Optimierungspotentiale der Studienabbruchsvorhersage durch Frühwarnsysteme – Werden die „richtigen“ Studierenden gewarnt?

Robert Pelz, Franziska Schulze-Stocker, Johannes Winter, Wolfgang Haag

Zusammenfassung: Um Studierende bereits vor der Verfestigung von Problemlagen im Studienverlauf und somit weit vor einem potenziellen Studienabbruch unterstützen zu können, sind an einigen deutschen Hochschulen Frühwarnsysteme eingeführt worden. Dabei unterscheiden sich die Warnsysteme in der Art und Weise, wie sie konzipiert sind und wie sie die Studierenden erkennen, die möglichweise einen problematischen Studienverlauf haben könnten. Für den vorliegenden Beitrag dient das Frühwarnsystem der Technischen Universität Dresden als Beispiel dafür, wie durch Variation von Identifizierungsregeln eine Verbesserung der Erkennung problematischer Studienverläufe erfolgen kann. Es wird dargestellt, wie ein Weiterentwicklungsprozess eines aus der Praxis entstandenen Systems erfolgen kann und welche Daten dafür zur Verfügung stehen müssen. Die Ergebnisse zeigen, dass durch die vorgestellten Optimierungsmethoden eine Verbesserung der Identifizierungsleistung des Frühwarnsystems erreicht werden kann, sodass im Testdatensatz sechs bis acht Prozentpunkte mehr Studierende korrekt bestimmt werden können.

Schlüsselwörter: Frühwarnsystem, Studienabbruch, Studienverlauf, Optimierung, Identifizierungsmerkmale

Potential for optimizing prediction of academic dropout through early warning systems – Are the "right" students being warned?

Abstract: In order to support students before problematic situations manifest in the course of their studies and before they potentially drop out, early warning systems have been implemented at a number of German universities. The warning systems differ in the way that they have been designed and how they recognize students who may have a problematic course of study. For this paper, the early warning system of the TUD Dresden University of Technology is used as an example of how the variation of identification principles can improve the detection of problematic study trajectories. It shows how an extended development process could be implemented and which data must be available for this. The results of this study show that the application of the presented optimization methods can achieve the improvement of the identification performance of the early warning system, so that six to eight percentage points more students could be correctly determined in the test data set.

Keywords: Early warning system, Dropout, Study progress, Optimization, Identifying indicators

1 Einleitung

Von Studienabbruch wird gesprochen, wenn Studierende das Hochschulsystem ohne das Erreichen eines (ersten) Abschlusses endgültig verlassen (Heublein & Wolter, 2011, S. 216). Dabei ist Studienabbruch ein mehrfaktorieller Prozess, der auf multiple Motivlagen zurückzuführen ist (Blüthmann et al., 2008; Heublein & Wolter, 2011; Pelz et al., 2020). Neuere Studien zeigen, dass dieses Ausscheiden aus dem Studium zumindest in gewissen Berufsfeldern nicht unbedingt zu einer *individuellen* Schlechterstellung der Studienabbrecher*innen auf dem Arbeitsmarkt führen muss (Daniel et al., 2019). Sie werden in Einzelfällen auch „als Chance für den individuellen Lebensverlauf erlebt" (Hambrinker, 2021, S. 2) oder können „eine gewinnbringende Phase der Neuorientierung" anstoßen (Neugebauer et al., 2019, S. 1019). Trotzdem sind diese Fehlallokationen „mit hohen Kosten in Form von nicht für den Bildungsweg genutzter Lebenszeit, einem späteren Einstieg in den Arbeitsmarkt sowie möglichen psychologischen Kosten verbunden" (Isphording & Wozny, 2018, S. 3). Eine Studie von Klein, Mishra und Müller (2021) kann sogar zeigen, dass im Lebensverlauf von Studienabbrecher*innen bezüglich der subjektiven Arbeits- und Lebenszufriedenheit Nachteile zu erwarten sind. Auf *gesellschaftlicher* Ebene entstehen – besonders vor dem Hintergrund des regionalen und branchenspezifischen Fachkräftemangels bei gleichzeitig gesamtgesellschaftlichen demografischen Veränderungen, die seit 2019 zu einem Rückgang der Zahl der Studienanfänger*innen beigetragen haben – hohe volkswirtschaftliche Kosten (Statistisches Bundesamt, 2023). Für Deutschland lagen diese bereits im Jahr 2007 im Milliardenbereich (Stifterverband für die Deutsche Wissenschaft, 2007). Die Sorge um den Studienabbruch hat aber auch auf *institutioneller* Ebene der Hochschulen in den letzten Jahren stark an Bedeutung gewonnen. Durch die Einführung einer leistungsorientierten Mittelvergabe mit Zielvereinbarungen, die den Studienerfolg in den Blick nehmen, über die Qualitätssicherungsbestrebungen in den Studiengängen im Rahmen von Akkreditierungsprozessen bis hin zum Image der Hochschulen oder dem Erfolg in Rankings können hohe Studienabbruchquoten deutlich größere Auswirkungen haben, als dies noch vor wenigen Jahren der Fall war. Durch unterstützende bundesweite Programme wie dem Qualitätspakt Lehre oder dem Zukunftsvertrag *Studium und Lehre stärken* haben die Hochschulen zudem weitere Anreize dafür erhalten, den Anteil an abbruchgefährdeten Studierenden zu minimieren (siehe hierzu auch Konrad & Wildner, 2020).

Eine wichtige Maßnahme, die zur Vermeidung des Studienabbruchs eingesetzt werden kann, sind sogenannte Frühwarnsysteme (Berens, 2021; Konrad & Wildner, 2020; Schulze-Stocker et al., 2017; Opitz, 2021). Aus der Forschung ist bekannt, dass beispielsweise in den Bachelorstudiengängen an Universitäten fast die Hälfte der Studienabbrecher*innen bis zum Studienabbruch nur maximal zwei Semester (Heublein et al., 2017) verbringen, unabhängig von der Fächergruppe. Zusammen mit der Erkenntnis aus der Praxis der Studienberatungen, dass Studierende meist zu spät auf Hilfestrukturen zurückgreifen, werden die vier Ziele von Frühwarnsystemen schnell plausibel: Studierende sollen noch vor dem Ansammeln von Problemen im Studienverlauf erkannt (1) und über ihren sogenannten „Identifizierungsstatus" benachrichtigt werden (2), damit noch die Chance besteht, den eigenen Studienverlauf zu reflektieren (3) und rechtzeitig, d.h. mit einem möglichst großen Interventionszeitraum, zu den passenden Beratungs- und Unterstützungsangeboten geleitet zu werden (4; siehe auch Ahles et al., 2016; Heublein et al., 2017; Berens, 2021; Konrad & Wildner, 2020). Bei der

Konzeption der Merkmale zur Identifizierung von problematischen Studienverläufen können unterschiedliche Datengrundlagen zum Tragen kommen:

1. Befragungsdaten und
2. administrative Studierendendaten (von aktuellen oder ehemaligen Studierenden).

Eine weitere Unterscheidung besteht bezüglich der Entwicklung der Merkmale: Die Merkmale abbruchgefährdeter Studierender können dabei entweder durch Expertise gesetzt (exogen; bspw. Westerholt et al., 2018; Anastasio et al., 2020) oder anhand historischer Studierendendaten ermittelt werden (endogen; Berens, 2021, S. 12, 90).

Der Vergleich von Berens und Schneider (2019) beziehungsweise Berens (2021) hinsichtlich drei verschiedener Frühwarnsysteme in Deutschland ergab, dass die Prognosen besser sind, wenn das Frühwarnsystem administrative Daten aktueller *und* ehemaliger Studierender nutzt und wenn die verwendeten Informationen komplexer sind. Da beim Frühwarnsystem PASST?! bei der Konzeption der Merkmale nicht auf bestehende administrative Studienverlaufsdaten zurückgegriffen werden konnte, die einen Studienerfolg oder Studienabbruch eindeutig erfassen, blieb es zunächst unklar, wie akkurat die Identifizierung studienabbruchsgefährdeter Studierender und die Nicht-Identifizierung erfolgreicher Studierender gelingt. Das Ziel dieses Artikels ist es daher, auf Basis von administrativen Studienverlaufsdaten von ausgewählten Studiengängen[1] an der Technischen Universität Dresden der Forschungsfrage(n) nachzugehen, ob die Genauigkeit dieser Identifizierung mit dem vorliegenden Gesamtmodell an Merkmalen verbessert werden kann und ob Variationen der bisher genutzten Merkmale zu (noch) besseren Modellen führen können.

In den folgenden Kapiteln werden verschiedene Ansätze der Optimierung von Frühwarnsystemen beschrieben. Daran anschließend wird im methodischen Teil das Frühwarnsystem der Technischen Universität Dresden und insbesondere die verwendeten Identifizierungsmerkmale und deren Optimierungspotenziale vorgestellt. Diese fünf Identifizierungsmerkmale erlauben es, studienabbruchsgefährdete Studierende zu kontaktieren und ihnen bereits beim Anschreiben passende Angebote zu unterbreiten. Im fünften Kapitel werden die Ergebnisse des Vorgehens dargestellt, bevor abschließend die Diskussion erfolgt.

2 Frühwarnsystem PASST?!

Das erste an einer Hochschule etablierte Frühwarnsystem in Deutschland (Berens, 2021) gibt es seit 2016 an der Technischen Universität Dresden. Dieses System prüft die individuellen Studienverläufe der am Programm teilnehmenden Studierenden hinsichtlich fünf sogenannter Identifizierungsmerkmale. Die Bezeichnung "Merkmal" darf dabei nicht als stochastisch, d. h. im Sinne des Ausgangs eines Zufallsexperiments, verstanden werden, sondern als Beschreibung entsprechender Zustände von Studierenden im Studienverlauf. Die Identifizierungsmerkmale wurden aus der Beratungserfahrung der Zentralen Studienberatung der Technischen Universität Dresden, den von der Universität verfolgten hochschulpolitischen Ziel-

1 Die Auswahl der Studiengänge (vgl. Kapitel 4) kann nicht als repräsentativ für die gesamte Universität angesehen werden, da aus forschungspragmatischen Gründen der Fokus auf vergleichsweise große Studiengänge gelegt werden musste, die über einen gewissen Zeitraum nur wenige Änderungen in den Prüfungs- und Studienordnungen vorweisen können.

setzungen (Technische Universität Dresden, 2025) und den Daten aus einer Auftaktbefragung des PASST?!-Programms (mehr Informationen zur Erhebung siehe Schulze-Stocker et al., 2017) „exogen“ (Berens, 2021, S. 90) ermittelt. Ausgangspunkt war eine vom Programm gesetzte Definition von Studienerfolg[2], die grundlegend die Studienleistungen als nötig, aber nicht hinreichend beschreibt (Schulze-Stocker et al., 2020, S.190). Daraus ergaben sich folgende fünf Identifizierungsmerkmale, die zusammen Problemlagen im gesamten Studienverlauf erfassen:

- *Identifizierungsmerkmal 1* ist mit dem Ziel entwickelt worden, Startschwierigkeiten im Studium aufzudecken und den Studierenden bei diesen Studieneingangsproblemen frühzeitig passende Beratungs- und Unterstützungsangebote unterbreiten zu können. Studierende, die aufgrund dieses Merkmals nach dem ersten Fachsemester angeschrieben werden, haben das Kriterium erfüllt, zu diesem Zeitpunkt noch keine oder nur eine Prüfungsleistung erworben zu haben.
- *Identifizierungsmerkmal 2* bildet die Situation der letzten beiden (aktiven) Fachsemester ab. Unterschreiten die Studierenden in Summe der beiden Semester 30 Creditpoints, werden sie vom Frühwarnsystem benachrichtigt. Somit können Leistungseinbrüche ab dem dritten Fachsemester im gesamten Studienverlauf ermittelt werden.
- *Identifizierungsmerkmal 3* führt zu einem Anschreiben der Studierenden, wenn diese eine Prüfung im zweiten Versuch nicht bestanden haben. Ziel ist es, in diesem Fall frühzeitige Hilfe in einer Akutsituation zu geben, da ein gescheiterter dritter Versuch die Exmatrikulation im studierten Studiengang zur Folge hat.
- *Identifizierungsmerkmal 4* wurde aufgrund der Erfahrungen aus der Beratung eingeführt. Es prüft, ob in den Prüfungsmanagementsystemen für die Studierenden im letzten Semester drei oder mehr Rücktritte von Prüfungsleistungen vermerkt sind. Somit wird eine Situation der Überlastung erfasst, die ganz unterschiedliche Ursachen haben kann.
- *Identifizierungsmerkmal 5* nimmt Studienabschlussprobleme in den Blick und steht damit mit den hochschulpolitischen Studienerfolgszielen stark in Verbindung. Die Studierenden werden kontaktiert, wenn sie bereits im zweiten Semester der Regelstudienzeitüberschreitung sind. Bei der Optimierung kommt diesem Merkmal eine Sonderrolle zu, da Studierende, die die Regelstudienzeit um mehr als zwei Fachsemester überschreiten, mit diesem Identifizierungsmerkmal immer vor Ende des (erfolgreichen) Studiums identifiziert werden (vgl. Szenarien der Prognose in Tabelle 1).

2 „Der Studienabschluss (Zeugnis) allein ist notwendig, aber nicht hinreichend für den Studienerfolg. Die weiteren Kriterien in der folgenden Betrachtungsweise sind ebenfalls wichtig. Hinreichend sind das Zeugnis und mindestens zwei der folgenden fünf Kriterien: (1) Der Abschluss wird nach maximal zwei Semestern über der Regelstudienzeit erreicht. (2) Das Studium führt zu einem Kompetenzerwerb in fachlicher, sozialer und methodischer Hinsicht. (3) Studienerfolg geht mit persönlicher Zufriedenheit mit dem eigenen Bildungsweg einher. (4) Das Studium trägt zur Persönlichkeitsentwicklung (einschließlich Korrekturfähigkeit der Studienwahl) bei. (5) Das Studium befähigt Studierende zur Ergreifung eines Berufes.“

3 Ansätze zur Optimierung der Identifizierungsmerkmale eines Frühwarnsystems: Hypothesengenerierung

Auch wenn keine einheitliche Definition von *Optimum* vorhanden ist (D'Angelo, 2019), wird darunter im Allgemeinen das beste erreichbare Resultat verstanden, welches unter einer Reihe von vorgegebenen Rahmenbedingungen und Voraussetzungen erreicht werden kann. Es grenzt sich dabei vom Ideal ab, das das beste denkbare Ergebnis bezeichnet, welches den höchsten Vorstellungen entspricht (Dudenredaktion, o. J.). Der Prozess, der zum Erreichen dieses Optimums führt, wird als Optimierung bezeichnet. In diesem Artikel geht es um die optimierte Genauigkeit der Identifizierung. Die Kriterien hierfür werden in Kapitel 4.2 beschrieben.

Die fünf oben genannten Identifizierungsmerkmale bilden im Verständnis des vorliegenden Artikels ein Modell. Das Ideal von Frühwarnsystemen besteht darin, mit Hilfe eines solchen Modells alle studienabbruchsgefährdeten Studierenden rechtzeitig im Studienverlauf zu kontaktieren, um ihnen Unterstützungsangebote zu unterbreiten. Auf diesem Weg sollen unnötige Studienabbrüche verhindert werden. Dieser Zustand kann durch eine starke Modellreduktion sehr einfach erreicht werden: Als einziges Merkmal könnte die Immatrikulation gelten, sodass alle Studierenden ohne vorherige Selektion kontaktiert werden. Dieses unspezifische Vorgehen hätte aber den Nachteil, dass weder ein genauer Grund der Kontaktaufnahme benannt werden kann noch präzise Handlungsempfehlungen gegeben werden können, die eine stark begrenzte Wirkung entfalten dürften. Außerdem würde das Kontaktieren von Personen, die nicht studienabbruchsgefährdet sind, kontraproduktiv wirken. Aus der Surveyforschung ist bekannt, dass die Bereitschaft an Befragungen teilzunehmen mit zunehmender Kontaktierung sinkt (Hauss & Seyfried, 2019). Folglich kann auch bei Frühwarnsystemen die Gefahr bestehen, dass durch die Zusendung überflüssiger, diffuser Anschreiben die weitere Teilnahmebereitschaft und Akzeptanz unter den Studierenden sinkt.

Eine Methode, um datentechnisch zwischen mutmaßlich studienabbruchsgefährdeten Studierenden und nicht studienabbruchsgefährdeten Studierenden zu unterscheiden, ist die Anwendung von Machine-Learning-Verfahren (Berens et al., 2019; Kemper et al., 2020; Theune, 2021), wie zum Beispiel neuronalen Netzwerken (Vandamme et al., 2007), Support-Vektor-Maschinen (Theune, 2021; Mohd et al., 2023) oder Random-Forrest-Algorithmen (Theune, 2021). Unter Nutzung eines administrativen Datenpools können verschiedenste individuelle und überindividuelle Daten ausgewertet werden, um daraus Modelle zu erstellen, die auf Basis dieser Trainingsdaten wiederum Prognosen für die Zukunft von aktuell Studierenden erstellen können. Diese Verfahren können unter Aufwendung großer Datenmengen etwas besser als praxeologische Frühwarnsysteme zwischen studienabbruchsgefährdeten und nicht gefährdeten Studierenden unterscheiden (Berens, 2021). Werden keine „White-Box-Modelle" genutzt, ist es allerdings aufgrund der Komplexität der Modelle bei Machine-Learning-Verfahren schwierig, den ermittelten Studierenden konkrete Gründe für die individuellen Prognosen, in diesem Falle also die für einen Studienabbruch, zu benennen (Waltl, 2019).

Insofern Identifizierungsmerkmale quantifizierbare Grenzen enthalten (wie z.B. Anzahl von Prüfungsleistungen oder Anzahl von Creditpoints), die in der Größe veränderbar sind, können solche Verfahren aber auch eingesetzt werden, um durch die Variation der Grenzen der Identifizierungsmerkmale selbst sowie durch die Veränderung der Parameter der Grenzen eine Optimierung der aus der Praxis entwickelten Identifizierungsmerkmale beziehungs-

weise des Gesamtmodells zu erreichen. Es sind eine Reihe von Verfahren bekannt, mittels derer sich mit mehr oder weniger Aufwand Parameter finden, die nahezu optimal sind. Deswegen muss eine Auswahl des geeigneten Machine-Learning-Modells basierend auf dem Problem und den vorhandenen Daten erfolgen. Dies kann beispielsweise ein Supervised-Learning-Modell wie neuronale Netzwerke, Entscheidungsbäume oder Support-Vector-Machines sein oder ein Unsupervised-Learning-Modell wie Cluster-Analyse oder Dimensionsreduktion. Oftmals fungiert auch der erwartete notwendige hohe Rechenaufwand als Entscheidungsgrundlage, auf dessen Basis die Entscheidung für ein bestimmtes Verfahren gefällt werden muss.

Da der Rechenaufwand für den vorhanden Datensatz nicht zu hoch war, konnte für die geplanten Analysen ein vergleichsweise rechenintensives Verfahren (vgl. Kapitel 4.2) gewählt werden, dem die exogenen Identifizierungsmerkmale als Basis zugrunde gelegt werden konnten. In diesem Zusammenhang war es für die Umsetzung der Optimierung und auch für die im Projekt geplante spätere Entscheidung für oder gegen eine Anpassung des bestehenden Modells wichtig, Schlussfolgerungen anhand der bisher genutzten Merkmale ziehen zu können. Da bei einzelnen Machine-Learning-Verfahren dieser Bezug weniger im Fokus steht, wurde sich daher dagegen entschieden, diese Ansätze, wie z.B. neuronale Netzwerke, für die anstehenden Analysen zu verwenden.

Ein zentraler Schritt stellt das sogenannte Training des Modells dar. Dabei werden die Modellparameter an die Trainingsdaten angepasst, um die Leistung des Modells zu optimieren beziehungsweise die Fehlerquote zu verringern. Dies wird in der Regel durch Optimierungsalgorithmen wie beispielsweise dem Grid-Search-Algorithmus, dem Bergsteigerverfahren oder dem Gradientenabstiegsverfahren durchgeführt. Die Evaluation des Modells erfolgt dann anhand der Bewertung der Leistung des trainierten Modells (Trainingsdaten) im Umgang mit Daten, die nicht für das Training genutzt wurden (Testdaten). Die Maße Accuracy, Recall, Precision, F1 und Specifity dienen zur Beurteilung der entsprechenden Modelle und werden im Verlauf des vorliegenden Artikels genauer vorgestellt.

Wendet man diese Verfahren der Verwendung von Optimierungsalgorithmen auf veränderbare Identifizierungsmerkmale an, sollten die Parameter so bestimmt werden, dass die Genauigkeit der Analyse unter dem vorhandenen Identifizierungsmerkmal maximiert wird. Im hier vorliegenden Fall sollte das Verfahren die Vorhersagekraft des Frühwarnsystems verbessern. Die Hypothese dazu lautet daher:

1. Die Identifizierungsmerkmale des Frühwarnsystems der Technischen Universität Dresden lassen sich mittels der Optimierungsverfahren – wie später noch ausführlicher erläutert – im hier vorliegenden Fall des Grid-Search-Algorithmus analysieren. Auf diese Weise lassen sich optimale Parameter für die Identifizierungsmerkmale bestimmen, für welche die Anzahl der korrekt identifizierten Studierenden erhöht wird.

Die bisherigen Merkmale des Frühwarnsystems der Technischen Universität Dresden werden unabhängig vom Studiengang, in dem sich die Studierenden befinden, formuliert (Schulze-Stocker et al., 2020). In der Praxis zeigt sich aber, dass in den an der Universität angebotenen Studiengängen unterschiedliche Rahmenbedingungen und Regelungen existieren und beispielsweise die Anzahl an Prüfungen, die pro Fachsemester abzulegen sind, variieren (Schulze-Stocker & Schäfer-Hock, 2020). Ein Beispiel wäre die Prüfungslastverteilung im Studienverlauf: Im Zuge der Bologna-Reform war es zwar ein Ziel, die Last über alle Semester gleich zu verteilen, in der Praxis, meist geregelt über ältere Studien- und Prüfungsordnun-

gen, werden aber auch beispielsweise Module über mehrere Semester angeboten. Dies hat zur Folge, dass in den „Startsemestern“ dieser Module nur Teil- oder Vorleistungen zu erbringen sind. Im Prüfungssystem werden aber nur die Leistungspunkte von vollständig absolvierten Modulen erfasst und den Studierenden formal gutgeschrieben. Somit sind die Leistungspunkte nicht in jedem Studiengang und über alle Semester gleich verteilt. So schwanken die im ersten Fachsemester laut Studienordnung zu erwerbenden Leistungspunkte bei den betrachteten Studiengängen zwischen vier und 30. Zudem sind aufgrund vieler semesterübergreifender Module in einzelnen Studiengängen in einem Fachsemester laut Studienordnung null Leistungspunkte und im folgenden Fachsemester hingegen 56 Leistungspunkte formal zu erbringen. Daher könnte es auch sinnvoll sein, summativ über den Studienverlauf Leistungspunkte zum Zeitpunkt X als Schranken für Abbruchhinweise anzunehmen. Zudem variieren zwischen Studiengängen die zu absolvierenden Prüfungsleistungen pro Modul. Daher wird in einem zweiten Schritt angenommen, dass Identifizierungsmerkmale, die mit studiengang- und semesterabhängigen Grenzen arbeiten, die Gefährdungen der Studierenden besser erfassen als solche, die studiengang- und semesterunabhängig arbeiten. Aus diesen Gedanken heraus wird folgende zweite Hypothese formuliert:

2. Durch die Variation der *Identifizierungsmerkmale 1* (Anzahl der Prüfungsleistungen im ersten Fachsemester) und *2* (Anzahl der erworbenen Creditpoints) kann die Genauigkeit der Identifizierung von Studierenden verbessert werden. Insbesondere wirken
   a. studiengangabhängige Identifizierungsmerkmale, d.h. im konkreten Fall sich über den Betrachtungszeitraum mehrerer Semester erstreckende aufsummierbare Merkmale, besser als studiengangunabhängige Identifizierungsmerkmale und
   b. fachsemesterabhängige Identifizierungsmerkmale, d.h. auf bestimmte Semester bezogene Merkmale, besser als fachsemesterunabhängige Identifizierungsmerkmale.

Diese zwei Hypothesen werden im Folgenden überprüft, dazu wird im nächsten Abschnitt zunächst genauer auf die verwendete Methode eingegangen.

4 Methode

4.1 Datengrundlage

Zur Überprüfung der genannten Thesen werden die Studienverlaufsdaten von Studierenden aus fünf Studiengängen der Technischen Universität Dresden verwendet, die ihr Studium zwischen Wintersemester 2017/18 und dem Wintersemester 2019/20 mit oder ohne Studienabschluss beendet haben. Die fünf Studiengänge lauten wie folgt:

- Bauingenieurwesen (Diplom),
- Forstwissenschaften (Bachelor),
- Physik (Bachelor),
- Psychologie (Bachelor),
- Wirtschaftsingenieurwesen (Diplom).

Für die Nutzung historischer Studienverlaufsdaten musste, zu Lasten der Repräsentativität, bei der Auswahl der Studiengänge der Fokus auf Studiengänge mit vergleichsweise großer

Anzahl an Studierenden gelegt werden, die zudem in den letzten drei bis fünf Jahren (Regelstudienzeiten der Studiengänge) nur wenige fundamentale Änderungen in den Prüfungs- und Studienordnungen ausweisen.

Berücksichtigt werden Vollzeit-Studierende in modularisierten Bachelor- und Diplomstudiengängen, die keine Studienverlaufslücken aufwiesen (beispielsweise durch Überspringen von Fachsemestern). Die Daten stammen aus dem Studierendenverwaltungssystem der Technischen Universität Dresden HISQIS. Insgesamt werden die Studienverlaufsdaten von 1.912 Studierenden betrachtet. Zur Analyse (siehe Kapitel 4.2) werden der Studiengang, die Fachsemester und der Abschlussstatus (abgeschlossen, Abschlussprüfung endgültig nicht bestanden, kein Status) einbezogen. Zusätzlich werden pro Semester und Studierender bzw. Studierendem folgende Daten genutzt, um die Studienverläufe der Studierenden zu konstruieren und die fünf oben beschriebenen Identifizierungsmerkmale prüfen zu können:

- Fachsemester
- Anzahl der im Semester abgelegten Prüfungsleistungen
- Anzahl der im Semester erworbenen Leistungspunkte
- Einschreibestatus (im Studiengang eingeschrieben/nicht im Studiengang eingeschrieben)
- Anzahl der im Semester nicht bestandenen Prüfungsleistungen
- Anzahl der im Semester zum zweiten Mal nicht bestandenen Prüfungsleistungen
- Anzahl der Prüfungsrücktritte im Semester (d.h. ausschließlich Anzahl der Personen mit Abmeldung mittels eines Attests in einer Frist von weniger als drei Tagen vor der Prüfung)

4.2 Vorgehen/Methode

Um die Qualität von Prognosen zu beurteilen, wird der Datensatz gewöhnlich in zwei Teile, einen Trainings- und einen Testdatensatz, unterteilt. Dies kommt insbesondere bei der Nutzung von Verfahren des Machine-Learnings zum Einsatz, ist aber allgemein für die Überprüfung der Wirksamkeit von mittels mathematischer Verfahren optimierter Prognosemodelle anwendbar (Murphy, 2012). Die Parameter werden auf dem Trainingsdatensatz optimiert und die Prognose auf den Testdatensatz angewendet und dort evaluiert. Der Gedanke dahinter ist, dass das Optimierungsverfahren potentiell zu einer Überanpassung des Modells auf die Trainingsdaten führt und über den Validierungsdatensatz, also die Daten, die das Modell noch nie zuvor gesehen hat, die wirkliche Qualität eines Modells ermittelt werden kann (Marsland, 2014; Russel & Norvig, 2010). Da hier zusätzlich noch unterschiedliche, trainierte Modelle verglichen werden, wird eine Dreiteilung genutzt, bei der ein Teil als Trainingsdatensatz, ein Teil als Validierungsdatensatz für den Modellvergleich und ein Teil als Testdatensatz zur Angabe des finalen Fehlers zum Einsatz kommt. Um Verzerrungen möglichst zu vermeiden, wird zusätzlich eine 5x3-Kreuzvalidierung durchgeführt. Das heißt der Ursprungsdatensatz wird fünfmal unterteilt, wobei jeweils 80 Prozent der Daten als Trainings-/Validierungsset und 20 Prozent als Testset verwendet werden. Jeder Trainings-/Validierungsdatensatz wird zusätzlich dreimal unterteilt, wobei in jeder Teilung 66,6 Prozent des größeren Datensatzes als Trainings- und 33,3 Prozent des größeren Datensatzes als Validierungsdatensatz zum Einsatz kommen (siehe Abb. 1). Es müssen somit im Training 15 = 5 * 3 Datensätze mit 53,3

Prozent der Daten trainiert werden und 5 * 3 mit 26,7 Prozent der Daten validiert werden. Zur Bewertung der Qualität werden jeweils die Durchschnittswerte angegeben.

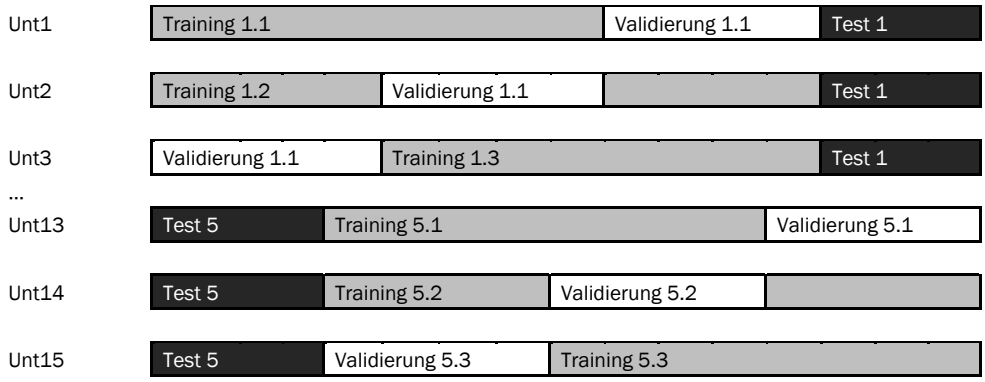

Abbildung 1: Datenunterteilung in Trainings-, Validierungs- und Testdatensatz (insges. 5 * 3 = 15 Unterteilungen (Unt))

Für die Optimierung der Parameter auf dem Trainingsdatensatz können verschiedene Verfahren zum Einsatz kommen. Das sind insbesondere der Grid-Search-Algorithmus, das Bergsteigerverfahren und das Gradientenabstiegsverfahren. Der Grid-Search-Algorithmus stellt das einfachste und gleichzeitig umfassendste Verfahren dar. Er wird auch Rastersuche genannt und basiert darauf, dass der gesamte zu untersuchende Parameterraum in gleiche Teile aufgeteilt wird (Liashchynskyi & Liashchynskyi, 2019). Es wird also für jedes Parametertupel (d.h. eine geordnete Menge von Parametern) anhand der vorgegebenen Regeln vorhergesagt, ob eine Person das Studium innerhalb der Regelstudienzeit plus zwei Fachsemester abschließt oder nicht. Dies bedeutet, dass für jedes Parametertupel die Anzahl der durch dieses Parametertupel identifizierten Studierenden (Studierende, die über bestimmte Eigenschaften, wie beispielsweise die gleiche Anzahl an erbrachten Leistungspunkten, identifizierbar sind) mit den zu identifizierenden Studierenden abgeglichen wird. Zum Schluss wird das Parametertupel ausgewählt, bei dem auf dem Trainingsdatensatz der Anteil der Falschvorhersagen am geringsten ist. Ein großes Problem bei diesem Verfahren ist der enorme Zeitbedarf bei der Berechnung, den das Verfahren mit steigender Anzahl an Parametern hat (Syarif et al., 2016). Dieser steigt exponentiell. Fügt man also in einem bestehenden zu optimierenden Regelwerk mit Parametern für eine Vorhersage einen weiteren Parameter hinzu, so wird der Zeitbedarf für die Optimierung mit der Anzahl der untersuchten Ausprägungen des Parameters multipliziert. Ein zusätzlicher Parameter mit zehn Ausprägungen verzehnfacht also den Zeitbedarf. Fügt man zwei Parameter mit zehn Ausprägungen hinzu, dann verhundertfacht sich der Zeitbedarf sogar. Im Folgenden handelt es sich beispielsweise um 3.000 mögliche Kombinationen, wenn der erste Parameter 10, der zweite 30 und der dritte 10 verschiedene Ausprägungen aufweist.

Die beiden weiteren Verfahren sollen für den Fall einer höheren Anzahl an Parametern, die eine Berechnung durch den Grid-Search-Algorithmus erschweren oder unmöglich machen, an dieser Stelle noch kurz erläutert werden: Ein Verfahren, das diesem Problem begegnen kann, ist das Bergsteigerverfahren (Schneider, 2022). Dabei werden ausgehend von einem Parametertupel die Vorhersagen für dieses und alle umliegenden Parametertupel in dem

Trainingsset getroffen. Im nächsten Schritt wird das Tupel gewählt, bei dem der Anteil der richtigen Vorhersagen am höchsten ist und dessen Vorhersage gleichzeitig genauer ist als die Vorhersage des bisherigen Parametertupels. Für dieses werden wiederum die Ergebnisse für alle anliegenden Parametertupel berechnet. Der Algorithmus endet, wenn kein umliegendes Parametertupel mehr eine bessere Vorhersagequalität liefert, sozusagen der Gipfel vom Bergsteiger erklommen ist. Ein Problem dieses Verfahrens ist, dass der höchste gefundene Parametertupel nicht tatsächlich die beste Vorhersagequalität haben muss, daher wird das Verfahren auch immer von mehreren zufällig gewählten Parametertupeln gestartet (Russel & Norvig, 2010). Außerdem ist auch die Berechnung der Ergebnisse aller umliegenden Parametertupel immer noch sehr aufwendig. Ein Verfahren, das diesen Problemen wiederum begegnet, ist das Gradientenabstiegsverfahren (Lückehe, 2015). Bei diesem wird eine mathematische Besonderheit ableitbarer Funktionen genutzt. Das Verfahren ist derart angelegt, dass der Gradient (die Ableitung einer Funktion an einer bestimmten Stelle) hierbei immer in Richtung des nächsten Maximums zeigt. Ziel bei diesem Verfahren ist es, in der Regel ein Minimum (den minimalen Fehler) zu finden, daher verschiebt man den Parametertupel einfach in die entgegengesetzte Richtung (Lückehe, 2015). Das Problem ist aber auch bei diesem Ansatz, dass der Algorithmus nur das nächstgelegene, nicht notwendigerweise das absolute Minimum findet (Dubey et al., 2022). Daher wird auch bei diesem Verfahren mehrmals von unterschiedlichen Parametertupeln aus gestartet. Außerdem muss sich der optimierende Sachverhalt eigentlich als differenzierbare Funktion darstellen lassen, damit das Verfahren anwendbar ist. In der Praxis, insbesondere in der Anwendung in neuronalen Netzen, werden aber auch oftmals Netzwerke mit fast überall differenzierbaren Aktivierungsfunktionen, wie die Rectified Linear Unit Funktion (ReLU), mit dem Gradientenabstiegsverfahren optimiert (Zou et al., 2020).

Da bei der vorliegenden Studie auf einem kleinen Parameterraum gearbeitet wurde, wurde das Grid-Search Verfahren angewendet, das im Fall der Auswertung aller möglicher Parameterkombinationen immer zu einem optimalen Ergebnis kommt. Geprüft werden insbesondere folgende Parameter und Parameterwerte, wobei für die *Identifizierungsmerkmale 1* und *2* nicht nur die fest erbrachte Zahl an Prüfungsleistungen oder Leistungspunkten einzelner Studierender herangezogen werden, sondern beispielsweise auch ein Vergleich mit dem Wert des Medians über alle Merkmalsausprägungen erfolgen soll. Die Abkürzungen (fix, med, stdo, sum_fix, sum_med, sum_stdo) verweisen jeweils auf die entsprechende Ausgestaltung der Identifizierungsmerkmale (id1 bis id4):

Für das Identifizierungsmerkmal 1:

id1_fix:	weniger als 1 bis 10 Prüfungsleistungen im ersten Fachsemester in Einerschritten
id1_med:	weniger als 1/10 bis 10/10 des Medians der von Absolventinnen und Absolventen abgelegten Prüfungsleistungen im ersten Fachsemester in Schritten von 1/10

Für das Identifizierungsmerkmal 2:

id2_(2fs)_fix:	weniger als 1/30 bis 30/30 von 60 Leistungspunkte in zwei Fachsemestern erbracht
id2_2fs_med:	weniger als 1/30 bis 30/30 des Medians der erbrachten Leistungspunkte in zwei Fachsemestern in Schritten von 1/30
id2_2fs_stdo:	weniger als 1/30 bis 30/30 der gemäß Studienordnung zu erbringenden Leistungspunkte in zwei Fachsemestern in Schritten von 1/30
id2_sum_fix:	weniger als 1/30 bis 30/30 * Anzahl Fachsemester * 30 Leistungspunkte in allen bisherigen Fachsemestern in Schritten von 1/30

id2_sum_med: weniger als 1/30 bis 30/30 der Summe des Medians der erbrachten Leistungspunkte von Absolventen und Absolventinnen in den bisherigen Fachsemestern in Schritten von 1/30

id2_sum_stdo: weniger als 1/30 bis 30/30 der Summe der gemäß Studienordnung in allen bisherigen Fachsemestern zu erbringenden Leistungspunkte in Schritten von 1/30

Für das Identifizierungsmerkmal 4:

id4: 1 bis 9 Rücktritte von Prüfungsleistungen in einem Fachsemester

Beim Training bleibt das Identifizierungsmerkmal 5 in der Bestimmung des Fehlers unberücksichtigt. Die Studierenden, die die Regelstudienzeit um mehr als zwei Fachsemester überschreiten, werden mit diesem Identifizierungsmerkmal immer vor Ende des Studiums identifiziert, jedoch eben erst zwei Fachsemester nach Überschreiten der Regelstudienzeit und somit erst sehr spät im Studienverlauf. Eine Identifizierung von potentiellen Studienabbrechern und Studienabbrecherinnen ist jedoch möglichst früh im Studienverlauf wünschenswert, damit das Zusenden passender Beratungs- und Unterstützungsangebote tatsächlich noch einen Interventionszeitraum ermöglichen kann. Würde man den Optimierungsalgorithmus unter Berücksichtigung dieses Merkmales verwenden, würden die Studierenden, die die Regelstudienzeit um mehr als zwei Fachsemester überschreiten, schlimmstenfalls tatsächlich erst zu diesem Zeitpunkt entdeckt. Da aber die Studierenden dennoch auch mit diesem Merkmal identifiziert werden, wird dieses für den Vergleich der Ergebnisse auf dem Validierungsset berücksichtigt.

Zur Evaluation der berechneten Modelle werden die im Machine-Learning üblicherweise verwendeten und weiter oben bereits genannten Maße Accuracy, Recall, Precision, F1 and Specifity (u.a. Berens, 2021; Powers, 2010) über alle Modelle angegeben und verglichen. Die Zielgröße definiert, ob Studierende das Studium innerhalb der Regelstudienzeit plus zwei Fachsemester abschließen (RSZ+2) oder nicht und geht damit – aufgrund der oben bereits genannten und im Programm genutzten Definition des Studienerfolgs – über die Berücksichtigung des reinen Studienabbruchs hinaus. Somit ergeben sich folgende in Tabelle 1 aufgeführten Szenarien der Prognose:

Tabelle 1: Szenarien der Prognose von Frühwarnsystemen sowie deren Bewertung des Eintretens

Prognose der Studierenden	Realität der Studierenden	Bewertung der Prognose
Brechen das Studium ab oder absolvieren erst nach RSZ+2	Brechen das Studium ab oder absolvieren erst nach RSZ+2	Richtig positiv (r+)
Absolvieren das Studium in RSZ+2	Brechen das Studium ab oder absolvieren erst nach RSZ+2	Falsch negativ (f-)
Brechen das Studium ab oder absolvieren erst nach RSZ+2	Absolvieren das Studium in RSZ+2	Falsch positiv (f+)
Absolvieren das Studium in RSZ+2	Absolvieren das Studium in RSZ+2	Richtig negativ (r-)

Anmerkung: RSZ+2 = Regelstudienzeit plus zwei Fachsemester

Die Accuracy gibt den Anteil aller korrekt vorhergesagten problematischen bzw. unproblematischen Studienverläufe unter allen betrachteten Studienverläufen an. Wie aus der programminternen Definition des Studienerfolgs (Schulze-Stocker et al., 2020, S. 190) geschlossen werden kann, gilt ein Studienverlauf als problematisch, wenn er zu einem Abbruch oder zu einer Studienzeitüberschreitung von mehr als zwei Semestern führt. Als unproblematisch

gelten alle Studienverläufe, bei denen das Studium innerhalb der Regelstudienzeit und zusätzlichen zwei Fachsemestern abgeschlossen wird.

Die Accuracy ergibt sich aus der Summe aller Richtig-Positiven und Richtig-Negativen geteilt durch alle einbezogenen Einheiten: $(r_{+}+r_{-})/(r_{+}+f_{-}+f_{+}+r_{-})$. Ein hoher Accuracy-Wert bedeutet, dass insgesamt viele Studienverläufe korrekt vorhergesagt werden.

Die Precision gibt den Anteil der Richtig-Positiven unter allen Positiven an, in diesem Fall also den Anteil derjenigen, die korrekt mit einem problematischen Studienverlauf identifiziert worden sind, unter allen, für die einen problematischen Studienverlauf vorhergesagt wurde: $r_{+}/(r_{+}+f_{+})$. Für die Identifizierung von Personen mit problematischen Studienverläufen ist diese Zahl zwar weniger von Bedeutung, allerdings muss auch darauf geachtet werden, dass Personen mit normalen Studienverläufen nicht zu oft fälschlicherweise angeschrieben werden. Daher wird auch diese Zahl jeweils mit geprüft.

Der Recall gibt den Anteil der Richtig-Positiven unter allen Positiven an, also den Anteil der Personen mit richtig prognostiziertem problematischem Studienverlauf, unter allen, die tatsächlich einen problematischen Studienverlauf aufweisen: $r_{+}/(r_{+}+f_{-})$. Sie beschreibt somit die Trefferquote, mit der ein problematischer Studienverlauf auch als solcher erkannt wird.

F1 ist das harmonische Mittel aus Precision und Recall. Dieser Wert wird nur dann groß, wenn sowohl die Genauigkeit als auch die Trefferquote groß werden und zeigt somit an, ob beide Maße hinreichend berücksichtigt werden.

Tabelle 2: Übersicht der Variation der Identifizierungsmerkmale je untersuchtem Modell im Vergleich zum Ursprungsmodell

Modell-nummer	Identifizierungs-merkmal 1	Identifizierungs-merkmal 2	Identifizierungs-merkmal 3	Identifizierungs-merkmal 4	Identifizierungs-merkmal 5
Ursprung	id1_fix	id2 _fix	id3	id4	id5
M1	id1_fix	id2_sum_fix	id3	id4	id5
M2	id1_fix	id2_sum_stdo	id3	id4	id5
M3	id1_fix	id2_sum_med	id3	id4	id5
M4	id1_fix	id2_2fs_fix	id3	id4	id5
M5	id1_fix	id2_2fs_stdo	id3	id4	id5
M6	id1_fix	id2_2fs_med	id3	id4	id5
M7	id1_med	id2_sum_fix	id3	id4	id5
M8	id1_med	id2_sum_stdo	id3	id4	id5
M9	id1_med	id2_sum_med	id3	id4	id5
M10	id1_med	id2_2fs_fix	id3	id4	id5
M11	id1_med	id2_2fs_stdo	id3	id4	id5
M12	id1_med	id2_2fs_med	id3	id4	id5

Die Specifity ist der Anteil der richtig Negativen unter allen Negativen: $r_{-}/(r_{-}+f_{+})$. Sie gibt somit den Anteil der Personen an, die richtigerweise nicht mit einem problematischen Studienverlauf identifiziert werden, unter allen, die tatsächlich keinen problematischen Studienverlauf aufweisen. Dieser Wert ist wichtig, da man die Studierenden, deren Studienverlauf unauffällig ist, nicht unnötig alarmieren möchte, steht aber in seiner Bedeutung etwas hinter den anderen Maßen zurück.

Für die Untersuchung werden insgesamt zwölf Modelle überprüft, bei denen alle möglichen Kombinationen von *Identifizierungsmerkmal 1* und *2* eingearbeitet werden (d.h. 2 * 6

Modelle). Die Modelle sind in Tabelle 2 dargestellt und berücksichtigen jeweils sowohl studiengangabhängige (summative) als auch fachsemesterabhängige Identifizierungsmerkmale:

5 Ergebnisse

In Tabelle 3 sind die Ergebnisse der Optimierung dargestellt. Spalte 1 enthält die Modellnummer, deren Spezifika aus Tabelle 2 abgelesen werden können. Die Spalten 2 – 6 von Tabelle 3 enthalten die Gütemaße der Optimierung.

Tabelle 3: Ergebnisse der Optimierung – Übersicht der Werte der fünf Gütemaße je Modell im Vergleich zum Ursprungsmodell

Modellnummer	Accuracy	Precision	Recall	F1	Specifity
Ursprung	0,72	0,79	0,71	0,75	0,75
M1	0,80	0,84	0,80	0,82	0,79
M2	0,79	0,83	0,81	0,82	0,77
M3	0,79	0,84	0,80	0,82	0,78
M4	0,79	0,84	0,80	0,82	0,78
M5	0,80	0,84	0,80	0,82	0,79
M6	0,80	0,84	0,80	0,82	0,79
M7	0,78	0,85	0,76	0,80	0,81
M8	0,78	0,82	0,80	0,81	0,76
M9	0,78	0,83	0,78	0,81	0,74
M10	0,78	0,85	0,76	0,80	0,82
M11	0,78	0,85	0,76	0,80	0,82
M12	0,79	0,85	0,76	0,81	0,82

Zunächst wird die Accuracy der Modelle auf den Testdatensets betrachtet. Es zeigt sich, dass das Ursprungsmodell eine Accuracy von 0,72 hat. Dies bedeutet, dass bereits 72 Prozent aller Personen richtig zugeordnet werden. Die optimierten Modelle haben eine Accuracy zwischen 0,78 – 0,80. Es werden in den optimierten Modellen somit sechs bis acht Prozentpunkte mehr Studierende korrekt vorhergesagt als bei dem rein auf Annahmen basierenden Ursprungsmodell. Zwischen den Modellen gibt es hingegen deutlich geringere Unterschiede: Die höchsten Werte der Accuracy sind mit jeweils einem Wert von 0,80 für die Modelle M1, M5 und M6 feststellbar. Geringere Verbesserungen werden mit den Modellen M7 – M11 mit einer Accuracy von 0,78 erreicht.

Auch hinsichtlich der anderen Gütemaße funktionieren die optimierten Modelle durchweg besser als das Ursprungsmodell. Die Precision liegt je nach Modell zwischen 0,82 – 0,85 (Ursprungsmodell: 0,79), der Recall zwischen 0,76 – 0,81 (Ursprungsmodell: 0,71), der F1-Wert zwischen 0,80 – 0,82 (Ursprungsmodell: 0,75) und die Specifity zwischen 0,74 – 0,82 (Ursprungswert: 0,75). Hinsichtlich der Precision gibt es zwischen den optimierten Modellen wieder nur geringe Abweichungen. Die Abweichungen liegen im Bereich von 0,03. Am besten passen die Modelle M7, M10 bis M12 mit einem Wert von 0,85. Bei ihnen wird der größte Anteil der Identifizierenden auch richtigerweise identifiziert.

Allerdings wirken die letztgenannten Modelle deutlich schlechter in einem für die Identifizierung von Studierenden mit problematischem Studienverlauf deutlich relevanteren Güteparameter, dem Recall. Dieser gibt an, welcher Anteil der zu Identifizierenden auch identifiziert worden ist. Das Modell M2 schneidet mit 0,81 am besten ab. Im F1-Wert wirken die Modelle M1 bis M6 mit jeweils 0,82 am besten. Alle anderen Modelle liegen mit 0,80 – 0,81 nur etwas darunter. Beim letzten dargestellten Gütemaß, der Specifity, erreichen die Modelle M10 bis M12 mit 0,82 die höchsten Werte. Das Modell M2 mit 0,77, das Modell M8 mit 0,76 und das Modell M9 mit 0,74 wirken am schlechtesten. Die Modelle M5 und M6, die hinsichtlich Accuracy und F1 die besten Gütemaße aufweisen, liegen auch hinsichtlich der anderen Gütemaße im Durchschnitt.

Tabelle 4: Ergebnisse der Optimierung – Übersicht der Durchschnittswerte der optimierten Parameter je Modell im Vergleich zum Ursprungsmodell

Modellnummer	Parameter ID1	Parameter ID2	Parameter ID4
Ursprung	id1_fix = 2	id2_fix = 0,5	id4 = 3
M1	Ø id1_fix = 6	Ø id2_sum_fix = 0,1	Ø id4 = 7
M2	Ø id1_fix = 6	Ø id2_sum_stdo = 0,5	Ø id4 = 7
M3	Ø id1_fix = 5	Ø id2_sum_med = 0,6	Ø id4 = 7
M4	Ø id1_fix = 6	Ø id2_2fs_fix = 0,3	Ø id4 = 7
M5	Ø id1_fix = 6	Ø id2_2fs_stdo = 0,2	Ø id4 = 7
M6	Ø id1_fix = 6	Ø id2_2fs_med = 0,2	Ø id4 = 7
M7	Ø id1_med = 0,9	Ø id2_sum_fix = 0,1	Ø id4 = 8
M8	Ø id1_med = 0,7	Ø id2_sum_stdo = 0,6	Ø id4 = 7
M9	Ø id1_med = 0,6	Ø id2_sum_med = 0,7	Ø id4 = 7
M10	Ø id1_med = 0,9	Ø id2_2fs_fix = 0,3	Ø id4 = 7
M11	Ø id1_med = 0,9	Ø id2_2fs_stdo = 0,3	Ø id4 = 7
M12	Ø id1_med = 0,8	Ø id2_2fs_med = 0,4	Ø id4 = 7

Interessant ist zudem eine Betrachtung der Parameter, die in den Modellen berechnet wurden. Diese sind in Tabelle 3 als Durchschnittswerte der berechneten Parameter über alle Modelle dargestellt. Bei der Interpretation der Parameter ist zu berücksichtigen, dass die Höhe des Parameters zu den Identifizierungsmerkmalen keine direkte Aussage über deren Einfluss im Modell zulässt und auch ein Vergleich der Höhe der Parameter insbesondere der *Identifizierungsmerkmale 1* (ID1) und *2* (ID2) schwierig ist, da dort teilweise vollkommen unterschiedliche Identifizierungsmerkmale angesetzt wurden, die nur im Kontext der Gesamtbetrachtung des Identifizierungsmerkmals verglichen werden können. Zum Beispiel bedeutet, dass der Parameter ID1 bei M1 einen Wert von 6 annimmt, nicht, dass dieser Parameter mehr Einfluss im Gesamtmodell hat als bei M7 (0,9). Es werden vollkommen unterschiedliche Bezüge verwendet, so sind die Parameter ID1 der Modelle M1 bis M6 Absolutwerte, die sich auf studiengangübergreifende Grenzen von bestandenen Abschlussprüfungen im ersten Fachsemester beziehen, während die Parameter ID1 der Modelle M7 bis M12 Relativwerte sind, die sich auf den Anteil von bestandenen Prüfungsleistungen im ersten Fachsemester an den durchschnittlich bestandenen Prüfungsleistungen von Absolventinnen und Absolventen beziehen. Für das Ziehen von Vergleichen ist es also wichtig, nicht nur die Parameter selbst, sondern auch den gesamten Kontext des entsprechenden Identifizierungsmerkmals zu betrachten.

Im Vergleich zum Ursprungsmodell ergeben sich im optimierten Modell 1 (M1) folgende Änderungen der Parameter der Identifikationsmerkmale:

- Identifikationsmerkmal 1: Die Schwelle für „weniger als X Prüfungsleistungen im ersten Fachsemester“ wurde von X = 2 auf im Mittel X = 6 Prüfungsleistungen angehoben.
- Identifikationsmerkmal 2: Anstelle der bisherigen Prüfung, ob Studierende in zwei Fachsemestern weniger als 30 Leistungspunkte erreicht haben, wird nun ein summatives Merkmal verwendet. Dieses prüft, ob Studierende weniger als 10 % der vorgesehenen Leistungspunkte des Fachsemesters erlangt haben.
- Identifikationsmerkmal 4: Die Anzahl der zulässigen Rücktritte wurde von ursprünglich weniger als 2 auf im Mittel weniger als 7 deutlich erhöht.

Würde das Ursprungsmodell so verändert, ergeben sich die Verbesserung der Gütekriterien des Gesamtmodells, die in Tabelle 3 dargestellt sind.

Zunächst zeigt sich bei Betrachtung von Tabelle 4, dass bei allen Modellen der Wert der *Identifizierungsmerkmale* ID1, die sich auf die im ersten Semester erworbenen Prüfungsleistungen beziehen, deutlich zugenommen hat. In den Modellen M1 bis M6 liegt die Grenze jeweils bei ca. sechs abgeschlossenen Prüfungsleistungen im ersten Fachsemester. Die Absolventinnen und Absolventen der betrachteten Studiengänge haben außerdem im Median im ersten Fachsemester fünf bis sechs Prüfungsleistungen erbracht. Daher liegen auch die Werte der Modelle M7 bis M12 mit minimal 60 Prozent * 5 Prüfungsleistungen = 3 oberhalb des Ursprungsmodells, das 0 oder eine Prüfungsleistung als Prüfkriterium verwendet hat.

Die Interpretation der Entwicklung des *Identifizierungsmerkmals 2* zu den erworbenen Leistungspunkten ist deutlich schwieriger, da sehr unterschiedliche Modelle zum Einsatz gekommen sind. Es zeigt sich zunächst, dass bei den Modellen, die mit einer Leistungspunktgrenze dergestalt arbeiten, dass diese angibt, inwiefern weniger als Parameter * 60 Leistungspunkte in zwei Fachsemestern erbracht worden sind, diese auch hinsichtlich der Parameter deutlich niedriger als im Ursprungsmodell ausfallen. Deutlich niedriger ist der Parameter nur in den Modellen M1 und M7, bei welchen überprüft wird, ob Parameter * Anzahl der Fachsemester * 30 Leistungspunkte erbracht worden sind.

Bemerkenswert ist auch die Betrachtung des Identifizierungsmerkmals 4. Dieses liegt bei allen betrachteten Modellen oberhalb eines Wertes von 7, mehr als doppelt so hoch wie im Ursprungsmodell. Es werden also nur Personen identifiziert, die von mehr als sieben Prüfungsleistungen innerhalb eines Fachsemesters zurückgetreten sind. Dies tritt gerade mal bei weniger als fünf der insgesamt betrachteten 1.992 Studierenden auf.

6 Diskussion

Frühwarnsysteme eignen sich zur Studienabbruchsvorhersage (Berens, 2021; Konrad & Wildner, 2020; Schulze-Stocker et al., 2017). Die vorliegende Studie analysiert die Genauigkeit der fünf exogen bestimmten Identifizierungsmerkmale (Berens, 2021), die im Rahmen des PASST?!-Programms der Technischen Universität Dresden ein Gesamtmodell bilden, und betrachtet mögliche Optimierungspotentiale dieses Modells. Im Ergebnis zeigt sich, dass ohne Optimierung bereits etwas mehr als 70 Prozent der Studierenden (Accuracy) richtig identifiziert werden. Für die Hochschulpraxis ist diese Form der Frühwarnsysteme bezie-

hungsweise deren Modelle nach Klärung der datenschutzrechtlichen Möglichkeiten leicht zu konstruieren beziehungsweise zu übernehmen und umzusetzen (siehe auch Konrad & Wildner, 2020).

Durch den Einsatz eines sehr einfachen Optimierungsverfahrens für spezifische Parameter auf Basis administrativer Studienverlaufsdaten konnte die Passgenauigkeit der Identifizierung insgesamt verbessert werden. Der Vorteil des geringen Aufwands von exogen im Vergleich zu endogen entwickelten Systemen (Berens, 2021, S. 146) wird dadurch nicht hinfällig. Betreffend der Accuracy sind Steigerungen um bis zu acht Prozentpunkte (auf 80 %) möglich. Durch Variation der Identifizierungsmerkmale kann zudem erreicht werden, dass mehr Studierende, die abbruchsgefährdet sind, notwendige Informationen erhalten, als auch, dass weniger Studierende, die nicht erfolgsgefährdet sind, unnötig vom Frühwarnsystem kontaktiert werden. Die Untersuchung bestätigt somit Hypothese 1. Durch die aus der Praxis entwickelten Merkmale ist es möglich, ein optimiertes Modell zur Identifizierung einzusetzen, ohne die Vorteile der Regelbasiertheit zu verlieren. Es bleibt möglich, Studierenden präzise Gründe für eine Kontaktaufnahme (beispielsweise auch Mehrfachidentifizierungen) zu nennen und beim Anschreiben auch gleich passende Angebote zu unterbreiten.

Die Angebotsvorschläge könnten sich betreffend ihrer Passung beispielsweise so gestalten, dass den Teilnehmenden, die vor einer zweiten Wiederholungsprüfung stehen, bereits im Anschreiben Hinweise zu einem Prüfungscoaching gegeben oder Studierende mit langen Studiendauern gezielt zu einer sogenannten „Endspurtberatung“ (Technische Universität Dresden, 2019) geleitet werden. Vorteil von diesem Vorgehen ist, dass sich Studierende bereits selbstbestimmt über mögliche weitere Unterstützungsschritte informieren können und nicht zwingend den Weg über eine Studienberatung gehen müssen. Da es sich beim Studienabbruch allerdings, wie bereits dargestellt (siehe auch Blüthmann et al., 2008; Heublein & Wolter, 2011; Pelz et al., 2020), um einen komplexen, multifaktoriellen Prozess handelt, werden allen Studierenden auch die Kontaktmöglichkeiten zur Studienberatung aufgezeigt. Werden die Angebote als wenig passend empfunden, können bei einer individuellen Beratung personzentriert Lösungen für mögliche Problemstellungen gefunden werden.

Hervorzuheben ist, dass (entgegen der Annahme aus Hypothese 2) in der vorliegenden Untersuchung die Modelle mit studiengang- und fachsemesterabhängigen Identifizierungsmerkmalen etwa gleich gute Ergebnisse liefern wie Modelle mit studiengang- und fachsemesterunabhängigen Merkmalen. Studiengang- und fachsemesterunabhängige Merkmale lassen sich leichter an Studierende kommunizieren und müssen bei Veränderungen in einzelnen Studiengängen nicht permanent angepasst werden. Ersterer Punkt kann dazu beitragen, dass Studierende eher bereit sind, sich an einem Studienerfolgsprogramm zu beteiligen, während letzterer die Umsetzung eines Frühwarnsystems in der Hochschulpraxis deutlich vereinfacht.

Auffällig in den Ergebnissen der Modellvergleiche sind die geringen Unterschiede in den Gütekriterien. Dies könnte damit in Verbindung stehen, dass die Modelle stets alle Merkmale in einer bestimmten Form enthalten. Bei zukünftigen Untersuchungen könnten die Vorhersagegüte der Einzelmerkmale genauer betrachtet werden und zudem geprüft werden, ob sich die identifizierten Studierendengruppen nach bestimmten Eigenschaften unterscheiden.

Obwohl diese Studie erste vielversprechende Ergebnisse liefert, ist ihre Übertragbarkeit durch die Beschränkung auf fünf Studiengänge und ein spezifisches Set an getesteten Parametern limitiert. Weitere Studien mit einer größeren Menge an Studiengängen sind notwendig, um noch genauere Ergebnisse zu liefern und die Generalisierbarkeit zu erhöhen. Aller-

dings wird auch deutlich, dass in der Praxis gerade an kleinen Hochschulen oder in kleinen Studiengängen mit häufigen Veränderungen in den Studien- und Prüfungsordnungen die Optimierung erschwert sein kann, da keine relevanten Fallzahlen für Trainings- und Testdatensätze erreicht werden. Für die Diskussion der Generalisierbarkeit ist an dieser Stelle hervorzuheben, dass die durchschnittliche Anzahl der von Absolventen und Absolventinnen eines Studiengangs im ersten Fachsemester abgelegten Prüfungsleistungen in allen fünf betrachteten Studiengängen sehr homogen (6 bis 7) ist. Nimmt man mehr Studiengänge hinzu, kann es hier aber zu einer größeren Variabilität kommen. Wenn beispielsweise die Anzahl der in den Studiengängen im ersten Fachsemester abzulegenden Prüfungsleistungen sehr unterschiedlich ist, dann ist auch anzunehmen, dass die Gütemaße für Modelle, mit für allen Studiengängen einheitlichen Grenzen für die abgelegten Prüfungsleistungen eine niedrigere Qualität anzeigen als für Modelle mit studiengangabhängigen Grenzen, da jeweils entweder die Studierenden von Studiengängen mit niedriger durchschnittlicher Prüfungszahl zu häufig identifiziert oder die Studierenden von Studiengängen mit niedriger abzulegender Prüfungszahl zu selten identifiziert werden. In den hier betrachteten Modellen gilt abweichend zu dieser Erwartung zumindest für die Accuracy das Gegenteil. Diese ist jeweils in den Modellen mit studiengangunabhängiger Grenze etwas besser als in den Modellen mit studiengangabhängiger Grenze. Allerdings bleibt die Frage offen, ob Ergebnisse, die an einer Teilmenge an Studiengängen optimiert wurden, sich auch auf andere Studiengänge anwenden lassen. Auch Überprüfungen mit unterschiedlichen Kohorten, also inwiefern mit den optimierten Daten einer bestimmten Menge an Jahrgängen auf den Erfolg von einer Menge von anderen Jahrgängen geschlossen werden kann, sind wichtig zu analysieren, da es auch Studienverlaufsaspekte geben kann, die sich mit den Kohorten ändern und somit die Qualität der Studienabbruchsvorhersage beeinflussen können.

Im Sinne des Datenschutzes sowie eines möglichst geringen Aufwands bei der Datenbereinigung und -aufarbeitung ist die Datensparsamkeit, also die – möglichst genaue – Vorhersage des Studienabbruchs der Studierenden unter Nutzung möglichst weniger Studienverlaufsdaten, ein wichtiges Qualitätsmerkmals der Studienabbruchsvorhersage. Allein zur Auswertung des Identifizierungsmerkmals 4 werden Daten zu Prüfungsrücktritten erhoben. Bei der Optimierung des Merkmals 4 wurde in den Modellen die Grenze der Prüfungsrücktritte so weit angehoben, dass nur noch sehr wenige Studierende durch das Merkmal identifiziert werden. Durch das Weglassen des Identifizierungsmerkmals 4 wird die Qualität der Vorhersage daher kaum beeinflusst. Für eine Verallgemeinerung dieses Ansatzes ist zu beachten, dass in den Daten der Technischen Universität Dresden für die betrachteten Studiengänge nur sehr späte Abmeldungen von Prüfungsleistungen, die nur wenige Tage vor der Prüfung mit Attest erfolgen, als Rücktritt erfasst werden. Interessant wäre es, anhand von Datensätzen von anderen Hochschulen zu überprüfen, ob sich die gefundenen Ergebnisse bestätigen lassen, wenn auch frühere Prüfungsabmeldungen in die Analyse einbezogen werden.

Literatur

Ahles, L., Köstler, U., Vetter, N. & Wulff, A. (2016). *Studienabbrüche an deutschen Hochschulen. Stand der Thematisierung und strategische Ansatzpunkte* (*Studien zum sozialen Dasein der Person, Bd. 20*). Nomos Verlagsgesellschaft. https:/doi.org/10.5771/9783845276816.Anastasio, S., Holthusen, L.,

Konrad, N., Lietz, S., Magnum, C., Wendler, G., Wildner, F. & Kiepenheuer-Drechsler, B. (2020). *Studienabbrecher/innen als Zielgruppe der Beratung und* Öffentlichkeitsarbeit. Beiträge aus dem Projekt „Queraufstieg Berlin“. Forschungsinstitut Betriebliche Bildung (f-bb) gGmbH. https://doi.org/10.3278/6004808w.

Berens, J. & Schneider, K. (2019). Drohender Studienabbruch: Wie gut sind Frühwarnsysteme. *Qualität in der Wissenschaft (QIW), 13*, 102–08.

Berens, J., Schneider, K., Görtz, S., Oster, S. & Burghoff, J. (2019). Early detection of students at risk: Predicting student dropouts using administrative student data from German universities and machine learning methods. *Journal of Educational Data Mining*, *11*, 1–41. https://doi.org/10.5281/zenodo.3594771.

Berens, J. (2021*). Maschinelle Früherkennung abbruchgefährdeter Studierender: Konzeption, Systemvergleich und Evaluation* (Dissertation). Bergische Universität Wuppertal. https://doi.org/10.25926/7jv7-2s67.

Blüthmann, I., Lepa, S. & Thiel, F. (2008). Studienabbruch und -wechsel in den neuen Bachelorstudiengängen. Untersuchung und Analyse von Abbruchgründen. *Zeitschrift für Erziehungswissenschaft, 11*, 406−429. https://doi.org/10.1007/s11618-008-0038-y.

D’Angelo, D. (2019). *Zeichenhorizonte: Semiotische Strukturen in Husserls Phänomenologie der Wahrnehmung*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-17468-2.

Daniel, A., Neugebauer, M. & Watermann, R. (2019). Studienabbruch und Einstellungschancen auf dem Ausbildungsmarkt: Ein faktorieller Survey mit Arbeitgeber/innen. *Zeitschrift für Erziehungswissenschaft*, *22*, 1147–1174. https://doi.org/10.1007/s11618-019-00905-0.

Dubey, S. R., Basha, S. S., Singh, S. K. & Chaudhuri, B. B. (2022). AdaInject: Injection based adaptive gradient descent optimizers for convolutional neural networks. *IEEE Transactions on Artificial Intelligence*. https://doi.org/10.1109/TAI.2022.3208223.

Dudenredaktion (o. J.). Optimum. In *Duden Online*. Abgerufen am 07. Mai 2023, von https://www.duden.de/rechtschreibung/Optimum.

Hambrinker, P. (2021). Studienabbruch? Was Studierende tun können, wenn es im Studium nicht „rund“ läuft. *API Magazin*, *2*. https://doi.org/10.15460/apimagazin.2021.2.1.53.

Hauss, K. & Seyfried, M. (2019). Hochschulen im Befragungsdilemma–Gedankenexperimente und organisationale Lösungen. *Zeitschrift für Hochschulentwicklung*, *14*, 9–27.

Heublein, U. & Wolter, A. (2011). Studienabbruch in Deutschland. Definition, Häufigkeit, Ursachen, Maßnahmen. *Zeitschrift für Pädagogik*, *57*, 214–236. https://doi.org/10.25656/01:8716.

Heublein, U., Ebert, J., Hutzsch, C., Isleib, S., König, R., Richter, J. & Woisch, A. (2017). Zwischen Studienerwartungen und Studienwirklichkeit. Ursachen des Studienabbruchs, beruflicher Verbleib der Studienabbrecherinnen und Studienabbrecher und Entwicklung der Studienabbruchquote an deutschen Hochschulen. *Forum Hochschule,* 1/2017. Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH.

Isphording, I. E. & Wozny, F. (2018). Ursachen des Studienabbruchs. Eine Analyse des Nationalen Bildungspanels. *IZA research report*, *82*. Abgerufen am 27. Juni 2023, von https://bit.ly/3f6Qj19.

Konrad, N. & Wildner, F. (2020) Studienzweifel frühzeitig erkennen und helfen: Frühwarnsysteme an Hochschulen. In Anastasio, S., Holthusen, L., Konrad, N., Lietz, S., Mangum, C., Wendler, G., Wildner, F. & Kiepenheuer-Drechsler, B.). *Studienabbrecher/innen als Zielgruppe der Beratung und Öffentlichkeitsarbeit. Beiträge aus dem Projekt „Queraufstieg Berlin“*. wbv, S. 50–69. (f-bb-online; 03/20). https://doi.org/10.25656/01:30402.

Kemper, L., Vorhoff, G. & Wigger, B. U. (2020). Predicting student dropout: A machine learning approach. *European Journal of Higher Education*, *10*, 28–47. https://doi.org/10.1080/21568235.2020.1718520.

Klein, D., Mishra, S. & Müller, L. (2021). Die langfristigen individuellen Konsequenzen des Studienabbruchs. In M. Neugebauer, H.-D. Daniel & A. Wolter (Hrsg.), *Studienerfolg und Studienabbruch* (S. 279–300). Springer VS. https://doi.org/10.1007/978-3-658-32892-4_12.

Liashchynskyi, P. & Liashchynskyi, P. (2019*). Grid search, random search, genetic algorithm: a big comparison for NAS*. https://doi.org/10.48550/*ARXIV.1912.06059.*

Lückehe, D. (2015). *Hybride Optimierung für Dimensionsreduktion: Unüberwachte Regression mit Gradientenabstieg und evolutionären Algorithmen*. *BestMasters.* Springer-Verlag. https://doi.org/10.1007/978-3-658-10738-3.

Marsland, S. (2014). *Machine learning: An algorithmic perspective* (2nd ed.). CRC Press (Taylor & Francis Group). https://doi.org/10.1201/b17476.

Mohd Talib, N. I., Abd Majid, N. A. & Sahran, S. (2023). Identification of Student Behavioral Patterns in Higher Education Using K-Means Clustering and Support Vector Machine. *Applied Sciences, 13, 3267.* https://doi.org/10.3390/app13053267.

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective.* The MIT Press.

Neugebauer, M., Heublein, U. & Hannover, B. (2019). Editorial „Studienabbruch*". Zeitschrift für Erziehungswissenschaften, 22,* 1019–1023. https://doi.org/10.1007/s11618-019-00918-9.

Opitz, L. (2021). *Dokumentations- und Evaluationsbericht des laufenden Projektes „Studienerfolg im Dialog"*. Frankfurt: Goethe-Universität. Online verfügbar unter https://www.uni-frankfurt.de/104680074.pdf.

Powers, D. M. (2010). *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.* https://doi.org/10.48550/ARXIV.2010.16061.

Pelz, R., Schulze-Stocker, F. & Gaaw S. (2020). *Determinanten der Studienabbruchneigung von Studierenden. Ergebnisse quantitativer Befragungen an der TU Dresden*. In F. Schulze-Stocker, C. Schäfer-Hock & H. Greulich (Hrsg.). *Wege zum Studienerfolg. Analysen, Maßnahmen und Perspektiven an der Technischen Universität Dresden 2016-2020* (S. 53–82). TUDpress.

Russel, S. J. & Norvig, P. (2010*). Artificial intelligence: A modern approach* (3rd ed.). Prentice Hall International.

Schneider, T. (2022). *Digitalisierung und Künstliche Intelligenz: Einsatz durch und im Controlling,* Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-33860-2.

Schulze-Stocker, F., Schäfer-Hock, C. & Pelz, R. (2017). Weniger Studienabbruch durch Frühwarnsysteme – Das Beispiel des PASST?!-Programms an der TU Dresden. *Zeitschrift für Beratung und Studium, 12*, 26−32.

Schulze-Stocker, F., Gallrein, A.-M. B., Blum, C., Rockstroh, M. & Ishig, A. (2020). PASST?! Partnerschaft – Studienerfolg – TU Dresden – ein Frühwarnsystem für Studierende. In F. Schulze-Stocker, C. Schäfer-Hock & H. Greulich (Hrsg*.), Wege zum Studienerfolg. Analysen, Maßnahmen und Perspektiven an der Technischen Universität Dresden 2016-2020* (S. 189–224). TUDpress.

Schulze-Stocker, F. & Schäfer-Hock, C. (2020). Transformation von Hochschulen–Transformation von Bildungswegen junger Erwachsener: Frühwarnsysteme an Hochschulen in Deutschland. In O. Dörner, C. Iller, I. Schüßler, H. von Heiden & S. Lerch (Hrsg.), *Erwachsenenbildung und Lernen in Zeiten von Globalisierung, Transformation und Entgrenzung* (S. 237–250). Schriftenreihe der Sektion Erwachsenenbildung in der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE). https://doi.org/10.25656/01:20589.

Statistisches Bundesamt. (2023). *6,5% weniger Studienanfängerinnen und -anfänger in MINT-Fächern im Studienjahr 2021*. Abgerufen am 27. Juni 2023, von https://www.destatis.de/DE/Presse/Pressemitteilungen/2023/01/PD23_N004_213.html.

Stifterverband für die Deutsche Wissenschaft. (2007). *Studienabbruch: Staat setzt jährlich 2,2 Mrd. Euro in den Sand*. Presseportal. Abgerufen am 27. Juni 2023, von https://www.presseportal.de/pm/18931/1057835.

Stubbs, R., Wilson, K. & Rostami, S. (2020). Hyper-parameter Optimisation by Restrained Stochastic Hill Climbing. In Z. Ju, L. Yang, C. Yang, A. Gegov & D. Zhou (Hrsg.), *Advances in Computational Intelligence Systems: Contributions Presented at the 19th UK Workshop on Computational Intelligence, September 4-6, 2019, Portsmouth, UK 19* (S. 189–200). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-29933-0_16.

Syarif, I., Prugel-Bennett, A. & Wills, G. (2016). SVM parameter optimization using grid search and genetic algorithm to improve classification performance. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *14*, 1502–1509. http://doi.org/10.12928/telkomnika.v14i4.3956.

Theune, K. (2021). Determinanten und Modelle zur Prognose von Studienabbrüchen. In M. Neugebauer, H.-D. Daniel & A. Wolter (Hrsg.), *Studienerfolg und Studienabbruch* (S. 19–40). Springer VS. https://doi.org/10.1007/978-3-658-32892-4_2.

Technische Universität Dresden. (2019). Endspurtberatung der Zentralen Studienberatung. Abgerufen am 21.02.2025, von https://tu-dresden.de/tu-dresden/veranstaltungen/veranstaltungskalender/endspurtberatung-der-zentralen-studienberatung/endspurtberatung-3.

Technische Universität Dresden. (2025). Vision und strategische Ziele der TU Dresden. Abgerufen am 16.04.2025, von https://tu-dresden.de/tu-dresden/profil/vision-und-strategische-ziele.

Vandamme, J.-P., Meskens, N. & Superby, J.-F. (2007). Predicting academic performance by data mining methods. *Education Economics*, *15*, 405–419. https://doi.org/10.1080/09645290701409939.

Waltl, D. B. (2019). Erklärbarkeit und Transparenz im Machine Learning. In K. Mainzer (Hrsg.), *Philosophisches Handbuch Künstliche Intelligenz* (1–23). Springer VS. https://doi.org/10.1007/978-3-658-23715-8_31-1.

Westerholt, N., Lenz, L., Stehling, V. & Isenhardt, I. (2018). *Beratung und Mentoring im Studienverlauf. Ein Handbuch.* Waxmann Verlag. https://doi.org/10.25656/01:16563.

Zou, D., Cao, Y., Zhou, D. & Gu, Q. (2018*). Stochastic gradient descent optimizes over-parameterized deep relu networks*. https://doi.org/10.48550/ARXIV.1811.08888.

Kontakt

Robert Pelz
Technische Universität Dresden
Zentrale Studienberatung
PASST?!-Programm
01062 Dresden
E-Mail: robert.pelz@tu-dresden.de

Technische Universität Dresden
Zentrum für Qualitätsanalyse
01062 Dresden:

Dr. Franziska Schulze-Stocker
E-Mail: franziska.schulze-stocker@tu-dresden.de

Johannes Winter
E-Mail: johannes.winter1@tu-dresden.de

Wolfgang Haag
E-Mail: wolfgang.haag@tu-dresden.de