

Commentarii informaticae didacticae | 13

Artikel erschienen in:

Jörg Desel, Simone Opel, Juliane Siegeris (Hrsg.)

Hochschuldidaktik Informatik HDI 2021

9. Fachtagung des GI-Fachbereichs Informatik und Ausbildung/Didaktik der Informatik 15.–16. September 2021 in Dortmund

(Commentarii informaticae didacticae (CID) ; 13)

2023 – 299 S.

ISBN 978-3-86956-548-4

DOI <https://doi.org/10.25932/publishup-56507>

Empfohlene Zitation:

Axel Böttcher; Veronika Thurner; Tanja Häfner; Sarah Ottinger: Erkenntnisse aus der Analyse von Studienverlaufsdaten als Grundlage für die Gestaltung von Beratungsangeboten, In: Hochschuldidaktik Informatik HDI 2021, Jörg Desel, Simone Opel, Juliane Siegeris (Hrsg.), Potsdam, Universitätsverlag Potsdam, 2023, S. 137–156.

DOI <https://doi.org/10.25932/publishup-61569>

Soweit nicht anders gekennzeichnet ist dieses Werk unter einem Creative Commons Lizenzvertrag lizenziert: Namensnennung 4.0. Dies gilt nicht für zitierte Inhalte anderer Autoren:

<https://creativecommons.org/licenses/by/4.0/legalcode.de>



Erkenntnisse aus der Analyse von Studienverlaufsdaten als Grundlage für die Gestaltung von Beratungsangeboten

Axel Böttcher¹, Veronika Thurner¹, Tanja Häfner¹, Sarah Ottinger¹

Abstract: Viele Studierende stoßen im Rahmen ihres Informatikstudiums auf Probleme und benötigen individuell bedarfsgerechte Unterstützung, um beispielsweise trotz gewisser Startschwierigkeiten ihr Studium erfolgreich zu Ende zu führen. In die damit verbundene Lern- bzw. Studienberatung fließen Empfehlungen zur weiteren Studienverlaufsplanung ein. Anhand einer Datenanalyse über den Prüfungsleistungsdaten der Studierenden überprüfen wir die hinter diesen Empfehlungen liegenden Hypothesen und leiten aus den dabei gewonnenen Erkenntnissen Konsequenzen für die Beratung ab.

Insgesamt zeigt sich, dass sich nach den ersten Semestern ein mittlerer Bereich von Studierenden identifizieren lässt, bei denen Studienabbruch und Studienerfolg etwa gleich wahrscheinlich sind. Für diese Personengruppe ist Beratungsbedarf dringend gegeben. Gleichzeitig stößt die Datenanalyse auch an gewisse Grenzen, denn es zeigen sich insgesamt keine echt trennscharfen Muster, die frühzeitig im Studium eindeutig Erfolg oder Misserfolg prognostizieren. Dieses Ergebnis ist jedoch insofern erfreulich, als es bedeutet, dass jede:r Studierende:r auch nach einem suboptimalen Start ins Studium noch eine Chance auf einen Abschluss hat.

Keywords: Learning Analytics; Datenanalyse; Studienverläufe; Erfolgsmessung; Datenschutz

¹ Hochschule München, Fakultät für Informatik und Mathematik, 80335 München, axel.boettcher@hm.edu  <https://orcid.org/0000-0002-0364-6355> | veronika.thurner@hm.edu  <https://orcid.org/0000-0002-9116-390X> | tanja.haefner@hm.edu | sarah.ottinger@hm.edu

1 Einleitung

Hohe Studienabbruchquoten sind ein Phänomen in vielen MINT-Studiengängen. Im Bereich der Informatik sind deutschlandweit Abbruchquoten von um die 45 % an Universitäten und knapp 40 % an Hochschulen für angewandte Wissenschaften verbreitet [Au18] (Tab. F4-1A). Gleichzeitig ist die Nachfrage nach Absolventinnen und Absolventen der Informatik auf dem Arbeitsmarkt seit Jahren ungebrochen hoch. Diejenigen Studierenden, die ihr Informatikstudium erfolgreich zu Ende führen, reichen bei Weitem nicht aus, um diesen Bedarf zu decken. Um dem entgegenzuwirken, haben viele Hochschulen Maßnahmen ergriffen, um die Studierenden bedarfsgerecht zu unterstützen und so insgesamt mehr Studierende der Informatik zu einem erfolgreichen Abschluss zu führen.

Auch an der Fakultät für Informatik und Mathematik der Hochschule München wurden vielfältige Angebote etabliert, mit besonderem Fokus auf die Studieneingangsphase, um den Studierenden den Einstieg in den Studierprozess zu erleichtern. In diesem Rahmen wurden auch die Beratungsangebote an der Fakultät auf- und ausgebaut. Die Möglichkeiten zur fachlichen Beratung durch die Lehrenden sowie zur psychosozialen Beratung durch eine entsprechend ausgebildete Mitarbeiterin werden nun ergänzt durch eine Lernberatung für Studierende. Diese wird durchgeführt durch zwei Referentinnen für Lehren und Lernen, deren Stellen aus Studienzuschüssen geschaffen wurden. Die Referentinnen für Lehren und Lernen haben keine Lehrverpflichtung, treffen also nicht selbst in der Rolle als Lehrende auf die Studierenden. So ist sichergestellt, dass im Beratungskontext erworbenes Wissen über persönliche Details der Lernenden nicht eine objektive Beurteilung von studentischen Leistungen verfälscht.

Diese Beratungsangebote werden zunehmend stark nachgefragt, insbesondere auch von Studierenden im ersten oder zweiten Semester, deren Einstieg ins Studium eher holprig verlaufen ist, die aber motiviert sind, ihr Studium weiterzuverfolgen und erfolgreich zu Ende zu bringen. Gemeinsame Ausgangsbasis der Beratungskonstellationen ist dabei die Existenz einer Problemsituation. Wie genau diese aussieht – und welche Lösungsmöglichkeiten ggf. offen stehen – ist dagegen hochgradig individuell, da jede Person als Hintergrund ihren individuellen Studienverlauf mit Erfolgen und Misserfolgen mitbringt, sowie ihr spezifisches Kompetenzprofil und ihre individuellen Persönlichkeitseigenschaften. Dieser Heterogenität müssen die Beratungsangebote Rechnung tragen und in Einzelgesprächen individuell passende Strategien für die nächsten Schritte im

Ausbildungsprozess erarbeiten. Deren Bandbreite reicht von der Entwicklung geeigneter Strategien für ein erfolgreiches Weiterstudieren des initial gewählten Studienganges über einen Wechsel des Studienfaches bis hin zu einem frühzeitigen Studienabbruch, ggf. mit Aufnahme einer kaufmännischen oder gewerblichen Berufsausbildung als alternativem Werdegang.

Aufgrund der Vielschichtigkeit der Beratungskonstellationen ist eine individuell bedarfsgerechte Beratung eine hochkomplexe Aufgabe. Damit die Beratungen einen möglichst großen Nutzen bewirken (d. h. den betroffenen Studierenden das Finden und Wählen des individuell für sie passenden und gangbaren Weges ermöglichen) ist eine inhaltliche Systematisierung der Beratungsangebote erforderlich.

Die vorliegende Arbeit ist eine Weiterentwicklung des Beitrages [Bö21b].

2 Ziele

Übergeordnetes Ziel ist also, diejenigen Studierenden mit Unterstützungsbedarf individuell bedarfsgerecht zu beraten und dadurch die Studienerfolgsquote zu verbessern.

Dazu sind zum einen diejenigen Studierenden zu identifizieren, die tatsächlich Unterstützungsbedarf haben. Zum anderen sind typische Erfolgsmuster und Best Practices herauszuarbeiten, die Studierende anwenden können, um ihr Studium wieder auf einen erfolgreichen Weg zu bringen. Darauf abgestimmt sind Heuristiken zu entwickeln, die in der Beratung helfen, aus der Menge der erarbeiteten Erfolgsmuster und Best Practices diejenigen als Empfehlung herauszufiltern, die für eine konkrete Person und deren konkrete Problemkonstellation tatsächlich individuell passend und zielführend sind und diese zu konkreten nächsten Schritten für die Gestaltung des eigenen Lernprozesses zu präzisieren.

Die bis dato etablierten Beratungsangebote adressieren auf der Mikroebene diverse „handwerkliche“ Tipps zur Lernorganisation, wie beispielsweise Zeitmanagement, Selbstorganisation oder das Identifizieren des eigenen Lerntyps. Diese sind nun zu ergänzen um Empfehlungen auf der Makroebene, d. h. der Gestaltung des individuellen Studienverlaufs derjenigen Studierenden, die wegen nicht angetretener bzw. nicht bestandener Prüfungen aus dem Regelverlauf herausfallen bzw. sich eine Welle von „Altlasten“ aufbauen, die rein mengenmäßig nicht zusätzlich zum „normal“ vorgesehenen Semesterprogramm zu

bewältigen sind. Hier ist also individuell bedarfsgerecht zu entscheiden, welche Module bzw. Prüfungen in welcher Reihenfolge fokussiert werden sollen und was verschoben werden muss, um so das im Semester zu absolvierende Pensum in einem schaffbaren Rahmen zu halten. Diese Überlegungen müssen die existierenden verwaltungs- und prüfungsrechtlichen Regeln hinsichtlich Fristen für das erste Antreten oder das Wiederholen von Prüfungen berücksichtigen.

Ergänzend ist zu überprüfen, inwieweit häufiger auftretende Problemkonstellationen strukturell bedingt sind und wie diese ggf. durch entsprechende Änderungen in der Organisation der Studienangebote vermieden oder zumindest entschärft werden können.

3 Stand der Forschung

Die im Studierprozess anfallenden Daten werden von vielen Institutionen mit unterschiedlichen Zielrichtungen und unterschiedlichen Kontexten analysiert. Eingebürgert hat sich dafür der Begriff „Learning Analytics“. Nach Schneidig und Holmeier werden „unter Learning Analytics [...] Möglichkeiten der Datenauswertung entwickelt und diskutiert, die ein tieferes Verständnis und eine Optimierung von Lernumgebungen und -prozessen in Aussicht stellen und sowohl Dozierende und Studierende als auch hochschulische Leitungspersonen und -gremien adressieren“ [SH21].

So werden detaillierte Informationen zum Verhalten der Studierenden in einzelnen Kursen erhoben und ausgewertet. Diese Analysen sind sehr feingranular und basieren meistens auf Daten aus Learning Management Systemen. Ziel ist einerseits die Steuerung von Lehr-/Lernprozessen auf Ebene eines Moduls. Außerdem wird versucht, Leistungen sowie die Gefahr eines vorzeitigen Abbruchs der Arbeit am Modul zu prognostizieren; siehe dazu z. B. [FY15].

Andererseits werden in vielen Studien auf einer abstrakteren Ebene ganze Studienverläufe analysiert. Die Datenquellen sind typischerweise die Campusmanagementsysteme [SH21]. Ziel ist dabei, Modelle für frühzeitige Prognosen des individuellen Studienerfolgs zu erstellen.

Die Auswertung von Studierendendaten ist nicht neu. So hat Meier zu Beginn des Jahrtausends den Zusammenhang zwischen Schulnoten und dem Abschneiden von Studierenden der Rechtswissenschaften im ersten juristischen Staatsexamen empirisch untersucht [Me03]. Seit etwa einer Dekade hat das

Thema mit der Etablierung von Data Science als eigenständiger Disziplin an Fahrt aufgenommen.

Die Arbeiten von Hinkelmann et al. [HMT16], Berens et al. [Be18] sowie von Kemper et al. [KVV20] sind für die vorliegende Arbeit insofern besonders relevant, als sie mit Blick auf die Zielsetzung vergleichbare Ansätze verfolgen und sich hinsichtlich rechtlicher und organisatorischer Rahmenbedingungen in einem ähnlichen Kontext bewegen. Die Autorinnen und Autoren wenden verschiedene statistische Analyseverfahren oder Machine-Learning-Ansätze an, um anhand soziodemografischer Daten und Daten über Studienverläufe frühzeitig individuelle Wahrscheinlichkeiten für Erfolg oder Misserfolg im Studium zu erstellen. In der aktuell vorliegenden Literatur werden bislang allerdings nur die tatsächlich erfolgten Prüfungen, nicht jedoch die vorgesehenen Prüfungen in die Analysen einbezogen.

Mit diesen Mechanismen lässt sich bereits anhand der bei der Immatrikulation vorhandenen Daten ein Drop-out mit 67 % Genauigkeit vorhersagen [Be18]. Darüber hinaus zeigt sich, dass nach dem ersten Fachsemester zu über 80 % zutreffende Vorhersagen über Erfolg bzw. Misserfolg getroffen werden können [Be18; KVV20].

Als aus den Ergebnissen abzuleitende Handlungen und Maßnahmen werden die Entwicklung von Frühwarnsystemen für Drop-out-gefährdete Studierende genannt [Be18], bzw. die Etablierung von Beratungsangeboten [HMT16].

Studienabbrüche können vielfältige Ursachen haben. Diese können sich zum einen auf die betroffene Person selbst beziehen, können zum anderen aber auch strukturell bedingt sein, beispielsweise durch die Gestaltung der Studiengänge [HSW06].

4 Vorgehensweise und Rahmenbedingungen

Um uns den in Abschnitt 2 definierten Zielen anzunähern, erschließen wir uns als Datenbasis das hochschulische System zur Verwaltung von Prüfungsleistungen, in dem auch Informationen zur Hochschulzugangsberechtigung der Studierenden hinterlegt sind. Generell erfordert die Auswertung personenbezogener Daten ein Datenschutzkonzept und entsprechende datenschutzrechtliche Freigaben. Des Weiteren treffen bei der Auswertung dieser personenbezogenen Daten aus dem Prüfungsverwaltungssystem die Grundrechte auf informationel-

le Selbstbestimmung der Studierenden und die Freiheit von Wissenschaft und Forschung aufeinander.

Zur Lösung dieses Spannungsfeldes haben wir unter Mitwirkung des Datenschutzbeauftragten die in [BTH20] detailliert beschriebene Vorgehensweise etabliert. Diese führt die Rolle „Datentreuhänder:in“ ein, die von einer Person eingenommen wird, die selbst nicht in die Lehre involviert ist. Der bzw. die Treuhänder:in bereitet die personenbezogenen Daten entsprechend der datenschutzrechtlichen Vorgaben durch Pseudonymisierung auf und stellt sie erst dann für die wissenschaftliche Auswertung zur Gewinnung allgemeiner Aussagen zur Verfügung. So wird sichergestellt, dass Lehrende im Rahmen der wissenschaftlichen Auswertung dieser Daten kein Zusatzwissen über solche Studierende erhalten können, mit denen sie in der Lehre noch konfrontiert sind.

Anhand dieses Datenbestandes aus dem Prüfungsverwaltungssystem analysieren wir Studienverläufe, um Muster zu identifizieren – sowohl für erfolgreiche Verläufe als auch für typische Problemkonstellationen. Mit Blick auf die Beratung steht dabei die Frage im Mittelpunkt, ob die Studienwege erfolgreicher Studierender Gemeinsamkeiten aufweisen, aus denen sich „Good Practices“ bzw. „Dos and Don’ts“ ableiten lassen, die als Empfehlung in Beratungsprozesse einfließen können. Ein besonderer Fokus liegt dabei auf den Studienwegen derjenigen Studierenden, die in der Startphase ihres Studiums selbst Schwierigkeiten hatten, ihr Studium jedoch trotzdem zu einem erfolgreichen Abschluss geführt haben.

Als Indikatoren für Schwierigkeiten bzw. die Existenz einer Problemsituation ziehen wir Prüfungen heran, die von den Studierenden entweder nicht bestanden oder zum im Studienplan vorgesehenen Zeitpunkt nicht angetreten wurden. Entsprechend lässt sich diese Betrachtung erst nach der ersten Prüfungsperiode eines Studienverlaufes durchführen. Ergänzend überprüfen wir, inwieweit sich bereits aus der prähochschulischen Bildungshistorie der Studierenden Indikatoren für potenziell auftretende Problemsituationen ergeben, beispielsweise anhand der Noten aus der Hochschulzugangsberechtigung oder aus dem Schultyp, an dem die Hochschulzugangsberechtigung erworben wurde.

Für diese Analysen fokussieren wir in dieser Arbeit den Bachelorstudiengang Informatik an der Hochschule München. Tabelle 1 listet die Module auf, die gemäß Studienplan in den ersten drei Semestern zu absolvieren sind. Die mit den Modulen verbundenen ECTS-Punkte sind jeweils in Klammern angegeben. Jedes Semester umfasst genau ein Wahlpflichtfach, alle anderen Fächer sind Pflichtfächer. Studienbeginn ist jeweils zum Wintersemester. Die

einzelnen Module werden einmal pro Jahr angeboten, die Module des ersten und dritten Semesters im Winter, die Module des zweiten Semesters im Sommer. Prüfungen zu den Pflichtmodulen werden in der Regel jedes Semester angeboten. Wiederholungsprüfungen laufen somit gegenläufig, also in einem Semester, in dem die Lehr-Lernveranstaltungen zum Modul nicht durchgeführt werden und lediglich eine Prüfung stattfindet.

Tab. 1: Module der ersten drei Semester. Credits (ECTS) jeweils in Klammern.

1. Semester	2. Semester	3. Semester
Softwareentwicklung I (8)	Softwareentwicklung II (8)	Algorithmen und Datenstrukturen (5)
Analysis (5)	Diskrete Mathematik (5)	Datenbanken I (5)
Lineare Algebra (5)	Angewandte Mathematik (5)	Software Engineering I (5)
Technische Informatik (5)	Theoretische Informatik (5)	Netzwerke I (5)
IT-Systeme Grundlagen (5)	IT-Systeme (5)	Wahlpflichtfach Mathe (5)
Allgemeinwissenschaftliches Wahlpflichtfach (2)	Allgemeinwissenschaftliches Wahlpflichtfach (2)	Wahrscheinlichkeitsrechnung und Statistik (5)

Grundsätzlich gilt die Regelung, dass alle Module des ersten (bzw. zweiten) Semesters spätestens im zweiten (bzw. dritten) Semester erstmals angetreten werden müssen² – ansonsten wird von Amts wegen die Note 5 vergeben. Die erste Wiederholung einer mit 5 bewerteten Prüfung muss im unmittelbar folgenden Semester angetreten werden; bis zu einem ggf. erforderlichen dritten Versuch darf dann ein weiteres Jahr gewartet werden. Ein nicht bestandener Drittversuch oder mehr als fünf notwendige Drittversuche im gesamten Studienverlauf führen zur Exmatrikulation.

Auf Grundlage einer speziellen Regelung darf das Modul Angewandte Mathematik ausschließlich dann besucht werden, wenn mindestens eines der Mathematikmodule des ersten Semesters bestanden wurde. Zu beachten ist

² Die Prüfungen des ersten und zweiten Semesters sind so genannte Grundlagen- und Orientierungsprüfungen. Dazu zählen nicht die Allgemeinwissenschaftlichen Wahlpflichtfächer (AW).

dabei, dass das Modul Angewandte Mathematik ausschließlich im Sommersemester angeboten wird. In der Folge erhalten alle diejenigen Studierenden, die im ersten Semester weder Analysis noch Lineare Algebra bestanden haben, automatisch im zweiten Semester eine fristbedingte 5 für Angewandte Mathematik.

5 Erkenntnisse

In diesem Abschnitt fassen wir die wichtigsten Ergebnisse unserer Analysen zusammen und leiten daraus zentrale Erkenntnisse ab.

5.1 Viele Studierende studieren nicht ernsthaft

Die ersten Analysen der Prüfungsdaten [BTH20] machten überraschend deutlich, dass in unserem Bachelorstudiengang Informatik etwa 20 % der Immatrikulierten nicht ernsthaft studieren. Studierende, die über mehrere Semester eingeschrieben sind, ohne je eine einzige Prüfungsleistung zu erbringen, bezeichnen wir als Geister (16,6 %), wie in Abbildung 1 dargestellt. Andere Studierende sind zwar im Studiengang eingeschrieben und erbringen dabei ggf. einige wenige Leistungen, wechseln jedoch nach wenigen Semestern in einen anderen Studiengang und lassen sich dabei die bisher erbrachten Leistungen vollständig anerkennen. Dies legt die These nahe, dass diese Studierenden in der Informatik parken, bis sie in ihrem eigentlichen Wunschstudiengang einen Platz bekommen (4,3 % an Parkstudierenden).

Für die weiteren Untersuchungen bereinigen wir diese „unechten“ Studierenden (also Geister und Parker) aus dem Datensatz und beschränken uns im Folgenden auf diejenigen Immatrikulierten, die wir als ernsthaft studierend erkannt haben.

Durch diese Bereinigung sinkt rechnerisch die Abbruchquote von ursprünglich knapp 59 % (bei vollem Datenbestand, also inklusive der „unechten“ Studierenden) auf eine „echte“ Abbruchquote von rund 48 % bei den tatsächlich ernsthaft Studierenden – was immer noch ein verhältnismäßig hoher Wert ist. Dennoch ist es gerade auch auf der politischen Ebene von Bedeutung, dass mehr als 10 % der in Informatik eingeschriebenen Studierenden ihren Studienplatz nicht aktiv nutzen, obwohl sie ihn über mehrere Semester belegen. Dies ist

doppelt ungünstig, da zum einen der Bedarf nach Informatikexpertise auf dem Arbeitsmarkt sehr hoch ist und zum anderen der Studiengang aus Kapazitätsgründen zulassungsbeschränkt ist, sodass jedes Jahr nicht alle Bewerber:innen zum Zuge kommen. Insbesondere die in Informatik eingeschriebenen „Geister“ nutzen also nicht nur ihre eigenen Studienoptionen nicht, sondern verwehren gleichzeitig ggf. anderen Personen den Zugang zum Studium.

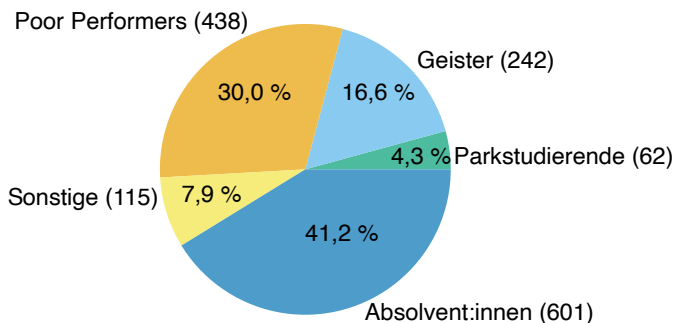


Abb. 1: Beobachtete Klassen von Studierenden aus den Immatrikulationsjahren 2006–2017 (in Summe 1458 Personen).

Als „Poor Performers“ kategorisieren wir diejenigen Studierenden, die aufgrund schlechter Leistungen ihr Studium aufgeben (müssen).

Bislang fließen in die Beratung von Studierenden in konkreten Problemsituationen auch Ratschläge ein, die aus einem Bauchgefühl der Lehrenden bzw. der Lernberatung heraus entstehen, oft basierend auf als typisch empfundenen Beobachtungen. Im Folgenden stellen wir Erkenntnisse bezüglich dieser vermuteten Einflüsse auf den Studienerfolg vor. Auf der Basis der Datenlage des Prüfungsverwaltungssystems untersuchten wir Fragestellungen, welche wir aus gängigen auf Bauchgefühl basierenden Hypothesen abgeleitet haben. Aus den dabei gewonnenen Erkenntnissen leiten wir Konsequenzen für die Beratung von Studierenden ab.

5.2 Personen bestimmter Herkunftsschultypen sind besonders gefährdet

Unser Bildungssystem ermöglicht viele unterschiedliche Eingangspfade in ein Hochschulstudium. Doch nicht jeder dieser Pfade bereitet auf jeden Studiengang gleich gut vor. Eine statistische Analyse hat ergeben, dass für unseren

Bachelorstudiengang Informatik die Studienerfolgsquoten der Studierenden unterschiedlicher Herkunftsschultypen zwischen rund 20 % (FOS Sozialer Zweig) und ca. 60 % (FOS Technischer Zweig) streuen (siehe Abbildung 2).

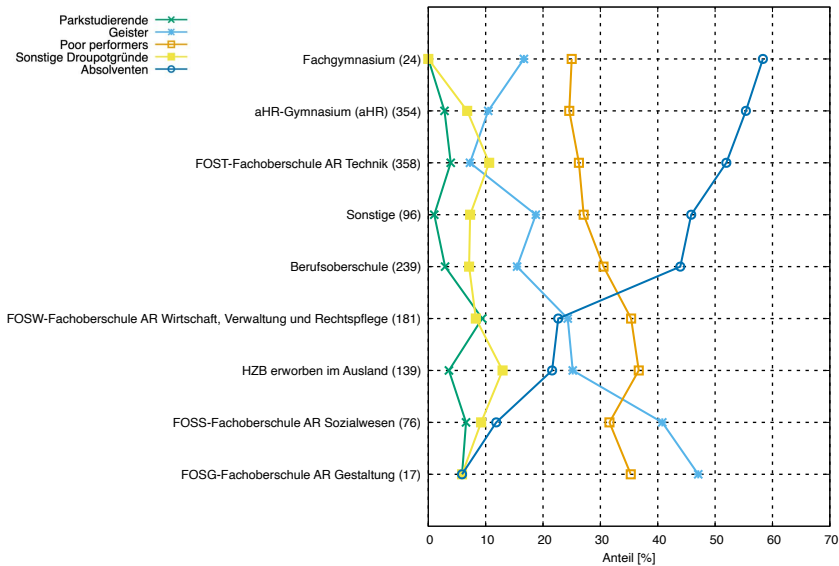


Abb. 2: Anteile der identifizierten Klassen von Studierenden aufgeschlüsselt nach Herkunftsschulen. Die Schularten sind so sortiert, dass die Werte der Klasse der Absolvent:innen monoton fällt.

Hier empfiehlt es sich, den Studierenden der Herkunftsschultypen mit niedrigerer Erfolgsquote frühzeitig und proaktiv – also bereits zu Beginn des ersten Semesters – Unterstützungsangebote nahe zu bringen mit dem Ziel, Potenziale zu wecken, *bevor* sich Überforderung und Misserfolge einstellen. Um jegliches Risiko für ein Bias in der Betreuung bzw. Bewertung durch die Lehrenden zu verhindern, ist dabei für die Umsetzung zwingend erforderlich, dass die Selektion der entsprechenden Studierenden über den bzw. die Datentreuhänder:in erfolgt und die Beratung ausschließlich der Lernberatung obliegt. Lehrende werden erst dann eingebunden, wenn der bzw. die Studierende das explizit anfragt.

5.3 Schulnoten sind weniger prädiktiv als die im ersten Semester erworbenen Credits

Eine Analyse verschiedener Variablen zur Bildungsbiografie und dem Verlauf der Studieneingangsphase hat gezeigt, dass *vor* dem ersten Semester die Schulnoten etwas prädiktiv sind – aber in deutlich geringerem Maße als die Anzahl vorheriger Immatrikulationen in anderen Studiengängen sowie die Anzahl der in dieser Vorgeschichte erworbenen und dann über Anerkennung eingebrachten Credits [Bö21a]. *Nach* dem ersten Semester ist die Anzahl der neu erworbenen Credits der aussagekräftigste Prädiktor für den Studienerfolg und dabei deutlich prädiktiver als die Durchschnittsnote der Hochschulzugangsberechtigung [Bö21a].

Das bedeutet, dass auch Studierende mit weniger guten Schulnoten sehr wohl eine Chance auf Studienerfolg haben. Wichtig ist dabei jedoch ein möglichst guter Einstieg ins erste Semester. Hier können frühzeitige Unterstützungsangebote einen Beitrag leisten.

Des Weiteren beobachten wir ein Phänomen, welches als „Murky Middle“ bezeichnet wird [EA15]. Dabei werden die Studierenden in der Anfangsphase des Studiums nach ihrer Leistungsfähigkeit betrachtet. Die Erfahrung zeigt, dass die Studierenden in der Spitzengruppe ihr Studium in der Regel erfolgreich bewältigen, während umgekehrt diejenigen Studierenden, die sich am unteren Ende des Leistungsspektrums bewegen, ihr Studium mit hoher Wahrscheinlichkeit nicht erfolgreich beenden werden. Das Mittelfeld der Studierenden wird als Murky Middle bezeichnet, weil sich gerade hier keine klare Prognose über den Studienerfolg treffen lässt: Aus dem Mittelfeld der Studierenden wird ein Teil erfolgreich sein, der andere Teil aber nicht – und beide Teile sind relativ groß. Bei [EA15] konnten ca. 2/5 der Studierenden in der Murky Middle ihr Studium am Ende *nicht* erfolgreich abschließen.

In der Arbeit von [EA15] werden die Leistungsdaten der Studierenden auf Grundlage des Grade Point Average (GPA) ausgewertet. Für unsere eigenen Analysen mussten wir aus Gründen des Datenschutzes von dieser Ebene der Noten abstrahieren auf die Ebene der erworbenen Credits (siehe Abbildung 3). Um die Ergebnisse unserer Datenanalyse leichter mit dem Bild aus [EA15] vergleichen zu können, haben wir unsere Farbwahl entsprechend an die von [EA15] angepasst.

Abbildung 3 zeigt die Anzahlen der Abbrecher:innen im bzw. nach dem ersten Studienjahr sowie die Anzahl der Absolvent:innen, jeweils in Abhän-

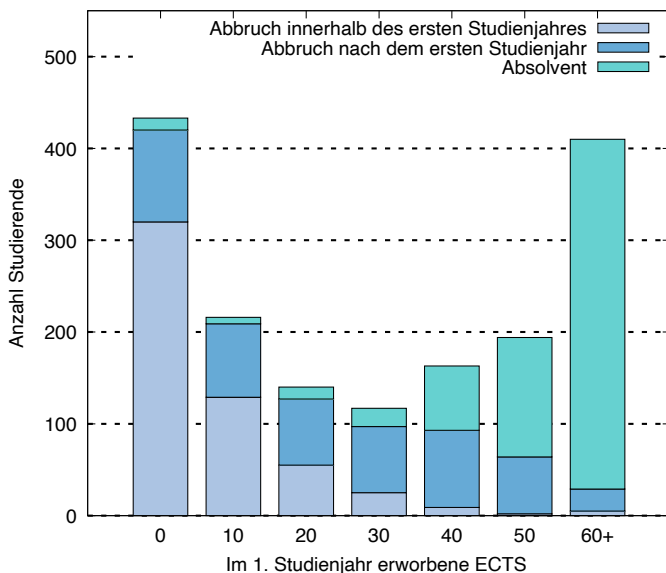


Abb. 3: Anzahl der erfolgreichen Studierenden sowie Dropout im ersten und nach dem ersten Studienjahr in Abhängigkeit der im ersten Studienjahr erzielten Credits.

gigkeit der innerhalb des ersten Studienjahres erwirtschafteten ECTS. Gemäß Studienplan wären innerhalb des ersten Studienjahres insgesamt 60 ECTS zu erarbeiten.

Die ganz linke Säule enthält viele Geister und Parkstudierende, die ganz rechte Säule dagegen diejenigen Studierenden, die nach dem ersten Studienjahr alle im Studienplan angesetzten ECTS erzielt haben. Unsere *Murky Middle* sind die etwa 20 %, die im ersten Studienjahr 20–40 Credits erworben haben und entweder nach einem Jahr ihr Studium abbrechen, oder aber erfolgreich absolvieren. Bei diesem Personenkreis ist also nach dem ersten Studienjahr noch alles offen.

5.4 Es ist nicht schädlicher, eine Prüfung zu schieben als sie nicht zu bestehen

Jede:r Studierende kann für jede Prüfung des ersten Semesters entscheiden, ob sie bzw. er an dieser teilnimmt oder nicht (die Prüfung also „schiebt“). Studierende, die im ersten Semester um eine Beratung bitten, stehen oft vor der

Frage, wie sie sich hinsichtlich der anstehenden Prüfungen konkret verhalten sollen. Das Aufschieben einer Prüfung ist auf den ersten Blick unschädlich. Diese Entscheidung hat allerdings ebenso wie das Nichtbestehen einer Prüfung eine Auswirkung auf die Arbeitslast in mindestens einem der Folgesemester. Dozierende neigen daher dazu, vom Schieben von Prüfungen abzuraten.

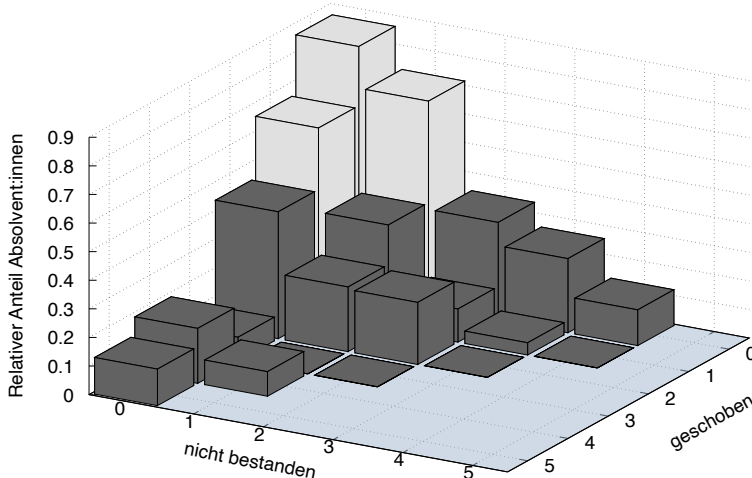


Abb. 4: Relativer Anteil erfolgreicher Studierender in Abhängigkeit der Anzahl an geschobenen und nichtbestandenen Prüfungen im ersten Semester.

Abbildung 4 zeigt den relativen Anteil der erfolgreichen Studierenden (d. h. Absolvent:innen), abhängig von der Anzahl geschobener und nichtbestandener Grundlagen- und Orientierungsprüfungen im ersten Semester. Von den Studierenden, die alle diese Prüfungsleistungen im ersten Semester erfolgreich erbracht haben (0 geschoben, 0 nicht bestanden) erreichen über 85 % ihren Abschluss. Bei einer fehlenden Prüfungsleistung sind es dagegen noch um die 70 %.

Insgesamt wird deutlich, dass der Grund, warum eine Prüfungsleistung nicht erbracht wurde (Nichtbestehen vs. Schieben), sich kaum auf den Studienerfolg auswirkt. Dies ist daran erkennbar, dass in beide Achsenrichtungen der Abfall der Säulen ähnlich verläuft. Dabei ist ein deutlicher Abfall der Erfolgsrate von einer zu zwei oder mehr nicht erfolgreich erbrachten Prüfungsleistungen erkennbar. Entsprechend ist es nicht ratsam, mehrere Prüfungen zu schieben. Der Datenpunkt bei fünf geschobenen Prüfungen deutet auf Studierende mit persönlichen Problemen oder ernsthaften Erkrankungen hin, die ihre Defizite

später aufholen und ihr Studium zu einem erfolgreichen Abschluss bringen konnten.

Als Erkenntnis lässt sich ableiten, dass wir den Studierenden, die sich im ersten Semester überfordert fühlen, durchaus raten dürfen, ggf. eine einzelne Prüfung zu schieben, um sich bei der Prüfungsvorbereitung nicht zu verzetteln. Offen bleibt dabei zunächst die Frage, auf welche Module dann vordringlich die Konzentration zu lenken ist.

5.5 Die große Hürde ist eher die Mathematik als die Softwareentwicklung

Zur Analyse, welche Module den Studierenden im ersten Semester die meisten Schwierigkeiten bereiten, wurde ein Entscheidungsbaum modelliert [HKP12; TA19]. Basis sind Prüfungsergebnisse des ersten Semesters, repräsentiert durch kategoriale Variablen mit jeweils drei Merkmalsausprägungen: *bestanden*, *5* sowie *geschoben*, wobei *5* die Bedeutung *nicht bestanden* repräsentiert. Die Zielvariable ist die dichotome Variable *Abschluss ja/nein*. Aus der Analyse herausgenommen wurden die als Geister und Parker identifizierten Personen sowie diejenigen, die alle Prüfungsleistungen aus Vorstudien anerkannt bekommen haben, was die Differenz der Gesamtzahl zu Abbildung 1 erklärt.

Das Ergebnis ist in Abbildung 5 dargestellt. Die Pfeile zu den beiden möglichen Werten *Kein Abschluss* bzw. *Abschluss* der Zielvariable sind beschriftet mit der Anzahl der korrekt (in Klammern: inkorrekt) klassifizierten Studierenden.

Auch in diesem Modell zeigt sich kein Unterschied, ob eine Prüfung nicht angetreten (*geschoben*) oder nicht bestanden (*5*) wurde. Die beiden Mathematikmodule erweisen sich als sehr prädiktiv, d. h. die Studierenden scheitern tendenziell eher an der Mathematik als am Modul Softwareentwicklung I. Das Modul IT-Systeme Grundlagen ist ebenfalls prädiktiv. Inhaltlich ist dieses Modul auch als programmiernah anzusehen, da dort intensiv auf Ebene von Maschinensprache gearbeitet wird.

Daraus lässt sich als Empfehlung an die Studierenden ableiten, sich bei Schwierigkeiten im ersten Semester auf eines der beiden Mathematik-Module sowie auf Softwareentwicklung I und IT-Systeme zu konzentrieren.

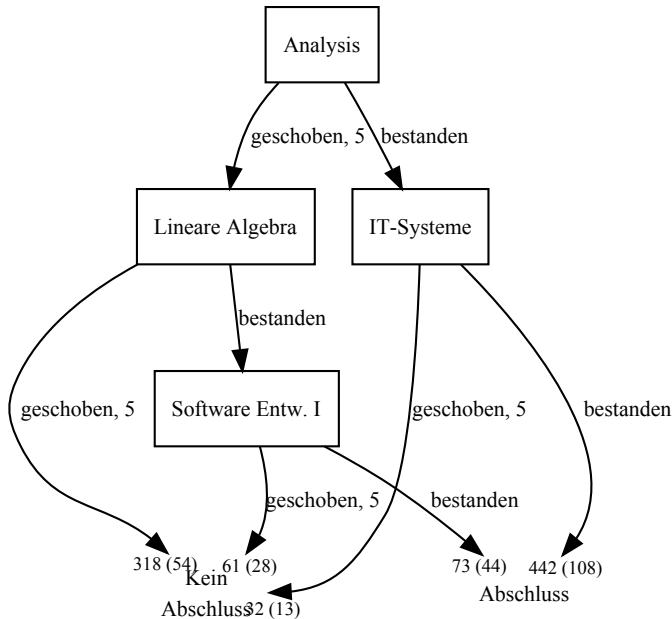


Abb. 5: Entscheidungsbaum für Prognose des Studienerfolgs auf Basis des Prüfungsverhaltens im ersten Semester.

5.6 Viele offene Prüfungsleistungen aus den ersten beiden Semestern sind gefährlich

Ein immer wieder beobachtbares Verhaltensmuster ist, dass sich Studierende mit Altlasten aus dem Bereich der Grundlagen- und Orientierungsprüfungen im dritten Semester zu viel vornehmen. Drittsemester-Module, die im dritten Semester zwar angetreten, aber nicht bestanden werden, müssen wegen der prüfungsrechtlich geltenden Fristen zwingend sofort im nächsten Semester als Wiederholungsprüfung angetreten werden. Studierende, die trotz bestehender Altlasten aus dem Bereich der Grundlagen- und Orientierungsprüfungen versuchen, „normal“ im dritten Semester weiterzustudieren und dort dann auch gemäß Regelstudienplan die Drittsemester-Prüfungen anzutreten, laufen also Gefahr, dass sich durch Nichtbestehen dieser Regelprüfungen des dritten Semesters unnötigerweise zusätzlicher Druck aufbaut, in Form von noch mehr einzuhaltenden Fristen für die dann sofort im nächsten (dem vierten) Semester anzutretenden Wiederholungsprüfungen dieser nicht bestandenem Drittsemester-Module.

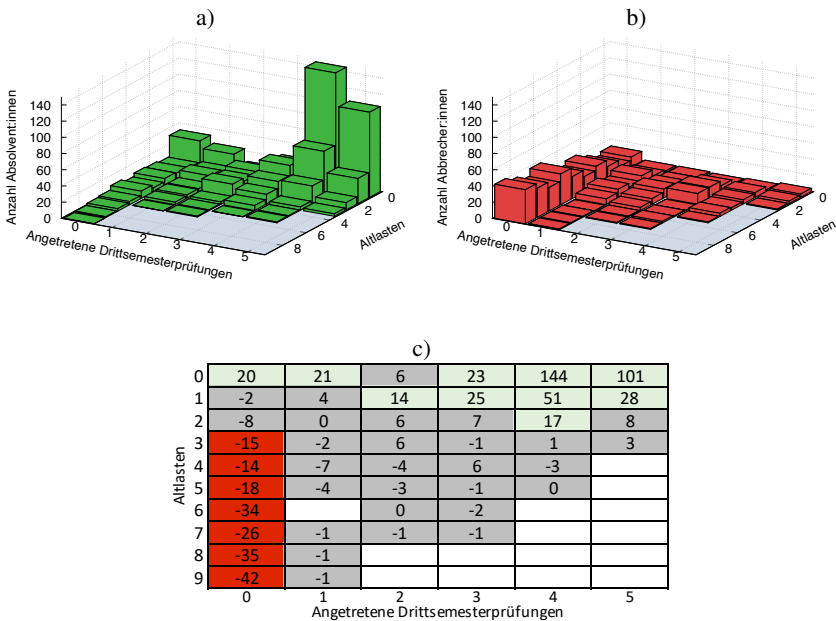


Abb. 6: Anzahl erfolgreicher (a) und erfolgloser Studierender (b) in Abhängigkeit davon, wie viele Prüfungen aus dem ersten und zweiten Semester sie im dritten Semester noch ablegen müssen und wie viele Prüfungen aus dem dritten Semester sie antreten. Teilgrafik (c) zeigt die Differenz als vereinfachte Heatmap. Farblich grau gekennzeichnet ist die „Murky Middle“.

Um herauszufinden, wie relevant diese Beobachtungen sind, stellen wir in Abbildung 6 die Anzahl der erfolgreichen (Diagramm a) und erfolglosen (Diagramm b) Studierenden einander gegenüber, in Bezug auf die Anzahl der offenen Grundlagen- und Orientierungsprüfungen („Altlasten“) und die bereits angetretenen Regelprüfungen des dritten Semesters. Diagramm c) zeigt die Differenz der Diagramme als vereinfachte, zweidimensionale Heatmap.

Auch hier wird wieder das bereits oben erwähnte Phänomen einer „Murky Middle“ deutlich [EA15]. Die Heatmap in Abbildung 6c) visualisiert eine Dreiteilung der Kohorten in

1. überwiegend erfolgreiche Studierende ohne oder mit höchstens einer offenen Prüfung aus den ersten beiden Semestern (grüner Bereich in Abbildung 6c); grün gefärbt sind dabei Felder mit mindestens 14 erfolgreichen Studierenden). Diese Studierenden benötigen keine besondere Betreuung.

2. diejenigen Studierenden, die den Studiengang absehbar bald verlassen werden; das sind im Wesentlichen diejenigen Studierenden, die im dritten Semester kaum reguläre Prüfungen antreten und mehrere Altlasten aus den ersten beiden Semestern haben (roter Bereich in Abbildung 6c)).
3. einen mittleren Bereich, in dem die Anzahl der erfolgreichen und erfolglosen Studierenden in etwa ausgeglichen ist. Dieser Bereich ist in Abbildung 6c) grau gefärbt.

Auf dieser Grundlage lassen sich sehr konkret Studierende mit Beratungsbedarf identifizieren: Gerade für die Personen im mittleren, grau eingefärbten Bereich ist statistisch noch offen, ob sie ihr Studium erfolgreich abschließen können werden oder nicht. Entsprechend ist genau für diese Studierenden eine gezielte Beratung ratsam, um ihre Chancen auf Studienerfolg zu erhöhen. Beispielsweise benötigen Studierende in diesem Bereich eine gezielte Unterstützung bei der Planung der als nächstes abzulegenden Prüfungen, damit sie sich nicht verzetteln.

6 Zusammenfassung und Ausblick

Die Datenanalyse hat hinsichtlich der Korrektheit gängiger Annahmen Klarheit geschaffen – teilweise wurde das Bauchgefühl widerlegt, teilweise bestätigt. Insgesamt lässt sich festhalten, dass Schulnoten weniger wichtig sind als gedacht, der Herkunftsschultyp jedoch eine große Rolle spielt für die Chance auf einen Studienerfolg. Des Weiteren ist es für den späteren Studienerfolg irrelevant, ob nicht erbrachte Leistungen im ersten Semester durch Schieben oder durch Nichtbestehen bedingt sind. Ferner wurde offensichtlich, dass die entscheidende Hürde in unserem Bachelorstudiengang Informatik die Mathematik darstellt.

Eine positive Erkenntnis ist, dass ein nicht optimaler Start ins Studium nicht zwingend bedeutet, dass das Studium erfolglos aufgegeben werden muss. Sofern ernsthaftes Interesse am Studium vorhanden ist (also wirklich aktiv studiert wird), kann nach dem ersten Semester kein eindeutiger Indikator für das Nichtbestehen identifiziert werden. Eine weitere Analyse von Studienverlaufsdaten derjenigen Studierenden, die nach einem suboptimalen Start einen Abschluss erhalten haben, förderte entsprechend auch keine klaren Muster zutage, auf deren Basis sich Empfehlungen aussprechen ließen. Um dennoch von dieser Personengruppe zu lernen und daraus Empfehlungen für eine Ausdehnung

bzw. zukünftig neu zu etablierende Beratung abzuleiten, führen wir derzeit qualitative Interviews durch, von denen wir uns weitere Einsichten erhoffen.

Mit Blick auf mögliche strukturelle Maßnahmen legen die Auswertungen der Prüfungsdaten nahe, das Modul Angewandte Mathematik auch im Wintersemester anzubieten. Das würde den Studierenden zum einen die Möglichkeit bieten, eine zwangsweise Fristüberschreitung zu vermeiden, und zum anderen nicht unnötig dazu verleiten, weitere Probleme durch das zu frühe Antreten von Prüfungen des dritten Semesters aufzubauen. Auch die Einführung strengerer Regeln zum Vorrücken in höhere Semester würde die Problemgenerierung durch das zu frühe Antreten von Drittsemester-Prüfungen wenigstens teilweise unterbinden.

Die Erfahrung zeigt, dass insbesondere die zentrale Studierendenvertretung massiv gegen die Einführung derartiger Regeln vorgeht – während die Studierendenvertretung der eigenen Fakultät durchaus deren Sinnhaftigkeit erkennt und der Einführung gegenüber offen ist. Die Einführung bzw. Handhabung entsprechender Regelungen ist also nicht nur ein inhaltliches, sondern insbesondere auch ein hochschulpolitisches Thema.

Literaturverzeichnis

- [Au18] Autorengruppe Bildungsberichterstattung: Bildung in Deutschland 2018 – Ein indikatorengestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung, 2018, URL: <https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2018>.
- [Be18] Berens, J.; Schneider, K.; Görtz, S.; Oster, S.; Burghoff, J.: Early Detection of Students at Risk – Predicting Student Dropouts Using Administrative Student Data and Machine Learning Methods, (CESifo Working Paper Series 7259), CESifo, München, 2018.
- [Bö21a] Böttcher, A.; Thurner, V.; Häfner, T.; Hertle, J.: A Data Science-based Approach for Identifying Counseling Needs in first-year Students. In: 2021 IEEE Global Engineering Education Conference (EDUCON). Institute of Electrical und Electronics Engineers, S. 425–434, 2021.
- [Bö21b] Böttcher, A.; Thurner, V.; Häfner, T.; Ottinger, S.: Adaptierung von Beratungsangeboten auf der Basis von Erkenntnissen aus der Analyse von Studienverlaufsdaten. In: (Desel, J.; Opel, S.; Siegeris, J., Hrsg.):

9. Fachtagung Hochschuldidaktik Informatik (HDI) 2021, 15.–16. September 2021 in Dortmund. Vorabdruck der Konferenzbeiträge; Fern-Universität in Hagen. S. 57–64, 2021.
- [BTH20] Böttcher, A.; Thurner, V.; Häfner, T.: Applying Data Analysis to Identify Early Indicators for Potential Risk of Dropout in CS Students. In: 2020 IEEE Global Engineering Education Conference (EDUCON). Institute of Electrical und Electronics Engineers, S. 827–836, 2020.
- [EA15] EAB Global, Inc.: The „Murky Middle“. Profiling campus segments to determine where schools should focus their efforts. 15.01.2015. 2015, URL: <https://eab.com/technology/infographic/student-success/the-murky-middle/>, Stand: 19.03.2022.
- [FY15] Fei, M.; Yeung, D.-Y.: Temporal Models for Predicting Student Dropout in Massive Open Online Courses. In: 2015 IEEE International Conference on Data Mining Workshop (ICDMW). Institute of Electrical und Electronics Engineers, S. 256–263, 2015.
- [HKP12] Han, J.; Kamber, M.; Pei, J.: Data mining concepts and techniques. Morgan Kaufmann Publishers, 2012.
- [HMT16] Hinkelmann, M.; Maucher, J.; Tobias, S.: Softwaregestützte Studienverlaufsanalyse zur frühzeitigen gezielten Studienberatung. Die Hochschullehre 2/, 2016, URL: http://www.hochschullehre.org/?dl_id=131.
- [HSW06] Holdt, U.; Schneider, H.; Wagner, B.: Analyse von Studienverläufen und Studienabbrüchen in den Bachelorstudiengängen Informatik an der Leibniz Universität Hannover. In (Forbrig, P.; Siegel, G.; Schneider, M., Hrsg.): HDI 2006: Hochschuldidaktik der Informatik – Organisation, Curricula, Erfahrungen. Gesellschaft für Informatik e.V., Bonn, S. 115–126, Dez. 2006.
- [KVV20] Kemper, L.; Vorhoff, G.; Wigger, B. U.: Predicting student dropout: A machine learning approach. European Journal of Higher Education 10/1, S. 28–47, 2020.
- [Me03] Meier, B.-D.: Ist der Erfolg im Jurastudium vorhersagbar? Empirische Befunde zum Zusammenhang zwischen Schulnoten und Abschneiden im Ersten Juristischen Staatsexamen. Beiträge zur Hochschulforschung 25/4, S. 18–35, 2003.

- [SH21] Scheidig, F.; Holmeier, M.: Learning Analytics aus institutioneller Perspektive: Ein Orientierungsrahmen für die hochschulische Datennutzung. In: Digitalisierung in Studium und Lehre gemeinsam gestalten: Innovative Formate, Strategien und Netzwerke. Springer Fachmedien, Wiesbaden, S. 215–231, 2021.
- [TA19] Therneau, T.; Atkinson, B.: rpart: Recursive Partitioning and Regression Trees, R package version 4.1-15, 2019.