension der Merkmalsvektoren. In den ersten beiden Aufgaben sollen die Schüler zunächst einen Eindruck der neuen Datenform bekommen. Sie sollen, ähnlich wie beim ersten Arbeitsblatt, durch die Eingabe verschiedener RGB-Werte gelbe Ampeln erzeugen. Da Gelb eine Mischfarbe ist, ist die Aufgabe jedoch schwieriger als beim ersten Arbeitsblatt. An dieser Stelle wurde bewusst eine bereits bekannte Aufgabenstellung in ihrem Niveau erweitert. Auf diese Weise soll der Eindruck verstärkt werden, dass das bisherige Modell durch die Verbesserung komplexer wird. Die Schüler müssen sich selbständig über additive Farbmischung informieren um mit den RGB Werten gelbe Ampeln zu simulieren. Aus einer Studie von Heinz (2018) zum Einsatz von Tablets im Unterricht geht hervor, dass Schüler durch selbständige Recherche eine stärkere Eigenverantwortung zur Aneignung des Lernstoffs fühlen (vgl. Heinz, 2018, S. 233). Dieser Gedanke wird an verschiedenen Stellen des Workshops umgesetzt.
Ebenfalls in Anknüpfung an das erste Arbeitsblatt, wird den Schülern in der zweiten Aufgabe wieder ein größerer Datensatz aus Verkehrsampeln grafisch präsentiert. Sie sollen überlegen, wie die Datenpunkte im dreidimensionalen Fall getrennt werden können. Dadurch, dass die ersten Arbeitsschritte der beiden ersten Blätter sich stark ähneln, soll der Übergang von zwei auf drei Dimensionen erleichtert werden. Besonders für Mathematik Kurse, welche noch keine Ebenen behandelt haben könnte das zweite Aufgabenblatt ansonsten eine zu große Hürde darstellen.

Auf dem ersten Arbeitsblatt haben die Schüler bereits anhand der Normalenform für Geraden eine Formel für die Trennfunktion und den Abstand aufgestellt. In der dritten Aufgabe des zweiten Arbeitsblattes wird auch dieser Schritt erweitert, indem die Schüler die Entscheidungsfunktion aufstellen sollen. Es wird wieder auf den Ansatz Heymanns zurückgegriffen, indem neuer mit bereits erarbeitetem Lernstoff verknüpft wird (vgl. Heymann, 2012, S. 8). Dazu erhalten die Schüler ähnlich wie beim ersten Blatt zunächst eine kurze Einführung in die Darstellungsform von Ebenen mit Hilfe des Normalenvektors. Die Entscheidungsfunktion sollen sie über eine Fallunterscheidung für das Skalarprodukt aufstellen. Für diesen Auftrag gibt es wieder gestufte Hilfekarten, sowie eine unterstützende Abbildung, die die drei zu betrachtenden Fälle zeigt. Auf diese Weise soll gewährleistet werden, dass auch die leistungsschwächeren Schüler die komplexere Aufgabe lösen können. Die allgemein aufgestellte Fallunterscheidung für das Skalarprodukt soll am Ende von Aufgabe 3 auf die Ebenengleichung in Normalenform übertragen werden. Die daraus entstehende Fallunterscheidung ist für den weiteren Verlauf des Workshops von großer Bedeutung. Sie wird in den weiteren Aufgaben benötigt um unbekannte Datenpunkte einzuordnen. Um dafür zu sorgen, dass alle Schüler die Bedeutung der Fallunterscheidung verstehen, soll auf dem Antwortzettel eine Art Merksatz ausgefüllt werden. Dieser wurde bereits vorformuliert und beinhaltet noch einige Lücken, welche die Schüler ausfüllen müssen. An dieser Stelle formulieren die Schüler zum ersten Mal die wirkliche Entscheidungsfunktion, mit der dreidimensionale Datenpunkte eingeordnet werden können. Das mathematische Modell, welches zu Beginn des ersten Aufgabenblattes aufgestellt wurde wird auf das neue komplexere Problem angewendet.

In Aufgabe 4 soll die neu aufgestellte Entscheidungsfunktion direkt eingesetzt und so ihre Bedeutung unterstrichen werden. Indem die Schüler sie nutzen, um unbekannte Datenpunkte in die entsprechende Klasse einordnen, wird der Nutzen dieser Formel deutlich. Didaktisch wird durch die Anwendung wieder der Gedanke Malles umgesetzt, nach dem die Schüler das Aufstellen einer Formel als sinnvoll erleben sollen (vgl. Bürger et al., 2013, S. 56). Auch hier wurde eine direkte Rückmeldung im Code eingebaut, anhand derer die Schüler erfahren, ob ihr Ergebnis korrekt ist. Die vierte Aufgabe ist gekennzeichnet als Vereinfachung, da zunächst nur die zwei Klassen gelb/rot und grün betrachtet werden. Gelbe und rote Ampel sollen also in die selbe Klasse eingeordnet werden. Dieser Schritt wird eingefügt, damit die Schüler nicht direkt den Transfer von zwei- auf dreidimensionale Daten und gleichzeitig noch von zwei zu drei Klassen leisten müssen. Es soll zunächst getestet werden ob das Verfahren von Blatt 1, welches zwei Klassen mit zweidimensionalen Daten voneinander trennt auch zwei Klassen mit dreidimensionalen Daten trennen kann. Indem die Schüler erkennen, dass ihr Modell auch für dreidimensionale Daten verwendet werden kann, soll der Transfer auf die höheren Dimensionen bei der Gesichtserkennung leichter fallen. Mit Beenden von Aufgabe 4, haben die Schüler im Sinnbild der Modellierungsspirale aus Abbildung 25 wieder eine volle Aufwärtsdrehung gemacht. Sie haben ihr erstes Modell von Aufgabenblatt 1 erweitert, indem sie die Vereinfachungen und Annahmen geändert und dementsprechend das mathematisches Modell auf ein komplexeres Problem angewendet haben. Sie haben dreidimensionale Ampeln klassifiziert und aus der mathematischen Lösung durch die Fallunterscheidung die entsprechende Ampelfarbe bestimmt.

In der fünften Aufgabe wird das mathematische Problem noch anspruchsvoller. Die gelben, roten und grünen Ampeln sollen nun in drei Klassen aufgeteilt werden. Dazu sollen sich die Schüler zunächst ein Verfahren überlegen, mit dem drei Klassen separiert werden können. An dieser Stelle müssen die Schüler kognitive Arbeit leisten um das bereits erstellte mathematische Modell nutzen zu können. Als Unterstützung können sie das Internet nutzen. Gängige Verfahren, auf die die Schüler im Internet stoßen

werden, sind das *One-Versus-One* (OVO) und das *One-Versus-All* (OVA) Verfahren. Sie sollen eines der Verfahren wählen und es selbständig in MATLAB umsetzen. Der Arbeitsauftrag ist die Klassifizierung eines unbekannten Datenpunktes. Es wird eine Situation simuliert, welche der Computer in einem autonom fahrenden Auto lösen müsste. Er erhält das Bild einer Ampel und soll entscheiden, ob sie rot, gelb oder grün ist. Die Schüler erhalten nicht das Bild der Ampel, sondern direkt den zugehörigen Merkmalsvektor mit den Grauwerten. Sie wählen entweder das OVO oder das OVA Verfahren um den Merkmalsvektor in eine der drei Klassen einzuordnen. Als Unterstützung für die Aufgabe werden gestufte Hilfen und zwei Codebausteine angeboten. Mithilfe des ersten Codebausteines können die Schüler nach dem bereits bekannten Verfahren für jeweils zwei Klassen die Trennfunktion berechnen lassen. Sie erhalten dann eine Grafik, die die beiden ausgewählten Klassen, sowie die separierende Ebene zeigt. Außerdem werden die Parameter der Trennebene in einer Tabelle angezeigt. Im zweiten Codebaustein können die Schüler die in Aufgabe 3 aufgestellte Fallunterscheidung eingeben und so das Ergebnis der einzelnen Klassifikationen auswerten. Haben die Schüler den Punkt in die Richtige der drei Klassen eingeordnet, so erhalten sie vom Code eine Animation, in der ein Auto vor einer roten Ampel anhält. Geben sie hingegen die falsche Klasse an, so erhalten sie die Animation eines Unfalls. Auf diese Weise wird die Arbeitsatmosphäre ein wenig aufgelockert. Die Lockerung der Arbeitsatmosphäre durch einen Scherz oder ein Lachen stellt nach Helmke einen wichtigen Bestandteil für ein lernförderlichen Klima dar (vgl. Helmke, 2007, S. 8). Durch die fünfte Aufgabe des zweiten Blattes haben die Schüler sich noch weiter nach oben auf der Modellierungsspirale bewegt. Sie haben die Vereinfachung verworfen, dass nur zwei Klassen betrachtet werden müssen. Anschließend haben sie das mathematische Modell auf das komplexe Problem von drei Klassen übertragen und die Einordnung einer Ampel vorgenommen. Die mathematische Lösung des Verfahrens mussten sie dann interpretieren, um eine Aussage über die Ampelphase treffen zu können.
Der Computer wird in diesem Modellierungsschritt einerseits zum Berechnen, als auch zum Visualisieren und Kontrollieren genutzt. Obwohl die Methode durch das OVO bzw. das OVA Verfahren vorgegeben ist, wird versucht die Aufgabe möglichst offen zu gestalten. Die Schüler sollen selbständig überlegen, wie sie die Informationen aus dem Internet verarbeiten können.

Auch am Ende des zweiten Arbeitsblattes gibt es Bonusaufgaben, welche von den leistungsstärkeren, beziehungsweise schneller arbeitenden Schülern bearbeitet werden können. Hintergrund der Aufgaben ist es, einen Nachteil des OVO Verfahrens zu untersuchen. Durch die drei Trennebenen gibt es meistens einen Bereich, indem die Datenpunkte nicht eindeutig zugeordnet werden können (vgl. Abb 20, Abschnitt 3.2). Die Schüler sollen sich eine Möglichkeit überlegen, dieses Problem zu lösen oder versuchen ein alternatives Klassifizierungsverfahren zu entwickeln. Der offen gestaltete Arbeitsauftrag lässt Spielraum für verschiedene Lösungen. Auf diese Weise ist die Bearbeitungszeit der Bonusaufgabe sehr dehnbar. Dies ermöglicht, dass Schüler je nach Motivation und Tempo unterschiedlich lange an der Aufgabe arbeiten. Zur Recherche über alternative Klassifizierungsmöglichkeiten können sie wieder das Internet nutzen.

Schnelle Schüler können durch die Bonusaufgabe noch eine dritte Modellverbesserung durchführen und somit eine weitere Stufe auf der Modellierungsspirale erreichen, indem sie die Schwäche des OVO Verfahrens beheben. Für das gesamte Lernmodul stellt das Blatt eine wichtige Gelenkstelle dar. Es soll den Übergang zur hochdimensionalen Bildklassifikation auf Blatt 3 ermöglichen. Die Schüler erfahren, dass ihr eigens entwickeltes Verfahren für zweidimensionalen Daten ohne Weiteres auf drei Dimensionen übertragbar ist. Dies soll die Übertragung auf höhere Dimensionen rechtfertigen.

4.3.5. Aufgabenblatt 3 | *Anwendung auf eigenen Datensatz: Gesichtserkennung*

Das dritte Aufgabeblatt stellt eine Anwendung des bisher entwickelten Klassifizierungsverfahrens in einem neuen Kontext dar. Es ist unter A.1.3 im Anhang aufgeführt. Die Schüler arbeiten mit einem selbst erstellten Datensatz. Dieser besteht aus 100 gelabelten 1782×1-Merkmalsvektoren jeder Schülergruppe. Die Vektoren beinhalten die Informationen von 100 Fotos des Gesichts eines der Gruppenmitglieder. Jede Schülergruppe wird als eigene Klasse gesehen, sodass abhängig von der Größe der Lerngruppe 10-15 verschiedene Klassen entstehen. Die Fotos wurden schon vor der Mittagspause des Workshops aufgenommen und in die Vektoren umgewandelt (vgl. Abschnitt 4.3.8). In Bezug auf den Modellierungskreislauf ist das reale Problem nun nicht mehr die Erkennung von Ampelphasen, sondern von Gesichtern.
Zu Beginn des Aufgabenblattes geht es, wie schon bei den ersten beiden Arbeitsblättern darum, Verständnis für die neue Situation zu erlangen. Das mathematische Modell bleibt zwar das Gleiche, jedoch wird das mathematische Problem durch Erhöhen der Dimension der Merkmalsvektoren komplexer. Dazu übertragen die Schüler in der ersten Aufgabe vier vorgegebene Bilder von der Pixel in die Vektorschreibweise. Durch mehrere Fragen zu den Vektoren sollen die Schüler sich ausführlicher mit dem neuen Problem beschäftigen. Es wird beispielsweise gefragt, welche Vorteile und Nachteile Vektoren höher Dimension gegenüber Vektoren mit weniger Einträgen haben.

Um das komplexere Problem zu lösen, sollen die Schüler das erarbeitete Klassifizierungsverfahren von Arbeitsblatt 2 anwenden. Dazu müssen sie als erstes die vorgegebenen Schritte in die richtige Reihenfolge bringen. Die Schritte sind dabei vorgegeben. Durch diesen Auftrag soll abermals betont werden, dass das Vorgehen für den komplexen, hochdimensionalen Gesichter-Datensatz dem für die dreidimensionalen Verkehrsampeln gleicht. Ob die Schritte richtig sortiert wurden, erfahren die Schüler direkt durch eine Rückmeldung im Code. Nachdem die Schüler die Schritte in die richtige Reihenfolge gebracht haben, öffnen sie zu jedem Schritt ein neues Live-Script. Im ersten Schritt teilen sie den Datensatz in Trainings- und Testdaten auf. Im Zweiten nutzen sie die Trainingsdaten um mit dem OVO Verfahren die Trennfunktionen herauszufinden. Der letzte Schritt ist die Einordnung der Testdaten, um eine erste Aussage über den Klassifizierungserfolg treffen zu können. Die Schüler müssen in dieser Arbeitsphase relativ wenige Eingaben in den Code machen, da sie bereits ein Trennverfahren Verfahren für mehrere Klassen erarbeitet haben. Die genauen Eingaben in MATLAB wären an dieser Stelle außerdem kompliziert und nicht zielführend. Im Fokus soll eher das Verständnis für den Ablauf des Verfahrens stehen. Didaktischer Hintergedanke ist an dieser Stelle das bereits erwähnte Black-Box Prinzip (vgl. Greefrath & Siller, 2018, S. 7). Mit dem trainierten Klassifizierer können Gesichter automatisch erkannt werden.

Nachdem der Klassifizierer für den Gesichterdatensatz schrittweise erarbeitet worden ist, sollen die Schüler unbekannte Fotos einordnen lassen. Durch diese Aufgabe soll getestet werden, wie gut der Algorithmus sich auf neue Daten anwenden lässt, die nicht im Test- beziehungsweise Trainingsdatensatz enthalten waren. Von jeder Gruppe wurde ein Foto aufgenommen, auf dem der fotografierte Schüler eine Grimasse schneiden sollte. Diese Fotos werden als unbekannte Daten genutzt. Die Schüler können eine Gruppe auswählen und deren Grimassen-Bild einordnen lassen. Als Ausgabe erhalten sie zunächst das eingelesene Bild im Original und als Pixelgrafik. Außerdem erhalten sie den Namen der zugewiesenen Klasse und ein Beispielbild dieser Klasse. Durch einen Vergleich der Bilder beziehungsweise der Gruppennamen, können die Schüler prüfen ob das *unbekannte Gesicht* korrekt eingeordnet wurde. Abermals wird also eine grafische Rückmeldung genutzt, die die Schüler interpretieren müssen. Werden Bilder falsch zugeordnet, so liefert diese Aufgabe einen guten Anhaltspunkt für die Abschlussdiskussion mit den Schülern. Sie können Schwachstellen des Modells identifizieren und überlegen, wie das Verfahren verbessert werden könnte. Dieser Schritt stellt die letzte Stufe der Modellierungsspirale für das Problem der Gesichtserkennung dar. Die Schüler müssen die Lösung, welche das Programm liefert, interpretieren.

In der letzten Aufgabe des dritten Aufgabenblattes, sollen die Gesichter von Prominenten mit dem erarbeiteten Algorithmus klassifiziert werden. Da die Gesichter der Prominenten eigentlich keiner Klasse korrekt zugeordnet werden können, müssen die Schüler hier Interpretationsarbeit leisten. Sie sollen überlegen, was es bedeutet, wenn ein Prominenter der Klasse eines Mitschülers zugeordnet wird. An dieser Stelle ist der Schritt im Modellierungskeislauf umgedreht. Es wird bereits ein Ergebnis in Form von einem Bild geliefert und dieses soll mathematisch interpretiert werden. Der Interpretationsschritt läuft rückwärts ab. Eine solche invertierte Aufgabenstellung ist angelehnt an das Prinzip einer Umkehraufgabe (vgl. Bruder et al., 2015, S. 527). Bei solchen Aufgaben, wird das gegeben, was bei den 'normalen' Aufgaben immer gesucht ist. Indem die Schüler die inverse Denkoperation vollziehen, kann ein tieferes Verständnis erzeugt werden und der Lerninhalt zugänglicher gemacht werden (vgl. Wiernicki-Krips, 2014, S. 1312). Der Arbeitsauftrag ist offen gestaltet und soll viel Raum für Interpretationen geben. Die Offenheit wird dadurch gegeben, dass keine näheren Fragen oder Hinweise gegeben werden. Die Schüler sollen das Resultat des Computers selbständig interpretieren. Es können interessante Resultate aufkommen, wenn zum Beispiel Heidi Klumm in die Klasse eines männlichen Schülers eingeordnet wird. Solche Ergebnisse liefern viele Gesprächspunkte für die Abschlusspräsentation und die Bewertung des Modells. Die Schüler können überlegen, wie es möglich wäre die Prominenten besser einzuordnen. Außerdem wird die Arbeitsatmosphäre durch diese letzte Aufgabe des Tages etwas entspannt.

4.3.6. Bonus-Aufgabenblatt

Zuletzt soll das Bonus Aufgabenblatt erläutert werden. Das Arbeitsblatt findet sich unter A.1.4 im Anhang. Dieses Blatt kann Schülergruppen ausgehändigt werden, die schon früh mit dem zweiten Arbeitsblatt fertig sind und sich nicht weiter mit den Extraaufgaben am Ende des Blattes beschäftigen wollen. Inhaltlich geht es um eine weitere Modellverbesserung. Die Schüler werden zu Beginn des Blattes mit dem Problem konfrontiert, dass die Trainingsdaten nicht mehr durch lineare Funktionen getrennt werden können. Die Punktwolken der Klassen überschneiden sich (vgl. Abb. 12, Abschnitt 3.2.1). In der ersten Aufgabe sollen sie den bisherigen Algorithmus verwenden um die Daten zu trennen. Sie erhalten als Rückmeldung eine Grafik, in der die Datenpunkte und die berechnete Trennebene eingezeichnet sind. Die Lage der Ebene ist dabei offensichtlich falsch, da sie die Daten nicht voneinander trennt. Um dieses Problem zu lösen, werden die sogenannte *Schlupfvariablen* eingeführt. Die Schüler erhalten dazu Informationsfolien in Form einer Präsentation (vgl. Abschnitt A.6.2 im Anhang), die sie an ihren Laptops durchgehen. Auf den Folien wird die Bedeutung der Schlupfvariablen schrittweise erklärt. Mit Hilfe der Präsentation können die weiteren Aufgaben auf dem Bonusblatt bearbeitet werden. Ziel ist es, dass die Schüler experimentell die Auswirkung des Kostenfaktors C herausfinden, welcher die Schlupfvariablen reguliert. Aus didaktischer Perspektive, ist das Bonusblatt problemorientiert aufgebaut (vgl. Vollrath & Roth, 2011). Kennzeichen von problemorientiertem Lernen ist unter anderem, dass der Lernprozess durch ein Problem in Gang gesetzt wird und die Lernenden weitgehend selbständig daran arbeiten. Durch problemorientiertes Lernen können Begriffe, Sachverhalte und Verfahren gelernt werden (vgl. Vollrath & Roth, 2011, S. 61). Die Schüler werden vor eine Situation gestellt, die sie mit ihren bisherigen Methoden nicht zufriedenstellend lösen können. Im vorliegenden Fall soll durch das Problem der Begriff der Schlupfvariablen gelernt werden. Mit Hilfe der Schlupfvariablen können die Schüler eine Situation lösen, die vorher für sie nicht zu bewältigen war. Dies ist ein Ansatz nach Leisen (2007), wonach den Schülern zunächst eine Schwierigkeit präsentiert wird, dessen Lösung zu einer gewinnbringenden Einsicht führt (vgl. Leisen, 2007, S. 87). Außerdem soll das Verfahren, welches bereits auf den ersten beiden Arbeitsblättern erlernt wurde, verbessert werden. Ein weiteres Kennzeichen von problemorientiertem Lernen ist der Erkenntnisgewinn (vgl. Vollrath & Roth, 2011, S. 63). Der Erkenntnisgewinn ist in diesem Fall, dass nicht alle Trainingsdatenpunkte korrekt klassifiziert werden müssen, um einen brauchbaren Klassifizierer zu erhalten.
Durch das Bonusaufgabenblatt wird das zu Grunde liegende mathematische Problem ein weiteres Mal in seiner Komplexität gesteigert.

4.3.7. Abschlusspräsentation

Als letzter Unterabschnitt zum Aufbau des Lernmoduls soll die Abschlusspräsentation des Moduls vorgestellt werden. Ziel dieser Präsentation ist es die Ergebnisse des Workshops zusammen zu fassen und diese auch aus nicht-mathematischen Perspektive zu diskutieren. Sie ist unter A.6.3 im Anhang angefügt.

Auf den ersten beiden Folien wird den Schülern gezeigt, an welcher Stelle sie in den Workshop gestartet sind und was sie nach 6 Stunden Arbeit erreicht haben. Dazu wird die gleiche Folie wie in der Einstiegspräsentation genutzt. Anhand des Modellierungskreislaufs sollen die Schüler dann reflektieren, wie sie zu dem Ergebnis gekommen sind. Der Betreuer gibt die reale Situation zu Beginn des ersten Arbeitsblatts vor. Es sollte ein Verfahren entwickelt werden, mit dem Verkehrsampeln automatisch erkannt werden können. In einer offenen Gesprächsrunde, sollen die Schüler dann die weiteren Schritte zuordnen. Wurde der Kreislauf einmal durchlaufen, also die Hälfte des erste Aufgabenblatt reflektiert, so wird die Modellverbesserung diskutiert. Die Schüler sollen überlegen an welchen Stellen des Kreislaufs Verbesserungen und Änderungen durchgeführt worden sind. Durch diesen Prozess wird die zweite Hälfte des ersten und das zweite Arbeitsblatt reflektiert. In der Präsentation sind die entsprechenden Änderungen zum ersten Durchlauf farblich markiert, sodass transparent wird, an welchen Stellen die Änderungen gemacht wurden. Da das dritte Arbeitsblatt bereits unmittelbar vor dem Abschlussvortrag besprochen wurde, soll es hier nicht mehr umfassend reflektiert werden. Indem auf den Folien betont wird, dass die Schüler die Tagesziele des Workshops erreicht haben, soll die Arbeit der Schüler wertgeschätzt werden. Die Wertschätzung von ihren Beiträgen liefert einen wichtigen Beitrag zur Lernfreude und Motivation der Schüler (vgl. Gold, 2015, S. 80); (Börlin, 2012, S. 185). Die Teilnehmer sollen den Workshop mit einem positiven Gefühl beenden, weshalb diese Folien gegen Ende der Abschlusspräsentation gewählt wurden. Durch die beiden Zitate auf der letzten Folie sollen sich die Schüler kritisch mit dem Begriff des Machine Learnings beziehungsweise der künstlichen Intelligenz auseinandersetzen. Das erste Zitat zielt auf eine positive Sichtweise von Machine Learning ab: „In the next 10 years, data science and software will do more for medicine than all of the biological sciences together“. Es stammt von Vinod Khosla, einem US-amerikanischen Geschäftsmann und Mitgründer einer großen Softwarefirma. Das zweite Zitat hingegen beinhaltet eine kritische Sichtweise dieses Begriffs und stammt von Stephen Hawking: „The development of full artificial intelligence could spell the end of human race“. Die beiden Zitate, aus jeweils entgegengesetzten Positionen, sollen die Diskussion erleichtern. Die Schüler verlassen die mathematische und diskutieren eher aus einer ethischen, gesellschaftskritischen Perspektive. Es soll bewusst werden, dass der technische Fortschritt nicht nur einseitig betrachtet werden darf.

4.3.8. Schülerdaten für den Gesichterdatensatz

Für das dritte Aufgabenblatt (vgl. 4.3.5) sollen die Schüler mit einem eigens erstellten Gesichterdatensatz arbeiten. Dazu sollen während des Workshops Fotos der Schüler aufgenommen werden. In jeder Zweiergruppe wird ein Schüler ausgewählt von dem mit Hilfe der Webcams der Laptops 100 Fotos für einen Beispieldatensatz aufgenommen werden. Außerdem wird ein zusätzliches Foto aufgenommen, auf dem der Schüler eine Grimasse schneiden soll. Dieses 101. Foto ist nicht Teil des Beispieldatensatzes. Die 101 Fotos werden von den Schülern mittels MATLAB in eine `.mat`-Datei umgewandelt. Diese senden sie dann über AirDrop an den Laptop der Dozenten, die den Workshop durchführen. Für das Versenden und Empfangen der Dateien wird eine Anleitung über den Beamer gezeigt. Der Betreuer verarbeitet die empfangenen Dateien dann mit einem speziellen MATLAB Script. Dieses Script fügt alle Schülerdateien zusammen, sodass ein Beispieldatensatz aus 100 Fotos jeder Schülergruppe entsteht. Diesen Datensatz sendet der Betreuer per AirDrop an die Schülergruppen, sobald sie Aufgabenblatt 3 bearbeiten. Es wurde bewusst AirDrop ausgewählt, da die Fotos auf diese Weise nicht in einer Cloud gesichert werden müssen.

5. Durchführung und Evaluation des Lernmoduls

In diesem Abschnitt sollen die ersten beiden Durchführungen des Lernmoduls evaluiert werden. Dazu wird zum einen die von den Schülern ausgefüllte Online-Evaluation ausgewertet. Zum anderen wird auf Beobachtungen und Auffälligkeiten eingegangen, die während des Workshops gemacht worden sind. Eine dritte Quelle für die Diskussion der Ergebnisse, liefern die Rückmeldungen der anderen Dozenten, die die Durchführungen mit betreut haben, der Lehrkraft, die die Gruppen begleitet hat und schließlich meiner Betreuerin Sarah Schönbrodt, die der ersten Durchführung als Beobachterin beigewohnt hat. Alle Ergebnisse verstehen sich als Anhaltspunkte für die Weiterentwicklung und Verbesserung des Lernmoduls und nicht als empirisch belegbare Studie.

5.1. Online-Evaluation

Am Ende eines jeden CAMMP days füllen die Schüler einen Online-Evaluationsbogen aus. Ein unausgefülltes Exemplar befindet sich im Anhang unter A.7.1. Nach einem kurzen Abschnitt über personenbezogene Daten können die Schüler im zweiten Teil der Evaluation allgemeine Fragen zum Workshop beantworten. Es wird beispielsweise nach der Handhabung von MATLAB oder dem Schwierigkeitsniveau des Moduls gefragt. Die einzelnen Aspekte werden durch Aussagen repräsentiert. Die Schüler können für jede Aussage aus den vier Optionen *Trifft voll zu (++)*, *Trifft zum Teil zu (+)*, *Trifft eher nicht zu (-)* und *Trifft überhaupt nicht zu (- -)* wählen. Das Vorgehen orientiert sich somit an einer vierstufigen Likert-Skala.
Im dritten Teil werden spezifischere Fragen zum Workshop gestellt. Die Schüler können hier beispielsweise angeben, was ihnen besonders schwer gefallen ist, oder wo sie sich mehr Unterstützung gewünscht hätten. Neben diesen Fragen, bei denen Freitextfelder ausgefüllt werden können, werden wieder Aussagen vorgegeben. Zu diesen Aussagen geben die Schüler, wie schon im zweiten Teil der Evaluation den Grad ihrer Zustimmung an. Da es sich um die ersten beiden Durchführungen des entwickelten Moduls handelt, wurden in diesem Teil noch Fragen und Aussagen ergänzt, die standardmäßig nicht in der CAMMP day Evaluation vorkommen. Auf diese wird im folgenden Abschnitt 5.2 zu den leitenden Gesichtspunkten eingegangen.
Der vierte Teil der Evaluation beinhaltet die abschließende Bewertung des Lernmoduls. Die Schüler können dem Modul und den Betreuern jeweils eine Schulnote geben und in einem Freitextfeld Lob, Kritik, sowie Verbesserungsvorschläge äußern.

5.2. Leitende Gesichtspunkte der Evaluation

Vor der ersten Durchführung eines neuen CAMMP days ist noch ungewiss an welchen Stellen die Schüler Schwierigkeiten haben werden oder wo Aufgaben noch nicht präzise genug formuliert sind. Um eine genauere Rückmeldung zu erhalten, wurden deshalb im Vorhinein leitende Gesichtspunkte für die Beobachtung und die Online Evaluation ausgewählt, auf die ich während der Durchführung besonders geachtet habe. Diese Gesichtspunkte lauten:

- Motivation und Interesse
  Inwieweit waren die Schüler am Thema Machine Learning interessiert? Wie äußerte sich dieses Interesse? Inwieweit konnte das Interesse eventuell während des Tages gesteigert werden? Konnten die Schüler motiviert werden, sich mit dem Thema zu beschäftigen?

- Verständnis
  Inwiefern haben die Schüler den mathematischen Hintergrund der SVM verstanden? Ist ihnen klar, wie die Entscheidungsfunktion zu Stande kommt? Ist ihnen der Dimensionswechsel von zwei, auf drei, auf n Dimensionen leicht gefallen? Haben sie verstanden, wie mehrere Klassen mittels SVM getrennt werden können?

- Arbeitsmaterialien
  Ist der Schwierigkeitsgrad der Aufgaben angemessen? Inwiefern werden die leistungsstarken Schüler durch Extraaufgaben gefordert? Können die Aufgaben auch durch die leistungsschwächeren Schüler gelöst werden, indem sie die gestuften Hilfen benutzen? Inwieweit waren die Aufgaben verständlich formuliert, sodass wenig Nachfragen zu den Aufgabenstellungen aufkamen?

- MATLAB
  Inwiefern gibt es Schwierigkeiten im Umgang mit den Live-Scripts? Sind die zu Verfügung gestellten Eingabehilfen ausreichend?

- Modellierungskompetenz
  Inwieweit konnten die Schüler ihre Modellierungskompetenz durch das Lernmodul verbessern? Wird den Schülern durch das Modul bewusster, welch wichtige Rolle mathematische Modellierung spielt? In welchem Maße konnten sie während des Moduls ihre Position im Kreislauf zuordnen?

- Arbeitsweise
  Inwieweit haben die Schüler eigenständig gearbeitet? In welchem Umfang wurden die gestuften Hilfekarten verwendet?

5.3. Die erste Durchführung

5.3.1. Rahmenbedingungen

Das Lernmodul wurde am 29.10.2018 zum ersten Mal durchgeführt. Der teilnehmende Kurs war ein Mathematik Leistungskurs aus der Qualifikationsphase II. Die Schüler besuchen die Gesamtschule Niederzier-Merzenich. Insgesamt bestand die Gruppe aus zwei Schülerinnen, 16 Schülern und einem Lehrer. Der Workshop wurde von Nils Steffen, einem Mitarbeiter aus dem CAMMP-Team, und mir durchgeführt. Als Beobachterin war zudem Sarah Schönbrodt anwesend.
Aufgrund von Verzögerungen bei der Anreise, konnte der CAMMP day erst um 9:15 Uhr, anstatt um 9:00 Uhr beginnen. Da die Gruppe jedoch flexibel in der Abreisezeit war, konnte der Workshop um eine Viertelstunde nach hinten verlängert werden, sodass

wie geplant sechs Stunden für die Durchführung zu Verfügung standen. In den sechs Stunden hatten die Schüler eine längere Mittagspause von 12:15 Uhr bis 13:15 Uhr. Für den Rest der Zeit haben sie sehr diszipliniert und konzentriert gearbeitet. Dies äußerte sich zum einen durch konstruktive Diskussionen über die Aufgaben in den Zweiergruppen. Zum anderen war die Gesprächslautstärke angemessen, sodass konzentriert gearbeitet werden konnte. Außerdem musste kein Schüler ermahnt werden, weil er sich mit seinem Handy oder Ähnlichem abgelenkt hat. An dieser Stelle kann bereits festgehalten werden, dass es keine großen Probleme während des Workshops gab und die erste Durchführung sehr positiv verlief. Der MATLAB Code lief bei korrekter Eingabe fehlerfrei durch und im Falle von kleineren Problemen bezüglich der Handhabung des Codes, konnten die Betreuer schnell weiterhelfen. Ein Knackpunkt bei der ersten Durchführung war jedoch die Zeit. Obwohl bereits im Vorfeld auf die *Powerwall*[11] verzichtet wurde, welche normalerweise Teil der CAMMP days ist, musste zusätzlich noch der Studieninformationsvortrag gestrichen werden. Daraus ergibt sich als Kernziel für die zweite Durchführung, den Aufgabenumfang ein wenig zu kürzen, um die Bearbeitungsdauer zu reduzieren. Die Materialien für die zweite Durchführung weichen deshalb leicht von denen in Abschnitt 4.3 vorgestellten ab. Darauf wird jedoch ausführlicher im Abschnitt 5.3.3 eingegangen.
Zum Vorwissen der Schülergruppe, welche die erste Durchführung absolviert haben, muss angemerkt werden, dass Ebenen bereits in der Schule behandelt wurden. Die Schüler des Q2 Leistungskurses haben folglich schon Kenntnisse über die Normalenform einer Ebene. Dies ist besonders bei der Bearbeitung des zweiten Arbeitsblatts (vgl. Abschnitt 4.3.4) hilfreich.

5.3.2. Beobachtungen anhand der leitenden Gesichtspunkte

Die genauen Ergebnisse der Online-Evaluation können im Anhang unter A.7.2 nachgesehen werden. Im Folgenden sollen die Ergebnisse der Beobachtungen und der Online Evaluation zu den leitenden Gesichtspunkten vorgestellt werden.

Motivation und Interesse:
Während der Arbeitsphasen konnte beobachtet werden, dass die Schüler konzentriert gearbeitet haben. Dies äußerte sich durch intensive und angeregte Diskussionen über die Aufgaben in den Zweiergruppen. Die Motivation der Schüler wird nicht zuletzt durch die Lern- und Arbeitsatmosphäre bestimmt. Im Fragebogen stimmten 94 % der Befragten der Aussage zu (++), dass diese angenehm war. Dies deutet darauf hin, dass das Interesse am Thema hoch war. Vor allem in der Zeit bis zur Mittagspause konnte dies festgestellt werden. Nach der Pause ließ die Motivation ein wenig nach. In dieser Arbeitsphase ließen sich die Schüler schneller ablenken und redeten über andere Themen, als die Aufgaben. Auch die Arbeitsgeschwindigkeit ließ ein wenig nach.

[11] Eine $7.5m^2$ große Projektionsfläche der AICES Graduiertenschule, auf der 3D-Modelle visualisiert werden können. Schüler sollen so einen Einblick in das Forschungsgebiet der Virtuellen Realität erhalten.

In der Online-Evaluation wurde dieser Gesichtspunkt durch die beiden Aussagen *„Das Thema des Kurses finde ich interessant.“* und *„Der Kurs hat mein Interesse an Themen der Naturwissenschaft und Technik gesteigert.“* abgefragt. Erstere bewertete nur ein Schüler mit *Trifft eher nicht zu (-)*. Der Rest des Kurses teilt sich gleichmäßig auf *Trifft zum Teil zu (+)* und *Trifft voll zu (++)* auf. Dies bestätigt die Beobachtungen, die während der Arbeitsphasen gemacht wurden. Auf die zweite Frage hin, ob das Interesse durch das Moduls noch gesteigert wurde, gaben 75 % der Teilnehmer eine Rückmeldung in den oberen beiden Bereichen (+) und (++).

Verständnis:
Zum Gesichtspunkt Verständnis wurden dem Standard-Evaluationsbogen spezielle Fragen hinzugefügt, die das Verständnis der Schwerpunkte der einzelnen Aufgabenblätter abfragen sollten. Auf dem ersten Blatt lag der Schwerpunkt auf dem Abstandsbegriff. Lediglich zwei Schüler gaben bei der Aussage *„Ich habe auf Blatt 1 verstanden, dass durch den Abstandsbegriff die optimale Trennfunktion bestimmt wurde.“* an, dass dies eher nicht zu treffe (-). Der Rest der Befragten gab an dass dies teilweise (+) oder voll (++) zutreffe. Im Zusammenhang mit dem Abstandsbegriff, gab außerdem ein Schüler an, dass ihm die Herleitung besonders schwierig gefallen ist. Darauf soll unter dem nächsten Gesichtspunkt noch genauer eingegangen werden. Knackpunkte auf dem zweiten Arbeitsblatt sind die Modellverbesserung durch den Übergang von zwei auf drei Dimensionen, die Fallunterscheidung für die Entscheidungsfunktion und die weitere Modellverbesserung, durch die Einteilung in drei verschiedene Klassen. Dreiviertel der Schüler gaben an, dass ihnen der Dimensionswechsel leicht gefallen sei. Inwiefern dieser Aussage bei der zweiten Durchführung zugestimmt wird, wenn die Gruppe noch nicht über ein solches Vorwissen über Ebenen verfügt, wird in Abschnitt 5.4.2 deutlich. Alle Befragten gaben an, dass sie die Fallunterscheidung für die Entscheidungsfunktion verstanden hätten. Es wurden ausschließlich die Rückmeldungen (+) und (++) gegeben, sodass davon ausgegangen werde kann, dass dieser entscheidende Punkt auf Blatt 2 gut vermittelt werden konnte. Ähnlich positiv fällt die Auswertung für die Aussage *„Ich habe auf Arbeitsblatt 2 verstanden, wie mehrere Klassen voneinander getrennt werden können.“* aus. Hier gab nur ein Schüler eine negative Rückmeldung (-). Der Rest der Lerngruppe teilt sich gleichmäßig auf die oberen beiden Bereiche auf. Im Rahmen des dritten Blattes, sollten die Schüler erkennen, dass ihre bisherigen Methoden für zwei, beziehungsweise dreidimensionale Datensätze einfach auf höhere Dimensionen übertragen werden können. Über 80 % der Befragten gaben an, dass sie dies durch den Wechsel von zwei auf drei Dimensionen verstanden hätten. Alle Schüler gaben an, dass ihnen klar geworden sei, wie Bilder mit einem Computer automatisch klassifiziert werden können. Dies spricht dafür, dass die Schüler ein hohes Verständnis für das gesamte Lernmodul entwickeln konnten. Der Eindruck wird dadurch unterstützt, dass lediglich zwei Schüler angaben, sie hätten das Prinzip der SVM eher nicht verstanden (-).

Arbeitsmaterialien:
Die bei der ersten Durchführung genutzten Arbeitsblätter wurden bereits ausführlich im Abschnitt 4.3 vorgestellt. Aufgrund des positiven Feedbacks zu diesem Gesichts-

punkt mussten sie nicht in ihrer Formulierung geändert werden. Lediglich der Umfang wurde etwas reduziert. Um den Schwierigkeitsgrad der Aufgaben bewerten zu können, wurden die beiden Aussagen *„Die Aufgaben waren zu schwierig.“* und *„Die Aufgaben waren zu einfach.“* genutzt. Alle Schüler gaben an, dass die Aufgaben überhaupt nicht (- -), beziehungsweise eher nicht (-) zu einfach wären. Im Gegenteil: gab etwa die Hälfte des Kurses an, die Aufgaben seien eher zu schwierig gewesen. Dieses Urteil konnte jedoch etwas relativiert werden, da über 70 % der Teilnehmer angaben, dass ihnen nichts besonders schwierig gefallen sei.
Auch der Einsatz der gestuften Hilfekarten muss unter diesem Gesichtspunkt berücksichtigt werden. Alle Schüler gaben an, dass diese hilfreich gewesen seien und nur zwei Teilnehmer hätten sich noch mehr Unterstützung gewünscht. Dies lässt darauf schließen, dass die Aufgaben tendenziell etwas über dem Leistungsniveau der Schüler angeordnet waren, sie jedoch mit den zu Verfügung stehenden Hilfen gut bearbeitet werden konnten. Genau dies ist das Ziel hinter dem Einsatz gestufter Hilfen und daher als positiv zu bewerten. Der Eindruck wurde durch die Antworten im Freitextbereich und durch mündliche Rückmeldungen bestätigt. Ein Schüler gab beispielsweise an, dass ihm die Hilfestellungen besonders gut gefallen hätten. Der begleitende Lehrer, welcher ebenfalls das Modul bearbeitet hat, merkte an, dass er die gestuften Hilfen als sehr hilfreich und lohnenswert empfunden habe. Inwieweit die leistungsstärkeren Schüler durch die Extraaufgaben gefördert wurden, lässt sich nach der ersten Durchführung nicht beurteilen. Aufgrund der bereits beschriebenen Zeitnot, schaffte es keine Schülergruppe bis zu den Extraaufgaben. Auch das Bonus-Aufgabenblatt (vgl. 4.3.6) wurde nicht bearbeitet und kann daher nicht evaluiert werden.
Der Einsatz der Antwortzettel in Kombination mit den Live-Scripts wurde von der begleitenden Lehrkraft und den Dozenten sehr positiv empfunden. Während der Besprechung der Blätter konnten die Laptops zugeklappt werden, was ein besseres Fokussieren auf die Besprechung ermöglichte. Leider konnten die Schülerfolien nicht eingesetzt werden, um die Ergebnisse der Gruppen zu präsentieren. Dies war dem Zeitmangel geschuldet und wurde bei der zweiten Durchführung verbessert.

MATLAB:
Während den Arbeitsphasen konnten kleinere Probleme im Umgang mit MATLAB beobachtet werden. Besonders die Eingabe der mathematischen Operatoren stellte sich als Schwierigkeit heraus. Über die Hälfte der Schüler gab in der Evaluation an, dass sie teilweise Schwierigkeiten mit dem Programm hatten. Um dem entgegen zu wirken, wurde bei der zweiten Durchführung während der Arbeitsphasen eine Übersicht über die Operatoren und einige Befehle in MATLAB per Beamer gezeigt. Außerdem soll eine kurze Einführung in das Programm zu Beginn des Workshops erfolgen. Eine weitere Veränderung wurde auf Aufgabenblatt 1 gemacht. Hier sollten die Schüler bei der ersten Durchführung eine Formel für die Grauwert-Berechnung eingeben. Um bei MATLAB eine Formel einzugeben, müssen die unabhängigen Variablen, beispielsweise x,y und z, mit `@(x,y,z)` erklärt werden. Obwohl dies anhand eines Beispiels vorgeben war, führte es bei vielen Schülern zu Unklarheiten. Diese, für das Verständnis unnötige, Hürde wurde für die zweite Durchführung entfernt. Die Variablen, welche benutzt

werden sollen, werden nun bereits in der Klammer vorgegeben, sodass nur noch die eigentliche Formel eingegeben werden muss.

Modellierungskompetenz:
Die Förderung der Modellierungskompetenz stellt eines der Kernziele eines CAMMP days dar. Die Verwendung, des in Abschnitt 2.2.2 beschriebenen didaktischen Modellierungskreislaufs (vgl. Abb. 3), sollte den Schülern eine bessere Orientierung während des Modellierungsprozesses ermöglichen. Ziel war es, dass die Schüler zu jeder Zeit wissen, an welcher Stelle des Kreislaufs sie sich befinden. Dieses Ziel konnte erreicht werden. Im Rahmen der Abschlusspräsentation, sollten die Schüler den Modellierungstag reflektieren und die einzelnen Arbeitsschritte an Stellen des Kreislaufs einordnen. Es zeigte sich dabei, dass die Schüler die einzelnen Schritte vollständig einordnen konnten. So stellte sich bei der Besprechung heraus, dass die Phasen von den Schülern genau so zugeordnet wurden, wie es im didaktisch-methodischen Konzept in Abschnitt 4.1 geplant worden ist. Außerdem war auffällig, wie viele Schüler sich zu dieser Diskussion gemeldet haben. Der Eindruck wird dadurch verstärkt, dass Dreiviertel der Teilnehmer in der Evaluation angegeben haben, dass sie Mathematische Modellierung durch das Modul besser verstanden hätten. Weitere 20 % der Schüler haben dieser Aussage teilweise zugestimmt. Nur ein Schüler stimmte dieser Aussage eher nicht zu (-). Außerdem empfand der gesamte Kurs den Modellierungsvortrag der Doktorandin zu Beginn des Workshops als hilfreich.

Arbeitsweise:
Während der Arbeitsphasen haben die Schüler weitestgehend selbständig gearbeitet. Die meisten Fragen, bei denen ein Betreuer helfen musste, bezogen sich auf MATLAB. Der Lehrer gab die Rückmeldung, dass sich auch solche Schüler in den Besprechungsphasen gemeldet hätten, welche im Schulunterricht eher still sind. Dies lässt sich möglicherweise auf die Arbeitsweise im Team und die Verwendung der gestuften Hilfen zurückführen. Einzelne Schüler meldeten zurück, dass sie durch die Kontrollmöglichkeiten mit Hilfekarten und MATLAB mehr Vertrauen in die Korrektheit ihrer Lösung hatten. Auch die Arbeitsweise, mit einem Partner zusammen zu arbeiten, wurde gelobt. Was bei der ersten Durchführung auffiel, war die Art und Weise, wie die Hilfekarten genutzt wurde. Die Idee hinter dem gestuften System ist es, dass die Schüler sich zunächst die untersten Stufe an Unterstützung nehmen. Sollte diese nicht ausreichen, können sie sich die nächste Stufe nehmen. Viele Schüler haben jedoch direkt mehrere Hilfe-Stufen auf einmal genommen. In der zweiten Durchführung sollte die Nutzung der Hilfekarten deshalb zu Beginn besser erklärt werden.

5.3.3. Resultierende Verbesserungen und Fazit

Wie im vorherigen Abschnitt zu den leitenden Gesichtspunkten schon bemerkt wurde, verlief die erste Durchführung des Moduls durchaus positiv und aus meiner Sicht als Dozent sehr zufriedenstellend. Aus den Beobachtungen und Rückmeldungen resultieren nur kleine Änderungen in den Materialien. Das größte Problem stellte der zeitliche

Rahmen dar, weshalb der Arbeitsumfang reduziert werden musste. Infolgedessen wurde Aufgabe 7 des ersten Blattes mit in den Bonusteil aufgenommen. In Aufgabe 7 sollten die Schüler die zu optimierende Funktion so umschreiben, dass sie alle Optimierungsparameter enthält und bei MATLAB eingegeben werden kann. Dieser Teil fällt nun bei den Pflichtaufgaben weg. Das Optimierungsverfahren wird somit stärker als Black-Box eingesetzt. Für interessierte Schüler bietet sich weiterhin die Möglichkeit im Rahmen der Bonusaufgaben genauer zu untersuchen, wie das Verfahren funktioniert. Aufgabe 2 des ersten Blattes wurde in ihrem Umfang reduziert. Inhaltlich sollen die Schüler hier die recherchierte Grauwert-Formel nutzen, um verschiedene Ampeln in ihren Grauwerten darzustellen und die Merkmalsvektoren als Punkte in einem Koordinatensystem visualisieren. Dazu sollten sie bei der ersten Durchführung acht Ampeln umwandeln. Die Anzahl wurde für die zweite Durchführung auf vier gesenkt. Bei der ersten Durchführung ist außerdem aufgefallen, dass der Dateiaustausch via AirDrop (vgl. 4.3.8) nicht gut funktioniert hat. Infolgedessen, wird auf zwei USB-Sticks zurück gegriffen, mit denen die Bilder der Schüler gesammelt werden. Außerdem wurden die bereits beschriebenen Änderungen bezüglich MATLAB eingebaut.

Die Schüler bewerteten den CAMMP day größtenteils mit der Schulnote gut und sehr gut. Lediglich drei der 18 Teilnehmer bewerteten ihn mit befriedigend. Nur ein Schüler gab an, dass er so einen Workshop eher nicht noch einmal machen würde. In den persönlichen Gesprächen mit den Schülern während des Moduls kam zum Ausdruck, dass ihnen das Modul Spaß machen würde. Auch der begleitende Lehrer zog ein positives Fazit des gesamten Workshops.
Aus der Online-Evaluation resultiert, dass die Schüler am Ende des Moduls nach eigener Aussage etwas Neues gelernt haben. Wie unter dem entsprechenden Gesichtspunkt *Verständnis* bereits beschrieben wurde, haben fast alle Schüler die entscheidenden mathematischen Inhalte der Aufgabenblätter verstanden und so das Prinzip der SVM zur automatischen Klassifikation von Bildern verinnerlicht.
Ebenfalls positiv zu bewerten ist die Tatsache, dass dem zweiten Dozenten Nils Steffen die Einarbeitung in das entwickelte Modul sehr leicht fiel.
Besonders die Ergebnissen hinsichtlich der Modellierungskompetenz sind zufriedenstellend. Offensichtlich war der Modellierungskreislauf, welcher sich als roter Faden durch den CAMMP day zieht, für die Schüler ersichtlich und eingängig. Ich denke, dass allen Schülern das Thema der mathematischen Modellierung näher gebracht worden ist und sie einen interessanten Tag im Schülerlabor verbracht haben, bei dem sie viel Neues gelernt haben.

5.4. Die zweite Durchführung

5.4.1. Rahmenbedingungen

Das Lernmodul wurde am 14.11.2018 zum zweiten Mal durchgeführt. Der teilnehmende Kurs war ein Mathematik Leistungskurs aus der Qualifikationsphase I. Wie schon der erste Kurs, besuchen die Schüler die Gesamtschule Niederzier-Merzenich. Insgesamt

bestand die Gruppe aus 4 Schülerinnen, 13 Schülern und einem Lehrer. Der Workshop wurde von Alisa Peters, einer Mitarbeiterin des CAMMP Teams, und mir durchgeführt. Da der Lehrer bereits bei der ersten Durchführung dabei war, konnte er bei Fragen während den Arbeitsphasen helfen.
Wie schon bei der ersten Durchführungen, wurde das Modul in 6 Stunden durchgeführt. Darin enthalten war eine einstündige Mittagspause und eine zehnminütige Frühstückspause.
Auch die zweite Durchführung des CAMMP days verlief reibungslos. Das Zeitproblem, welches bei der ersten Durchführung aufgetreten ist konnte stellenweise behoben werden, indem bestimmte Aufgabenteile reduziert wurden. Da der MATLAB Code nicht verändert, sondern lediglich gekürzt wurde, gab es hier wenig Fehlermeldungen. Im Falle eines kleineren Problems, konnten die Betreuer schnell weiterhelfen. Auch bei der zweiten Durchführung wurde auf die *Powerwall* verzichtet. Im Unterschied zum ersten Mal konnte jedoch der Studieninformationsvortrag gehalten werden. Für zukünftige Veranstaltungen könnte die Bearbeitungsdauer weiter reduziert werden, damit die Powerwall wieder in den CAMMP day integriert werden kann.
Zum Vorwissen der zweiten Schülergruppe muss angemerkt werden, dass Ebenen noch nicht thematisiert wurden. Die Schüler kennen jedoch schon das Skalarprodukt und die Beschreibung von Geraden mittels Parameterform.

5.4.2. Beobachtungen anhand der leitenden Gesichtspunkte

Die genauen Ergebnisse der Online-Evaluation können im Anhang unter A.7.3 nachgelesen werden. Es werden die gleichen Gesichtspunkte wie bei der ersten Durchführung verwendet. Zusätzlich soll noch der Zeitaspekt diskutiert werden, da dieser beim ersten Mal das Kernproblem war.

Motivation und Interesse:
Bei der Beobachtung der Lerngruppe während der Arbeitsphasen fiel bereits auf, dass die Schüler sehr motiviert und interessiert waren. Besonders auffällig war, dass die Gruppen sogar während der Frühstückspause noch weiter gearbeitet haben. In Gesprächen mit der Lehrkraft wurde dieser Eindruck bestätigt. Sie merkte an, dass die Schüler im Mathematikunterricht selten über einen längeren Zeitraum so konzentriert arbeiten würden wie während des Workshops. Auch die Ergebnisse der Online-Evaluation decken sich mit diesen Beobachtungen. Alle Teilnehmer stimmten der Aussage, dass das Thema sie interessiert, teilweise (+), beziehungsweise voll (++) zu. Außerdem gaben über 80 % an, dass der Workshop ihr Interesse noch gesteigert habe. Die Rückmeldungen in den Freitextfeldern der Evaluation vervollständigen das positive Bild dieses Gesichtspunktes: „*Sehr hilfreich und **interessant**.*“, „Wenn es Fragen gab wurde immer geholfen und die Aufgaben waren verständlich und **interessant**. Also alles super.“, „*Ich muss sagen, dass der Vortrag sehr gut war, **der Workshop interessant** und sonst auch alles gut und glatt gelaufen ist.*“

Verständnis:
Dieser Gesichtspunkt ist im Vergleich zur ersten Durchführung besonders interessant, da der Q1 Kurs das Thema Ebenen in der Schule noch nicht behandelt hat, wie bereits in den Rahmenbedingungen angemerkt wurde (vgl. Abschnitt 5.4.1). Inwiefern sich dies auf das Verständnis beim zweiten Arbeitsblatt ausgewirkt hat, konnte einerseits in der entsprechenden Arbeitsphase beobachtet und andererseits wieder durch die zusätzlichen Fragen in der Evaluation betrachtet werden. Die drei Aussagen, die auf das Verständnis der Modellverbesserung durch die Verwendung von Ebenen abzielen sind „*Die Modellverbesserung durch den Wechsel von zwei auf drei Dimensionen ist mir leicht gefallen.*“, „*Die Fallunterscheidung auf Blatt 2 habe ich verstanden.*“, sowie „*Ich habe auf Blatt 2 verstanden, wie mehrere Klassen voneinander getrennt werden können.*“. Der ersten Aussage stimmten über 85 % teilweise (+) oder voll (++) zu. Lediglich zwei Schüler gaben an, dass sie Schwierigkeiten dabei hatten. Dies konnte auch während der Arbeitsphase festgestellt werden. Die ersten Aufgaben des zweiten Arbeitsblattes (vgl. Abschnitt 4.3.4) bereiteten den Schülergruppen keine großen Schwierigkeiten. Anhand der entsprechenden Abbildung im Code zu den dreidimensionalen Merkmalsvektoren schien es für viele Gruppen intuitiv Ebenen als Trennfunktionen zu nutzen. Mit Hilfe der gegebenen Ebenengleichung, welche für die Schüler neu war, musste anschließend die Fallunterscheidung aufgestellt werden. Bei der Bearbeitung der Aufgabe fiel auf, dass die Schüler viel mehr Zeit benötigten als der Q2 Kurs bei der ersten Durchführung. Dies könnte darauf zurückzuführen sein, dass der ältere Kurs Ebenen bereits im Unterricht thematisiert hat. Trotzdem kann festgehalten werden, dass die Q1 Schüler die Fallunterscheidung verstanden haben, da der Aussage, welche auf das Verständnis der Fallunterscheidung abzielt, nur ein Teilnehmer mit „*trifft eher nicht zu (-)*“ antwortete. Die Fallunterscheidung stellt einen entscheidenden Aspekt für den weiteren Verlauf des zweiten Arbeitsblattes dar, somit ist das Verständnis an dieser Stelle essenziell. Ein weiterer Knackpunkt auf dem zweiten Arbeitsblatt ist die Unterteilung in mehr als zwei Klassen. Lediglich zwei Schüler gaben an, dass sie die Einteilung teilweise nicht verstanden hätten (-). Wiederum ca. 70 % stimmten der Aussage zu, dass sie das Trennverfahren für mehrere Klassen vollständig verstanden hätten (++).
Der Gesichtspunkt Verständnis kann als durchaus positiv evaluiert werden. Nur ein Schüler stimmte der Aussage, dass er das Prinzip der SVM verstanden habe, eher nicht zu (-). Die übrigen Schüler gaben an, die Funktionsweise dieses Algorithmus verstanden zu haben. Ähnliche Werte ergaben sich für die Aussage „*Mir ist klar geworden, wie Bilder mit Hilfe eines Computers klassifiziert werden können.*“. Hier gaben nur zwei Schüler an, dass dies eher nicht der Fall sei (-).

Arbeitsmaterialien:
Im Vorfeld der zweiten Durchführung wurden die Arbeitsmaterialien, welche im Abschnitt 4.3 ausführlich beschrieben wurden geringfügig verändert. Die Änderungen wurden in Abschnitt 5.3.3 erklärt. Der Schwierigkeitsgrad der Aufgaben wurde durch die Veränderungen nicht reduziert, sondern lediglich ihr Umfang. Daher ist es erstaunlich, dass 80 % der Teilnehmer angaben die Aufgaben seien nicht (-) oder überhaupt nicht

(- -) zu schwer gewesen. Bei der ersten Durchführung hatte über die Hälfte der Schüler angegeben, dass die Aufgaben teilweise zu schwierig wären. Das Anforderungsniveau war für den Q1 Kurs offenbar genau richtig gewählt, da über 80 % die Aussage verneinten, die Aufgaben seien zu einfach gewesen. An dieser Stelle muss außerdem betont werden, dass alle Schüler die Hilfekarten in Anspruch genommen haben und diese laut Evaluation als sehr hilfreich empfunden haben.
Die Bonusaufgaben wurden bei der zweiten Durchführung von einigen Teilnehmer bearbeitet. Drei Schülergruppen suchten beispielsweise auf dem zweiten Arbeitsblatt nach der Schwachstelle des OVO Verfahrens. Die Extraaufgaben auf dem ersten Blatt wurden hingegen von keiner Gruppe bearbeitet. Die Möglichkeit durch dieses Zusatzangebot schnelle Gruppen weiter zu fordern, soll trotzdem beibehalten werden. Durch die erforderlichen weiteren Kürzungen in den Pflichtteilen der Aufgabenblätter, könnten die Bonusaufgaben in zukünftigen Durchführungen verstärkt genutzt werden.
Eine weitere Verbesserung im Vergleich zur ersten Durchführung, ist der Einsatz der Schüler-Besprechungsfolien (vgl. Abschnitt 4.2.3). Die Folien wurden während der ersten Arbeitsphase an verschiedene Gruppen ausgeteilt. Jeweils ein Schüler in jeder dieser Gruppen hat die erarbeitete Lösung des entsprechenden Arbeitsauftrags nicht nur auf dem eigenen Antwortzettel, sondern auch auf der Folie notiert. Auf diese Weise konnten die Schülerlösungen in der Besprechung des ersten Arbeitsblattes übersichtlicher diskutiert werden.

MATLAB:
In der Evaluation der ersten Durchführung merkten viele Schüler an, sie hätten Probleme bei der Bearbeitung der Aufgaben mit MATLAB gehabt. Über 40 % gaben an der Umgang mit dem Programm sei ihnen schwer gefallen. Als Reaktion darauf, wurde bei der zweiten Durchführung zu Beginn eine kurze MATLAB Einführung gemacht. Außerdem wurden einige typische Befehle oder Operatoren in MATLAB für alle Schüler sichtbar an die Tafel geschrieben. Diese beiden Maßnahmen konnten den Anteil der Schüler, die Schwierigkeiten mit MATLAB hatten auf 25 % reduzieren. Die kleinen Änderungen, welche in Abschnitt 5.3.3 genannt wurden, führten unter anderem dazu, dass weniger Fehler bei der Bearbeitung des Codes aufgetreten sind. Indem beispielsweise vorgegeben wurde, welche Variablen die Schüler für die Grauwertfunktion auf Blatt 1 nutzen sollen, verlief dieser Schritt reibungsloser. Ein Schüler machte in der Online-Evaluation außerdem die Aussage, der CAMMP day habe sein grundlegendes Verständnis von Programmen erweitert.

Modellierungskompetenz:
Wie schon bei der ersten Durchführung sprechen die Beobachtungen und die Evaluation durchaus dafür, dass die Modellierungskompetenz der Schüler verbessert wurde. 31 % der Schüler stimmten der Aussage, dass sie mathematisches Modellieren durch den Workshop besser verstanden hätten zum Teil zu (+). Die restlichen 69 % stimmten sogar voll zu (++). Bei den Besprechungsphasen, speziell in der Abschlusspräsentation, wurde deutlich, dass die Schüler den Modellierungskreislauf als roten Faden des Workshops erkannt haben. Sie konnten jeden Arbeitsschritt an der richtigen Stelle im

Kreislauf einordnen, was auch schon bei der ersten Durchführung funktioniert hat. In der Evaluation konnten die Schüler angeben, was sie persönlich durch die Teilnahme am Workshop gelernt haben. Ein Schüler gab dabei an, er habe gelernt, dass vieles mit Mathematik zusammen hängt. Diese Aussage passt zu den Beobachtungen und den Schülermeldungen während der Abschlusspräsentation. Die Materialien, welche im Rahmen der vorliegenden Arbeit entwickelt wurden, tragen folglich dazu bei eines der Ziele von CAMMP zu erreichen: die Entwicklung der Modellierungskompetenz zu fördern.

Arbeitsweise:
Im Vergleich zum ersten Durchlauf (vgl. Abschnitt 5.3.2) sollte bei der zweiten Durchführung besonders auf den Einsatz der Hilfekarten geachtet werden. Bei der ersten Durchführung war aufgefallen, dass einige Gruppen sehr schnell den höchsten Grad an Unterstützung ausgewählt haben. Deshalb wurde beim zweiten Mal verstärkt darauf hingewiesen, dass die Schüler zunächst das unterste Niveau der gestuften Hilfen auswählen und ausgehend davon schrittweise ein Niveau höher gehen sollen. Infolgedessen wurden die Hilfekarten bewusster verwendet. Die Schülergruppen entwickelten den Ehrgeiz mit möglichst wenig Hilfe auszukommen. Dies ergab auch die Reflexion des Workshops mit der begleitenden Lehrkraft. Sie sprach davon, die Schüler seien motiviert gewesen die Aufgaben möglichst selbständig zu lösen.

5.4.3. Resultierende Verbesserungen und Fazit

Abschließend lässt sich sagen, dass die ersten beiden Durchläufe des neu entwickelten CAMMP days nahezu reibungslos und zufriedenstellend verlaufen sind. Dies wurde in den Abschnitten 5.3.2 und 5.4.2 zu den leitenden Gesichtspunkten verdeutlicht. Das größte Problem aus der ersten Durchführung war die Zeitnot. Durch das Reduzieren des Aufgabenumfangs (vgl. Abschnitt 5.3.3) konnte dies jedoch teilweise behoben werden. Die zweite Lerngruppe arbeitete in der ersten Arbeitsphase bis zur Mittagspause sehr konzentriert und die Kürzung der Aufgaben ermöglichte eine umfangreichere Besprechung. Da auf dem zweiten Arbeitsblatt nicht ohne weiteres Kürzungen vorgenommen werden können, wurde Auftrag 2.2 auf dem ersten Arbeitsblatt gestrichen. In diesem Auftrag sollten die Schüler verschiedene [r,g,b]-Werte mit der recherchierten Formel in die entsprechenden Grauwerte umformen. In der aktuellen Version der Materialien wird dieser Schritt übersprungen. Die Anwendung der Grauwertformel beschränkt sich auf das Umwandeln einer der [r,g,b] Ampeln aus Aufgabe 1. Durch das Streichen von Auftrag 2.2 soll mehr Zeit zu Verfügung stehen, um schon vor der Mittagspause mit dem zweiten Aufgabenblatt zu beginnen. Geplant ist, dass die Schüler in dieser Zeit schon die ersten beiden Aufgaben des zweiten Aufgabenblattes bearbeiten.
Bei der zweiten Durchführung wurden die USB Sticks benutzt, um die Daten der Schüler zu sammeln. Die Methode erwies sich als zuverlässig und kann von den Betreuern während der ersten Viertelstunde der Mittagspause erledigt werden. Der Verzicht auf AirDrop sorgte für einen reibungslosen Ablauf. Die Betreuer können außerdem kontrollieren, ob alle Gruppen ihre Dateien richtig benannt haben und gegebenenfalls schnelle

Umbenennungen vornehmen.
Alisa Peters, welche bei der zweite Durchführung als Dozentin beteiligt war, konnte sich sehr schnell in das Material einarbeiten. Wie schon Nils Steffen, welcher bei der ersten Durchführung assistiert hat, konnte sie die Fragen der Schüler zum Workshop bereits bei ihrer ersten Durchführung beantworten. Alisa gab an, die Vorbereitung auf den CAMMP day sei ihr mit den entwickelten Materialien leicht gefallen.
Die Änderungen, welche aus dem ersten Tag resultierten, erwiesen sich alle als lohnenswert. Mein positiver Eindruck vom Lern-Modul spiegelt sich auch in der Bewertung der Teilnehmer wieder. Über die Hälfte der Teilnehmer bewertet den Tag abschließend mit der Note sehr gut, während die Übrigen die Schulnote gut verteilten. In den persönlichen Kommentaren zum Workshop wurden Rückmeldungen wie: „*Spitzen Arbeit*“, „*Sehr hilfreich und interessant.*“, „*War grandios.*“ oder „*War ein sehr guter Workshop.*“ gegeben.
Des Weiteren könnten sich alle bis auf einen Teilnehmer vorstellen einen ähnlichen Workshop noch ein mal zu besuchen. Neben der Förderung der Modellierungskompetenz, wurde in Kapitel 2 dieser Arbeit beschrieben, dass CAMMP auch das Ziel hat Schüler für ein mathematisches Studium zu begeistern. Neben dem Studien- Informationsvortrag trägt auch die Anwendungsnähe und die Relevanz des Modulthemas zum Erreichen des Ziels bei. Über 70 % der Teilnehmer des zweiten Workshops gab an, dass sie sich vorstellen könnten ein Studium oder eine Ausbildung im Bereich der Naturwissenschaften zu absolvieren. Viele Teilnehmer zeigten Interesse am Studiengang CES und nahmen Informationsbroschüren mit.
Neben dem Faktor, dass das erarbeitete Modul den Teilnehmern Spaß machen soll, ist es für die Bewertung des Workshops entscheidend, was die Schüler gelernt haben. Bei beiden Durchführungen gaben die Schüler mit großer Mehrheit an, dass sie die zentralen mathematischen Ideen hinter der Bildklassifikation verstanden hätten. Dies äußerte sich sowohl in den Ergebnissen der Online Evaluation, als auch in den Schülermeldungen in den Sicherungsphasen, welche bereits oben beschrieben wurden (vgl. Abschnitte 5.3.2 und 5.4.2). Es wurde bewusst darauf verzichtet, die ausgefüllten Antwortzettel der Schüler einzusammeln. Indem sie ihre Ergebnisse mit nach Hause nehmen können wird ermöglicht, dass die Schüler das Material noch weiter nutzen können. Durch den starken Bezug zum Lehrplan kann der Lehrer bei der Behandlung der Themen im Schulunterricht auf die Arbeitsmaterialien des Workshops verweisen.

6. Ausblick

Im letzten Kapitel dieser Arbeit werden Ideen für Weiterentwicklungsmöglichkeiten des Workshops vorgestellt.
Wie in der Einleitung und in Kapitel 3.1 beschrieben, ist Machine Learning beziehungsweise künstliche Intelligenz ein aktuelles Thema, das zahlreiche Anwendungsmöglichkeiten bietet. Der für das erarbeitete Modul ausgewählte Bereiche der Bilderkennung ist durchaus austauschbar und durch andere Themen zu ersetzen. Eine Möglichkeit wäre das Thema Spam-Filter für E-Mail Postfächer zu wählen. Das reale Problem zu Beginn des Workshops wäre dann die automatische Erkennung von Spam E-Mails. Wie in Abschnitt 3.1 unter dem Unterpunkt überwachtes Lernen beschrieben, basiert ein Spam-Filter auch auf einem Klassifizierungsproblem. Das Modul würde in seiner Struktur beibehalten werden. Man könnte den Schülern zunächst den Auftrag geben zwei Merkmale festzulegen, die eine Spam E-Mail von einer Ham E-Mail unterscheiden. Ausgehend von dieser vereinfachten Situation müssten sie sich überlegen, wie die Merkmale mathematisch im Merkmalsvektor gespeichert werden können. Darauf aufbauend ließe sich der Abstandsbegriff und das primale Problem formulieren. Am Ende des ersten Blattes testen die Schüler dann, wie gut ihr Filter funktioniert. Als Modellerweiterung könnte ein drittes Merkmal hinzugezogen werden und der Übergang von zwei zu drei Dimensionen würde hergestellt. Je nach Datensatz, würde sich der Filter durch Hinzunahme eines dritten Merkmals wahrscheinlich nur gering verbessern. Die Idee mehrere Merkmale zu benutzen könnte dann genutzt werden, um einen hochdimensionalen Merkmalsvektor zu definieren. Bei dieser Variante des Moduls wäre es möglich auf die OVO bzw. OVA Methode zu verzichten, da es nur die zwei verschieden Klassen Spam und Ham gibt. Wenn auch die Mehrklassen Einteilung berücksichtigt werden soll, könnte man noch eine dritte Kategorie von E-Mails, wie zum Beispiel Werbemails definieren. Ein weiterer Vorteil liegt darin, dass die Bilder für die Gesichtserkennung an sich schon komplex sind. Diese Komplexität fällt beim E-Mail Problem weg.
Eine weitere Alternative für die Entwicklung eines Machine Learning Workshops wäre das Thema Käuferverhalten. Unternehmen wie *Amazon* oder *Netflix* teilen ihre Kunden in verschiedene Klassen ein. Die Algorithmen der Unternehmen geben Kaufempfehlungen oder schlagen ihren Nutzern spezifische Produkte vor, die sie interessieren könnten. Als dritte Möglichkeit soll noch die Idee gegeben werden, das Thema autonomes Fahren für den gesamten Workshop zu verwenden. Auf dem dritten Aufgabenblatt (vgl. Abschnitt 4.3.5) werden bis jetzt die Gesichter der Schüler für die Bilderkennung genutzt. Es besteht jedoch auch die Möglichkeit Verkehrsschilder zu klassifizieren. Es müsste dann im Vorhinein ein Datensatz aus verschiedenen Schildern erstellt werden. Jedem Schild würde eine Klasse zugewiesen.

In manchen Fällen ist es nötig, dass ein CAMMP day verkürzt werden muss, da für die Durchführung keine sechs Stunden zu Verfügung stehen. In einem solchen Fall könnte das erstellte Lehr-Lern-Modul ebenfalls genutzt werden. Der Aufbau des Moduls ermöglicht es, auf das dritte Arbeitsblatt zu verzichten. Die Schüler arbeiten nur mit den ersten beide Aufgabenblättern und die Anwendung beschränkt sich auf die Bilderken-

nung für Bilder mit zwei, beziehungsweise drei Pixeln. Am Ende des zweiten Aufgabenblattes (vgl. Abschnitt 4.3.4) haben die Schüler bereits ein funktionierendes Modell für die automatische Erkennung von drei verschiedenen Ampelphasen erarbeitet. Dies kann als zufriedenstellendes Ergebnis und als Lösung des realen Problems verstanden werden. In einem solchen Fall müssten die Präsentationen angepasst werden. Außerdem sollte am Ende des zweiten Blattes noch eine Aufgabe eingebaut werden, in der mehrere unbekannte Ampeln mit dem aufgestellten Modell klassifiziert werden.
Weiterhin kann das erarbeitete Modul durchaus für einen zweitägigen CAMMP day verwendet werden. In den beiden in Kapitel 5 beschriebenen Durchführungen wurde bereits erwähnt, dass keine Gruppe das Bonusaufgabenblatt (vgl. Abschnitt 4.3.6) bearbeitet hat. Die Aufgaben auf diesem Blatt lassen sich ohne Weiteres als Pflichtaufgaben festlegen, sodass der Workshop noch verlängert werden kann. Des weiteren ließe sich das primale Optimierungsproblem und dessen Lösungsstrategie an dieser Stelle genauer diskutieren. Soll der Workshop noch weiter verlängert werden, könnte zusätzlich eine der oben beschriebenen Anwendungen ausgewählt werden. Den Schülern würde auf diese Weise verdeutlicht, dass die gleichen Algorithmen in verschiedensten Anwendungsbereichen genutzt werden. Neben dem SVM Algorithmus gibt es natürlich noch weitere interessante Machine Learning Algorithmen, welche einbezogen werden könnten. Denkbar wäre beispielsweise ein Modul zur Entwicklung neuronaler Netze zu entwerfen.

Durch den starken Bezug zur Schulmathematik, welche unter anderem durch das Skalarprodukt und den Abstandsbegriff gegeben ist, könnte das entwickelte Material sehr gut im Unterricht eingesetzt werden. Eine Möglichkeit wäre es in diesem Zusammenhang auf die Optimierung zu verzichten und keine optimale Trennfunktion zu bestimmen, sondern lediglich eine Mögliche. Anhand dieser könnten bereits Daten klassifiziert werden. Auf diese Weise würde die Thematisierung der Normalenform in der Schule in einen interessanten Anwendungsfall eingekleidet. Diese Unterrichtsidee könnte in zukünftigen Arbeiten noch weiter ausgeschärft werden.
Eine weitere Möglichkeit zur Weiterentwicklung, die sich durch den starken Schulbezug ergibt, wäre die Entwicklung einer Abituraufgabe im Inhaltsfeld der analytischen Geometrie.

Literatur

Alt, W. & Seydenschwanz, M. (2013). *EAGLE-STARTHILFE Optimale Steuerung: Theorie und numerische Verfahren.* Edition am Gutenbergplatz Leipzig.

Benesty, J., Sondhi, M. & Huang, Y. (2007). *Springer Handbook of Speech Processing.* Springer Berlin Heidelberg.

Bishop, C. M. (2006). *Pattern Recognition And Machine Learning* (M. Jordan, J. Kleinberg & B. Schölkopf, Hrsg.). Springer-Verlag New York, Inc.

Blum, W. (1985). *Anwendungsorientierter Mathematikunterricht in der didaktischen Diskussion.* GhK. Zugriff auf `https://books.google.de/books?id=nrYrGwAACAAJ`

Blum, W. (2006). Modellierungsaufgaben im Mathematikunterricht. In A. Büchter, H. Humenberger, S. Hußmann & S. Prediger (Hrsg.), *Realitätsnaher Mathematikunterricht: vom Fach aus und für die Praxis ; Festschrift für Hans-Wolfgang Henn zum 60. Geburtstag.* Franzbecker.

Blum, W. & Leiß, D. (2005). Modellieren im Unterricht mit der "TankenAufgabe. *mathematik lehren* (128), 18-21.

Börlin, J. (2012). *Das Experiment als Lerngelegenheit: vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität.* Logos Verlag Berlin.

Brand, S. (2014). *Erwerb von Modellierungskompetenzen: Empirischer Vergleich eines holistischen und eines atomistischen Ansatzes zur Förderung von Modellierungskompetenzen.* Springer Fachmedien Wiesbaden.

Bruder, R., Linneweber-Lammerskitten, H. & Reibold, J. (2015). Individualisieren und Differenzieren. In R. Bruder, L. Hefendehl-Hebeker, B. Schmidt-Thieme & H. G. Weigand (Hrsg.), *Handbuch der Mathematikdidaktik* (S. 513–534). Springer Spektrum.

Bruder, R. & Reibold, J. (2010, October). Weil jeder anders lernt - Ein alltagstaugliches Konzept zur Binnendifferenzierung. *Mathematik Lehren*, *162*, 2–9.

Bruner, J. (1974). *Entwurf einer Unterrichtstheorie.* Berlin Verlag. Zugriff auf `https://books.google.de/books?id=SVPtAQAACAAJ`

Brzezinska, A. (2009). *Lernpsychologie und Mnemotechniken beim Fremdsprachenlernen: am Beispiel des Artikellernens im DaF-Unterricht.* StudienVerlag.

Büchter, A. & Leuders, T. (2005). *Mathematikaufgaben selbst entwickeln: Lernen fördern - Leistung überprüfen.* Cornelsen Scriptor.

Bürger, H., Wittmann, E. & Malle, G. (2013). *Didaktische Probleme der elementaren Algebra* (G. Malle, Hrsg.). Vieweg+Teubner Verlag.

Duden. (2018, October). Zugriff am 19.10.2018 auf `https://www.duden.de/rechtschreibung/modellieren`

Ernst Klett Verlag. (2017). *Lambacher Schweizer. 5. Schuljahr.* Klett Ernst /Schulbuch.

Ferri, R. & Greefrath, G. (2013). *Mathematisches Modellieren für Schule und Hochschule: Theoretische und didaktische Hintergründe* (G. Kaiser, Hrsg.). Springer Fachmedien Wiesbaden.

Frank, M., Richter, P., Roeckerath, C. & Schönbrodt, S. (2018). Wie funktioniert eigentlich GPS? - ein computergestützter Modellierungsworkshop. In G. Greefrath & H. Siller (Hrsg.), *Digitale Werkzeuge, Simulationen und mathematisches Modellieren: Didaktische Hintergründe und Erfahrungen aus der Praxis* (S. 137–164). Springer Fachmedien Wiesbaden.

Franke-Braun, G., Schmidt-Weigand, F., Stäudel, L. & Wodzinski, R. (2008). Aufgaben mit gestuften Lernhilfen – ein besonderes Aufgabenformat zur kognitiven Aktivierung der Schülerinnen und Schüler und zur Intensivierung der sachbezogenen Kommunikation. In Kasseler Forschergruppe Empirische Bildungsforschung (Hrsg.), *Lernumgebungen auf dem Prüfstand: Zwischenergebnisse aus den Forschungsprojekten.* Kassel University Press.

Frochte, J. (2018). *Maschinelles Lernen: Grundlagen und Algorithmen in Python.* Carl Hanser Verlag GmbH & Company KG.

Géron, A. (2018). *Praxiseinstieg Machine Learning mit Scikit-Learn und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme.* O'Reilly.

Gold, A. (2015). *Guter Unterricht: Was wir wirklich darüber wissen.* Vandenhoeck & Ruprecht.

Greefrath, G. & Siller, H. (2018). *Digitale Werkzeuge, Simulationen und mathematisches Modellieren: Didaktische Hintergründe und Erfahrungen aus der Praxis.* Springer Fachmedien Wiesbaden.

Greefrath, G. & Weitendorf, J. (2013). Modellieren mit digitalen Werkzeugen. In G. Kaiser (Hrsg.), *Mathematisches Modellieren für Schule und Hochschule: Theoretische und didaktische Hintergründe.* Springer Fachmedien Wiesbaden.

Griesel, H., Gundlach, A., Postel, H. & Suhr, F. (Hrsg.). (2011). *Elemente der Mathematik. Schülerband. Nordrhein-Westfalen: Qualifikationsphase: Grundkurs und Leistungskurs. Sekundarstufe 2 - Ausgabe 2011.* Schroedel. Zugriff auf `https://books.google.de/books?id=fZwWYAAACAAJ`

Guido, S. & Müller, A. (2017). *Einführung in Machine Learning mit Python: Praxiswissen Data Science.* O'Reilly.

Hamel, L. (2011). *Knowledge Discovery with Support Vector Machines.* Wiley.

Hänze, M., Schmidt-Weigand, F. & Stäudel, L. (2010). Gestufte Lernhilfen. In S. Boller & R. Lau (Hrsg.), *Innere Differenzierung in der Sekundarstufe II. Ein Praxishandbuch für Lehrer/innen.* (S. 63-73). Beltz.

Heinz, S. (2018). *Mobile Learning und Fremdsprachenunterricht: Theoretische Verortung, Forschungsüberblick und Studie zum Englischunterricht in Tablet-Klassen an Sekundarschulen in Bayern.* Julius Klinkhardt.

Helmke, A. (2007, October). Was wissen wir über guten Unterricht? Wissenschaftliche Erkenntnisse zur Unterrichtsforschung und Konsequenz für die Unterrichtsentwicklung. In *Lehren und Lernen für die Zukunft* (S. 1–17).

Hennen, M. (2008). Mit Unterschieden rechnen - Binnendifferenzierung im Mathematikunterricht. In I. Scholz (Hrsg.), *Der Spagat zwischen Fördern und Fordern: Unterrichten in heterogenen Klassen.* Vandenhoeck & Ruprecht.

Heymann, H. W. (1998). Üben und Wiederholen. *Pädagogik 50* (10), 7–11.

Heymann, H. W. (2012). Schüler beim Aufbau von Kompetenzen unterstützen. Üben, Anwenden, Vertiefen - Gelingsbedingungen für nachhaltiges Üben. *Pädagogik 64* (12), 6–11.

Hinrichs, G. (2008). *Modellierung im Mathematikunterricht.* Spektrum Akademischer Verlag.

Hornsteiner, C. (2010). *Simulation und Klassifikation von Dopplerradarsignalen.* Logos Verlag Berlin.

Jörissen, S. & Schmidt-Thieme, B. (2015). Darstellen und Kommunizieren. In R. Bruder, L. Hefendehl-Hebeker, B. Schmidt-Thieme & H. G. Weigand (Hrsg.), *Handbuch der Mathematikdidaktik* (S. 385–410). Springer Spektrum.

Kircher, E., Girwidz, R. & Häußler, P. (2009). *Physikdidaktik Theorie und Praxis* (Bd. 2). Springer Berlin Heidelberg.

Kress, K. & Pappas, M. (2014). *Binnendifferenzierung in der Grundschule: Profi-Tipps und Materialien aus der Lehrerfortbildung (1. bis 4. Klasse).* Auer Verlag.

Kultusministerkonferenz. (2012, October). *Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife.* Institut zur Qualitätsentwicklung im Bildungswesen. Unter den Linden 6 10099 Berlin.

Kultusministerkonferenz. (2015, June). *Förderstrategie für leistungsstarke Schülerinnen und Schüler.* Institut zur Qualitätsentwicklung im Bildungswesen. Unter den Linden 6 10099 Berlin.

Lampert, C. & Lampert, C. (2009). *Kernel Methods in Computer Vision.* Now Publishers.

Leisen, J. (2007). Problemorientierter Unterricht und Aufgabenkultur. In S. Mikelskis-Seifert & T. Rabe (Hrsg.), *Physik-Methodik: Handbuch für die Sekundarstufe I und II* (2. Aufl.). Cornelsen Scriptor.

Meyer, H. (2015). *UnterrichtsMethoden II - Praxisband* (Bd. 15). Cornelsen Verlag GmbH.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2014). *Kernlehrplan Mathematik für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen.* Völklinger Straße 49, 40221 Düsseldorf.

Padberg, F. & Greefrath, G. (2010). *Didaktik des Sachrechnens in der Sekundarstufe.* Spektrum Akademischer Verlag.

Pollak, H. O. (1977). The Interaction between Mathematics and Other School Subjects (Including Integrated Courses). In H. Anthen & H. Kunle (Hrsg.), *Proceedings of the Third International Congress on Mathematical Education* (S. 255–264). Zentralblatt für Didaktik der Mathematik.

Rabe, T. (2014). Bitte einsteigen - Funktionen, Ziele und Dimensionen von Unterrichtseinstiegen. *Naturwissenschaften im Unterricht Physik*, *139* (25), 4–9.

Raschka, S. & Mirjalili, V. (2017). *Machine Learning mit Python und Scikit-Learn und TensorFlow: Das umfassende Praxis-Handbuch für Data Science, Predictive Analytics und Deep Learning.* mitp-Verlag.

Reit, X. (2016). *Denkstrukturen in Lösungsansätzen von Modellierungsaufgaben: Eine kognitionspsychologische Analyse schwierigkeitsgenerierender Aspekte.* Springer Fachmedien Wiesbaden.

Roddeck, W. (2008). Mechatronik. In A. Böge (Hrsg.), *Vieweg Handbuch Maschinenbau: Grundlagen und Anwendungen der Maschinenbau-Technik* (18. Aufl., S. 551–584). Vieweg+Teubner Verlag.

Schmitt, G. (1999). *Lernen und Verhaltensänderungen.* Vorlesungsskript. Universität Essen.

Schmitz, A. (2017). *Beliefs von Lehrerinnen und Lehrern der Sekundarstufen zum Visualisieren im Mathematikunterricht.* Springer Fachmedien Wiesbaden.

Schneuwly, G. (2014). *Differenzierungskonzepte sichtbar gemacht: Eine qualitative Fallstudie zur inneren Differenzierung im Mathematikunterricht der Primarschulstufe* (Nr. 18). Waxmann Verlag GmbH.

Schölkopf, B. & Smola, A. J. (2001). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.* MIT Press.

Schönbrodt, S. (2018). *Vergleich zweier Methoden zur Bildklassifizierung auf Basis maschineller Lernalgorithmen und ihre Anwendbarkeit in der Vermittlung mathematischer Modellierung* (Unveröffentlichte Diplomarbeit). RWTH Aachen University.

Schukajlow, S. (2011). *Mathematisches Modellieren.* Waxmann Verlag.

Statnikov, A., Aliferis, C., Hardin, D. & Guyon, I. (2011). *A Gentle Introduction to Support Vector Machines in Biomedicine: Volume 1: Theory and Methods.*

Storz, R. (2009). *Fachdidaktik - Seminar Mathematik.* Pro Business.

Süße, H. & Rodner, E. (2014). *Bildverarbeitung und Objekterkennung: Computer Vision in Industrie und Medizin.* Springer Fachmedien Wiesbaden.

Thuselt, F. & Gennrich, F. (2014). *Praktische Mathematik mit MATLAB, Scilab und Octave: für Ingenieure und Naturwissenschaftler.* Springer Berlin Heidelberg.

Vollrath, H. & Roth, J. (2011). *Grundlagen des Mathematikunterrichts in der Sekundarstufe* (Bd. 2; F. Padberg, Hrsg.). Spektrum Akademischer Verlag.

von Aufschnaiter, C. (2003). Prozessbasierte Detailanalysen der Bildungsqualität von Physik-Unterricht: Eine explorative Studie. *Zeitschrift für Didaktik der Naturwissenschaften*, *9*, 105–124.

von Aufschnaiter, C. & von Aufschnaiter, S. (2001). Eine neue Aufgabenkultur für den Physikunterricht. Was fachdidaktische Lernprozess-Forschung zur Entwicklung von Aufgaben beitragen kann. *Der mathematische und naturwissenschaftliche Unterricht*, *54* (7), 409-416.

Whitenack, D. (2017). *Machine Learning With Go: Implement Regression, Classification, Clustering, Time-series Models, Neural Networks, and More using the Go Programming Language.* Packt Publishing.

Wiernicki-Krips, T. (2014). Invertieren als fundamentale Idee in der Mathematik. In J. Roth & J. Ames (Hrsg.), *Beiträge zum Mathematikunterricht* (S. 1311–1314). WTM-Verlag.

Winter, H. (1975). *Allgemeine Lernziele für den Mathematikunterricht.* (Leicht bearbeiteter Nachdruck aus Zentralblatt für Didaktik der Mathematik 7 (1975), H. 3, S. 106 – 116.)

Winter, H. (1996). Mathematikunterricht und Allgemeinbildung. *Mitteilung der Gesellschaft für Didaktik*, *61*, 37–46.

Wodzinski, R. (2013). Lernen mit gestuften Hilfen. *Physik Journal*, *12* (3), 45–49.

Wroble, S., Joachims, T. & Morik, K. (2013). Maschinelles Lernen und Data Mining. In G. Görz, J. Schneeberger & U. Schmid (Hrsg.), *Handbuch der Künstlichen Intelligenz* (Bd. 5, S. 405–471). De Gruyter.

Zhou, Y. (2018). *Moderne Go-Eröffnung: Der Einfluss von AlphaGo auf Profi-Partien.* Brett und Stein Verlag.

Zimbardo, P. (2013). *Psychologie* (Bd. 4; W. Angermeier, J. Brengelmann & T. Thiekötter, Hrsg.). Springer Berlin Heidelberg.

A. Anhang

Der Anhang ist ausschließlich in digitaler Form auf der beiliegenden CD gespeichert. Die Seitenzahlen stimmen mit den im Inhaltsverzeichnis angegebenen Zahlen überein.

A.1. Aufgabenblätter (Live-Scripts)

Die Aufgabenblätter sind in ihrer Endversion beigefügt. Durch das Exportieren der Live-Scripts in PDF Dateien entstehen Fehler in der Darstellung bei Blatt 1 (S. 7-8, S. 14-15) und auf Blatt 2 (S. 5).

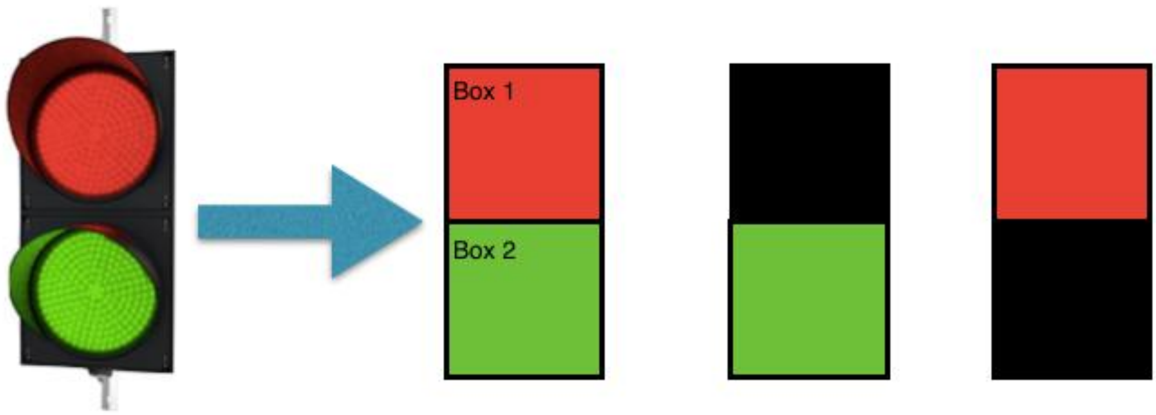
A.1.1. Aufgabenblatt 1

Aufgabenblatt 1 I Autonomes Fahren

Aufgabe 1

Wir wollen zunächst ein vereinfachtes Modell der Ausgangssituation herstellen. Lest dazu den folgenden Text:

Ein Automobilhersteller will autonomes Fahren für seine neuen Autos entwickeln. Die Autos sollen erkennen, ob eine voraus liegende Ampel rot oder grün ist. Für ein vereinfachtes Modell gehen wir davon aus, dass eine Ampel aus zwei untereinander angebrachten Boxen besteht:



Ampel in der Realität | Ampel als Modell | Ampel steht auf grün | Ampel steht auf rot

Farben werden üblicherweise als <u>RGB</u>-Wert in der Form $[r, g, b]$ angegeben. Dabei steht *r* für den Anteil der *roten*, *g* für den Anteil der *grünen* und *b* für den Anteil der *blauen* Farbe. Die Werte gehen jeweils von $0-255$. Ein rein blauer Farbton würde beispielsweise so aussehen: $[0, 0, 120]$.

<u>Für schwarz müssen alle Einträge gleich gewählt werden und sollten maximal bis 50 gehen!</u>

<u>Auftrag 1.1:</u> Erzeugt nun selber verschiedene **rote** Ampeln, indem ihr verschiedene Werte bei NaN für r1, g1 und b1 für Box 1 und für r2, g2 und b2 für Box 2 eingebt

Überprüft euer Ergebnis indem ihr den *Run Section* Button drückt:

```
clear
%Box 1
r1 = NaN;
g1 = NaN;
b1 = NaN;
%Box 2
r2 = NaN;
g2 = NaN;
b2 = NaN;
```

```
Ampel(r1,g1,b1,r2,g2,b2);
```

<u>Auftrag 1.2:</u> Erzeugt nun selber verschiedene **grüne** Ampeln, indem ihr verschiedene Werte für r,g und b eingebt:

```
clear
%Box 1
r1 = NaN;
g1 = NaN;
b1 = NaN;
%Box 2
r2 = NaN;
g2 = NaN;
b2 = NaN;
Ampel(r1,g1,b1,r2,g2,b2);
```

Aufgabe 2

In Aufgabe 1 habt ihr gesehen, dass immer drei Werte $[r, g, b]$ für eine Box abgespeichert werden müssen. Für eine Ampel mit 2 Boxen hat man also immer insgesamt 6 Werte. Bei einem großen Datensatz mit vielen Ampeln sammeln sich dann sehr viele Daten an und die Rechenzeit steigt. Um dem entgegen zu wirken, wollen wir **für jede Box nur einen Wert** abspeichern, der sich aus den drei Werten ergibt. Auf diese Weise können wir auch das **vereinfachte, mathematische** Modell für das **reale** Problem aufstellen. Enthält jede Box nur einen Wert, so kann man die zwei Boxen als einen Vektor mit zwei Einträgen darstellen, wie in Abbildung 2 zu sehen ist.

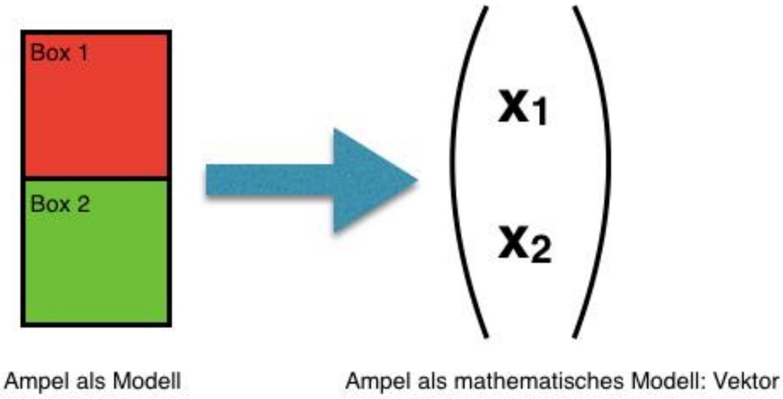


Abbildung 2: mathematisches Modell

Dafür darf jede Box jedoch **nur einen Wert** und nicht die drei Werte $[r, g, b]$ enthalten.

Auftrag 2.1: Informiert euch bei Wikipedia über den *Grauwert* in der Bildverarbeitung und gebt die Formel zu Berechnung des *Grauwertes* im Code ein:

Hinweis: Das '@(r,g,b)' in der Formel x = @(r,g,b) NaN bedeutet, dass die Zeichen r,g und b in der Klammer als Variablen der Gleichung gesehen werden.

*Hinweis: Komma-Zahlen werden in Matlab mit einem Punkt geschrieben: **3.6 nicht 3,6***

```
x = @(r,g,b) NaN; %hier Gleichung eingeben

Grauwertermittlung(x);
AmpelSW(r1,g1,b1,r2,g2,b2);
```

Drückt auf *Run Section* und euch wird angezeigt, ob euer Ergebnis korrekt ist. Außerdem erhaltet ihr ein Grauwert-Bild der letzten Ampel, die ihr in Aufgabe 1 eingegeben habt.

Aufgabe 3

Auf dem Antwortblatt seht ihr ein Koordinatensystem in dem die Grauwerte der Ampelphasen abgebildet werden. Es sind die Einträge der Vektoren von 200 Ampeln eingezeichnet.

Diese Punkte sind die **Trainingsdaten** für den Lernalgorithmus, mit dem **unbekannte Test-**Ampeln **später** automatisch erkannt werden sollen. Im Einstiegsvortrag habt ihr bereits etwas von einer **Trennfunktion** gehört.

Auftrag 3.1: Zeichnet eine möglichst einfache Funktion in das Koordinatensystem, welche die beiden Punktwolken voneinander trennt.

Dies ist eure Trennfunktion. Punkte, die oberhalb liegen werden zur Klasse der grünen Ampeln zugeordnet. Punkte, die darunter liegen werden zur Klasse der roten Ampeln zugeordnet.

!!!!! ERST AUFTRAG 3.1 BEANTWORTEN, DANN RUNTER SCROLLEN ZU AUFGABE 3.2 !!!!!!

Auftrag 3.2: Die Kamera am Auto erkennt eine Ampel P mit den Werten $x_1 = 20$ und $x_2 = 50$.

Zeichnet den zugehörigen Punkt in das Koordinatensystem. Zu welcher Klasse ordnet eure Gerade den Punkt zu?

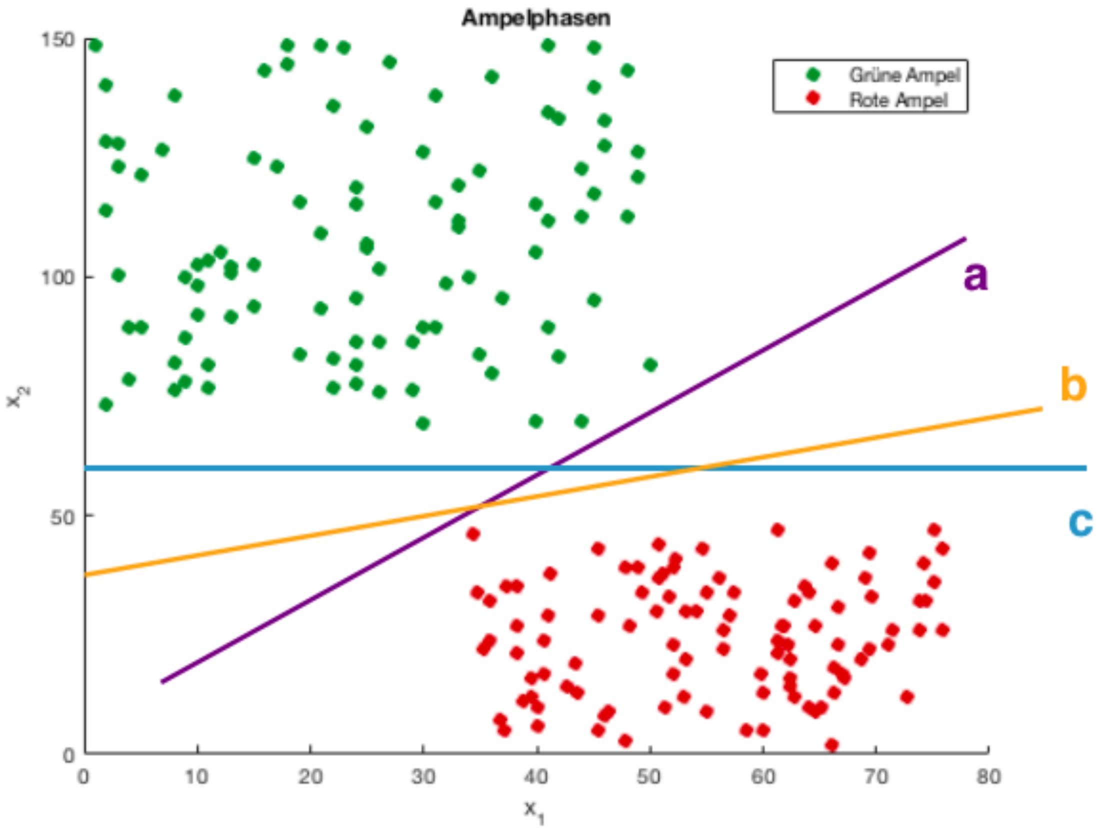


Abbildung 3: verschiedene Trennfunktionen

Wie ihr in Abbildung 3 sehen könnt, gibt es viele verschiedene Trennfunktionen für einen Datensatz. Die Frage ist nun, welche die Daten *am besten* trennt.

Auftrag 3.3: Um zu testen, wie gut eure Gerade und die drei Geraden a,b und c zum Klassifizieren geeignet sind, sollt ihr die folgenden Ampeln einordnen lassen. Die Daten, welche die Trennfunktion testen nennt man **Testdaten**.

Vergleicht die zugeteilte Klasse mit dem darüber angezeigtem Bild der Ampel und notiert in der Tabelle 1 auf dem Antwortblatt, ob die Gerade die Ampel richtig oder falsch klassifiziert. **Wiederholt dies für alle 6 Ampeln!**

Ampel 1: $x_1 = 30$ und $x_2 = 50$ Ampel 2: $x_1 = 40$ und $x_2 = 120$ Ampel 3: $x_1 = 45$ und $x_2 = 58$ Ampel 4: $x_1 = 70$ und $x_2 = 30$

Gebt im Code unter der Geradengleichung in der Form $y = m * x + n$ die Werte für die Steigung m und den y-Achsenabschnitt **eurer Gerade** ein.

Hinweis: Um die Steigung eurer Gerade zu bestimmen, könnt ihr das Steigungsdreieck nutzen.

```
y = @(x) m * x + n ;
m = NaN ;      %   Wert für die Steigung eurer Gerade
n = NaN ;        %   Wert für y-Achsenabschnitt eurer Gerade
```

Gebt hier die Ampelnummer und die x1, x2 Werte der Ampel ein, die ihr testen wollt:

```
Ampel = NaN; %Nummer der Ampel hier eintragen
x1 = NaN;  %Wert für die obere Box
x2 = NaN;  %Wert für die untere Box
Ampeltest(x1,x2,y,m,n,Ampel);
```

Drückt auf *Run Section* .

Euch wird dann eine Tabelle angezeigt, in der die jeweilige Gerade, die zugeteilte Klasse gelistet wird. Unter 'e' wird eure eigene Gerade gelistet. Außerdem seht ihr ein Bild der Ampel.

Frage 3.4: Welche Gerade klassifiziert die Testdaten am besten? Könnt ihr das auch schon am Verlauf der Geraden in Abbildung 3 erklären?

Aufgabe 4

In Frage 3.3 habt ihr anhand von 6 Testdaten entschieden, welche Gerade die Beste ist. Es gibt noch eine andere Möglichkeit, **ohne die Nutzung der Testdaten** zu entscheiden, welche die beste Trennfunktion für einen Datensatz ist.

Auftrag 4.1: Auf dem Antwortblatt sind vier verschiedene Trainingsdatensätze dargestellt. Zeichnet in jedes Koordinatensystem die Trennfunktion ein, **welche die Daten eurer Meinung nach am besten trennt**. Umkreist außerdem jeweils die Datenpunkte, die ihr als Orientierung genutzt habt, um die Trennfunktion zu zeichnen.

Frage 4.2: Worauf habt ihr beim Zeichnen geachtet? Was zeichnet für euch die *beste Gerade* aus?

!!!!! ERST FRAGE 4.2 BEANTWORTEN, DANN RUNTER SCROLLEN ZU AUFGABE 5 !!!!!!

Aufgabe 5

Beim Zeichnen der Geraden, habt ihr vermutlich darauf geachtet, dass die Gerade nicht zu nah an den Trainingsdatenpunkten liegt. Der Bereich um die Gerade, in dem keine Datenpunkte liegen wird **Margin** genannt. Die *beste Gerade* zeichnet sich dadurch aus, dass der Margin möglichst groß ist.

In der folgenden Abbildung sind noch vier weitere Datensätze mit jeweils 200 verschiedenen Trainingsdaten abgebildet. Es sind jeweils die Punkte umkreist, **die den kürzesten Abstand** zu den jeweiligen Geraden haben. Könnt ihr anhand der Abstände zu den nächst gelegenen Punkten entscheiden, welche jeweils die beste Gerade ist?

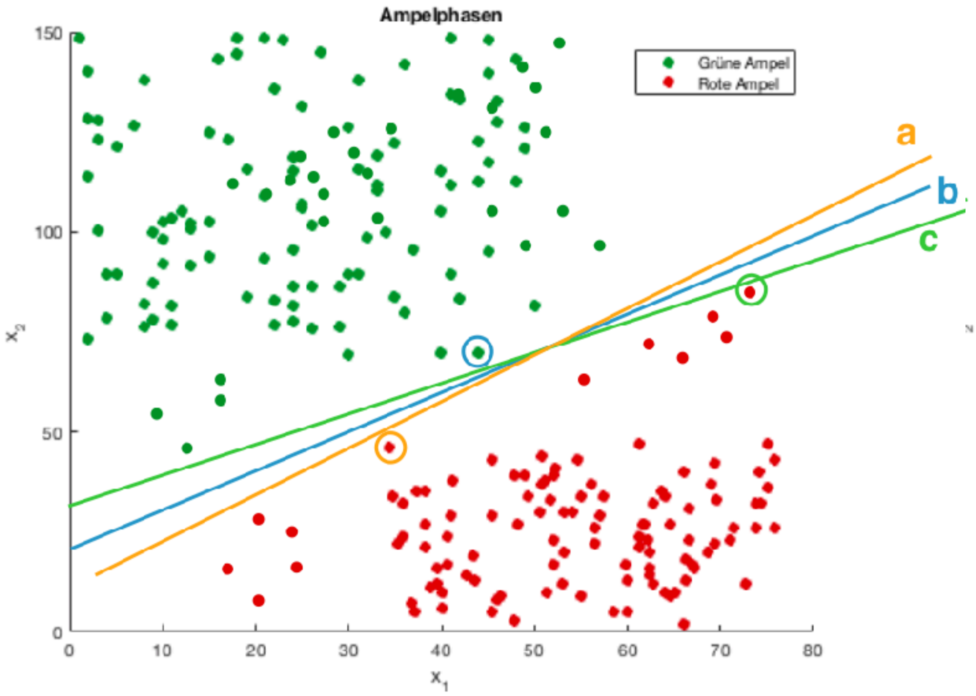
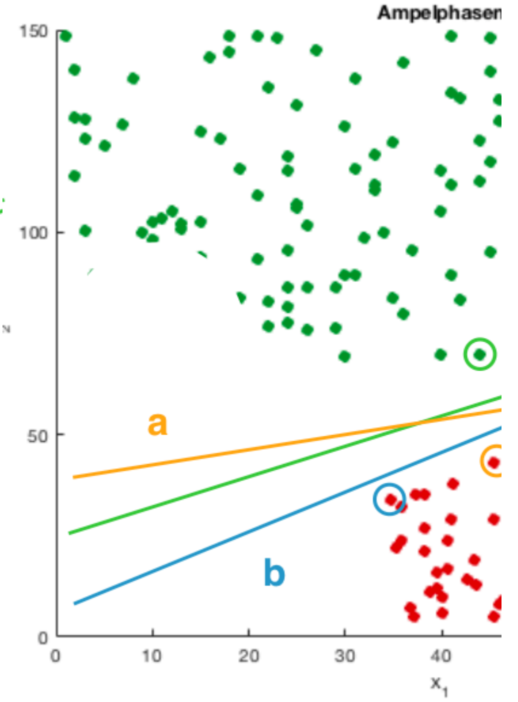


Abbildung 1



Abbildun

kürzester Abstand zu Gerade a: 12.5

kürzester Abstand zu Gerade b: 5.07

kürzester Abstand zu Gerade c: 11.37

kürzester Abstand zu Gerade a: 3.75

kürzester Abstand zu Gerade b: 5.01

kürzester Abstand zu Gerade c: 2.53

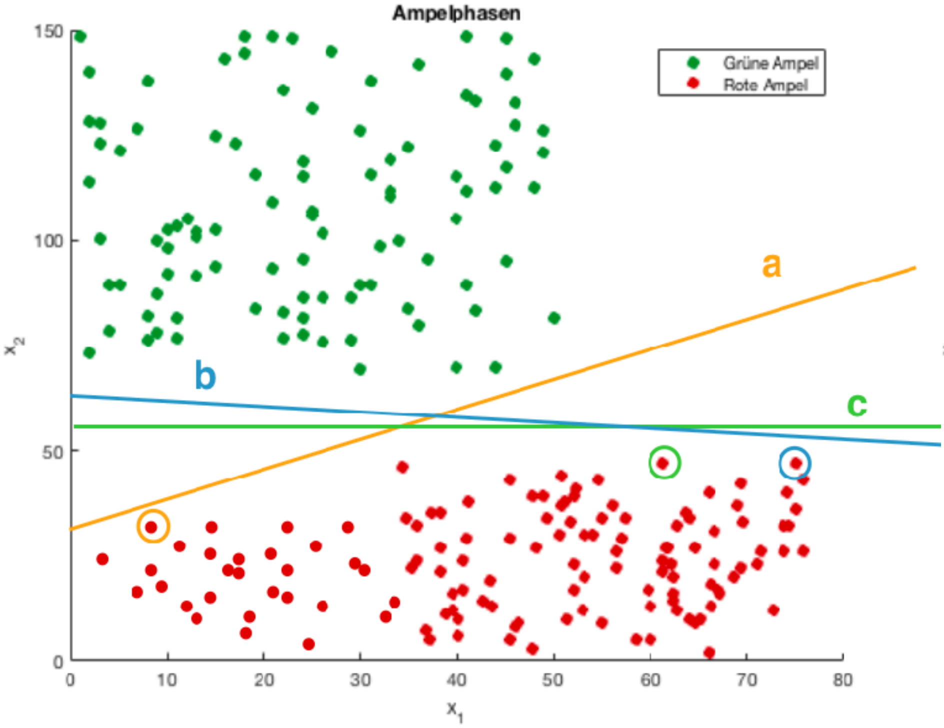
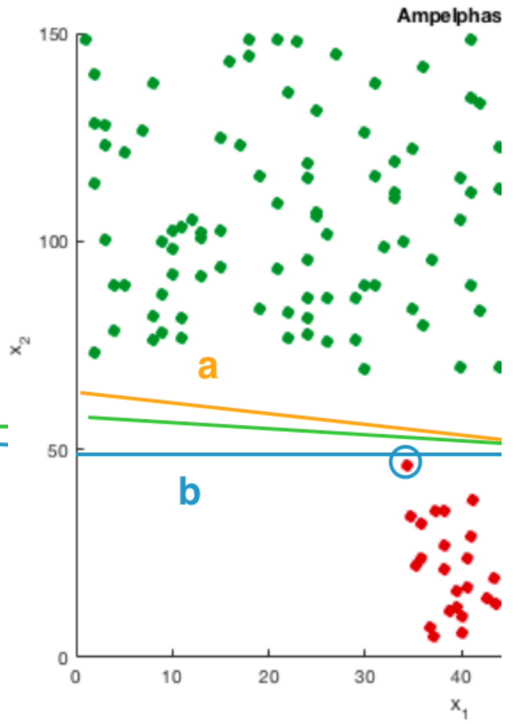


Abbildung 3



Abbil

kürzester Abstand zu Gerade a: 6.31

kürzester Abstand zu Gerade b: 2.57

kürzester Abstand zu Gerade c: 3.78

kürzester Abstand zu Gerade a: 5.12

kürzester Abstand zu Gerade b: 6.25

kürzester Abstand zu Gerade c: 8.78

<u>Auftrag 5.1:</u> Tragt jeweils die *beste* Gerade zur passenden Abbildung ein: **Ändert nur das NaN und lasst die ' ' stehen!**

```
Abbildung1 = 'NaN';
Abbildung2 = 'NaN';
Abbildung3 = 'NaN';
Abbildung4 = 'NaN';
Abstandtest(Abbildung1,Abbildung2,Abbildung3,Abbildung4);
```

Drückt nun den *Run Section* Button um euer Ergebnis zu überprüfen.

Auftrag 5.2: Angenommen es liegt ein neuer Datensatz aus Trainingsdaten vor. Formuliert einen wörtlichen Befehl an den Computer, mit dem er die beste Gerade finden soll: *"Finde die Gerade, die..."*

Damit der Computer die Trennfunktion berechnen kann müsst ihr zunächst eine mathematische Formel für die Gerade und dann für den Abstand aufstellen.

Eine Gerade kann man in der Mathematik auf verschiedene Weisen beschreiben. Aus der Schule kennt ihr eventuell schon die **Parameterform**:

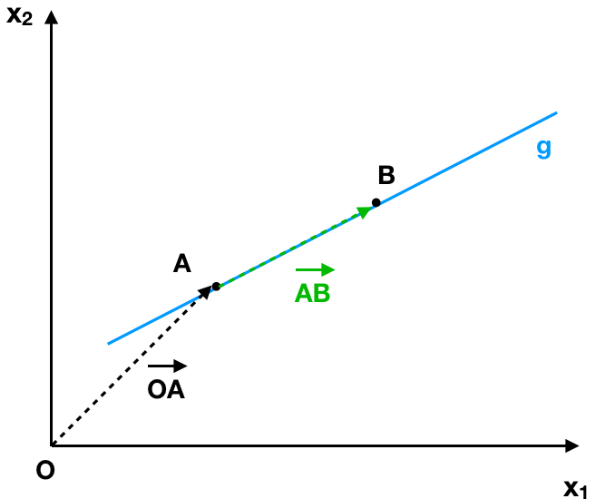


Abbildung 4: Skizze zur Paramterform mit Stütz- und Richtungsvektor

Hier wird die Gerade durch einen *Stützvektor* $\overrightarrow{OA}$ (schwarz) und einen Richtungsvektor $\overrightarrow{AB}$ (grün) beschrieben. Es werden also nur zwei Punkte A und B benötigt um die Gleichung für die Gerade aufzustellen:

$$g: \vec{x} = \overrightarrow{OA} + s \cdot \overrightarrow{AB} \quad . \tag{1}$$

Eine andere Darstellungsmöglichkeit, welche für unsere Anwendung besser geeignet ist, ist die Normalenform:

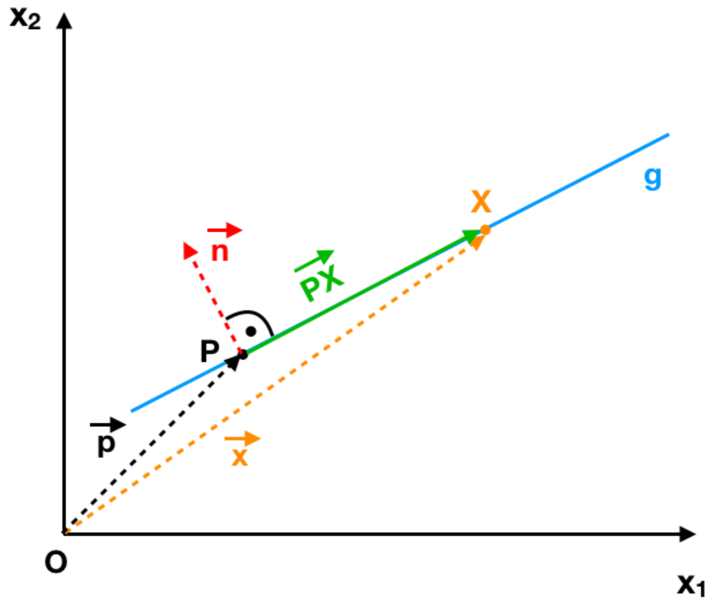


Abbildung 5: Skizze zur **Normalenform** mit Stütz- und Normalenvektor

Hier wird die Gerade g durch einen *Stützvektor* $\vec{p}$ (schwarz) und den Normalenvektor $\vec{n}$ (rot) beschrieben.

Jeder Punkt X mit einem Ortsvektor $\vec{x}$ (orange) liegt genau dann auf der Geraden g, wenn folgende Bedingung gilt:

$$\vec{n} * (\vec{x} - \vec{p}) = 0 \ , \qquad (2)$$

dabei steht $*$ für das *Skalarprodukt* zweier Vektoren. Das Skalarprodukt zweier Vektoren ist folgendermaßen definiert:

$$\vec{a} * \vec{b} = cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}| \ , \qquad (3)$$

Was auch wie folgt umgeschrieben werden kann:

$$cos(\alpha) = \frac{\vec{a} * \vec{b}}{|\vec{a}| \cdot |\vec{b}|} \ .$$

wobei α für den Winkel zwischen den beiden Vektoren und $|\vec{a}|$ für den Betrag, also die Länge des Vektors steht.

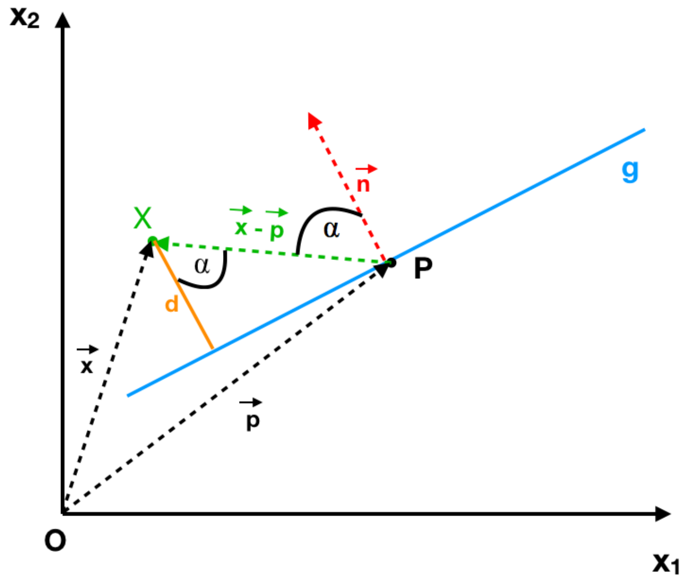
Stehen zwei Vektoren senkrecht zueinander, so ist das Skalarprodukt der beiden 0. Damit ist Gleichung (2) anschaulich nachvollziehbar: Der Normalenvektor steht nach seiner Definition senkrecht auf der Geraden.

Alle Verbindungsvektoren $\overrightarrow{PX}$ (grün) von Punkt P zu Punkten X auf der Gerade verlaufen auf bzw. parallel zu der Geraden. Deshalb steht der Normalenvektor auch senkrecht auf den Verbindungsstrecken.

Bedingung (2) stellt also ebenfalls eine Geradengleichung dar:

$$g : \vec{n} * (\vec{x} - \vec{p}) = 0 \ . \qquad (4)$$

Auftrag 5.3: Stellt mit Hilfe der Abbildung 6, und den obigen Informationen eine Formel für den Abstand d eines Punktes X von einer Geraden g auf. Nutzt dazu den Platz auf dem Antwortzettel.



MATLAB TOOLBOX

Betrag (Länge) eines Vektors	norm(x)
Wurzel	sqrt(x)
Skalarprodukt	dot(x,y)

Abbildung 6: Links: Skizze zur Berechnung des Abstandes d (orange) eines Punktes X (grün) von der Gerade g (blau) mit Winkel α, Normalenvektor n (rot) und Punkt P (schwarz). Rechts: MATLAB Eingabehilfe.

Hinweis: Nutzt die Definition des Skalarproduktes und stellt zwei verschiedene Formeln für $\cos(\alpha)$ ***auf!***

Nutzt bei Bedarf die Hilfekarten.

Auftrag 5.4: Tragt eure Lösung in den Code ein und überprüft eure Eingabe mit *Run Section.*

```
d = @(x,p,n) NaN;  %Tragt hier eure Formel für den Abstand ein
Abstand(d);
```

Auftrag 5.5: In Abbildung 7 sind 4 Trainingsdaten graphisch dargestellt. Bestimmt den Abstand der Punkte zur Trennfunktion g.

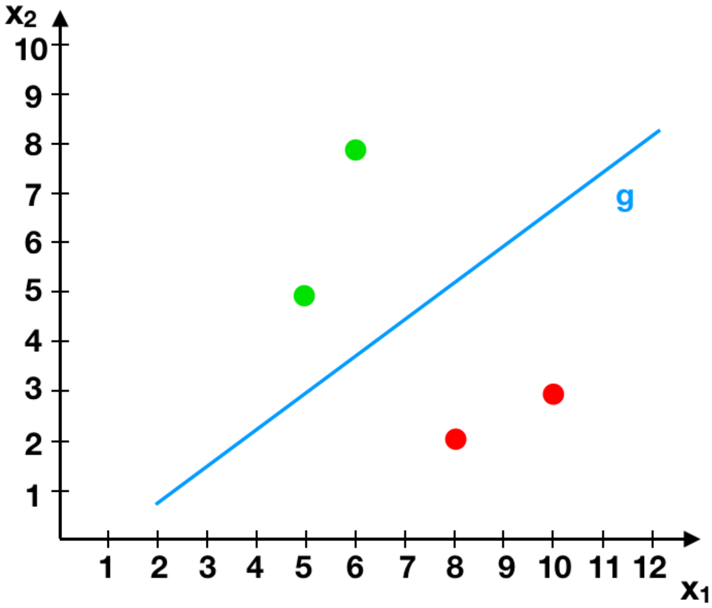


Abbildung 7: 4 Trainingsdaten mit Trennfunktion g (blau).

Geradengleichung **in Normalenform**: $g : \begin{pmatrix} -3 \\ 4 \end{pmatrix} * \left(\begin{pmatrix} x1 \\ x2 \end{pmatrix} - \begin{pmatrix} 4 \\ 2 \end{pmatrix} \right) = 0$

Gebt dazu die Werte für den Stützvektor, den Normalenvektor und den jeweiligen Datenpunkt in den Code ein und drückt *Run Section.*

Tragt die Abstandswerte in die Tabelle auf dem Antwortzettel ein.

```
x = [NaN;NaN];  %Tragt hier den jeweiligen Trainingsdatenpunkt  ein (aus A
p = [NaN;NaN];  %Tragt hier den Stützvektor ein
n = [NaN;NaN];  %Tragt hier den Normalenvektor ein

abstand = d(x,p,n) %hier wird eure Formel für d von oben benutzt
```

Frage 5.6: Was fällt euch auf, wenn ihr die Abstandswerte mit der Lage des jeweiligen Punktes und seiner Klasse (rot, grün) vergleicht? Erkennt ihr eine Gemeinsamkeit?

!!!!! ERST FRAGE 5.6 BEANTWORTEN, DANN RUNTER SCROLLEN ZU AUFGABE 6 !!!!!!

Aufgabe 6

Eure Abstandsformel liefert für Trainingspunkte der grünen Ampelklasse ein positives und für Trainingspunkte der roten Ampelklasse ein negatives Vorzeichen. Der Abstand ist in der Mathematik jedoch immer als positiver Wert definiert. Ihr müsst also noch eine Änderung an der Abstandsformel vornehmen.

Wie im Einstiegsvortrag schon beschrieben, ist für die Trainings- und Testdaten schon bekannt, welcher Klasse sie angehören. In unserem Fall haben wir einen Datensatz von Ampeln, von denen wir wissen, ob sie grün oder rot sind. Dementsprechend werden die Daten gekennzeichnet. Alle grünen Ampeldaten erhalten das Kennzeichen $t = +1$ und alle roten Ampeln erhalten das Kennzeichen $t = -1$.

Auftrag 6.1: Ändert eure Abstandsformel `d(x,p,n)` so um, dass sie ausschließlich positive Werte ausgibt.

Hinweis: Nutzt die neuen Informationen über die Kennzeichnungen t. Denkt daran, wieder alle Variablen in `@(...)` zu schrieben.

```
d = @(x,p,n,t) NaN;
Abstand2(d);
```

Aufgabe 7

Nun habt ihr eine Formel für den Abstand aufgestellt, sodass ihr schon einen Teil des wörtlichen Befehsl aus Auftrag 5.2 für den Computer mathematisch ausdrücken könnt. Vermutlich habt ihr einen ähnlichen Befehl formuliert wie:

"Der Computer soll die Trenngerade finden, für die der **Abstand** zu den **nächst gelegenen** Datenpunkten **maximal** wird."

Den Abstand könnt ihr bererits mit der neuen Formel ausdrücken, fehlt nur noch der Rest!

Das Problem ist, dass nun gleichzeitig die Parameter n und p der Trenngeraden gesucht werden und zu dieser Trenngeraden die Abstände bestimmt werden müssen. Es handelt sich um ein doppeltes Optimierungsproblem, welches mit MATLAB gelöst werden kann:

Befolgt die folgenden Schritte zur Lösung des Problems.

Schritt 1

Auftrag 7.1: Lasst euch zunächst den Trainingsdatensatz anzeigen, er besteht aus zwei Klassen mit je zehn Punkten.

Druckt dazu *Run Section.*

```
clear
Punkte;
```

Schritt 2

Für den Datensatz soll nun die Gerade gesucht werden, die den maximalen Margin hat. Der **Abstand** zu den **nächst gelegenen** Datenpunkten soll **maximal** werden.

In MATLAB kann dies wie folgt programmiert werden:

```
% Optimization: ABSCHNITT NICHT BEARBEITEN %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
options.MaxFunEvals = 1e5;
[u] = ga(@(u)fcn(u,x1,x2,t),4);          %%hier....
[v] = fminsearch(@(u)fcn(u,x1,x2,t),u); %%... und hier wird der optimale Al

n1 = u(1); n2 = u(2); p1 = u(3); p2 = u(4);
xplot = linspace(-5,5,100);
%plot(xplot,-(n1*xplot-(n1*p1+n2*p2))/n2);
xlim([min(x1) max(x1)]);
ylim([min(x2) max(x2)]);
axis equal
hold on;
n1 = v(1);
n2 = v(2);
p1 = v(3);
p2 = v(4);
plot(x1(1:N1),x2(1:N1),'ro',x1(N1+1:end),x2(N1+1:end),'bo');
```

```
plot(xplot,-(n1*xplot-(n1*p1+n2*p2))/n2);
legend('Klasse 1','Klasse 2', 'Trennfunktion');
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Drückt auf *Run Section* und MATLAB bestimmt die Gerade mit maximalem Margin.

Abschnitt Funktion

```
function y = fcn(u,x1,x2,t)

%%%%% zu optimierende Parameter %%%%
n1 = u(1); % x Koordinate Normalenvektor
n2 = u(2); % y Koordinate Normalenvektor
p1 = u(3); % x Koordinate Punkt P
p2 = u(4); % y Koordinate Punkt P
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%% Optimierungsfunktionen %%%%
f = t.*(n1*x1+n2*x2-(n1*p1+n2*p2))/sqrt(n1^2+n2^2);    % Minimierung umgescl
y = max(f);     %Bestimmung des Maximums der minimalen Abstände
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

end
```

Für Schnelle: Wählt aus folgenden Aufgaben aus

Auftrag S.1:

Ihr sollt euch in dieser Aufgabe genauer mit der Suche nach der optimalen Gerade beschäftigen. In Schritt 2 von Aufgabe 7 wurde die optimale Gerade mit MATLAB bestimmt. Dafür musste eure Abstandsfunktion jedoch noch umgeschrieben werden. Wie oben beschrieben, soll der **Abstand der Gerade zu den Datenpunkten optimiert werden**. Die Abstandsformel wird durch die Vektoren $\vec{n}$ und $\vec{p}$ beschrieben. Diese beiden Vekoren haben jeweils 2 Eintrage: eine x- und eine y-Koordinate. Insgesamt gibt es also vier Parameter **die während der Optimierung bestimmt** werden müssen:

n1: x-Koordinate von $\vec{n}$ n2: y-Koordinate von $\vec{n}$ p1: x-Koordinate von $\vec{p}$ p2: y-Koordinate von $\vec{p}$

Bis jetzt hat eure Abstandsformel die Gestalt:

```
d = t*(dot((x-p),n)/norm(n));
```

Damit alle 4 Parameter optimiert werden können, muss die Formel umgeformt werden.

Formt die Formel für d so um, dass die vier Optimierungsparamter vorkommen. Nutzt dazu den Platz auf dem Antwortzettel und die folgenden Hinweise.

Hinweis: Das Skalarprodukt von zwei Vektoren kann folgendermaßen aufgelöst werden:

$$\vec{a} = \begin{pmatrix}1\\2\end{pmatrix}, \vec{b} = \begin{pmatrix}3\\4\end{pmatrix} \quad \vec{a} * \vec{b} = \begin{pmatrix}1\\2\end{pmatrix} * \begin{pmatrix}3\\4\end{pmatrix} = 1 \cdot 3 + 2 \cdot 4 = 11$$

Hinweis: Beim Skalarprodukt gilt das Distributivgesetz!

Hinweis: Der Betrag eines Vektors kann folgendermaßen aufgelöst werden:

$$\vec{a} = \begin{pmatrix}1\\2\end{pmatrix} \qquad |\vec{a}| = \left|\begin{pmatrix}1\\2\end{pmatrix}\right| = \sqrt{1^2 + 2^2} = \sqrt{5}$$

Die Wurzel wird mit `sqrt()` *in MATLAB eingegeben.*

Hinweis: Für `t` *muss* `t.` *geschrieben werden.*

Nutzt bei Bedarf die Hilfekarten.

Vergleicht eure Lösung mit der Funktion `f` im obigen Abschnitt Funktion (orange).

Auftrag S.2:

Die Geraden, die ihr am Anfang in Aufgabe 3 eingegeben habt, haben die Form $y = m * x + b$. Formt diese Darstellungsform in die Normalenform um.

Nutzt bei Bedarf die Hilfekarten.

A.1.2. Aufgabenblatt 2

Aufgabenblatt 2 | Modellverbesserung: drei Ampelphasen

Auf dem ersten Übungsblatt wurden Ampeln mit nur zwei Phasen, rot und grün, betrachtet. Um die Realität besser abbilden zu können nehmen wir nun noch die Phase gelb dazu.

Aufgabe 1

Abbildung 1: gelbe Ampel

Der Code um eine grüne bzw. eine rote Ampel mit zwei Boxen zu erzeugen sah so aus:

```
clear

%Box 1

r1 = 180;

g1 = 0;

b1 = 0;

%Box 2

r2 = 70;

g2 = 70;

b2 = 70;

Ampel(r1,g1,b1,r2,g2,b2);
```

Auftrag 1.1: Verändert den Code so, dass eine gelbe Ampel wie in Abbildung 1 dargestellt wird. Tragt den neuen Code in das untere Feld ein.

Falls ihr Probleme habt die gelbe Farbe zu erzeugen, sucht im Internet nach *additiver Farbmischung.*

```
clear
%Box 1
r1 = NaN;
g1 = NaN;
b1 = NaN;
%Box 2
```

```
r2 = NaN;
g2 = NaN;
b2 = NaN;
%Box 3
r3 = NaN;
g3 = NaN;
b3 = NaN;
Ampel3(r1,g1,b1,r2,g2,b2,r3,g3,b3);
```

Aufgabe 2

Aus den Vektoren mit zwei Einträgen vom ersten Modell werden nun Vektoren mit drei Einträgen, welche die zusätzliche Ampelphase berücksichtigen.

Auftrag 2.1: Drückt den *Run Section* Button.

```
neuerDatensatz;
```

Euch wird ein neuer Datensatz aus 100 grünen, 100 roten und 100 gelben Ampeln angezeigt. Wenn ihr mit dem Zeiger über die **Abbildung** geht, erscheint oben das Symbol . Durch einen Klick darauf könnt ihr die Abbildung drehen.

Benutzt die Drehfunktion um die Abbildung zu bewegen.

Frage 2.2: Auf dem ersten Arbeitsblatt habt ihr die Daten mit Geraden separiert. Was würdet ihr hier benutzen um sie zu trennen?

!!!!! ERST FRAGE 2.2 BEANTWORTEN, DANN RUNTER SCROLLEN ZU AUFGABE 3 !!!!!!

Aufgabe 3

Eine Ebene kann man in der Mathematik auf verschiedene Weisen beschreiben. Aus der Schule kennt ihr eventuell schon die Parameterform:

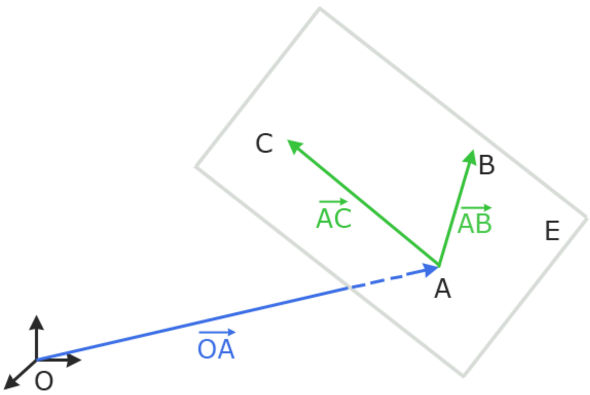


Abbildung 2: Ebene in Paramterform

Hier wird die Ebene durch einen *Stützvektor* $\overrightarrow{OA}$ und die beiden

Richtungsvektoren $\overrightarrow{AC}$ und $\overrightarrow{AB}$ beschrieben. Es werden also drei Punkte A, B und C benötigt um die Gleichung für die Ebene aufzustellen:

$$E : \vec{x} = \overrightarrow{OA} + s \cdot \overrightarrow{AC} + r \cdot \overrightarrow{AB} \ . \tag{1}$$

Wie auf Blatt 1 bereits bei einer Gerade gesehen, kann man auch eine Ebene durch die Normalenform ausdrücken:

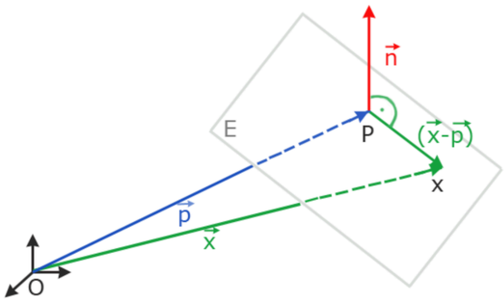


Abbildung 3: Ebene in Normalenform

Hier wird die Ebene durch einen *Stützvektor* $\vec{p}$ und den Normalenvektor $\vec{n}$ beschrieben. Jeder Punkt mit einem Ortsvektor $\vec{x}$ liegt genau dann in der Ebene, wenn folgende Bedingung gilt:

$$\vec{n} * (\vec{x} - \vec{p}) = 0 \ , \qquad (2)$$

dabei steht $*$ für das *Skalarprodukt* zweier Vektoren. Das *Skalarprodukt* kennt ihr schon vom ersten Arbeitsblatt:

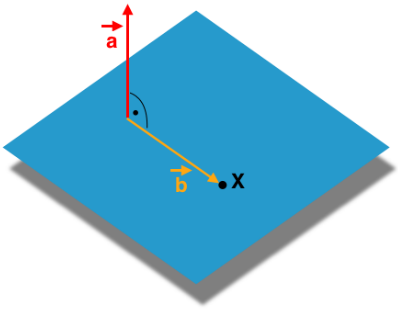
Stehen zwei Vektoren senkrecht zueinander, so ist das *Skalarprodukt* der beiden Vektoren 0. Damit ist Gleichung (2) anschaulich nachvollziehbar: Der Normalenvektor steht nach seiner Definition senkrecht auf der Ebene. Alle Verbindungsvektoren $\vec{p} - \vec{x}$ von Punkt P zu Punkten X, die in der Ebene liegen, verlaufen ebenfalls in der Ebene. Deshalb steht der Normalenvektor auch senkrecht auf den Verbindungsstrecken. Die Bedingung (2) beschreibt also eine Ebene

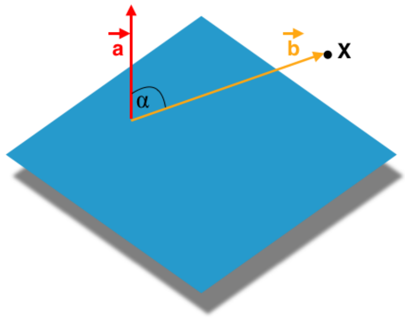
$$E : \vec{n} * (\vec{x} - \vec{p}) = 0 \ , \qquad (3)$$

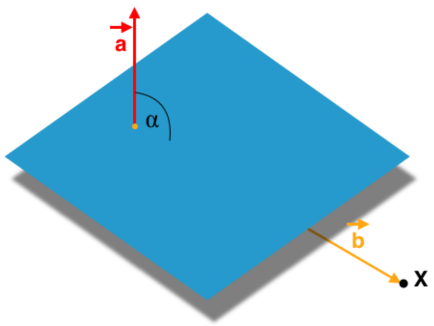
Ihr sollt nun untersuchen, was mit Punkten X passiert, die nicht in der Ebene liegen. Dazu seht ihr hier wieder die Definition des *Skalarproduktes* von Blatt 1.

$$\vec{a} * \vec{b} = cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}| \ . \qquad (4)$$

In der folgenden Abbildung sind drei Fälle dargestellt. Einmal liegt der Punkt in, einmal oberhalb und einmal unterhalb der Ebene.







Fall 1: Punkt in der Ebene

Fall 2: Normalenvektor zeigt in die Richtung,
in der der Punkt liegt.

Fall 3: Normalenvektor zeigt entgegen
der Richtung, in der der Punkt liegt.

Auftrag 3.1: Stellt auf dem Antwortzettel unter 3.1 eine allgemeine Fallunterscheidung für Formel (4) für die drei oben gezeigten Fälle 1-3 auf. Überlegt euch, warum es ausreicht **nur Winkel zwischen 0° und 180° zu betrachten!**

Nutzt bei Bedarf die Hilfekarten.

Auftrag 3.2: Übertragt eure Fallunterscheidung auf die Ebenengleichung (3): Vervollständigt dazu die Fallunterscheidung und den Satz bei Auftrag 3.2.

Modellierungsschritt: Vereinfachtes Modell

In der folgenden Aufgabe gehen wir zunächst davon aus, dass die Ampeln nur in zwei Klassen aufgeteilt werden. Die eine Klasse setzt sich aus den roten und gelben Ampeln zusammen, die andere aus den grünen Ampeln. Diesen Modellierungsschritt machen wir um zu testen, ob sich unser Verfahren für zweidimensionale Daten von Blatt 1auf dreidimensionale Daten übertragen lässt.

Aufgabe 4

Nun sollen die Datenpunkte aus Aufgabe 2 durch die *beste Ebene* getrennt werden. Im Grunde wird das gleiche Verfahren angesetzt, dass ihr zur Berechnung der *besten* Gerade auf Blatt eins erarbeitet habt. Da es nun jedoch mehr Datenpunkte gibt und diese auch noch dreidimensional sind, wird die Funktion, die optimiert werden, soll noch ein wenig verändert.

Auftrag 4.1: Drückt den *Run Section* Button.

```
BerechneEbene;
```

In der Tabelle wird unter Abstandsminimum der Wert für den Abstand des Datenpunktes angegeben, der am nächsten an der Ebene liegt. Unter n wird der Normalenvektor angegeben. Unter b wird der Wert für $\vec{n} * (-\vec{p})$ aus der Ebenengleichung angegeben.

Auftrag 4.2: Klassifiziert folgende 3 Testdaten mit der Entscheidungsfunktion und tragt die Ergebnisse in die Tabelle auf dem Antwortzettel unter 4.2 ein.

1. $A_1(50|20|20)$
2. $A_2(20|200|100)$
3. $A_3(50|10|130)$

Hinweis: Es ist wichtig, dass ihr euch zunächst anseht in welche Richtung der Normalenvektor $\vec{n}$ *verläuft!*

Rechnung:

MATLAB TOOLBOX

Betrag (Länge) eines Vektors	norm(x)
Wurzel	sqrt(x)
Skalarprodukt	dot(x,y)

```
n = [NaN ; NaN ; NaN];  %Normalenvektor aus 4.1
b = NaN; %Wert für b aus 4.1
x = [NaN ; NaN ; NaN]; %Koordinaten Ampel A
Klassifizierung = dot(n,x)+b      %Formel Entscheidungsfunktion SIEHE KASTEN AUFGABE 3
```

Ihr könnt eure Lösung hier überprüfen, indem ihr für A1-A5 die Zahlen 1 oder 2 einsetzt. **Dabei steht 1 für grün und 2 für rot/gelb**.

```
A1 = NaN;
A2 = NaN;
A3 = NaN;
checkAmpeln;
```

Modellierungsschritt: Modellverbesserung

In Aufgabe 4 haben wir gesehen, dass das Verfahren für zweidimensionale Daten auf dreidimensionale Daten übertragen werden kann. Ihr habt nun die Ebene gefunden, welche die Daten für zwei Klassen am besten separiert.

Das Problem ist, dass der Computer bis jetzt immer nur zwei Klassen voneinander trennt. In Aufgabe 4 trennt er die Klasse der grünen von der Klasse der gelben und roten Ampeln.

Ihr sollt in dieser Aufgabe einen Code schreiben, mit dem man die Ampeln in drei Klassen rot, gelb und grün unterteilen kann.

Aufgabe 5

Auftrag 5.1: Überlegt euch ein Verfahren, das die Daten in drei Klassen klassifiziert. Formuliert dazu eine Art Anleitung/Rezept auf dem Antwortblatt. *Wendet höchstens fünf Minuten für diesen Auftrag auf!*

Hinweis: Orientiert euch am bereits gelernten Verfahren zur Klassifizierung in 2 Klassen. Nutzt das Internet!

Nutzt bei Bedarf die Hilfekarten.

Im Folgenden sollt ihr mit einem gängigen Verfahren weiterarbeiten. Dazu könnt ihr entweder das *One-Versus-One* oder das *One-Versus-All Verfahren* auswählen. Solltet ihr in 5.1 ein anderes Verfahren entwickelt haben, diskutiert mit einem Betreuer darüber.

Auftrag 5.2: Informiert euch über die Verfahren und setzt eines der beiden in MATLAB um. Klassifiziert dazu den Punkt $X\ (60 \mid 80 \mid 40)$.

Die Daten der drei Klassen sind schon im folgenden Code unter ListeGruen3, ListeRot3 und ListeYellow3 gespeichert!

- Mit dem Befehl: `Separation3(class_1, class_2);` könnt ihr **je** zwei Klassen voneinander trennen lassen.
- mehrere Klassen können mit `class_1 = [NaN;NaN];` eingegeben werden

Nutzt bei Bedarf die Hilfekarten.

Hier könnt ihr einzelne Klassen voneinander trennen lassen. Gebt für class_1 bzw. class_2 die jeweilige(n) Klasse(n) ein.

```
clear
load ListeGruen3.mat
load ListeRot3.mat
load ListeYellow3.mat
```

Gebt dafür für NaN die jeweilige Klasse ein:

Grün = ListeGruen3

Rot = ListeRot3

Gelb = ListeYellow3

```
class_1 = [NaN;NaN]; %mehrere Klassen können mit [NaN;NaN] eingegeben werden !!
class_2 = [NaN];
Separation3(class_1, class_2);
```

In der Tabelle findet ihr wieder den **Normalenvektor** $\vec{n}$ und unter **b** den Wert für $\vec{n} * (-\vec{p})$ der jeweiligen Ebene.

Im nächsten Abschnitt könnt ihr die Werte aus der Tabelle nutzen, um den Punkt in die ausgewählten Klassen einzuordnen.

Hier könnt ihr den Punkt $X \;(60 \mid 80 \mid 40)$ mit der Entscheidungsfunktion einordnen lassen.

Hinweis: Die Formel für die Entscheidungsfunktion habt ihr schon in Aufgabe 3 und 4 aufgestellt.

Schaut euch an in welche Richtung jeweils der Normalenvektor zeigt und nutzt wieder die Fallunterscheidung aus Auftrag 3.2, um zu entscheiden in welcher Klasse der Punkt liegt.

- Falls ihr das *One-Versus-One Verfahren* nutzt: Auf dem Antwortblatt könnt ihr in die Tabelle eintragen welches Ergebnis jede der drei eins-gegen-eins Klassifikationen liefert.

```
n = [NaN; NaN ; NaN]; %Normalenvektor
b = NaN; %Wert für b
x = [NaN;NaN;NaN]; %Koordinaten der Ampel
Klassifizierung = dot(x,n)+b      %Formel für Klassifizierung
```

- Falls ihr das *One-Versus-All* Verfahren nutzt: Auf dem Antwortblatt könnt ihr in die Tabelle eintragen welche Werte sich für den Abstand ergeben und so nachher das Maximum suchen.

```
n = [NaN; NaN ; NaN]; %Normalenvektor
b = NaN; %Wert für b
x = [NaN;NaN;NaN]; %Koordinaten der Ampel
Abstand = (dot(x,n)+b)/norm(n)      %Formel für die Abstandsberechnung
```

Gebt im Code für `NaN` ein, welcher Klasse ihr den Punkt X mit eurem Verfahren insgesamt zugeordnet habt. Dabei gilt:

gelb = 1

grün = 2

rot = 3

Überprüft euer Ergebnis, indem ihr *Run Section* drückt.

```
Klasse = NaN;
Check3(Klasse);
```

Für Schnelle:

Nun soll das *One-Versus-One* Verfahren näher betrachtet werden. Für die, die in Aufgabe 7 das andere Verfahren benutzt haben, eine kurze Erklärung:

Beim *One-Versus-One* Verfahren werden immer zwei Klassen miteinander verglichen und jeweils eine separierende Funktion für die beiden Klassen gesucht. Dies wird für alle möglichen Kombinationen der Klassen gemacht. Ein Datenpunkt, der klassifiziert werden soll, wird dann der Klasse hinzugefügt, zu der er durch die einzelnen One-versus-One Klassifikationen am häufigsten hinzugefügt wurde.

Auftrag S.1: Drückt wieder *Run Section*

```
svm_3d_3klassen;
```

Auftrag S.2: Klassifiziert den Punkt $S = (5|105|43)$. Nutzt dafür die Formeln aus Aufgabe 5.

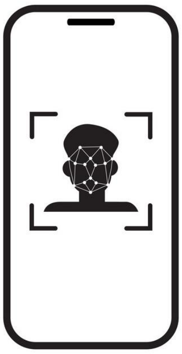
Wo ist das Problem des *One-Versus-One* Verfahrens? Benutzt die Drehfunktion um die Abbildung zu bewegen.

Stellt Überlegungen an, wie ihr das Verfahren erweitern könnt, um das Problem zu beheben.

A.1.3. Aufgabenblatt 3

Aufgabenblatt 3 | Anwendung auf eigenen Datensatz: Gesichtserkennung

Als letztes sollt ihr nun die erarbeitete Methode zur Support Vektor Machine auf einen neuen Kontext anwenden: die Gesichtserkennung.



Aufgabe 1

Übertragt zunächst das Modell der drei Ampelboxen auf das Foto eines Gesichtes.

Die Ampeln mit drei Boxen haben wir als Vektoren mit drei Einträgen dargestellt. Man schriebt auch **3x1 Vektoren,** da sie aus **drei Zeilen und einer Spalte** bestehen. Die Boxen der Ampeln kann man sich auch als Pixel eines Bildes vorstellen. So wird eine Ampel zu einem Bild aus drei Pixeln. Der oberste Pixel ist die Box für die rote Lampe, der mittlere Pixel die Box für die gelbe Lampe und der unterste Pixel ist die Box für die grüne Lampe.

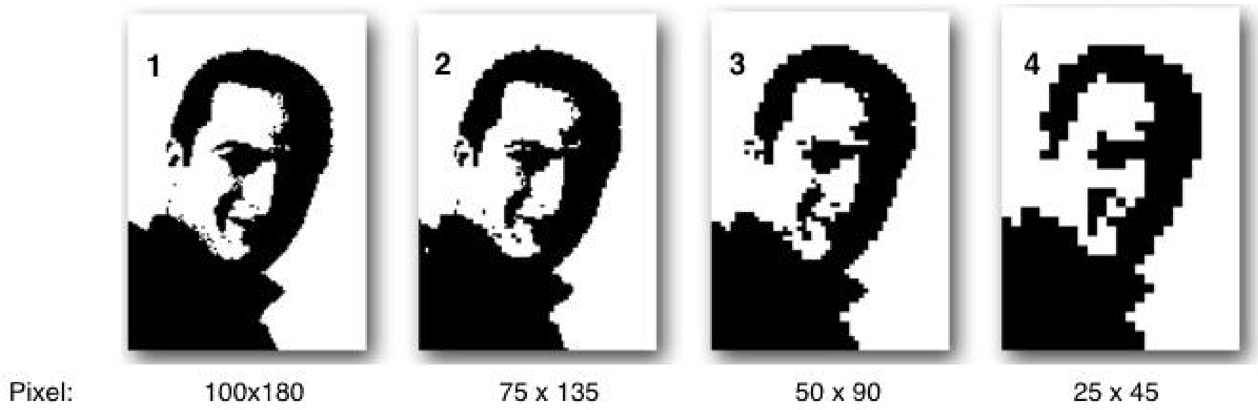


Abbildung 1: Foto eines Gesichtes in verschiedenen Auflösungen (Pixelzahlen)

Das Gesicht in Abbildung 1 ist in vier verschiedenen Auflösungen dargestellt. In Bild 1 sieht man eine hohe Auflösung mit 1**00 Pixeln in der Breite und 180 Pixeln in der Höhe**. Man schriebt auch **100x180**. Die Auflösung wird immer geringer, bis sie bei Bild 4 nur noch bei 25x45 liegt.

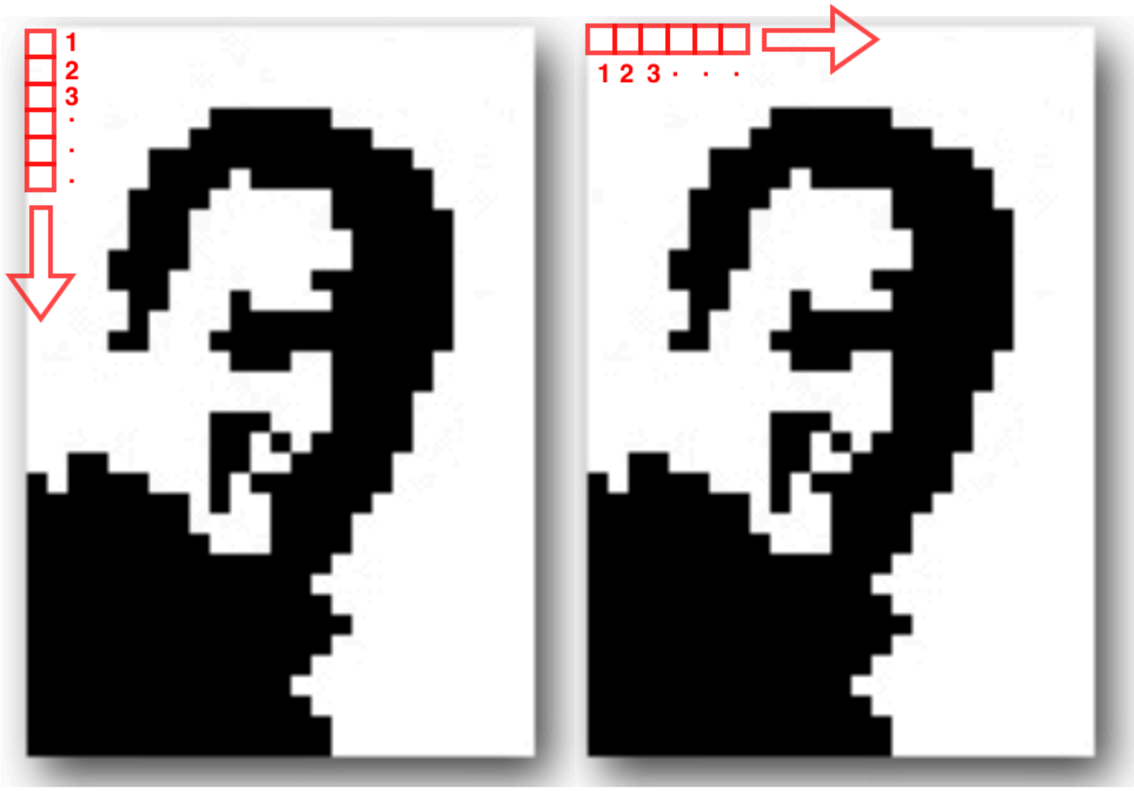
Frage/Auftrag 1.1: Die Bilder aus Abbildung 1 sollen analog zu den Ampeln auch als Vektoren **mit einer Spalte** geschrieben werden. Wie viele Zeilen hätten dann jeweils die neuen Vektoren für die Bilder 1 - 4?

Tragt die Werte für `'NaNxNaN'` in den Code ein und überprüft eure Lösung mit *Run Section.*

Hinweis: Ein Vektor mit 1024 Zeilen und einer Spalte würde mit '1024x1' eingegeben werden.

```
Bild1 = 'NaN';
Bild2 = 'NaN';
Bild3 = 'NaN';
Bild4 = 'NaN';
checkBild;
```

Frage 1.2: Betrachtet Abbildung 2. Wie würdet ihr die Pixel in den Bildern abzählen? Spielt die Art und Weise abzuzählen eine Rolle? Worauf ist zu achten, wenn man mehrere Bilder für einen Datensatz abspeichern will?



(1) (2)

Abbildung 2: 2 verschiedene Weisen die Pixel eines Bildes zu zählen.

Frage 1.3: Was könnten bei der Bilderkennung Vorteile von langen Vektoren z.b. 145300x1 gegenüber kürzeren mit z.B. 1376x1 sein? Was könnten Nachteile sein?

Frage 1.4: Die Trennfunktionen, die ihr bisher betrachtet habt, haben Daten im zwei- bzw. dreidimensionalen Raum separiert. Welche Dimension hätte der Raum von Bild 4 aus Abbildung 1?

Aufgabe 2

Ihr arbeitet nun mit eurem **eigenen Datensatz**. Der Datensatz besteht aus den Fotos, welche ihr vor der Mittagspause aufgenommen habt.

Auftrag 2.1: Bringt zunächst die nötigen Schritte in die richtige Reihenfolge:

- Trenn- und Entscheidungsfunktion suchen (1)
- Funktion mit Testdaten testen (2)

- Daten in Test- und Trainingsdaten einteilen (3)

Tragt die Richtige Reihenfolge in den Code ein und überprüft euer Ergebnis mit *Run Section.*

```
Schritt_1 = NaN;
Schritt_2 = NaN;
Schritt_3 = NaN;
checkSchritte;
```

Aufgabe 3

<u>Auftrag 3.1:</u> Wählt nun die entsprechenden Matlab Skripte (`Schritt_1 - Schritt_3)` aus und bearbeitet sie.

Beginnt mit `Schritt_1`!

Aufgabe 4

In Schritt drei habt ihr schon bemerkt, dass der Computer die Daten sehr gut trennen kann. Hier sollt ihr das Ganze an einem Beispiel überprüfen.

<u>Auftrag 4.1:</u> Lasst den Computer herausfinden, bei wem es sich auf einem Foto handelt.

Gebt dazu im Code unter `load` für `NaN` den Namen der Gruppe ein, also beispielsweise `'02_Max'` und drückt auf *Run Section.*

```
load 'data/NaN'; %hier anstelle von NaN einen fremden Gruppennamen eingeben!
zuordnungtest;
```

Überprüft, ob die Person auf dem eingelesenen Bild richtig zugeordnet wurde.

<u>Auftrag 4.2:</u> Lasst den Computer nun einen der folgenden Prominenten einordnen.

Folgende fünf Prominenten stehen zur Auswahl:

- Justin Bieber: `NaN = justin`

- Heidi Klumm: `NaN = heidi`
- Christiano Ronaldo: `NaN = christiano`
- Kylie Janner: `NaN = kylie`
- Will Smith: `NaN = will`

```
load 'promis/NaN'; %hier anstelle von NaN einen fremden Gruppennamen eingeben!
zuordnungtest;
```

Frage 4.3: Interpretiert das Ergebnis. Überlegt euch dazu, was "am ähnlichsten" mathematisch bedeutet? Wie könnte der Computer das festlegen?

Frage 4.4: Wie würdet ihr das Verfahren ändern, um eine bessere Zuordnung der Prominenten zu euren Mitschülern zu erhalten?

A.1.4. Bonusblatt

Bonusblatt | Modellverbesserung

Aufgabe 1

<u>Auftrag 1.1:</u> Ladet den neuen Datensatz, indem ihr *Run Section* drückt.

```
neueDatenSlack;
```

<u>Frage 1.2:</u> Was könnte hier problematisch werden? Nutzt wieder die Funktion, indem ihr mit der Maus über die Abbildung geht.

<u>Auftrag 1.3:</u> Bestimmt mit Hilfe des Computers und dem `BerechneEbene` Befehl von Aufgabe 5 die Ebene für die Datenpunkte, indem ihr auf Run Section klickt. *Hinweis: es wird das OVO Verfahren verwendet*

```
BerechneEbene;
```

<u>Auftrag 1.4:</u> Bewertet die Lösung des Computers. Nutzt wieder die -Funktion.

Öffnet die Datei Zwischenvortrag.pdf auf eurem Desktop und geht diese zu zweit durch. Hier erfahrt ihr, wie man das Problem lösen kann.

Aufgabe 2

Ihr habt in der Zwischenpräsentation erfahren, was passiert, wenn die Daten nicht sauber durch eine Gerade oder durch eine Ebene getrennt werden können.

Bei sich überschneidenden Klassen werden die *Slack Variablen* eingeführt, welche durch einen Faktor C reguliert werden können.

In der ersten Aufgabe sollt ihr die Auswirkung des Faktors untersuchen.

Auftrag 2.1: Ladet noch mal das Ergebnis von Aufgabe 1 indem ihr *Run Section* drückt.

```
    BerechneEbene;
```

Auftrag 2.2: Nun sollen die Slack Variablen genutzt werden. Gebt verschiedene Werte für C aus **dem Intervall** $\left[\ 1 \cdot 10^{-4} \quad 1 \cdot 10^{8}\ \right]$ ein und drückt *Run Section.*

```
C = NaN;
BerechneEbeneSlack;
```

Es werden nun drei Ebenen angezeigt. Die mittlere ist die Trennfunktion und die anderen beiden Ebenen geben die Größe des Margin an.

Unter dem Wert Genauigkeit in der Tabelle wird die Anzahl der korrekt eingeordneten Testdatenpunkte angezeigt. Dabei wird ihre Zuordnung durch die Trennebene mit ihrer wahren Klasse (gegeben durch die Labels) verglichen.

Auftrag 2.3: Was passiert für verschiedene Werte von C? Tragt eure Ergebnisse in die Tabelle unter 2.3 auf dem Antwortblatt ein. Wählt auch mal sehr große C wie 10^{20} !

A.2. Antwortblätter

A.2.1. Antwortblatt 1

- Antwortblatt zu Aufgabenblatt 1 -

Aufgabe 3

Auftrag 3.1

Trennfunktion für Trainingsdaten in Abbildung 1 einzeichnen.

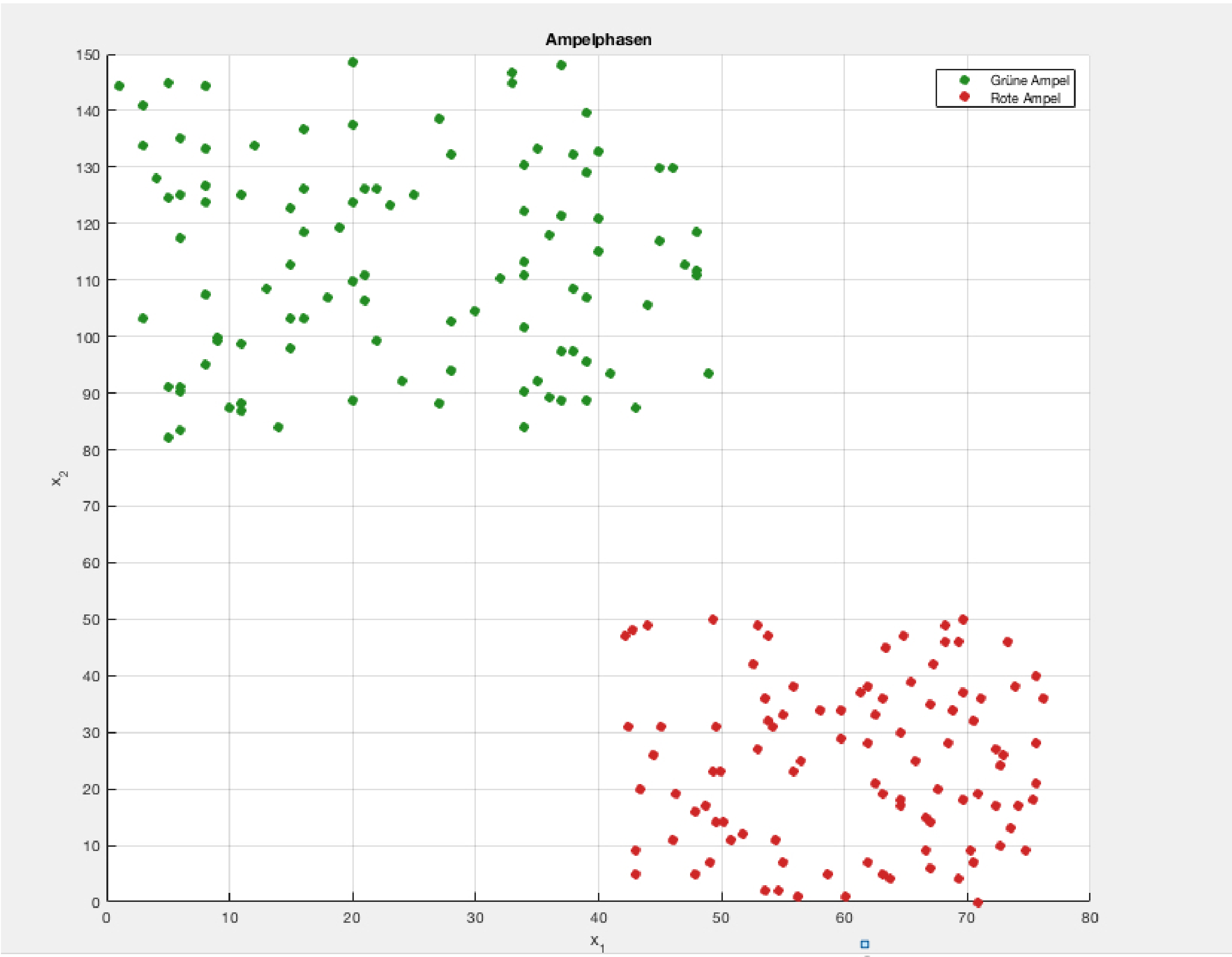


Abbildung 1: Trainingsdatensatz aus 100 grünen und 100 roten Ampeln

Auftrag 3.2

Ampel mit $x_1 = 20$ und $x_2 = 50$ in Abbildung 1 einzeichnen.
Zuordnung in Klasse: .

Auftrag 3.3

Tabelle für Zuordnung der Testdaten:

Tabelle 1: Klassifizierung der Testdatenpunkte

Ampel	Gerade a	Gerade b	Gerade c	eigene Gerade e
1				
2				
3				
4				
5				
6				
Anzahl richtiger Klassifizierungen				

Auftrag 3.4

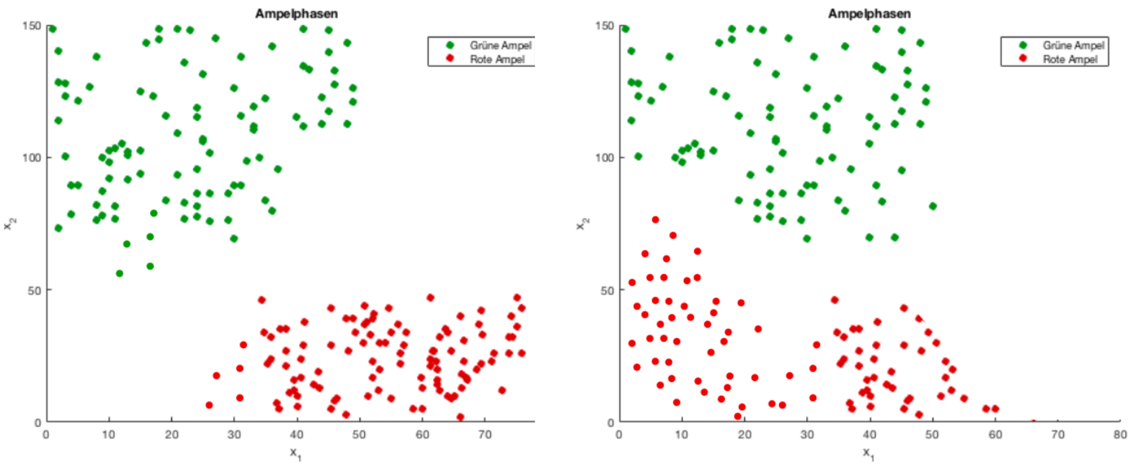
Welche Gerade klassifiziert die Testdaten am besten?

Antwort:

Aufgabe 4

Auftrag 4.1

Zeichne die Trennfunktion in Abbildung 2 ein, welche die Daten eurer Meinung nach am besten trennt. Umkreist Punkte, die ihr als Orientierung genutzt habt.



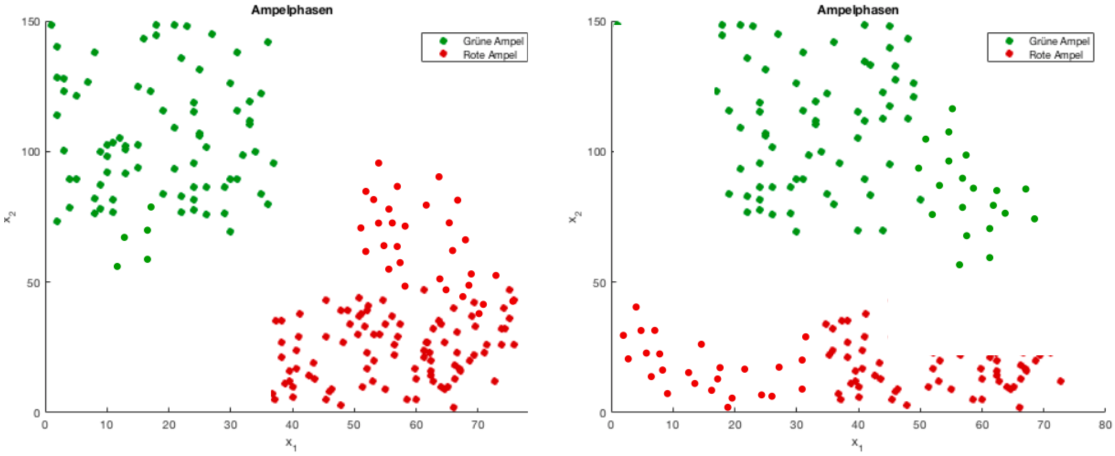


Abbildung 2: Weitere Trainingsdaten, die getrennt werden sollen

Frage 4.2

Worauf habt ihr beim Zeichnen geachtet? Was zeichnet für euch die *„beste Gerade"* aus?

Antwort:

Aufgabe 5

Auftrag 5.2

Wörtlicher Befehl an den Computer, um die *beste Gerade* zu finden:
„Finde die Gerade, die

Auftrag 5.3

Rechnung für Abstandsformel:

Auftrag 5.5
Tabelle für Abstandswerte:

Tabelle 2: Abstände der 4 Datenpunkte und ihre Zuordnung

Koordinaten Punkt	Klasse	Abstand

Frage 5.6
Was fällt auf, wenn ihr die Abstandswerte mit der Lage der Punkte und ihrer Klasse vergleicht?

Antwort:

Aufgaben für Schnelle

Auftrag S.1
Platz für Umstellung der Abstandsformel:

```
d = t * (dot((x-p),n) / norm(n))
```

Auftrag S.2
Umformung von $y = m * x + b$ in die Normalenform:

A.2.2. Antwortblatt 2

- Antwortblatt zu Aufgabenblatt 2 -

Aufgabe 2

Frage 2.2

Was würdet ihr nun verwenden, um die Daten zu trennen?

Antwort:

Aufgabe 3

Auftrag 3.1

Allgemeine Fallunterscheidung für $\vec{a} * \vec{b} = \cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}|$ für die drei Fälle:

$$\vec{a} * \vec{b} = \cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}| \begin{cases} > 0 & \text{für } \square \leqslant \alpha < \square \\ = 0 & \text{für } \qquad \alpha = \square \\ < 0 & \text{für } \square < \alpha \leqslant \square \end{cases}$$

Auftrag 3.2

Fallunterscheidung für Ebene:

$$\vec{n} * (\vec{x} - \vec{p}) = \boxed{\phantom{\vec{n} * (\vec{x} - \vec{p})}} \begin{cases} \square & \text{für } \square \leqslant \alpha < \square \\ = 0 & \text{für } \qquad \alpha = \square \\ \square & \text{für } \square < \alpha \leqslant \square \end{cases}$$

Zeigt der Normalenvektor **nach oben aus der Ebene**, ist der Ausdruck $\vec{n} * (\vec{x} - \vec{p})$ also negativ für Punkte ____ der Ebene und positiv für Punkte ____ der Ebene.

Zeigt der Normalenvektor **nach unten aus der Ebene**, ist der Ausdruck $\vec{n} * (\vec{x} - \vec{p})$ also negativ für Punkte ____ der Ebene und positiv für Punkte ____ der Ebene.

Für Punkte ____ der Ebene, ist der Ausdruck stets Null.
Somit ist $\vec{n} * (\vec{x} - \vec{p})$ die gesuchte **Entscheidungsfunktion**, mit der wir Datenpunkte klassifizieren können. Dazu ist es wichtig zunächst zu prüfen, in welche Richtung der Normalenvektor zeigt.

Im Folgenden wird $\vec{n} * (\vec{x} - \vec{p})$ in der Form $\vec{n} * \vec{x} + b$ benutzt, wobei b der Wert von $-\vec{n} * \vec{p}$ ist.

Lösungswörter: oberhalb - unterhalb - in

Aufgabe 4

Auftrag 4.2

Klassifikation der drei Ampeln $A_1 - A_3$:
Richtung des Normalenvektors bezüglich der Ebene:

Tabelle 1: Klassifizierung der drei Testampeln

Ampel	Wert aus Entscheidungsfunktion	Klasse
1		
2		
3		

Aufgabe 5

Auftrag 5.1

Anleitung/Rezept für Trennverfahren:

Auftrag 5.2

Für One-Versus-One Verfahren:

Tabelle 2: Ergebnisse der 3 einzelnen Klassifikationen

	SVM grün vs. rot	SVM gelb vs. rot	SVM gelb vs. grün
Wert Entsch. Funktion			
Zugeteilte Klasse			

Für One-Versus-All Verfahren:

Tabelle 3: Ergebnisse der 3 einzelnen Klassifikationen

	SVM grün & gelb vs. rot	SVM gelb & rot vs. grün	SVM rot & grün vs. gelb
Betrag Abstandsformel			
Multiplikation mit ±1			

Antwort

Zuordnung des Punktes X $(60 \mid 80 \mid 40)$ in Klasse:

Aufgaben für Schnelle

Auftrag S.2

Wo ist das Problem beim One-Versus-One Verfahren?

Antwort

A.2.3. Antwortblatt 3

- Antwortblatt zu Aufgabenblatt 3 -

Aufgabe 1

Frage 1.2

Wie würdet ihr die Pixel in den Bildern abzählen? Spielt die Art und Weise eine Rolle? Worauf ist zu achten, wenn man mehrere Bilder für einen Datensatz abspeichern will?

Antwort:

Frage 1.3

Was könnten Vorteile von langen bzw. kurzen Vektoren sein? Wo liegen Nachteile?

Antwort:

Frage 1.4

Welche Dimension hätte der Raum von Bild 4?

Antwort:

Aufgabe 4

Frage 4.3

Was bedeutet „am ähnlichsten“ mathematisch ? Wie könnte der Computer das festlegen?

Antwort:

Frage 4.4

Wie würdet ihr das Verfahren ändern, um eine bessere Zuordnung der Prominenten zu euren Mitschülern zu erhalten?

Antwort:

A.2.4. Antwortblatt Bonus

- Antwortblatt zum Bonusblatt -

Aufgabe 1

Frage 1.2

Was könnte problematisch werden?

Antwort:

Auftrag 1.4

Lösung des Computers bewerten:

Aufgabe 2

Auftrag 2.3

Ergebnisse für verschiedene C-Werte:

Tabelle 1: Auswirkungen verschiedener Werte für C

C	Korrekte Klassifizierung [%]	Ebenenabstand

Was passiert für die verschiedenen Werte? Was passiert für sehr große Werte wie $C = 10^{20}$?

A.3. Lösungen

A.3.1. Lösung zu Blatt 1

CAMMP day

Gesichtserkennung und Autonomes Fahren mit Mathematik?!

- Lösungsblatt zu Aufgabenblatt 1 -

Aufgabe 1

Auftrag 1.1

Verschiedene rote Ampeln erzeugen:
Um **rote** Ampeln zu erzeugen, muss für r1, also den Rotanteil für die obere Box ein Wert zwischen 1 und 255 gewählt werden. Die Werte g1 und b1 müssen bei 0 bleiben. Die Werte r2, g2 und b2 also die Farbwerte für die untere Box müssen alle gleich gewählt werden und können Werte zwischen 0 und 255 annehmen.
Eine beispielhafte Codeeingabe würde lauten:

```
clear
%Box 1
r1 = 155;
g1 = 0;
b1 = 0;
%Box 2
r2 = 10;
g2 = 10;
b2 = 10;
Ampel(r1,g1,b1,r2,g2,b2);
```

Auftrag 1.2

Verschiedene grüne Ampeln erzeugen:
Um **grüne** Ampeln zu erzeugen, muss für g2, also den Grünanteil für die untere Box ein Wert zwischen 1 und 255 gewählt werden. Die Werte r2 und b2 müssen bei 0 bleiben. Die Werte r1, g1 und b1 also die Farbwerte für die obere Box müssen alle gleich gewählt werden und können Werte zwischen 0 und 255 annehmen.
Eine beispielhafte Codeeingabe würde lauten:

```
clear
%Box 1
r1 = 0;
g1 = 50;
b1 = 00;
%Box 2
r2 = 0;
g2 = 25;
b2 = 0;
Ampel(r1,g1,b1,r2,g2,b2);
```

Aufgabe 2

Auftrag 2.1

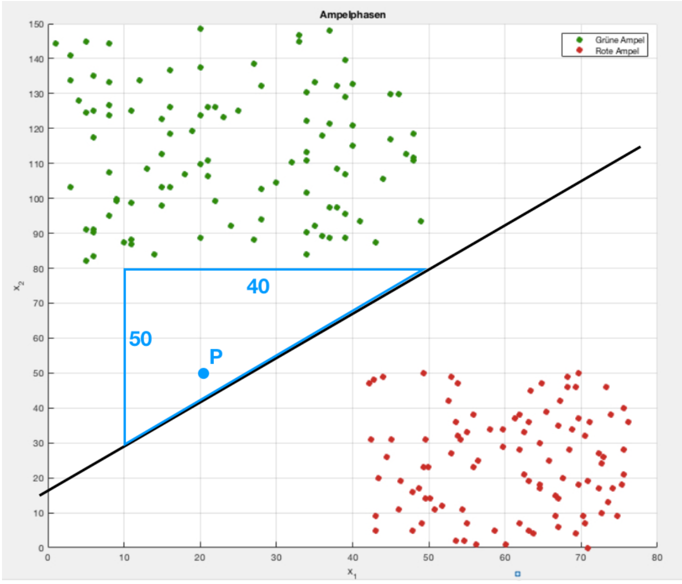
RGB-Tripel zu einem Wert zusammenfassen: Um aus den RGB-Tripeln einen Wert zu generieren empfiehlt es sich den Grauwert zu benutzen. Der Grauwert kann mit der Formel $0.299 \cdot r + 0.587 \cdot g + 0.114 \cdot b$ aus den RGB Werten berechnet werden. Diese Formel findet man im Internet und auch in der MATLAB help. Die Codeeingabe lautet:

```
x = @(r,g,b) 0.299*r+0.587*g + 0.114* b; %hier Gleichung eingeben
Grauwertermittlung(x);
```

Aufgabe 3

Auftrag 3.1

Es gibt unendlich viele Möglichkeiten, die Datenpunkte zu trennen. Wichtig ist, dass alle Trainingsdatenpunkte richtig zugeteilt werden. Eine mögliche Gerade mit Steigungsdreieck könnte so aussehen:



Auftrag 3.2
Die Ampel P wird durch die dargestellte Trennfunktion zu den grünen Ampeln zugeordnet.

Auftrag 3.3
Hier muss zunächst die eigene Gerade eingegeben werden. Steigung und y-Achsenabschnitt sollen die Schüler aus der Zeichnung bestimmen.
y = @(x) 50/40 * x + 16 ;%Geradengleichung eurer Gerade
m = 50/40 ; % Wert für Steigung
n = 16 ; % Wert für y-Achsenabschnitt

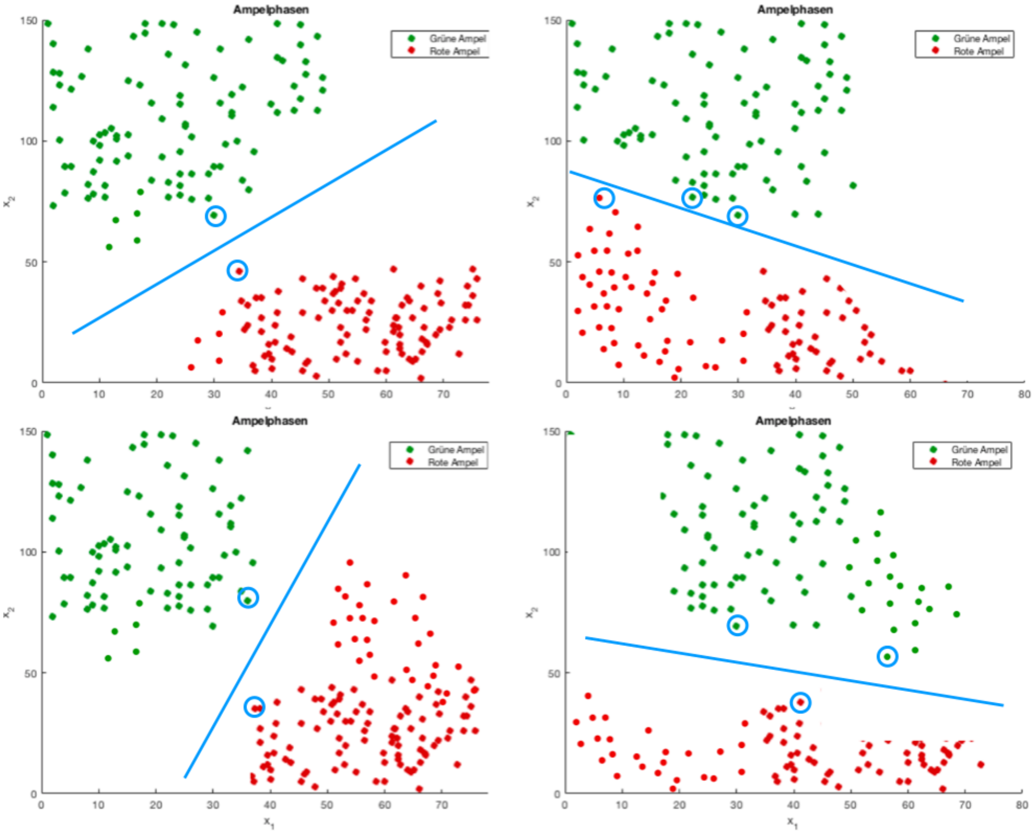
Ampel = 1; %Nummer der Ampel
x1 = 20;
x2 = 50;
Ampeltest(x1,x2,y,m,n,Ampel);

Die ausgefüllte Tabelle sollte folgendermaßen aussehen:

Ampel	Gerade a	Gerade b	Gerade c	eigene Gerade e
1	FALSCH	FALSCH	RICHTIG	RICHTG
2	RICHTIG	RICHTIG	RICHTIG	RICHTG
3	RICHTIG	FALSCH	RICHTIG	RICHTG
4	RICHTIG	RICHTIG	RICHTIG	RICHTG
Anzahl richtiger Klassifizierungen	**3**	**2**	**4**	**4**

Auftrag 3.4
Die Gerade c klassifiziert die Testdatenpunkte am besten. In Abbildung 3 ist zu sehen, dass die Gerade c am weitesten von den Datenpunkten entfernt ist. Dieser Punkt kann hier schon von den Schülern bemerkt werden. Es kann außerdem passieren, dass die eigene Gerade der Schüler die Testdaten genau so gut klassifiziert wie die Gerade c.

Aufgabe 4
Auftrag 4.1/4.2
Beim Zeichnen der Geraden sollte darauf geachtet werden, den Abstand zu den Datenpunkten möglichst groß zu wählen. Die Geraden könnten folgendermaßen aussehen:



Die beste Gerade wird dadurch ausgezeichnet, dass der Bereich um die Gerade, der frei von Datenpunkten ist, möglichst groß ist. Die Gerade sollte also so gewählt werden, **dass der Abstand zu den nächst gelegenen Datenpunkten maximal ist.**

Aufgabe 5

Auftrag 5.1

Es ist immer die Gerade zu wählen, bei der der kürzeste Abstand maximal ist. Die Eingabe im Code sollte dann so aussehen:

```
Abbildung1 = 'b';
Abbildung2 = 'a';
Abbildung3 = 'c';
Abbildung4 = 'a';
Abstandtest(Abbildung1,Abbildung2,Abbildung3,Abbildung4);
```

Auftrag 5.2

Spätestens hier sollten die Schüler dann den Befehl an den Computer formulieren können. Wichtig ist, dass sinngemäß vorkommt, dass der Margin maximiert werden soll. Eine mögliche Formulierung wäre:
„Finde die Gerade, die den maximalen Abstand zu den nächst gelegenen Trainingsdatenpunkten aufweist. “

Auftrag 5.3

Die Forme für den Abstand kann wie folgt hergeleitet werden:
Zunächst müssen zwei Formeln für α aufgestellt werden:

$$\cos(\alpha) = \frac{d}{|\overrightarrow{PX}|} \tag{1}$$

$$\cos(\alpha) = \frac{\vec{n} * \overrightarrow{PX}}{|\vec{n}| \cdot |\overrightarrow{PX}|} \tag{2}$$

Gleichsetzen liefert:

$$\begin{aligned} \frac{d}{|\overrightarrow{PX}|} &= \frac{\overrightarrow{PX} * \vec{n}}{|\overrightarrow{PX}| \cdot |\vec{n}|} && | \cdot |\overrightarrow{PX}| \\ d &= \frac{\overrightarrow{PX} * \vec{n}}{|\vec{n}|} && | \text{ ersetze } \overrightarrow{PX} = \vec{x} - \vec{p} \\ d &= \frac{(\vec{x} - \vec{p}) * \vec{n}}{|\vec{n}|} \end{aligned}$$

Auftrag 5.4

Die Eingabe in MATLAB lautet dann

```
d = @(x,p,n) dot((x-p),n)/norm(n); %Tragt hier eure Formel für den Abstand ein
```

Auftrag 5.5

In dieser Aufgabe sollen die Schüler die neu gelernte Formel anwenden. Die Eingabe im Code sollte so aussehen:

```
x = [5;5]; %Tragt hier den jeweiligen Trainingsdatenpunkt ein
p = [4;2]; %Tragt hier den Stützvektor ein
n = [-3;4]; %Tragt hier den Normalenvektor ein

abstand = d(x,p,n)
```

Auftrag 5.6

Es sollte auffallen, dass Punkte oberhalb der Geraden g, die zur grünen Klasse gehören ein negatives Vorzeichen für den Abstand erhalten. Punkte unterhalb der Geraden, die zur Klasse der roten Ampeln gehöre erhalten ein positives Vorzeichen für den Abstand. Hier liegt die Vermutung schon nah, dass man die Formel auch als Entscheidungsfunktion nutzen könnte. Sollten die Schüler diesen Vorschlag äußern kann er aufgenommen werden und auf Blatt 2 verwiesen werden. An dieser Stelle ist es jedoch wichtig, dass ein einheitlicher Abstandsbegriff ohne verschiedene Vorzeichen festgehalten wird. Die Schüler haben ja schon gefordert, das der Abstand zu den nächst gelegenen Punkten maximal sein soll. Dafür ist es wenig sinnvoll, wenn Punkte verschiedene Vorzeichen für den Abstand hätten.

Aufgabe 6

Auftrag 6.1

Hier sollen die Schüler die Labels t mit einbeziehen. Die Formel für den Abstand ändert sich zu

```
d=@(x,p,x,t) t*dot((x-p),n)/norm(n);
```

Es ist wichtig, dass Schülern erklärt wird, dass auch der Betrag für gleihce Vorzeichen sorgen würde, jedoch bezieht dieser nicht die Vorinformationen über die Daten ein. In Aufgabe 7 kann es dann passieren, dass der Punkt auf die falsche Seite der Geraden gelegt wird und so eine falsche Trennfunktion resultiert.

Aufgaben für Schnelle

Auftrag S.1

Die Formel aus MATLAB lautet umgeschrieben:

$$d = t \cdot \vec{n} * (\vec{x} - \vec{p}) \, / |\vec{n}|$$

Nun kann das Skalarprodukt und der Betrag des Vektors aufgelöst werden.

$$\begin{aligned} d &= t \cdot \begin{pmatrix} n1 \\ n2 \end{pmatrix} * \begin{pmatrix} x1 \\ x2 \end{pmatrix} - \begin{pmatrix} n1 \\ n2 \end{pmatrix} * \begin{pmatrix} p1 \\ p2 \end{pmatrix})/\sqrt{n1^2 + n2^2} \\ d &= t \cdot (n1 \cdot x1 + n2 \cdot x2 - (n1 \cdot p1 + n2 \cdot p2)) \, / \sqrt{n1^2 + n2^2} \end{aligned}$$

Der Abschnitt Funktion im Code lautet dann:

```
%%%%% zu optimierende Parameter %%%%
n1 = u(1); % x Koordinate Normalenvektor
n2 = u(2); % y Koordinate Normalenvektor
p1 = u(3); % x Koordinate Punkt P
p2 = u(4); % y Koordinate Punkt P

%%%%% Optimierungsfunktionen %%%%
f = t.*(n1*x1+n2*x2-(n1*p1+n2*p2))/sqrt(n1^2+n2^2);
% Minimierung umgeschriebene Abstandsfunktion
y = max(f); %Bestimmung des Maximums der minimalen Abstände
```

Auftrag S.2

Ersetzt man y durch $x2$ und x durch $x1$, erhält man

$$x2 = m \cdot x1 + b$$

Nun kann x2 auf die andere Seite bringen und die Formel mit der Normalenform vergleichen:

$$0 = \underbrace{n1 \cdot x1 + n2 \cdot x2}_{m \cdot x1 - 1 \cdot x2} - (n1 \cdot p1 + n2 \cdot p2)$$

Damit hat man den Normalenvektor schon gefunden:

$$\vec{n} = \begin{pmatrix} m \\ -1 \end{pmatrix}$$

Für den Vektor $\vec{p}$ kann man nun einfach den y-Achsenabschnitt benutzen. Der Punkt $(0|b)$ liegt auf jeden Fall auf der Geraden. Damit lautet die umgeformte Gleichung:

$$g : 0 = \begin{pmatrix} m \\ -1 \end{pmatrix} * \left(\begin{pmatrix} x1 \\ x2 \end{pmatrix} - \begin{pmatrix} 0 \\ b \end{pmatrix} \right) .$$

A.3.2. Lösung zu Blatt 2

CAMMP day

Gesichtserkennung und Autonomes Fahren mit Mathematik?!

- Lösungsblatt zu Aufgabenblatt 2 -

Aufgabe 1

Auftrag 1.1

Die Box für die gelbe Ampelphase wird in die Mitte zwischen der roten und der grünen Box gesetzt. Für gelbe Farben aus RGB-Tripeln muss der Rot- und der Grünanteil r2 und g2 jeweils gleich gewählt werden und kann Werte von 1-255 annehmen. Der Blauanteil b2 bleibt bei 0. Außerdem sollten die anderen beiden Boxen wieder schwarz sein. Die Werte r1,g1 und b1 sollten also gleich sein und einen Wert von 0-255 annehmen. Genau so auch die Werte r3,g3 und b3, wobei sie nicht denen der ersten Box entsprechen müssen.
Eine beispielhafte Eingabe könnte so aussehen:

```
clear
%Box 1
r1 = 70;
g1 = 70;
b1 = 70;
%Box 2
r2 = 180;
g2 = 180;
b2 = 0;
%Box 3
r3 = 70;
g3 = 70;
b3 = 70;
Ampel3(r1,g1,b1,r2,g2,b2,r3,g3,b3);
```

Aufgabe 2

Frage 2.2

Da die Daten nun dreidimensional sind, werden die Trennfunktionen durch Ebenen gegeben.

Aufgabe 3

Frage 3.1

Die Fallunterscheidung sieht wie folgt aus:

$$\vec{a} * \vec{b} = \cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}| = \begin{cases} > 0 & \text{für } 0^\circ \leqslant \alpha < 90^\circ \\ = 0 & \text{für } \alpha = 90^\circ \\ < 0 & \text{für } 90^\circ < \alpha \leqslant 180^\circ \end{cases}$$

Auftrag 3.2

Fallunterscheidung für Ebene:

$$\vec{n} * (\vec{x} - \vec{p}) = \begin{cases} > 0 & \text{für } 0^\circ \leqslant \alpha < 90^\circ \\ = 0 & \text{für } \alpha = 90^\circ \\ < 0 & \text{für } 90^\circ < \alpha \leqslant 180^\circ \end{cases}$$

Zeigt der Normalenvektor **nach oben aus der Ebene**, ist der Ausdruck $\vec{n} * (\vec{x} - \vec{p})$ also negativ für Punkte **unterhalb** der Ebene und positiv für Punkte **oberhalb** der Ebene.

Zeigt der Normalenvektor **nach unten aus der Ebene**, ist der Ausdruck $\vec{n} * (\vec{x} - \vec{p})$ also negativ für Punkte **oberhalb** der Ebene und positiv für Punkte **unterhalb** der Ebene.

Für Punkte **in** der Ebene, ist der Ausdruck stets Null.
Somit ist $\vec{n} * (\vec{x} - \vec{p})$ die gesuchte **Entscheidungsfunktion**, mit der wir Datenpunkte klassifizieren können. Dazu ist es wichtig zunächst zu prüfen, in welche Richtung der Normalenvektor zeigt.

Im Folgenden wird sie in der Form $\vec{n} * \vec{x} + b$ benutzt, wobei b der Wert von $-\vec{n} * \vec{p}$ ist.

Aufgabe 4

Frage 4.2

Die Werte für den Normalenvektor n und für b müssen aus der Ausgabe für Auftrag 4.1 gezogen werden. Die Eingabe lautet dann:

```
n = [0.0179 ; 0.0029 ; -0.0528]; %Normalenvektor aus 4.1
b = 2.7753; %Wert für b aus 4.1
x = [30 ; 40 ; 100]; %Koordinaten der Ampel A
Klassifizierung = dot(n,x)+b %Formel für Entscheidungsfunktion aus Aufgabe 3
```

Die Einordnung der Testdaten lautet:

Klasse \ Ampel	1	2	3
grüne Ampeln			x
rote/gelbe Ampeln	x	x	

Es könnte den Schülern hier schwer fallen die Richtung des Normalenvektors zu bestimmen. Man kann sie darauf hinweisen, dass sie sich lediglich die x_3-Koordinate ansehen müssen um zu entscheiden, ob der Vektor nach oben oder nach unten aus der Ebene zeigt.

Aufgabe 5

Frage 5.1

Es werden im Folgenden die Tabellen für das One-Versus-One Verfahren und das One-Versus-All Verfahren angegeben.

	SVM grün vs. rot	SVM gelb vs. rot	SVM gelb vs. grün
Wert Entsch. Funktion	-1.528	0.5413	0.3676
Zugeteilte Klasse	rot	rot	grün

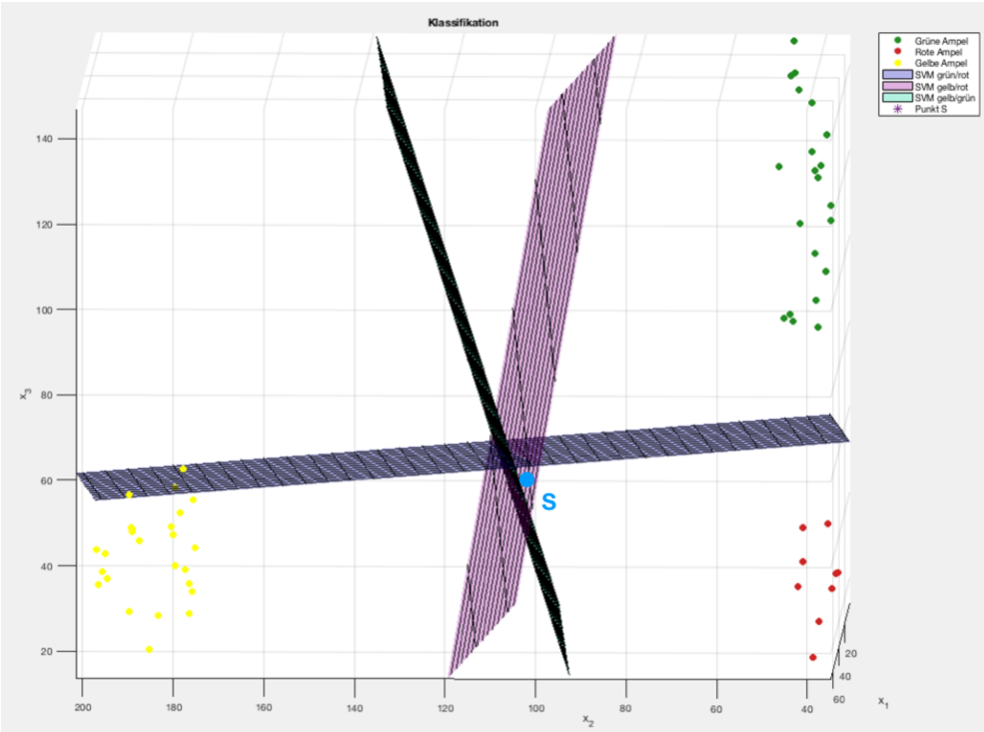
	SVM grün & gelb vs. rot	SVM gelb & rot vs. grün	SVM rot & grün vs. gelb
Betrag Abstandsformel	27.3976	36.6632	6.5798
Multiplikation mit ±1	-1	-1	1

Der Punkt $X\ (60 \mid 80 \mid 40)$ wird also in die Klasse der roten Ampeln zugeordnet.

Aufgaben für Schnelle

Auftrag S.2

Das Problem bei der One-Versus-One Methode ist, dass es einen Bereich gibt, indem die Datenpunkte nicht eindeutig zugeordnet werden. Genau in diesem Bereich liegt Punkt S:



Versucht man S mit dem OVO Verfahren einzuordnen, so wird er bei den 3 Einzel-Klassifikationen jeder Klasse einmal zugeordnet und es gibt keinen „Sieger“.
Das Verfahren könnte verbessert werden, indem die Abstände des Punktes zu den Trenngeraden betrachtet werden. Der Punkt wird dann der Klasse zugeordnet, zu deren Trennfunktion der Abstand minimal ist.

A.3.3. Lösung zu Blatt 3

- Lösungsblatt zu Aufgabenblatt 3 -

Aufgabe 1
Auftrag 1.1
Für die Spaltenvektoren ergeben sich folgende Werte im Code:

```
Bild1 = '18000x1';
Bild2 = '10125x1';
Bild3 = '4500x1';
Bild4 = '1125x1';
checkBild;
```

Frage 1.2
Hier sollte mit den Schülern diskutiert werden, dass es völlig egal ist, wie man die Pixel abzählt und in den Vektor überträgt. Wichtig ist, dass man es für alle Bilder gleich macht.

Frage 1.3
Auch hier kann wieder mit den Schülern diskutiert werden. Es sollte klar werden, dass der Vorteil von langen Vektoren in einem höheren Detailreichtum besteht. Es werden mehr Informationen abgespeichert. Dementsprechend ist jedoch auch der Rechenaufwand höher. Darin liegt der Vorteil der kurzen Vektoren.

Frage 1.4
Der Raum hätte eine Dimension von $25 \cdot 45 = 1125$.

Aufgabe 2
Auftrag 2.1
Die Reihenfolge der Schritte im Code lautet:

```
Schritt_1 = 3;
Schritt_2 = 1;
Schritt_3 = 2;
checkSchirtte;
```

Für die restlichen Schritte gibt es keine Musterlösung mehr. Die Schüler sollen einfach mit ihrem Datensatz rumspielen und verschiedene Bilder einordnen lassen. Die Bilder, die eingeordnet werden sollen, sind die Grimassen, welche vor der Pause aufgenommen wurden.

Aufgabe 4
Frage 4.3
Die Schüler sollten hier diskutieren, wie der Computer diese Daten klassifiziert. Die Promis sind nicht im Trainingsdatensatz enthalten gewesen und so gibt es keine *richtige* Zuordnung. Es kann durchaus passieren, dass männliche Promis Schülerinnen zugeordnet werden. Der Computer legt „am ähnlichsten“ fest, indem er den minimalen Abstand zu einer der Klassen ausrechnet.

Frage 4.4
Eine Möglichkeit das Verfahren zu verbessern wäre es den Computer mit vielen Promi-Bildern zu trainieren und dann die Bilder der Schüler als unbekannte Daten einordnen zu lassen.

A.3.4. Lösung zum Bonus Blatt

- Lösungsblatt zum Bonusblatt -

Aufgabe 1

Frage 1.2

Es fällt auf, dass die Trainingsdaten nicht mehr linear trennbar sind. Sie überlappen sich und es gibt keinen Trennbereich mehr, der frei von Datenpunkten ist.

Auftrag 1.4

Die Lösung des Computers erscheint wenig sinnvoll. Die Klassen werden durch die vorgeschlagene Ebene **nicht** voneinander getrennt.

Aufgabe 2

Auftrag 2.3

Für verschiedene Werte von C ergibt sich folgende Tabelle:

C	Korrekte Klassifizierung [%]	Ebenenabstand
10^(-04)	95	48.7326
10^(-03)	97	22.6202
10^(-02)	97	12.7551
10^(-01)	97	6.8832
10^0	97	4.7480
10^1	97	4.7480
10^2	97	4.7480
10^3	97	4.7480
10^4	97	4.7480
10^9	97	4.7480
10^10	97	4.7545
10^11	90	4.7480
10^12	92	3.1375

An den Werten ist zu erkennen, dass für kleinere C der Margin immer größer wird. Es wird immer mehr Trainingspunkten erlaubt falsch klassifiziert zu werden oder im Margin zu liegen. Für ganz große C ist man wieder in der Situation ohne Slack-Variablen. Für die großen Werte ist die Rechenzeit außerdem viel größer als bei den kleinen Werten. Dies erkennt man an den Rechenschritten des Optimierungsverfahrens.

A.4. Hilfekarten

A.4.1. Hilfekarte zu Blatt 1

- Gestufte Hilfekarten für Auftrag 5.2 auf Blatt 1 -

Niveau 1

Versucht zwei verschiedene Formeln für $\cos(\alpha)$ aufzustellen und diese dann gleichzusetzen.
Hinweis: Nutzt den Wechselwinkel!

Niveau 2

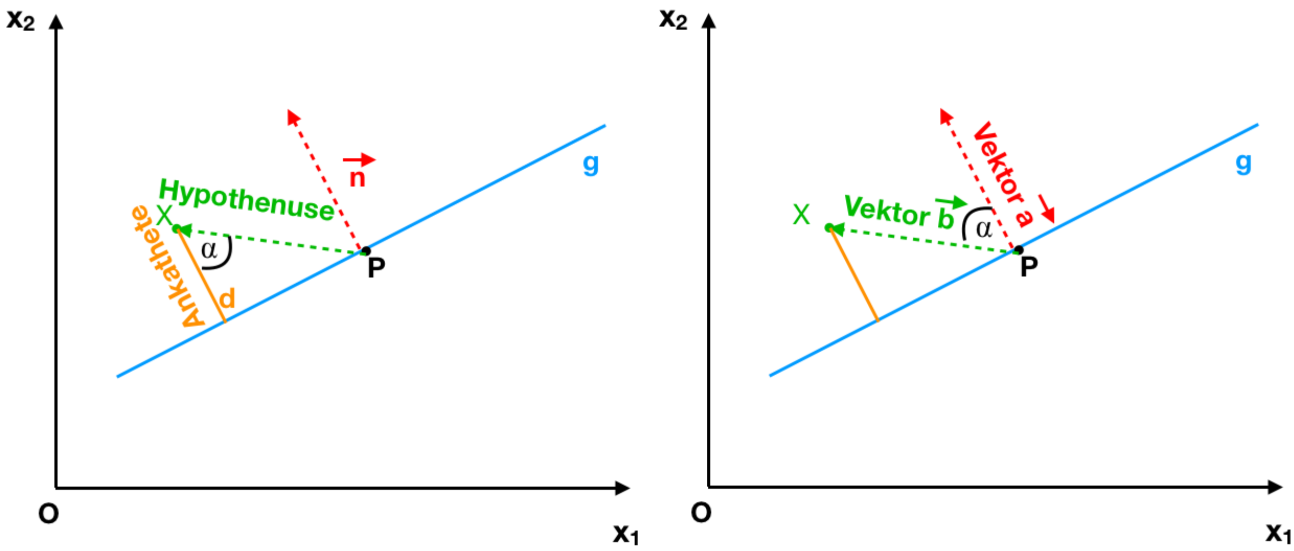
Die erste Formel für $\cos(\alpha)$ erhaltet ihr über

$$\cos(\alpha) = \frac{Ankathete}{Hypothenuse} \ .$$

Die zweite Formel erhaltet ihr aus der Definition des Skalarproduktes:

$$\vec{a} * \vec{b} = \cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}| \ .$$

Niveau 3



(a) Skizze für erste Formel

(b) Skizze für zweite Formel

Niveau 4 

Die beiden Formeln lauten:

$$\cos(\alpha) = \frac{d}{|\overrightarrow{PX}|} \tag{1}$$

$$\cos(\alpha) = \frac{\vec{n} * \overrightarrow{PX}}{|\vec{n}| \cdot |\overrightarrow{PX}|} \tag{2}$$

Niveau 5 

Setzt (1) und (2) gleich und löst nach d hin auf. Dabei könnt ihr $\overrightarrow{PX}$ noch umschrieben, sodass $\vec{x}$ und $\vec{p}$ vorkommen.

Niveau 6 

Gleichsetzen liefert:

$$\begin{aligned} \frac{d}{|\overrightarrow{PX}|} &= \frac{\overrightarrow{PX} * \vec{n}}{|\overrightarrow{PX}| \cdot |\vec{n}|} && | \cdot |\overrightarrow{PX}| \\ d &= \frac{\overrightarrow{PX} * \vec{n}}{|\vec{n}|} && | \text{ ersetze } \overrightarrow{PX} = \vec{x} - \vec{p} \\ d &= \frac{(\vec{x} - \vec{p}) * \vec{n}}{|\vec{n}|} \end{aligned}$$

- Gestufte Hilfekarten für Auftrag S.1 auf Blatt 1 -

Niveau 1

Die Ausgangsgleichung für den Abstand hat die Form

```
d = t*(dot((x-p),n)/norm(n))
```

Damit alle vier Optimierungsparameter vorkommen, müsst ihr die Befehle `dot()` und `norm` als mathematischen Ausdruck umschreiben. `dot()` steht für das Skalarprodukt * und `norm` für den Betrag $|\cdot|$

Niveau 2

Schreibt man die Befehle `dot()` und `norm` um, so ergibt sich:

$$d = t \cdot ((\vec{x} - \vec{p}) * \vec{n}) \,/\, |\vec{n}|$$

Nun müssen die Vektoren $\vec{x}$, $\vec{p}$ und $\vec{n}$ noch so umgeschrieben werden, dass ihre Einträge vorkommen.

Niveau 3

x,p und n sind Vektoren mit jeweils 2 Einträgen. $\vec{x}$ kann geschrieben werden als $\begin{pmatrix} x1 \\ x2 \end{pmatrix}$. Schreibt alle Vektoren in der Abstandsformel auf diese Weise um.

Niveau 4

Schreibt man die Vektoren mit Einträgen auf, so ergibt sich

$$d = t \cdot \left(\left(\begin{pmatrix} x1 \\ x2 \end{pmatrix} - \begin{pmatrix} p1 \\ p2 \end{pmatrix} \right) * \begin{pmatrix} n1 \\ n2 \end{pmatrix} \right) / \left| \begin{pmatrix} n1 \\ n2 \end{pmatrix} \right|$$

Löst nun das Skalarprodukt und den Betrag mit Hilfe der Hinweise unter der Aufgabenstellung auf.

Niveau 5 

Für das Skalarprodukt gilt das Distributivgesetz:

$$\begin{pmatrix} a1 \\ a2 \end{pmatrix} * \left(\begin{pmatrix} b1 \\ b2 \end{pmatrix} - \begin{pmatrix} c1 \\ c2 \end{pmatrix} \right) = \begin{pmatrix} a1 \\ a2 \end{pmatrix} * \begin{pmatrix} b1 \\ b2 \end{pmatrix} - \begin{pmatrix} a1 \\ a2 \end{pmatrix} * \begin{pmatrix} c1 \\ c2 \end{pmatrix}$$

Niveau 6 

Nutzt man den dritten Hinweis und löst das Skalarprodukt und den Betrag auf, so ergibt sich:

```
t.*(n1*x1+n2*x2-(n1*p1+n2*p2))/sqrt(n1^2+n2^2)
```

- Gestufte Hilfekarten für Auftrag S.2 auf Blatt 1 -

Niveau 1

Bei der Geradengleichung $y = m \cdot x + b$ kann y als $x2$ und x als $x1$ geschrieben werden.

Niveau 2

Ersetzt man y durch $x2$ und x durch $x1$, erhält man

$$x2 = m \cdot x1 + b \ .$$

Bringt nun $x2$ auf die andere Seite und vergleicht das Ergebnis mit der Normalenform

$$0 = \vec{n} * (\vec{x} - \vec{p})$$

Niveau 3

Für einen besseren Vergleich, müsst ihr das Skalarprodukt in

$$0 = \vec{n} * (\vec{x} - \vec{p})$$

ähnlich wie in Auftrag 7 auflösen.

Niveau 4

Löst man das Skalarprodukt auf, so ergibt sich für die Normalenform die Gleichung:

$$0 = n1 \cdot x1 + n2 \cdot x2 - (n1 \cdot p1 + n2 \cdot p2) \tag{3}$$

Die andere Gleichung aus $y = m \cdot x + b$ lautet:

$$0 = m \cdot x1 - 1 \cdot x2 + b \tag{4}$$

Vergleicht (1) und (2).

Niveau 5

Löst man das Skalarprodukt auf, so ergibt sich für die Normalenform die Gleichung:

$$0 = \underbrace{n1 \cdot x1 + n2 \cdot x2}_{m \cdot x1 - 1 \cdot x2} - (n1 \cdot p1 + n2 \cdot p2)$$

Damit hat man den Normalenvektor schon gefunden:

$$\vec{n} = \begin{pmatrix} m \\ -1 \end{pmatrix}$$

Es fehlt nun lediglich noch der Ortsvektor $\vec{p}$ eines Punktes P auf der Geraden. Nutzt dafür den y-Achsenabschnitt b.

Niveau 6

Der y-Achsenabschnitt b liefert einen Punkt P, der auf jeden Fall auf der Geraden liegt. Er hat die Koordinaten $P(0|b)$. Damit ergibt sich die Normalengleichung

$$g: 0 = \begin{pmatrix} m \\ -1 \end{pmatrix} * \left(\begin{pmatrix} x1 \\ x2 \end{pmatrix} - \begin{pmatrix} 0 \\ b \end{pmatrix} \right)$$

A.4.2. Hilfekarte zu Blatt 2

- Gestufte Hilfekarten für Auftrag 3.1 auf Blatt 2 -

Niveau 1

Die Fallunterscheidung soll

für $\vec{a} * \vec{b} > 0$

für $\vec{a} * \vec{b} = 0$

für $\vec{a} * \vec{b} < 0$

gemacht werden.

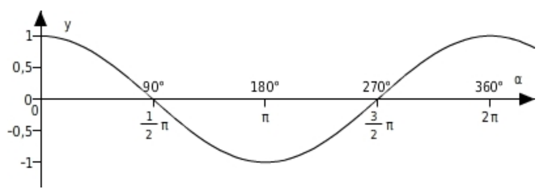
Niveau 2

Sucht dafür die Teile in $cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}|$ die ihr Vorzeichen ändern können.

Niveau 3

$|\vec{a}| \cdot |\vec{b}|$ ist immer > 0, da ja die Beträge genommen werden.
Benutzt die Abbildung zur Kosinusfunktion um zu entscheiden, für welche Winkel α $\cos(\alpha) >$, $<$ oder $= 0$ ist. Betrachtet nur Winkel von $0° - 180°$. Dies liefert euch die Fallunterscheidung für

$$\vec{a} * \vec{b} = cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}\ .$$



Niveau 4

Die Fallunterscheidung lautet:

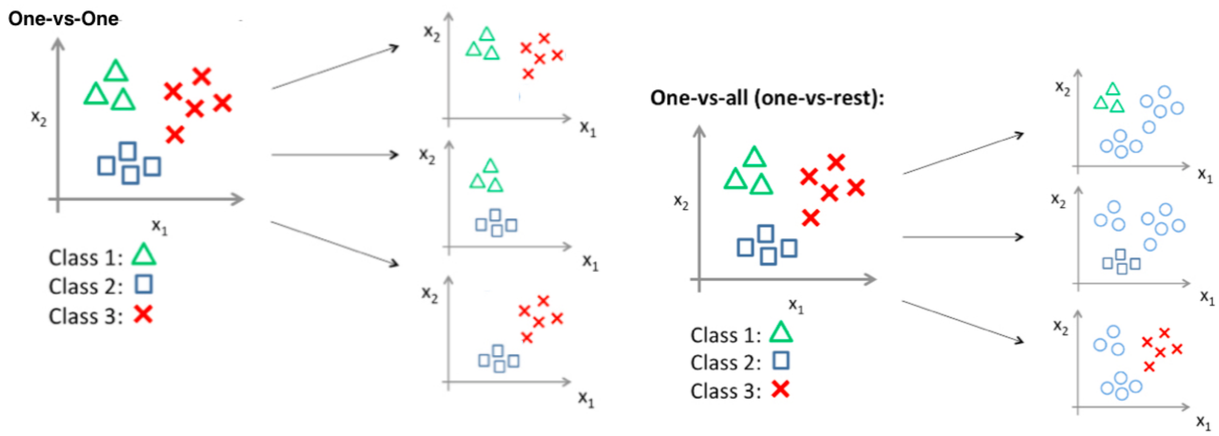
$$\vec{a} * \vec{b} = \cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}| = \begin{cases} > 0 & \text{für } 0° \leqslant \alpha < 90° \\ = 0 & \text{für } \alpha = 90° \\ < 0 & \text{für } 90° < \alpha \leqslant 180° \end{cases}$$

- Gestufte Hilfekarten für Auftrag 5.1/5.2 auf Blatt 2 -

Niveau 1

Gängige Verfahren sind das *One-Versus-One* und das *One-Versus-All* Verfahren.

Niveau 2



(a) One-Versus-One Verfahren

(b) One-Versus-All Verfahren

Abbildung 1: Schemata der Verfahren

Niveau 3

Beschreibung One-Versus-All Verfahren:
Bei diesem Verfahren wird immer eine Klasse mit einer größeren Klasse, bestehend aus allen anderen Klassen verglichen. Zum Beispiel vergleicht man zunächst die Klasse der roten Ampeln mit der Klasse aus grünen und gelben Ampeln. Es wird jeweils eine separierende Funktion gesucht. Die einzelne Klasse wird als positiv betrachtet und der Zusammenschluss der anderen Klassen als negativ. Dann wird für den Datenpunkt, der klassifiziert werden soll der **Abstand** zu jeder der Geraden ausgerechnet. Liegt der Punkt auf der Seite der separierenden Gerade, auf der die einzelne Klasse liegt, so wird der Abstand mit $+1$ multipliziert. Liegt er auf der Seite, auf der der Zusammenschluss der anderen Klassen liegt, wird der Abstand mit -1 multipliziert. Der Punkt wird dann der Klasse zugeteilt, bei dem dieser Wert maximal ist.
Hinweis: Die Formel, um den Abstand zu berechnen, habt ihr in Aufgabe 3 bereits aufgestellt!

Beschreibung One-Versus-One Verfahren:
Beim diesem Verfahren werden immer zwei Klassen miteinander verglichen und jeweils eine separierende Funktion für die beiden Klassen gesucht. Dies wird für alle möglichen Kombinationen der Klassen gemacht. Ein Datenpunkt, der klassifiziert werden soll, wird dann der Klasse hinzugefügt, zu der er durch die einzelnen eins-gegen-eins Separationen am häufigsten hinzugefügt wurde.

Niveau 4

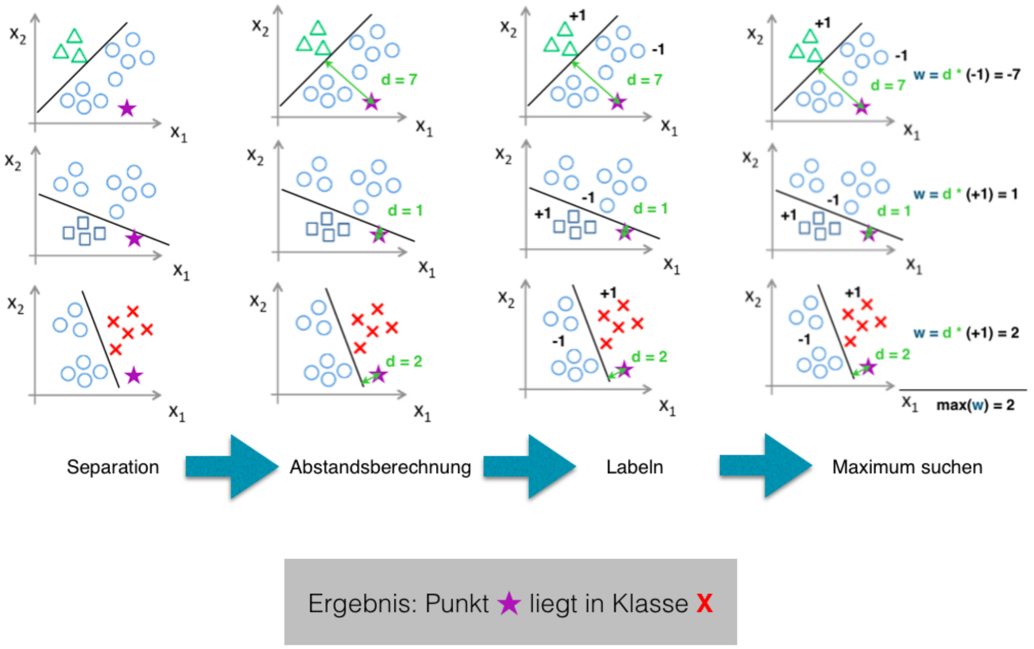


Abbildung 2: Ablauf One-Versus-All

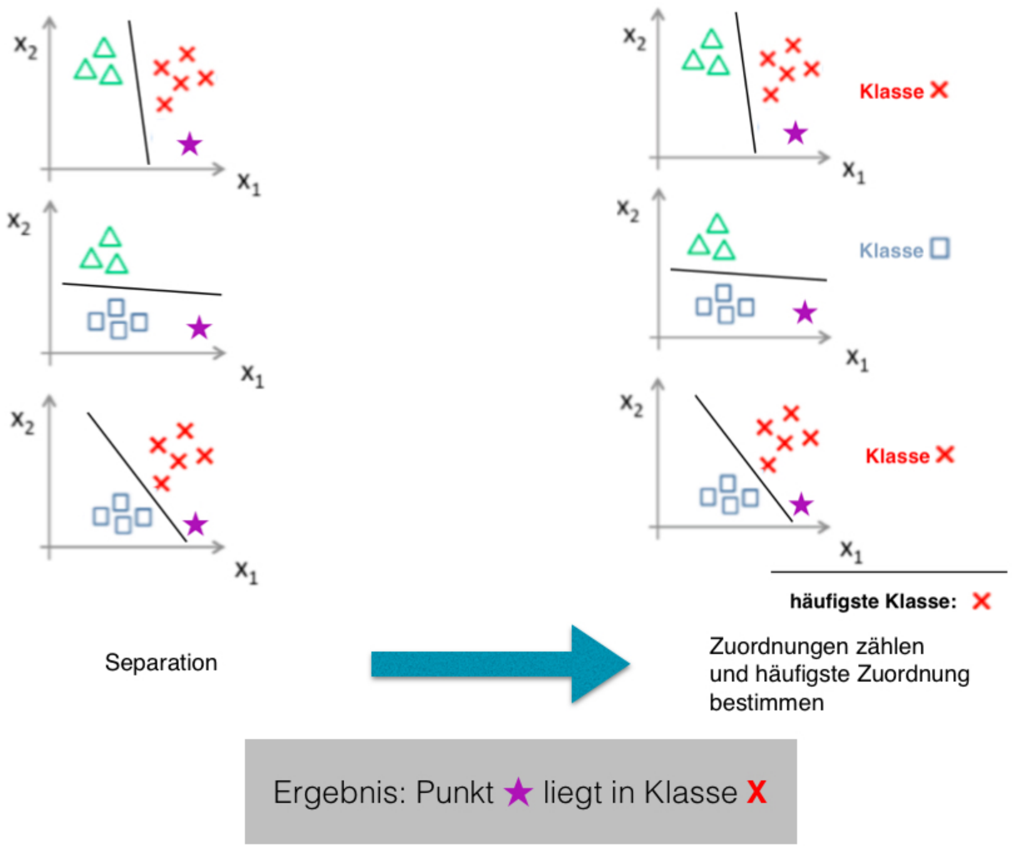


Abbildung 3: Ablauf One-Versus-One

Niveau 5

Vorgehen bei MATLAB für One-Versus-All:

1. Berechnet zunächst die Parameter der Trennebene mittels der `Separation` Funktion. Dazu müsst ihr für `class_1` die einzelne Klasse eingeben. Für `class_2` gebt ihr die beiden übrigen Klassen als Vektor ein z.B. `[ListeGruen3;ListeRot3]`.

2. Führt die Separation durch, indem ihr *Run Section* drückt.

3. Notiert euch den angezeigten Normalenvektor, den Wert für b.

4. Berechnet dann weiter unten im Code mit der Abstandsformel von Blatt 1 `d=dot(n,x)+b)/norm(n)` den Wert für Punkt X. Dafür benötigt ihr die Werte für $\vec{n}$ und b, sowie die Koordinaten des Punktes X.

5. Mit *Run Section* wird der Punkt eingeordnet. Notiert euch **den Betrag des Wertes!**

6. Liegt der Punkt X auf der Seite der einzelnen Klasse, multipliziert den Wert mit Label $+1$. Liegt er auf der Seite der beiden anderen Klassen, multipliziert ihn mit Label -1. Tragt das Ergebnis in die entsprechende Tabelle auf dem Antwortblatt ein.

7. Tragt das Ergebnis in die entsprechende Tabelle auf dem Antwortblatt ein und wiederholt die Schritte 1-7 mit den anderen 2 Kombinationen der Klassen.

8. Sucht den maximalen Wert aus den drei mit ± 1 multiplizierten Werten.

9. Die einzelne Klasse zu dem Ergebnis aus Schritt 8 ist die Lösung.

Vorgehen bei MATLAB für One-Versus-One:

1. Berechnet zunächst die Parameter der Trennebene mittels der `Separation` Funktion. Dazu müsst ihr für `class_1` und `class_2` jeweils eine der drei Klassen Grün, Gelb und Rot auswählen.

2. Führt die Sepration durch, indem ihr *Run Section* drückt.

3. Notiert euch den angezeigten Normalenvektor und den Wert für b.

4. Berechnet dann weiter unten im Code mit der Klassifizierungsformel aus Aufgabe 3 den Wert für Punkt X. Dafür benötigt ihr die Werte für $\vec{n}$ und b, sowie die Koordinaten des Punktes X.

5. Mit *Run Section* wird der Punkt eingeordnet.

6. Tragt das Ergebnis in die entsprechende Tabelle auf dem Antwortblatt ein.

7. Wiederholt die Schritte 1-6 für die anderen beiden Klassenkombinationen.

8. Zählt in welche Klasse X am häufigsten eingeordnet wurde.

9. Das Ergebnis aus Schritt 8 ist die Lösung.

A.5. Besprechungsfolien

A.5.1. Besprechungsfolien zu Blatt 1

- Besprechung zu Aufgabenblatt 1 -

Aufgabe 2

Auftrag 2.1

Möglichkeit die drei Werte zu einem zusammen zu fassen:

Aufgabe 3

Auftrag 3.4

Geradengleichung Anzahl korrekter Klassifizierungen

Aufgabe 5

Auftrag 5.2

„Finde die Gerade, die

A.5.2. Schülerfolien zu Blatt 1

- Schülerfolien zu Aufgabenblatt 1 -

Aufgabe 3

Auftrag 3.1

Trennfunktion für Trainingsdaten in Abbildung 1 einzeichnen.

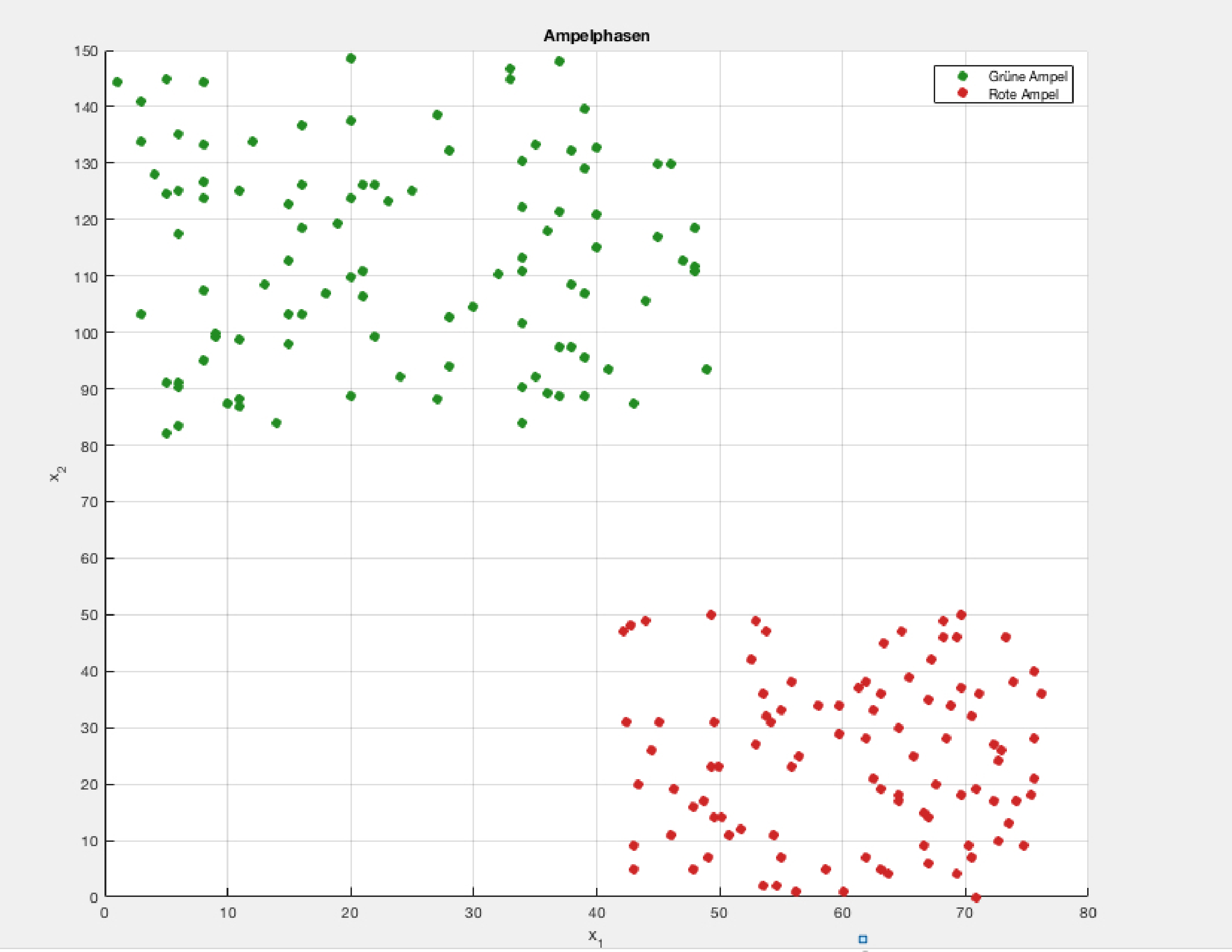


Abbildung 1: Trainingsdatensatz aus 100 grünen und 100 roten Ampeln

Geradengleichung:

Aufgabe 4

Auftrag 4.1

Zeichne die Trennfunktion in Abbildung 2 ein, welche die Daten eurer Meinung nach am besten trennt. Umkreist Punkte, die ihr als Orientierung genutzt habt.

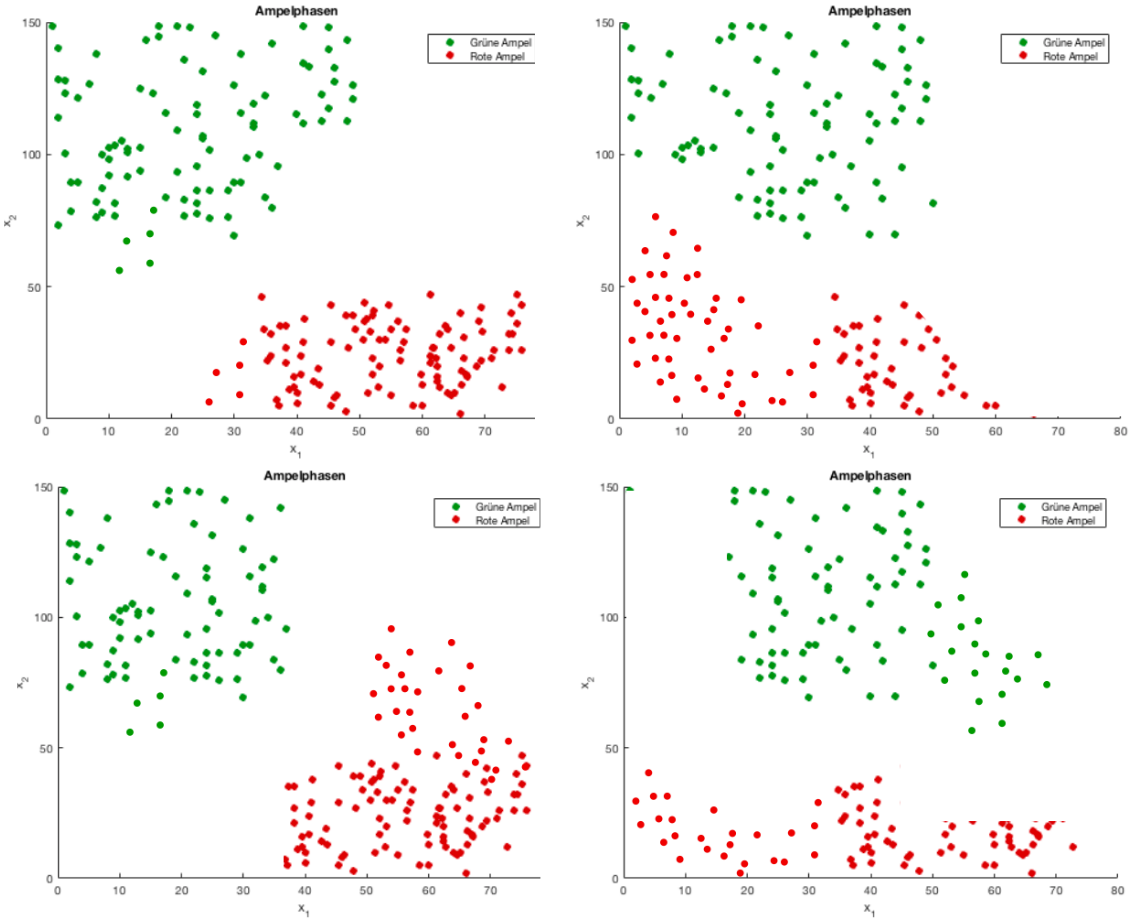


Abbildung 2: Weitere Trainingsdaten, die getrennt werden sollen

Frage 4.1

Worauf habt ihr beim Zeichnen geachtet? Was zeichnet für euch die „*beste Gerade*“ aus?

Antwort:

Aufgabe 5

Auftrag 5.3

Rechnung für Abstandsformel:

A.5.3. Besprechungsfolien zu Blatt 2

- Besprechung zu Aufgabenblatt 2 -

Aufgabe 5

Auftrag 5.2

Für One-Versus-One Verfahren:

Tabelle 1: Ergebnisse der 3 einzelnen Klassifikationen

	SVM grün vs. rot	SVM gelb vs. rot	SVM gelb vs. grün
Wert Entsch. Funktion			
Zugeteilte Klasse			

Für One-Versus-All Verfahren:

Tabelle 2: Ergebnisse der 3 einzelnen Klassifikationen

	SVM grün & gelb vs. rot	SVM gelb & rot vs. grün	SVM rot & grün vs. gelb
Betrag Abstandsformel			
Multiplikation mit ±1			

Antwort

Zuordnung des Punktes $X\ (60 \mid 80 \mid 40)$ in Klasse:

A.5.4. Schülerfolien zu Blatt 2

CAMMP day

Gesichtserkennung und Autonomes Fahren mit Mathematik?!

- Schülerfolien zu Aufgabenblatt 2 -

Aufgabe 3

Auftrag 3.1

Allgemeine Fallunterscheidung für $\vec{a} * \vec{b} = \cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}|$ für die drei Fälle:

$$\vec{a} * \vec{b} = \cos(\alpha) \cdot |\vec{a}| \cdot |\vec{b}| \begin{cases} \boxed{} & \text{für } \boxed{} \leqslant \alpha < \boxed{} \\ = 0 & \text{für } \alpha = \boxed{} \\ \boxed{} & \text{für } \boxed{} < \alpha \leqslant \boxed{} \end{cases}$$

Auftrag 3.2

Fallunterscheidung für Ebene:

$$\vec{n} * (\vec{x} - \vec{p}) = \boxed{} \begin{cases} \boxed{} & \text{für } \boxed{} \leqslant \alpha < \boxed{} \\ = 0 & \text{für } \alpha = \boxed{} \\ \boxed{} & \text{für } \boxed{} < \alpha \leqslant \boxed{} \end{cases}$$

Zeigt der Normalenvektor **nach oben aus der Ebene**, ist der Ausdruck $\vec{n} * (\vec{x} - \vec{p})$ also negativ für Punkte $\boxed{}$ der Ebene und positiv für Punkte $\boxed{}$ der Ebene.

Zeigt der Normalenvektor **nach unten aus der Ebene**, ist der Ausdruck $\vec{n} * (\vec{x} - \vec{p})$ also negativ für Punkte $\boxed{}$ der Ebene und positiv für Punkte $\boxed{}$ der Ebene.

Für Punkte $\boxed{}$ der Ebene, ist der Ausdruck stets Null.
Somit ist $\vec{n} * (\vec{x} - \vec{p})$ die gesuchte **Entscheidungsfunktion**, mit der wir Datenpunkte klassifizieren können. Dazu ist es wichtig zunächst zu prüfen, in welche Richtung der Normalenvektor zeigt.

Im Folgenden wird $\vec{n} * (\vec{x} - \vec{p})$ in der Form $\vec{n} * \vec{x} + b$ benutzt, wobei b der Wert von $-\vec{n} * \vec{p}$ ist.

A.5.5. Besprechungsfolien zu Blatt 3

CAMMP day

Gesichtserkennung und Autonomes Fahren mit Mathematik?!

- Besprechung zu Aufgabenblatt 3 -

Aufgabe 1

Frage 1.3

Was könnten Vorteile von langen bzw. kurzen Vektoren sein? Wo liegen Nachteile?

Frage 1.4

Welche Dimension hätte der Raum von Bild 4?

Aufgabe 4

Auftrag 4.1

Klassifizierung der Grimassen:

Tabelle 1: Zuordnung der Grimassen

Eingegebener Gruppenname	Zugeordneter Gruppenname	☑/☒

Frage 4.3

Was bedeutet „am ähnlichsten“ mathematisch? Wie könnte der Computer das festlegen?

Frage 4.4

Wie würdet ihr das Verfahren ändern, um eine bessere Zuordnung der Prominenten zu euren Mitschülern zu erhalten?

A.6. Präsentationen

A.6.1. Einstiegspräsentation

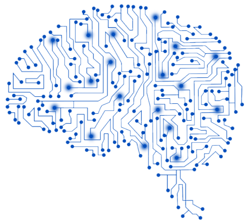
Künstliche Intelligenz: Gesichtserkennung und autonomes Fahren mit Mathematik?!

Lars Schmidt
Lehrstuhl für Mathematik
Center for Computational Engineering Science

Kickoff

Was verbindet ihr mit dem Begriff *Künstliche Intelligenz*?

Wo könnte sie eingesetzt werden?



Machine Learning, was ist das?

Definition von *Lernen* nach Zimbardo: "Lernen ist ein Prozess, der zu relativ stabilen Veränderungen im Verhalten führt und auf Erfahrung aufbaut." Zimbardo (1992, 227)

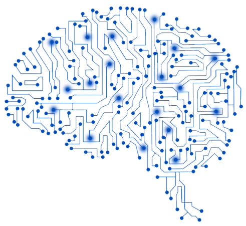
Machine Learning, was ist das?

Definition von *Lernen* nach Zimbardo: "Lernen ist ein Prozess, der zu relativ stabilen Veränderungen im Verhalten führt und auf Erfahrung aufbaut." Zimbardo (1992, 227)

Bei *Machine Learning*: "Verfahren, durch die Computersysteme befähigt werden, selbstständig Wissen aufzunehmen und zu erweitern, um ein gegebenes Problem besser lösen zu können als vorher." (Dr. Siepermann TU Dortmund)

Machine Learning, was ist das?

- Mit dem Computer Wissen aus einer Menge von Daten gewinnen
- dazu Nutzung verschiedener Algorithmen
- so erlangtes Wissen auf ein Problem anwenden



"Machine Learning, warum sollte ich das können?"

Machine Learning, warum sollte ich das können?

"Weil der Lehrer gesagt hat, wir gehen zu CAMMP und da gibt es so ein Lernmodul."

Machine Learning, warum sollte ich das können?

"Weil der Lehrer gesagt hat, wir gehen zu CAMMP und da gibt es so ein Lernmodul."

Schlechter Grund!

Machine Learning, warum sollte ich das können?

"Weil bedeutsame Persönlichkeiten sagen, dass es wichtig ist."

"A breakthrough in machine learning would be worth ten Microsofts."
- Billl Gates, Microsoft -
"Machine learning is the next Internet."
- Tony Tether, DARPA -
"Machine learning is the hot new thing."
- John Hennessy, Stanford University -

Machine Learning, warum sollte ich das können?

"Weil bedeutsame Persönlichkeiten sagen, dass es wichtig ist."

"A breakthrough in machine learning would be worth ten Microsofts."
- Billl Gates, Microsoft -
"Machine learning is the next Internet."
- Tony Tether, DARPA -
"Machine learning is the hot new thing."
- John Hennessy, Stanford University -

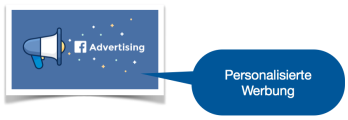
Etwas besserer Grund!

Machine Learning, warum sollte ich das können?

"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag".

Machine Learning, warum sollte ich das können?

"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag".



Machine Learning, warum sollte ich das können?

"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag".

Machine Learning, warum sollte ich das können?

"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag".

Machine Learning, warum sollte ich das können?

"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag".

Machine Learning, warum sollte ich das können?

"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag".

Machine Learning, warum sollte ich das können?

"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag".

Machine Learning, warum sollte ich das können?

"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag".

Machine Learning, warum sollte ich das können?

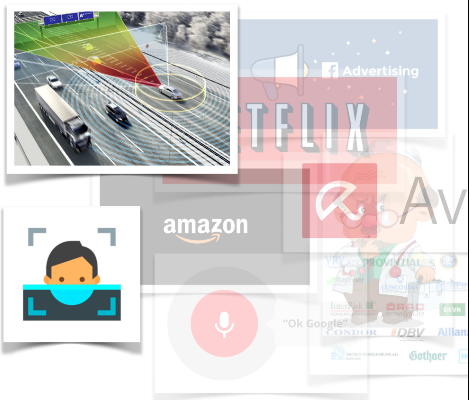
"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag".

Machine Learning, warum sollte ich das können?

"Wegen der unzähligen Anwendungsbereiche und ihrer Relevanz im Alltag". **Bester Grund!**

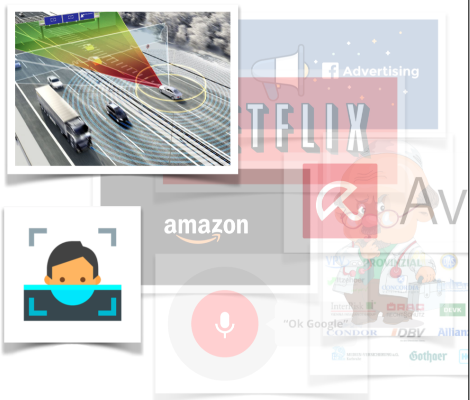
Anwendungsbereiche heute

- Autonomes Fahren
- Gesichtserkennung



Anwendungsbereiche heute: *Klassifizierungsprobleme*

- Autonomes Fahren
- Gesichtserkennung



Klassifizierungsprobleme

Was ist ein Klassifizierungsproblem?

Klassifizierungsprobleme

Mathematisch...

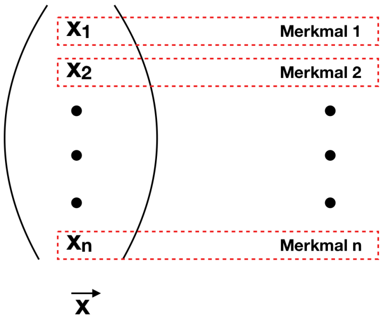
- zu klassifizierendes Objekt X
- gegeben durch *Merkmalsvektor* $\vec{x}$

$$\begin{pmatrix} x_1 \\ x_2 \\ \bullet \\ \bullet \\ \bullet \\ x_n \end{pmatrix}$$

$$\vec{x}$$

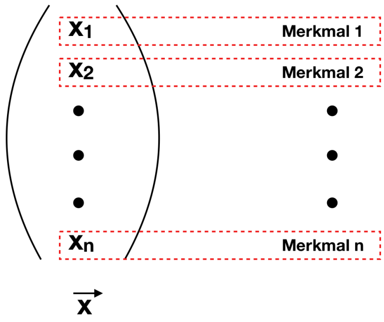
Klassifizierungsprobleme

- zu klassifizierendes Objekt X
- gegeben durch *Merkmalsvektor* $\vec{x}$
- $\vec{x}$ beinhaltet n verschiedene Merkmale



Klassifizierungsprobleme

- zu klassifizierendes Objekt X
- gegeben durch *Merkmalsvektor* $\vec{x}$
- $\vec{x}$ beinhaltet n verschiedene Merkmale
- Frage: Liegt X in einer bestimmten Klasse K?



Beispiel für Klassifizierungsproblem

Es soll eine Software für autonomes Fahren zu entwickelt werden, die Verkehrsampeln automatisch erkennt. Dafür sollen zunächst rote und grüne Verkehrsampeln voneinander getrennt werden. Bei einer Ampel soll also entschieden werden, ob sie grün oder rot ist.

Beispiel für Klassifizierungsproblem

Es soll eine Software für autonomes Fahren zu entwickelt werden, die Verkehrsampeln automatisch erkennt. Dafür sollen zunächst rote und grüne Verkehrsampeln voneinander getrennt werden. Bei einer Ampel soll also entschieden werden, ob sie grün oder rot ist.

Was ist hier Objekt X? Was ist Klasse K? Wie könnten die Merkmale aussehen?

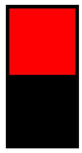
Beispiel für Klassifizierungsproblem

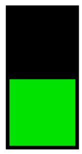
- zu klassifizierendes Objekt X: Verkehrsampel
- gegeben durch *Merkmalsvektor* $\vec{x}$
- mögliche Merkmale: Farbwerte der Lampen
- Klassen K: rote und grüne Ampeln

Beispiel für Klassifizierungsproblem

Vereinfacht...

- zu klassifizierendes Objekt X: Verkehrsampel aus zwei Boxen
- gegeben durch *Merkmalsvektor* $\vec{x}$
- mögliche Merkmale: Farbwert für jede Box
- Klassen K: rote und grüne Ampeln

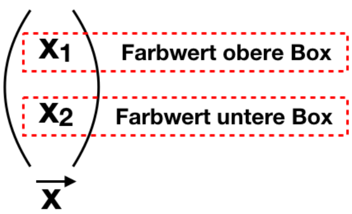




Beispiel für Klassifizierungsproblem

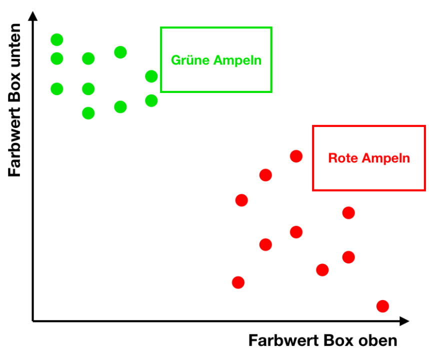
Mathematisch...

- zu klassifizierendes Objekt X: Verkehrsampel aus zwei Boxen
- gegeben durch *Merkmalsvektor* $\vec{x}$
- mögliche Merkmale: Farbwert für jede Box
- Klassen K: rote und grüne Ampeln



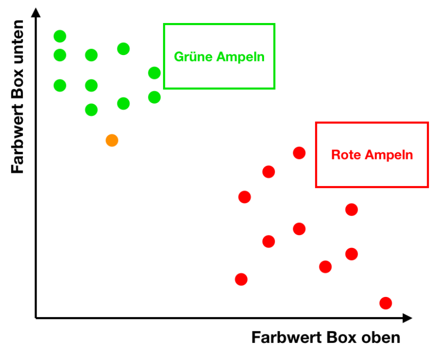
Beispiel für Klassifizierungsproblem

Graphisch dargestellt...



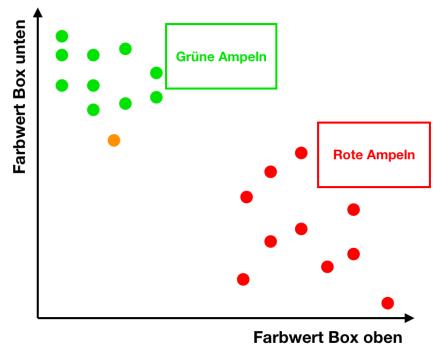
Beispiel für Klassifizierungsproblem

Welcher Klasse K gehört die Verkehrsampel an?



Beispiel für Klassifizierungsproblem

Welcher Klasse K gehört die Ampel an? Den grünen Ampeln!



Modellierungskreislauf

In den letzten fünf Minuten sind wir bereits einmal durch den Modellierungskreislauf gegangen!

Reales Problem — vereinfachen → **Vereinfachtes Problem**

Vereinfachtes Problem — mathematisch beschreiben → **Mathematisches Modell**

Mathematisches Modell — berechnen → **Mathematische Lösung**

Mathematische Lösung — interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem: Klassifizierung roter und grüner Ampeln — vereinfachen → **Vereinfachtes Problem**

Vereinfachtes Problem — mathematisch beschreiben → **Mathematisches Modell**

Mathematisches Modell — berechnen → **Mathematische Lösung**

Mathematische Lösung — interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem: Klassifizierung roter und grüner Ampeln — vereinfachen → **Vereinfachtes Problem**: Ampeln: 2 Boxen, Merkmale: Farbwerte

Vereinfachtes Problem — mathematisch beschreiben → **Mathematisches Modell**

Mathematisches Modell — berechnen → **Mathematische Lösung**

Mathematische Lösung — interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem: Klassifizierung roter und grüner Ampeln — vereinfachen → **Vereinfachtes Problem**: Ampeln: 2 Boxen, Merkmale: Farbwerte

Vereinfachtes Problem — mathematisch beschreiben → **Mathematisches Modell**

Mathematisches Modell — berechnen → **Mathematische Lösung**

Mathematische Lösung — interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem: Klassifizierung roter und grüner Ampeln — vereinfachen → **Vereinfachtes Problem**: Ampeln: 2 Boxen, Merkmale: Farbwerte

Vereinfachtes Problem — mathematisch beschreiben → **Mathematisches Modell**: Ampel als 2D Merkmalsvektor

Mathematisches Modell — berechnen → **Mathematische Lösung**

Mathematische Lösung — interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem: Klassifizierung roter und grüner Ampeln — vereinfachen → **Vereinfachtes Problem**: Ampeln: 2 Boxen, Merkmale: Farbwerte

Vereinfachtes Problem — mathematisch beschreiben → **Mathematisches Modell**: Ampel als 2D Merkmalsvektor

Mathematisches Modell — berechnen → **Mathematische Lösung**: Vekor als Punkt im Koordinatensystem

Mathematische Lösung — interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem
Klassifizierung roter und grüner Ampeln

— vereinfachen →

Vereinfachtes Problem
Ampeln: 2 Boxen, Merkmale: Farbwerte

mathematisch beschreiben

Mathematisches Modell
Ampel als 2D Merkmalsvektor

← berechnen —

Mathematische Lösung
Vekor als Punkt im Koordinatensystem

Ampel ist grün

Modellierungskreislauf

Was stimmt hier noch nicht ganz?

Reales Problem
Klassifizierung roter und grüner Ampeln

— vereinfachen →

Vereinfachtes Problem
Ampeln: 2 Boxen, Merkmale: Farbwerte

mathematisch beschreiben

Mathematisches Modell
Ampel als 2D Merkmalsvektor

← berechnen —

Mathematische Lösung
Vekor als Punkt im Koordinatensystem

interpretieren

Modellierungskreislauf

Was stimmt hier noch nicht ganz?

Reales Problem
Klassifizierung roter und grüner Ampeln

— vereinfachen →

Vereinfachtes Problem
Ampeln: 2 Boxen, Merkmale: Farbwerte

mathematisch beschreiben

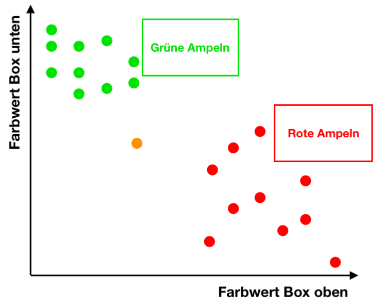
Mathematisches Modell
Ampel als 2D Merkmalsvektor

← berechnen "hingucken" —

Mathematische Lösung
Vekor als Punkt im Koordinatensystem

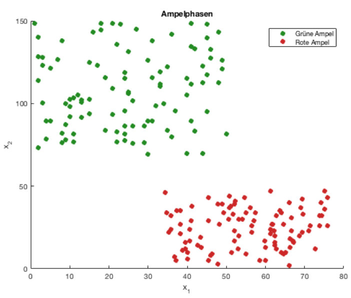
interpretieren

Problem: "Hingucken"



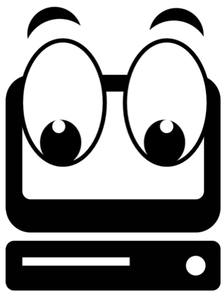
- "Hingucken" ist weder präzise, noch eindeutig.

Problem: "Hingucken"



- "Hingucken" ist weder präzise , noch eindeutig.
- Bei großen Datensätzen wird Computer benötigt

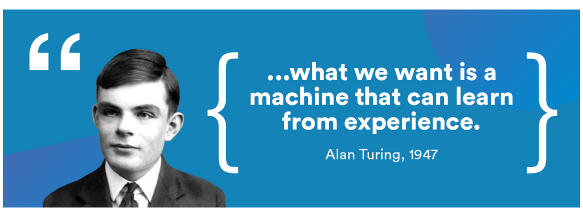
Problem: "Hingucken"



- "Hingucken" ist weder präzise , noch eindeutig.
- Bei großen Datensätzen wird Computer benötigt
- Computer kann nicht einfach "hingucken", benötigt mathematischen Befehl

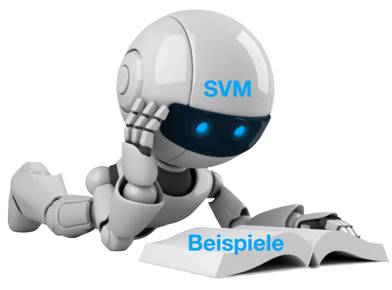
Lösung: Lernalgorithmus *Support Vector Machine*

Support Vector Machine (SVM)



- Algorithmus für das Lernen aus Beispielen
- Für Beispiele (*Trainingsdaten*) ist die Klassifikation schon bekannt

Support Vector Machine (SVM)



- Aus Trainingsdaten Klassifikationsregel ableiten

→ Klassifikationsregel an Testdaten prüfen

Support Vector Machine | Schritt 1

1. Einlesen des Datensatzes

2. Einteilung in Test- und Trainingsdaten

3. Suchen der Trenn- und Entscheidungsfunktion mit Trainingsdaten

Support Vector Machine | Schritt 2

1. Einlesen des Datensatzes

2. Einteilung in Test- und Trainingsdaten

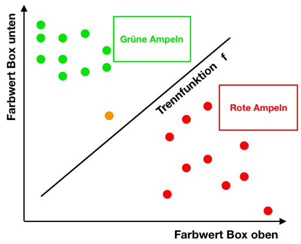
3. Suchen der Trenn- und Entscheidungsfunktion mit Trainingsdaten

Support Vector Machine | Schritt 3

1. Einlesen des Datensatzes

2. Einteilung in Test- und Trainingsdaten

3. Suchen der Trenn- und Entscheidungsfunktion mit Trainingsdaten



Support Vector Machine | Schritt 4

3. Suchen der Entscheidungsfunktion mit Trainingsdaten

4. Test der Entscheidungsfunktion mit Testdaten

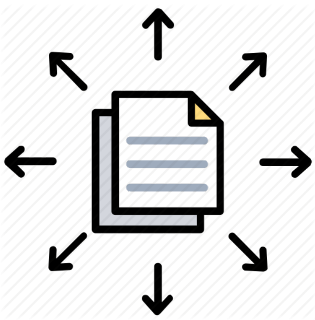
5. Verallgemeinern und auf neue Datensätze anwenden

Support Vector Machine | Schritt 5

3. Suchen der Entscheidungsfunktion mit Trainingsdaten

4. Test der Entscheidungsfunktion mit Testdaten

5. Verallgemeinern und auf neue Datensätze anwenden



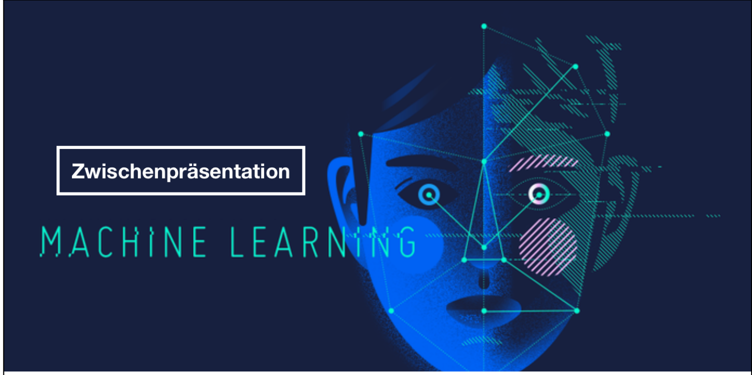
Euer Auftrag heute

1. Entwickelt eine SVM zur Erkennung von Verkehrsampeln für autonomes Fahren

2. Übertragt euer Wissen um eine automatische Gesichtserkennung für eure Klasse zu erstellen

- Bearbeite die Arbeitsblätter!
- Nutze MATLAB!
- Arbeite im Team mit deinem Partner!
- Nutze die Hilfekarten!
- Frag' die Betreuer!
- Nutze das Internet!
- Präsentiere Ergebnisse deinen Mitschülern!
- Hilf deinen Mitschülern!

A.6.2. Zwischenpräsentation

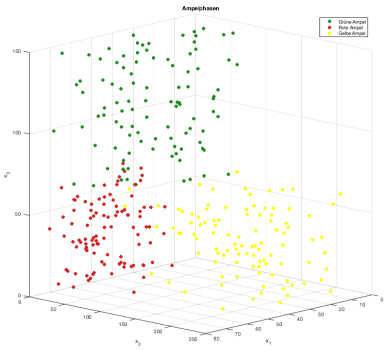


Künstliche Intelligenz: Gesichtserkennung und autonomes Fahren mit Mathematik?!

Lars Schmidt
Lehrstuhl für Mathematik
Center for Computational Engineering Science

Das Problem

Die verschiedenen Klassen überlappen. Zwischen den Datenpunkten der grünen und roten Ampelphasen ist kein Freiraum mehr, in dem die Trennfunktion verlaufen kann.

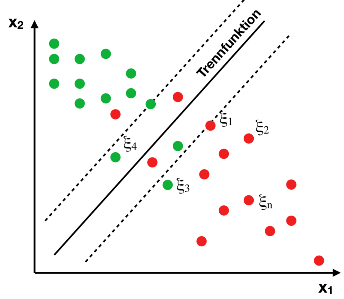


Möglicher Grund

- Lichteinstrahlung durch Sonne oder Scheinwerfer
  $\rightarrow$ eigentlich schwarze Boxen sind gefärbt
- Beispielsweise leuchtet bei eigentlich roter Ampel auch die grüne oder die gelbe Box

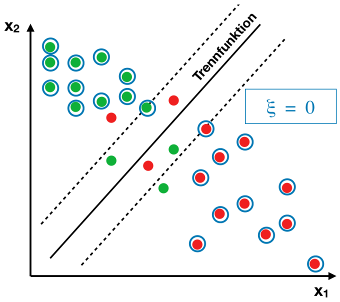
Die Lösung: *Slack Variablen* ξ

- wörtliche Übersetzung: *Schlupf-Variablen*
- jeder Trainingsdatenpunkt X_i erhält eine Slack Variable ξ_i
- Slack Variablen erlauben, dass Trainingsdaten falsch klassifiziert werden



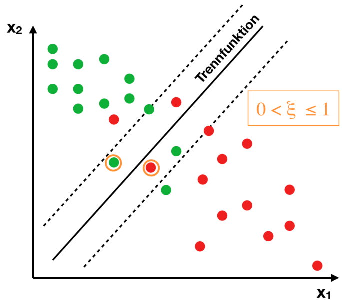
Slack Variablen ξ

- Trainingsdaten X_i außerhalb des Margin, auf der richtigen Seite der Trennfunktion werden korrekt klassifiziert und erhalten $\xi = 0$.



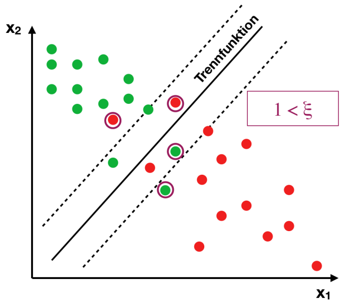
Slack Variablen ξ

- Trainingsdaten X_i außerhalb des Margin, auf der richtigen Seite der Trennfunktion werden korrekt klassifiziert und erhalten $\xi = 0$.
- Trainingsdaten X_i innerhalb des Margin, auf der richtigen Seite der Trennfunktion werden korrekt klassifiziert und erhalten $0 < \xi \leq 1$.



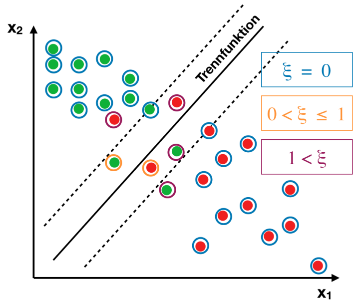
Slack Variablen ξ

- Trainingsdaten X_i außerhalb des Margin, auf der richtigen Seite der Trennfunktion werden korrekt klassifiziert und erhalten $\xi = 0$.
- Trainingsdaten X_i innerhalb des Margin, auf der richtigen Seite der Trennfunktion werden korrekt klassifiziert und erhalten $0 < \xi \leq 1$.
- Trainingsdaten X_i innerhalb und außerhalb des Margin, auf der falschen Seite der Trennfunktion werden falsch klassifiziert und erhalten $1 < \xi$.



Slack Variablen ξ

- Trainingsdaten X_i außerhalb des Margin, auf der richtigen Seite der Trennfunktion werden korrekt klassifiziert und erhalten $\xi = 0$.
- Trainingsdaten X_i innerhalb des Margin, auf der richtigen Seite der Trennfunktion werden korrekt klassifiziert und erhalten $0 < \xi \leq 1$.
- Trainingsdaten X_i innerhalb und außerhalb des Margin, auf der falschen Seite der Trennfunktion werden falsch klassifiziert und erhalten $1 < \xi$.



Veränderungen durch Einsatz von Slack Variablen

Vorher: Ohne Slack
- Trainingsdatensatz ist durch eine lineare Funktion trennbar

Jetzt: Mit Slack
- Trainingsdatensatz ist durch eine lineare Funktion trennbar

Veränderungen durch Einsatz von Slack Variablen

Vorher: Ohne Slack
- Trainingsdatensatz ist durch eine lineare Funktion trennbar
- Um Trennfunktion gibt es einen Datenpunkt-freien Bereich (Margin)

Jetzt: Mit Slack
- Trainingsdatensatz ist durch eine lineare Funktion trennbar
- Trainingspunkte liegen auch innerhalb des Margins

Veränderungen durch Einsatz von Slack Variablen

Vorher: Ohne Slack
- Trainingsdatensatz ist durch eine lineare Funktion trennbar
- Um Trennfunktion gibt es einen Datenpunkt-freien Bereich (Margin)
- Bedingung: **alle** Trainingsdaten müssen von Entscheidungsfunktion richtig eingeordnet werden

Jetzt: Mit Slack
- Trainingsdatensatz ist durch eine lineare Funktion trennbar
- Trainingspunkte liegen auch innerhalb des Margins
- Bedingung: **nicht alle** Trainingsdaten müssen von Entscheidungsfunktion richtig eingeordnet werden

Veränderungen durch Einsatz von Slack Variablen

Vorher: Ohne Slack
- Trainingsdatensatz ist durch eine lineare Funktion trennbar
- Um Trennfunktion gibt es einen datenpunktfreien Bereich (Margin)
- Bedingung: **alle** Trainingsdaten müssen von Entscheidungsfunktion richtig eingeordnet werden

  $\rightarrow$ überlappende Datensätze können **nicht** klassifiziert werden

Jetzt: Mit Slack
- Trainingsdatensatz ist durch eine lineare Funktion trennbar
- Trainingspunkte liegen auch innerhalb des Margins
- Bedingung: **nicht alle** Trainingsdaten müssen von Entscheidungsfunktion richtig eingeordnet werden

  $\rightarrow$ überlappende Datensätze können klassifiziert werden

Eure Aufgabe

Einsatz der Slack Variablen wird durch einen Faktor C reguliert. Findet auf dem Bonusblatt heraus, welche Auswirkungen verschiedene Werte von C auf die Klassifizierung haben!

A.6.3. Abschlusspräsentation

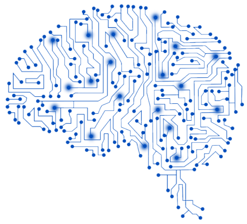
Künstliche Intelligenz: Gesichtserkennung und autonomes Fahren mit Mathematik?!

Lars Schmidt
Lehrstuhl für Mathematik
Center for Computational Engineering Science

Heute Morgen, ca. 9:30 Uhr

Was verbindet ihr mit dem Begriff
Künstliche Intelligenz?

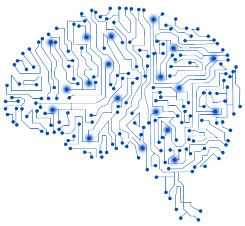
Wo könnte sie eingesetzt werden?



Jetzt

Was verbindet ihr mit dem Begriff
Künstliche Intelligenz?

Wo könnte sie eingesetzt werden?



Einsatz der SVM bei Klassifizierungsproblemen.

Was dazwischen passiert ist

Reales Problem — vereinfachen → **Vereinfachtes Problem**

Vereinfachtes Problem — mathematisch beschreiben → **Mathematisches Modell**

Mathematisches Modell — berechnen → **Mathematische Lösung**

Mathematische Lösung — interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem: Automatische Erkennung von Ampelphasen

— vereinfachen → **Vereinfachtes Problem**

— mathematisch beschreiben → **Mathematisches Modell**

— berechnen → **Mathematische Lösung**

— interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem: Automatische Erkennung von Ampelphasen

— vereinfachen → **Vereinfachtes Problem**: Unterscheidung Ampeln mit grüner und roter Box

— mathematisch beschreiben → **Mathematisches Modell**

— berechnen → **Mathematische Lösung**

— interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem: Automatische Erkennung von Ampelphasen

— vereinfachen → **Vereinfachtes Problem**: Unterscheidung Ampeln mit grüner und roter Box

— mathematisch beschreiben → **Mathematisches Modell**: Ampeln: 2x1 Vektor mit Grauwerten, Opt. Problem: Abstand

— berechnen → **Mathematische Lösung**

— interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem: Automatische Erkennung von Ampelphasen

— vereinfachen → **Vereinfachtes Problem**: Unterscheidung Ampeln mit grüner und roter Box

— mathematisch beschreiben → **Mathematisches Modell**: Ampeln: 2x1 Vektor mit Grauwerten, Opt. Problem: Abstand

— berechnen (SVM) → **Mathematische Lösung**: Wert aus Entscheidungsfunktion

— interpretieren → **Reales Problem**

Modellierungskreislauf

Reales Problem
Automatische Erkennung von Ampelphasen

— vereinfachen →

Vereinfachtes Problem
Unterscheidung Ampeln mit grüner und roter Box

mathematisch beschreiben

Mathematisches Modell
Ampeln: 2x1 Vektor mit Grauwerten, Opt. Problem: Abstand

← berechnen

Mathematische Lösung
Wert aus Entscheidungsfunktion

Ampel ist grün oder rot

Es gibt ja drei Ampelphasen!

Reales Problem
Automatische Erkennung von Ampelphasen

— vereinfachen →

Vereinfachtes Problem
Unterscheidung Ampeln mit grüner und roter Box

mathematisch beschreiben

Mathematisches Modell
Ampeln: 2x1 Vektor mit Grauwerten, Opt. Problem: Abstand

← berechnen

Mathematische Lösung
Wert aus Entscheidungsfunktion

Ampel ist grün oder rot

Modellverbesserung

Reales Problem
Automatische Erkennung von Ampelphasen

— vereinfachen →

Vereinfachtes Problem
Unterscheidung in Ampeln mit grüner, roter und gelber Box

mathematisch beschreiben

Mathematisches Modell
Ampeln: 3x1 Vektor mit Grauwerten, Opt. Problem: Abstand

← berechnen OVO/ OVA SVM

Mathematische Lösung
Wert aus Entscheidungsfunktion

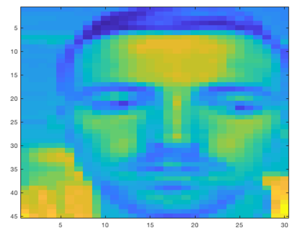
Ampel ist grün, gelb oder rot

Support Vector Machine | Schritt 5

3. Suchen der Entscheidungsfunktion mit Trainingsdaten

4. Test der Entscheidungsfunktion mit Testdaten

5. Verallgemeinern und auf neue Datensätze anwenden



Tagesziele

1. Entwickelt eine SVM zur Erkennung von Verkehrsampeln für autonomes Fahren
2. Übertragt euer Wissen um eine automatische Gesichtserkennung für eure Klasse zu erstellen

Tagesziele

1. Entwickelt eine SVM zur Erkennung von Verkehrsampeln für autonomes Fahren
2. Übertragt euer Wissen um eine automatische Gesichtserkennung für eure Klasse zu erstellen

Abschlussdiskussion

Nehmt Stellung zu den folgenden Aussagen:

Abschlussdiskussion

Nehmt Stellung zu den folgenden Aussagen:

"Machine Learning kann nur von großen Firmen mit riesigen Forschungs- und IT- Abteilungen gemacht werden."

Abschlussdiskussion

Nehmt Stellung zu den folgenden Aussagen:

"In the next 10 years, data science and software will do more for medicine than all of the biological sciences together."

- Vinod Khosla -

"The development of full artificial intelligence could spell the end of human race."

- Stephen Hawking -

A.7. Ergebnisse Evaluation

A.7.1. Evaluationsbogen

CAMMP Day Machine Learning SuS

Seite 1

Es besteht immer die Möglichkeit unsere Programme zu verbessern und wir würden gerne deine Meinung erfahren. Vielen Dank für deine Rückmeldung.

1. Persönliche Angaben

○ 8 ○ 9 ○ 10/EF ○ 11/Q1 ○ 12/Q2 ○ 13

Bitte gib dein Geschlecht an:

○ weiblich ○ männlich

Welche Leistungskurse besuchst du?

Seite 2

2. Bewertung des Workshops

	Trifft gar nich zu (--)	Trifft eher nicht zu (-)	Trifft zum Teil zu (+)	Trifft voll zu (++)
Durch den Workshop habe ich mathematisches Modellieren besser begriffen.	○	○	○	○
Der Vortrag über Modellierung war hilfreich.	○	○	○	○
Der einführende Kurzfilm war hilfreich.	○	○	○	○
Die Einführung in MATLAB war hilfreich.	○	○	○	○
Der Umgang mit MATLAB fiel mir schwer.	○	○	○	○
Die Aufgaben waren zu einfach.	○	○	○	○
Die Aufgaben waren zu schwierig.	○	○	○	○
Die gestuften Hilfekarten waren hilfreich.	○	○	○	○

3. Weiterführende Fragen
Wenn du über deine Erfahrungen mit diesem Kurs nachdenkst, wie sehr treffen die folgenden Aussagen auf dich zu?

	Trifft gar nicht zu (--)	Trifft eher nicht zu (-)	Trifft zum Teil zu (+)	Trifft voll zu (++)
Die Lern- und Arbeitsatmosphäre war angenehm.	○	○	○	○
Die Inhalte wurden klar vermittelt.	○	○	○	○
Die Lern- und Arbeitszeiten waren angemessen (nicht zu lang oder kurz mit ausreichenden Pausen).	○	○	○	○

Der Kurs hat mein Interesse an Themen der Naturwissenschaften und Technik gesteigert.	○	○	○	○
Ich habe nun eine genauere Vorstellung von Künstlicher Intelligenz bzw. Machine Learning als vor dem Workshop.	○	○	○	○
Das Thema des Kurses war für mich neu.	○	○	○	○
Das Thema des Kurses finde ich interessant.	○	○	○	○
Mir ist klar geworden, wie Bilder mit Hilfe eines Computers klassifiziert werden können.	○	○	○	○
Ich habe auf Blatt 1 verstanden, dass durch den Abstandsbegriff die optimale Trennfunktion bestimmt wurde.	○	○	○	○
Die Modellverbesserung durch den Wechsel von 2 auf 3 Dimensionen ist mir leicht gefallen.	○	○	○	○
Der Wechsel von 2 auf 3 Dimensionen hat mir gezeigt, dass die gleichen Methoden auch in höhere Dimensionen gelten.	○	○	○	○
Die Fallunterscheidung auf Blatt 2 habe ich verstanden.	○	○	○	○
Ich habe auf Blatt 2 verstanden, wie mehrere Klassen voneinander getrennt werden können.	○	○	○	○
Alles in allem habe ich das Prinzip der Support Vector Machine verstanden.	○	○	○	○
Durch den Kurs habe ich interessante Berufs- und Studienmöglichkeiten kennen gelernt.	○	○	○	○
Ich habe in diesem Kurs viel Neues gelernt, was mir für die Schule, für ein Studium, für einen Beruf weiterhelfen kann.	○	○	○	○
Ich kann mir vorstellen ein Studium oder eine Ausbildung im Bereich der Naturwissenschaften oder der Technik zu beginnen.	○	○	○	○
Ich würde so einen Kurs gerne noch einmal besuchen und würde ihn auch anderen weiterempfehlen.	○	○	○	○
Alles in allem hat mir die Veranstaltung gut gefallen.	○	○	○	○

Ist dir etwas besonders leicht gefallen?

○ Nein.

○ Ja, und zwar

Ist dir etwas besonders schwer gefallen?

○ Nein.

○ Ja, und zwar

Hättest du dir stellenweise noch mehr Unterstützung gewünscht?

○ Nein.

○ Ja, und zwar bei

Hat dir etwas an dem Kurs absolut nicht gefallen?

- Nein
- Ja, und zwar:

Hat dir etwas an diesem Kurs besonders gut gefallen?

- Nein
- Ja, und zwar:

Hättest du gerne noch etwas anderes gesehen oder erfahren?

- Nein
- Ja, und zwar:

Seite 3

4. Lernzuwachs
Was hast du für dich persönlich durch die Teilnahme am Workshop gelernt?

5. Abschließende Bewertung:
Ich gebe dem CAMMP day die Schulnote:

- 1 (sehr gut)
- 3 (befriedigend)
- 5 (mangelhaft)
- 2 (gut)
- 4 (ausreichend)
- 6 (ungenügend)

Ich gebe den Betreuern die Schulnote:

- 1 (sehr gut)
- 3 (befriedigend)
- 5 (mangelhaft)
- 2 (gut)
- 4 (ausreichend)
- 6 (ungenügend)

Abschließender persönlicher Kommentar (z.B. Lob, Kritik, Verbesserungsvorschläge):

» Umleitung auf Schlussseite von Umfrage Online

A.7.2. Ergebnisse Durchführung 1

CAMMP Day Machine Learning SuS

1. 1. Persönliche Angaben

   Anzahl Teilnehmer: 18

   - (0.0%): 8

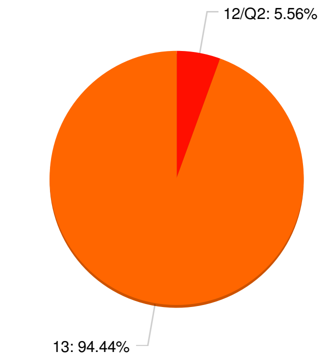
   - (0.0%): 9

   - (0.0%): 10/EF

   - (0.0%): 11/Q1

   1 (5.6%): 12/Q2

   17 (94.4%): 13

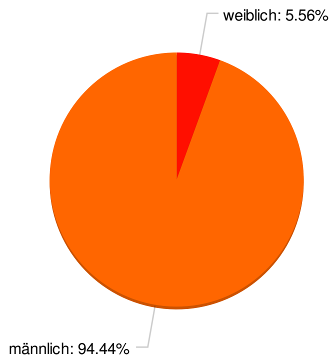


2. Bitte gib dein Geschlecht an:

   Anzahl Teilnehmer: 18

   1 (5.6%): weiblich

   17 (94.4%): männlich



3. Welche Leistungskurse besuchst du?

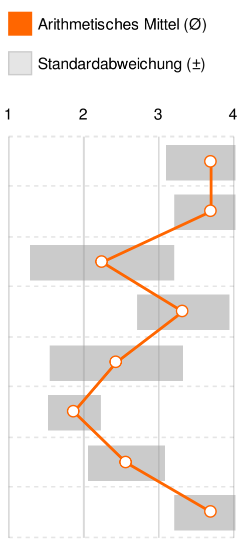
Anzahl Teilnehmer: 18

- Mathematik, Erdkunde
- Mathe englisch
- Mathe Biologie
- Mathe
- Mathe erdkunde
- Mathe und Erdkunde
- Bio, Chemie, Mathe, Sport
- Mathe und Biologie
- Mathe Erdkunde
- Mathematik Physik
- Mathe, Erdkunde
- Mathe Bio
- Mathe deutsch
- Mathe Biologie
- Mathe Erdkunde
- Mathe Erdkunde
- Mathe / Englisch
- mathe, englisch

4. 2. Bewertung des Workshops

Anzahl Teilnehmer: 16

	Trifft gar nich zu (--) (1)		Trifft eher nicht zu (-) (2)		Trifft zum Teil zu (+) (3)		Trifft voll zu (++) (4)			
	Σ	%	Σ	%	Σ	%	Σ	%	Ø	±
Durch den Workshop hab...	-	-	1x	6,25	3x	18,75	12x	75,00	3,69	0,60
Der Vortrag über Modelli...	-	-	-	-	5x	31,25	11x	68,75	3,69	0,48
Der einführende Kurzfilm...	1x	25,00	1x	25,00	2x	50,00	-	-	2,25	0,96
Die Einführung in MATLAB...	-	-	1x	6,67	8x	53,33	6x	40,00	3,33	0,62
Der Umgang mit MATLAB ...	2x	12,50	7x	43,75	5x	31,25	2x	12,50	2,44	0,89
Die Aufgaben waren zu e...	2x	13,33	13x	86,67	-	-	-	-	1,87	0,35
Die Aufgaben waren zu s...	-	-	7x	43,75	9x	56,25	-	-	2,56	0,51
Die gestuften Hilfekarte...	-	-	-	-	5x	31,25	11x	68,75	3,69	0,48

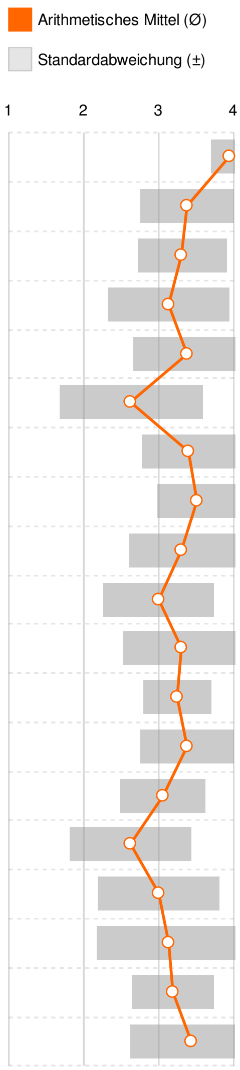


5. 3. Weiterführende Fragen

Wenn du über deine Erfahrungen mit diesem Kurs nachdenkst, wie sehr treffen die folgenden Aussagen auf dich zu?

Anzahl Teilnehmer: 16

	Trifft gar nicht zu (--) (1)		Trifft eher nicht zu (-) (2)		Trifft zum Teil zu (+) (3)		Trifft voll zu (++) (4)			
	Σ	%	Σ	%	Σ	%	Σ	%	Ø	±
Die Lern- und Arbeitsatmo...	-	-	-	-	1x	6,25	15x	93,75	3,94	0,25
Die Inhalte wurden klar ve...	-	-	1x	6,25	8x	50,00	7x	43,75	3,38	0,62
Die Lern- und Arbeitszeite...	-	-	1x	6,25	9x	56,25	6x	37,50	3,31	0,60
Der Kurs hat mein Interes...	-	-	4x	25,00	6x	37,50	6x	37,50	3,13	0,81
Ich habe nun eine genaue...	-	-	2x	12,50	6x	37,50	8x	50,00	3,38	0,72
Das Thema des Kurses wa...	1x	6,25	8x	50,00	3x	18,75	4x	25,00	2,63	0,96
Das Thema des Kurses fin...	-	-	1x	6,67	7x	46,67	7x	46,67	3,40	0,63
Mir ist klar geworden, wie...	-	-	-	-	8x	50,00	8x	50,00	3,50	0,52
Ich habe auf Blatt 1 versta...	-	-	2x	12,50	7x	43,75	7x	43,75	3,31	0,70
Die Modellverbesserung d...	-	-	4x	25,00	8x	50,00	4x	25,00	3,00	0,73
Der Wechsel von 2 auf 3 D...	-	-	3x	18,75	5x	31,25	8x	50,00	3,31	0,79
Die Fallunterscheidung au...	-	-	-	-	12x	75,00	4x	25,00	3,25	0,45
Ich habe auf Blatt 2 versta...	-	-	1x	6,25	8x	50,00	7x	43,75	3,38	0,62
Alles in allem habe ich da...	-	-	2x	12,50	11x	68,75	3x	18,75	3,06	0,57
Durch den Kurs habe ich i...	2x	12,50	3x	18,75	10x	62,50	1x	6,25	2,63	0,81
Ich habe in diesem Kurs v...	-	-	5x	31,25	6x	37,50	5x	31,25	3,00	0,82
Ich kann mir vorstellen ei...	-	-	6x	37,50	2x	12,50	8x	50,00	3,13	0,96
Ich würde so einen Kurs g...	-	-	1x	6,25	11x	68,75	4x	25,00	3,19	0,54
Alles in allem hat mir die ...	-	-	3x	18,75	3x	18,75	10x	62,50	3,44	0,81



6. Ist dir etwas besonders leicht gefallen?

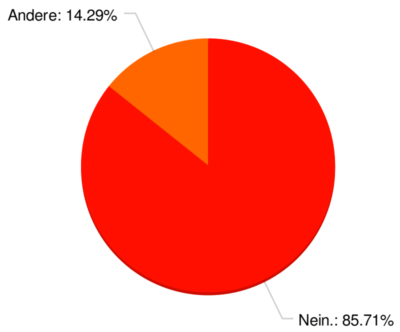
Anzahl Teilnehmer: 14

12 (85.7%): Nein.

2 (14.3%): Andere

Antwort(en) aus dem Zusatzfeld:

- Das räumliche Vorstellen
- One vs one



7. Ist dir etwas besonders schwer gefallen?

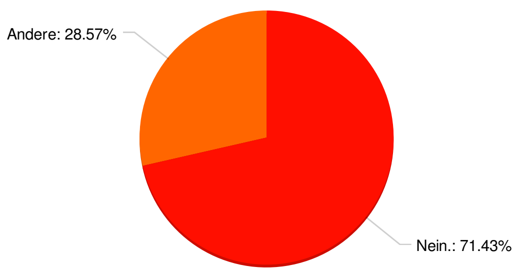
Anzahl Teilnehmer: 14

10 (71.4%): Nein.

4 (28.6%): Andere

Antwort(en) aus dem Zusatzfeld:

- Das direkte Verstehen der Aufagen
- Sich in das problem reinzudenken
- Die Herleitung der Abstandsformel
- Vieles



8. Hättest du dir stellenweise noch mehr Unterstützung gewünscht?

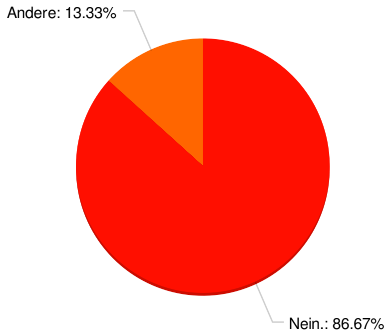
Anzahl Teilnehmer: 15

13 (86.7%): Nein.

2 (13.3%): Andere

Antwort(en) aus dem Zusatzfeld:

- Den nicht so einfachen Dingen
- Herleitung der Abstandsformel

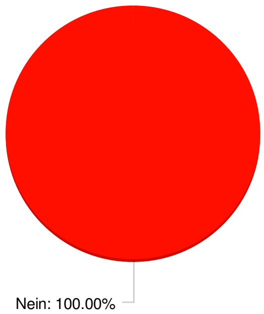


9. Hat dir etwas an dem Kurs absolut nicht gefallen?

Anzahl Teilnehmer: 15

15 (100.0%): Nein

- (0.0%): Andere

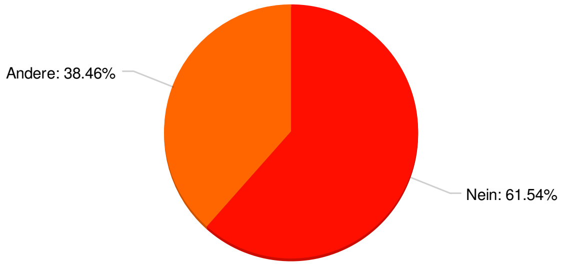


10. Hat dir etwas an diesem Kurs besonders gut gefallen?

Anzahl Teilnehmer: 13

8 (61.5%): Nein

5 (38.5%): Andere



Antwort(en) aus dem Zusatzfeld:

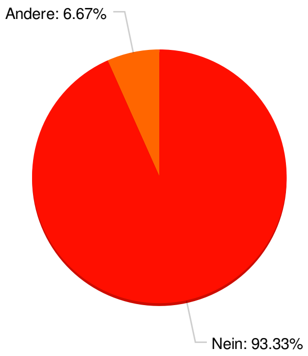
- Neffe und hilfsbereite Leute
- MacBooks zum arbeiten
- Die Kompetenz der Vortragende Studenten.War echt gut ▫
- Die Hilfe der präsentierenden
- Hilfestellungen

11. Hättest du gerne noch etwas anderes gesehen oder erfahren?

Anzahl Teilnehmer: 15

14 (93.3%): Nein

1 (6.7%): Andere



Antwort(en) aus dem Zusatzfeld:

- Eine kurze Erklärung zum Thema Sprachassistent.

12. 4. Lernzuwachs
Was hast du für dich persönlich durch die Teilnahme am Workshop gelernt?

Anzahl Teilnehmer: 3

- Besseres Verständnis für Künstlche Intelligenz
- Ich werde kein student
- Der Workshop war interessant, jedoch hat er mir gezeigt, dassIch mir keinen mathematischen Beruf vorstellen kann.

13. 5. Abschließende Bewertung:
Ich gebe dem CAMMP day die Schulnote:

Anzahl Teilnehmer: 14

4 (28.6%): 1 (sehr gut)

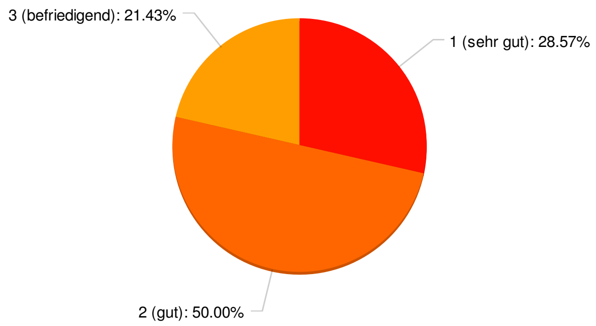
7 (50.0%): 2 (gut)

3 (21.4%): 3 (befriedigend)

- (0.0%): 4 (ausreichend)

- (0.0%): 5 (mangelhaft)

- (0.0%): 6 (ungenügend)



14. Ich gebe den Betreuern die Schulnote:

Anzahl Teilnehmer: 14

8 (57.1%): 1 (sehr gut)

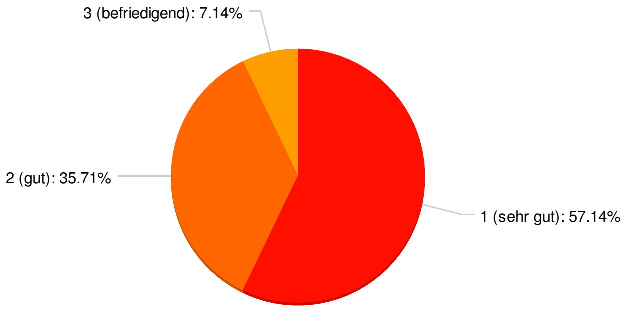
5 (35.7%): 2 (gut)

1 (7.1%): 3 (befriedigend)

- (0.0%): 4 (ausreichend)

- (0.0%): 5 (mangelhaft)

- (0.0%): 6 (ungenügend)



15. Abschließender persönlicher Kommentar (z.B. Lob, Kritik, Verbesserungsvorschläge):

Anzahl Teilnehmer: 2

- Coole sache
- Die Leiter des Workshops waren sehr nett. Jedoch sollte man mehr Zeit für den Workshop planen

A.7.3. Ergebnisse Durchführung 2

CAMMP Day Machine Learning SuS

1. 1. Persönliche Angaben

   Anzahl Teilnehmer: 16

   - (0.0%): 8

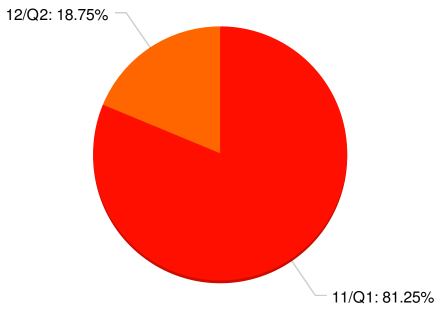
   - (0.0%): 9

   - (0.0%): 10/EF

   13 (81.3%): 11/Q1

   3 (18.8%): 12/Q2

   - (0.0%): 13

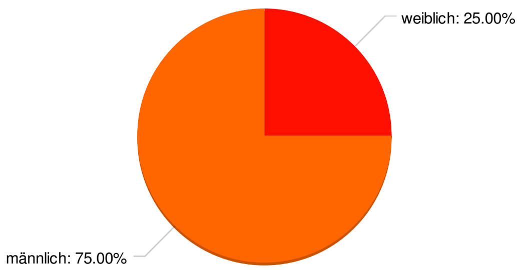


2. Bitte gib dein Geschlecht an:

   Anzahl Teilnehmer: 16

   4 (25.0%): weiblich

   12 (75.0%): männlich



3. Welche Leistungskurse besuchst du?

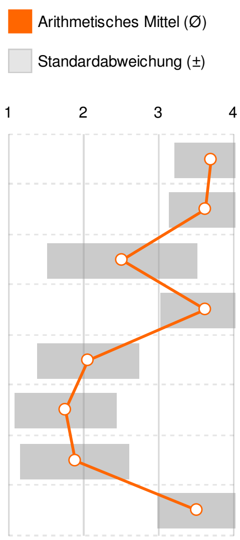
Anzahl Teilnehmer: 15

- Mathe und Erdkunde
- Mathe Erdkunde
- M,E
- Mathe , Erdkunde
- Mathe/Deutsch
- Mathe und Bio
- Mathe Bio
- Mathe/Erdkunde
- Mathe Englisch
- mathematik
- Mathe\Erdkunde
- Mathe, Erdkunde
- mathe, englisch
- Mathe Biologie
- Mathe englisch

4. 2. Bewertung des Workshops

Anzahl Teilnehmer: 16

	Trifft gar nich zu (--) (1)		Trifft eher nicht zu (-) (2)		Trifft zum Teil zu (+) (3)		Trifft voll zu (++) (4)			
	Σ	%	Σ	%	Σ	%	Σ	%	Ø	±
Durch den Workshop hab...	-	-	-	-	5x	31,25	11x	68,75	3,69	0,48
Der Vortrag über Modelli...	-	-	-	-	6x	37,50	10x	62,50	3,63	0,50
Der einführende Kurzfilm...	1x	25,00	-	-	3x	75,00	-	-	2,50	1,00
Die Einführung in MATLAB...	-	-	1x	6,25	4x	25,00	11x	68,75	3,63	0,62
Der Umgang mit MATLAB ...	3x	18,75	9x	56,25	4x	25,00	-	-	2,06	0,68
Die Aufgaben waren zu e...	6x	37,50	8x	50,00	2x	12,50	-	-	1,75	0,68
Die Aufgaben waren zu s...	5x	31,25	8x	50,00	3x	18,75	-	-	1,88	0,72
Die gestuften Hilfekarte...	-	-	-	-	7x	50,00	7x	50,00	3,50	0,52

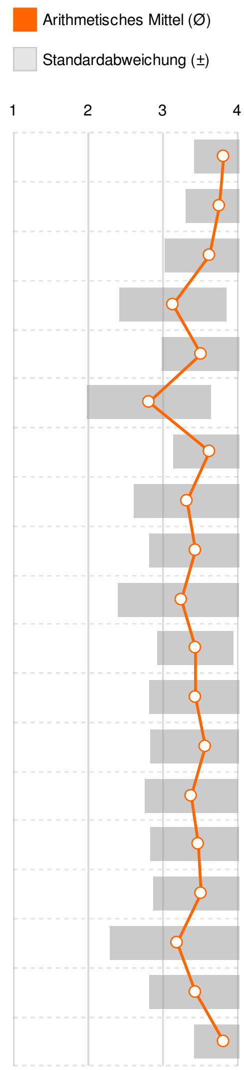


5. 3. Weiterführende Fragen

Wenn du über deine Erfahrungen mit diesem Kurs nachdenkst, wie sehr treffen die folgenden Aussagen auf dich zu?

Anzahl Teilnehmer: 16

	Trifft gar nicht zu (--) (1)		Trifft eher nicht zu (-) (2)		Trifft zum Teil zu (+) (3)		Trifft voll zu (++) (4)			
	Σ	%	Σ	%	Σ	%	Σ	%	Ø	±
Die Lern- und Arbeitsatmo..	-	-	-	-	3x	18,75	13x	81,25	3,81	0,40
Die Inhalte wurden klar ve..	-	-	-	-	4x	25,00	12x	75,00	3,75	0,45
Die Lern- und Arbeitszeite..	-	-	1x	6,25	4x	25,00	11x	68,75	3,63	0,62
Der Kurs hat mein Interes...	-	-	3x	18,75	8x	50,00	5x	31,25	3,13	0,72
Ich habe nun eine genaue...	-	-	-	-	8x	50,00	8x	50,00	3,50	0,52
Das Thema des Kurses wa...	1x	6,25	4x	25,00	8x	50,00	3x	18,75	2,81	0,83
Das Thema des Kurses fin...	-	-	-	-	6x	37,50	10x	62,50	3,63	0,50
Mir ist klar geworden, wie..	-	-	2x	13,33	6x	40,00	7x	46,67	3,33	0,72
Ich habe auf Blatt 1 verst...	-	-	1x	6,25	7x	43,75	8x	50,00	3,44	0,63
Die Modellverbesserung d...	1x	6,25	1x	6,25	7x	43,75	7x	43,75	3,25	0,86
Der Wechsel von 2 auf 3 D..	-	-	-	-	9x	56,25	7x	43,75	3,44	0,51
Die Fallunterscheidung au..	-	-	1x	6,25	7x	43,75	8x	50,00	3,44	0,63
Ich habe auf Blatt 2 verst...	-	-	2x	12,50	3x	18,75	11x	68,75	3,56	0,73
Alles in allem habe ich da...	-	-	1x	6,25	8x	50,00	7x	43,75	3,38	0,62
Durch den Kurs habe ich i...	-	-	1x	6,67	6x	40,00	8x	53,33	3,47	0,64
Ich habe in diesem Kurs v...	-	-	1x	6,25	6x	37,50	9x	56,25	3,50	0,63
Ich kann mir vorstellen ei...	1x	6,25	2x	12,50	6x	37,50	7x	43,75	3,19	0,91
Ich würde so einen Kurs g...	-	-	1x	6,25	7x	43,75	8x	50,00	3,44	0,63
Alles in allem hat mir die...	-	-	-	-	3x	18,75	13x	81,25	3,81	0,40



6. Ist dir etwas besonders leicht gefallen?

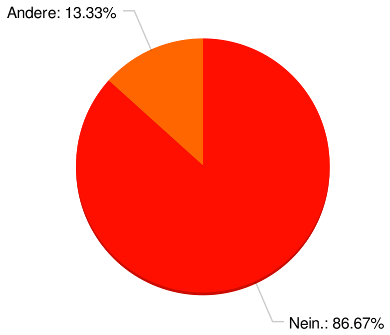
Anzahl Teilnehmer: 15

13 (86.7%): Nein.

2 (13.3%): Andere

Antwort(en) aus dem Zusatzfeld:

- Derz Übergang von 2auf 3 Dimensionen
- Daten abgleichen (analog zu digital)



7. Ist dir etwas besonders schwer gefallen?

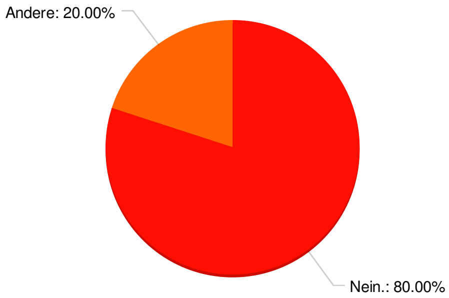
Anzahl Teilnehmer: 15

12 (80.0%): Nein.

3 (20.0%): Andere

Antwort(en) aus dem Zusatzfeld:

- Der Umstieg von 2D auf 3D
- Die Arbeit nach der Pause
- diverses

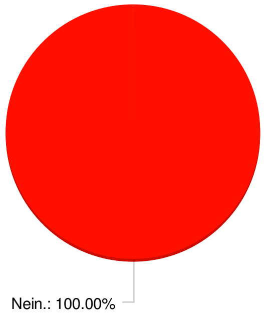


8. Hättest du dir stellenweise noch mehr Unterstützung gewünscht?

Anzahl Teilnehmer: 15

15 (100.0%): Nein.

- (0.0%): Andere

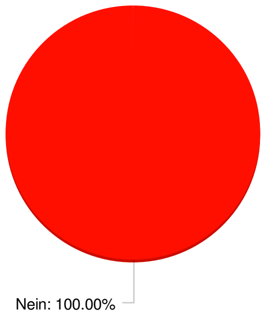


9. Hat dir etwas an dem Kurs absolut nicht gefallen?

Anzahl Teilnehmer: 16

16 (100.0%): Nein

- (0.0%): Andere

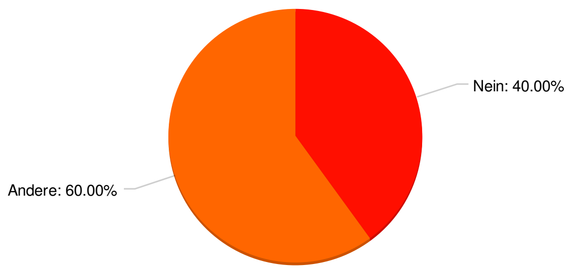


10. Hat dir etwas an diesem Kurs besonders gut gefallen?

Anzahl Teilnehmer: 15

6 (40.0%): Nein

9 (60.0%): Andere



Antwort(en) aus dem Zusatzfeld:

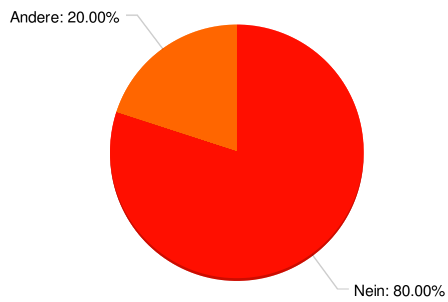
- Unterstützung von den Kursleitern
- Das Computerprogramm
- Infos zu studienmöglichkeiten
- Das Prinzip mit den Ampeln
- das arbeiten am Laptop
- Vortrag, Einführung
- Hilfsbereitschaft
- Dass gut geholfen wurde so dass man es versteht und sie sehr freundlich waren
- Die gute und aufschlussreiche Präsentation.

11. Hättest du gerne noch etwas anderes gesehen oder erfahren?

Anzahl Teilnehmer: 15

12 (80.0%): Nein

3 (20.0%): Andere



Antwort(en) aus dem Zusatzfeld:

- Mehr Möglichkeiten die es im Bereich gibt
- Mehr
- genauer die Materie

12. 4. Lernzuwachs

Was hast du für dich persönlich durch die Teilnahme am Workshop gelernt?

Anzahl Teilnehmer: 6

- Das vieles mit Mathematik zusammen hängt
- Das die neuen Geräte sehr komplexe Algorithmen haben.
- Die unterschiedlichen rechenarten
- Das das Thema Künstlicheintelligenz
- Das ist viel komplizierter als ich dachte.
- Ich habe um einiges besser verstanden wie Künstliche Intelligenzen funktionieren. Mein grundlegendes Verständnis von Programmen ist erweitert.

13. 5. Abschließende Bewertung:

Ich gebe dem CAMMP day die Schulnote:

Anzahl Teilnehmer: 15

8 (53.3%): 1 (sehr gut)

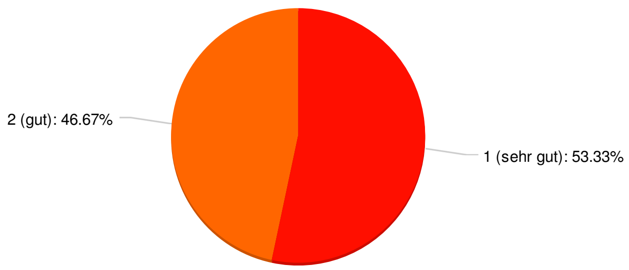
7 (46.7%): 2 (gut)

- (0.0%): 3 (befriedigend)

- (0.0%): 4 (ausreichend)

- (0.0%): 5 (mangelhaft)

- (0.0%): 6 (ungenügend)



14. Ich gebe den Betreuern die Schulnote:

Anzahl Teilnehmer: 16

14 (87.5%): 1 (sehr gut)

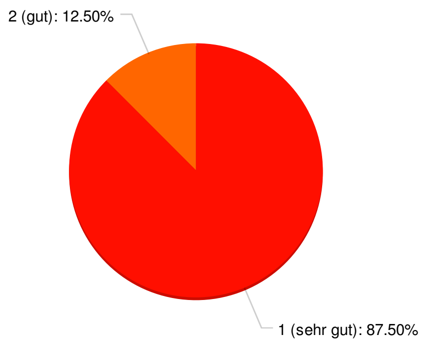
2 (12.5%): 2 (gut)

- (0.0%): 3 (befriedigend)

- (0.0%): 4 (ausreichend)

- (0.0%): 5 (mangelhaft)

- (0.0%): 6 (ungenügend)



15. Abschließender persönlicher Kommentar (z.B. Lob, Kritik, Verbesserungsvorschläge):

Anzahl Teilnehmer: 7

- Spitzen Arbeit
- Sehr hilfreich und interessant.
- Mehr Spontaneität
- War grandios
- Wenn es Fragen gab wurde immer geholfen und die Aufgaben waren verständlich und interessant . Also alles super
- War ein sehr guter workshop
- Ich muss sagen, dass der Vortrag sehr gut war, der Workshop Interessant und sonst auch alles gut und glatt gelaufen ist.

A.8. Eigenständigkeitserklärung

Zentrales Prüfungsamt/Central Examination Office

RWTH AACHEN UNIVERSITY

Eidesstattliche Versicherung
Statutory Declaration in Lieu of an Oath

Schmidt, Lars

Name, Vorname/Last Name, First Name

334227

Matrikelnummer (freiwillige Angabe)
Matriculation No. (optional)

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Arbeit/Bachelorarbeit/ Masterarbeit* mit dem Titel

I hereby declare in lieu of an oath that I have completed the present paper/Bachelor thesis/Master thesis* entitled

Machine Learning: automatische Bilderkennung mit Mathematik?!
Ein Lehr-Lern-Modul im Rahmen eines mathematischen Modellierungstages für Schülerinnen und Schüler der Sekundarstufe II

selbstständig und ohne unzulässige fremde Hilfe (insbes. akademisches Ghostwriting) erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Für den Fall, dass die Arbeit zusätzlich auf einem Datenträger eingereicht wird, erkläre ich, dass die schriftliche und die elektronische Form vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

independently and without illegitimate assistance from third parties (such as academic ghostwriters). I have used no other than the specified sources and aids. In case that the thesis is additionally submitted in an electronic format, I declare that the written and electronic versions are fully identical. The thesis has not been submitted to any examination body in this, or similar, form.

Aachen, 01.03.2019

Ort, Datum/City, Date

Lars Schmidt

Unterschrift/Signature

*Nichtzutreffendes bitte streichen

*Please delete as appropriate

Belehrung:
Official Notification:

§ 156 StGB: Falsche Versicherung an Eides Statt

Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

Para. 156 StGB (German Criminal Code): False Statutory Declarations

Whoever before a public authority competent to administer statutory declarations falsely makes such a declaration or falsely testifies while referring to such a declaration shall be liable to imprisonment not exceeding three years or a fine.

§ 161 StGB: Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt

(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.

(2) Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.

Para. 161 StGB (German Criminal Code): False Statutory Declarations Due to Negligence

(1) If a person commits one of the offences listed in sections 154 through 156 negligently the penalty shall be imprisonment not exceeding one year or a fine.

(2) The offender shall be exempt from liability if he or she corrects their false testimony in time. The provisions of section 158 (2) and (3) shall apply accordingly.

Die vorstehende Belehrung habe ich zur Kenntnis genommen:
I have read and understood the above official notification:

Aachen, 01.03.2019

Ort, Datum/City, Date

Lars Schmidt

Unterschrift/Signature