

Julia SPÖRK¹, Karl LEDERMÜLLER, Robert KRIKAWA, Gabriel WURZER & Shabnam TAUBÖCK (Wien)

Analyse von Studierbarkeit mittels Prognose- und Simulationsmodellen

Zusammenfassung

Ergebnisindikatoren von Studierbarkeit können als studienerefolgsrelevante Kenngrößen operationalisiert und dadurch modelliert und prognostiziert werden. Im vorliegenden Papier wird gezeigt, wie die Wahl eines passenden Machine-Learning-Verfahrens sowohl die Prognose individueller Studierbarkeit mit einer Treffsicherheit von fast 90%, als auch die Identifizierung von Einflussfaktoren auf individuelle Studierbarkeit ermöglicht. Weiters wird eine konzeptionelle Verschränkung des Prognosemodells mit einem Simulationsmodell diskutiert, um die strukturelle Dimension von Studierbarkeit analysieren zu können.

Schlüsselwörter

Studierbarkeit, Studienerfolg, Prognosemodelle, Simulationsmodelle, Machine Learning

¹ E-Mail: julia.spoerk@wu.ac.at



Analysis of studyability by means of prediction and simulation models

Abstract

Outcome indicators of studyability can be operationalized as parameters relevant to study success and thus modeled and predicted. This paper shows how choosing the appropriate machine learning method enables both the prediction of individual studyability with an accuracy of almost 90% and the analysis of factors influencing individual studyability. Furthermore, a conceptual intersection of the prediction model with a simulation model is discussed in order to analyse the structural dimension of studyability.

Keywords

studyability, student success, prediction models, simulation models, machine learning

1 Einleitung

Studierbarkeit ist ein vielschichtiges und mehrdimensionales Phänomen, welches eine individuelle wie auch eine strukturelle Ebene umfasst:

- Die individuelle Ebene ist abhängig von der aktuellen Lebenssituation, der sozialen und räumlichen Herkunft sowie vom Engagement im weitesten Sinn (ZUCHA et al., 2020). Sie definiert das Potenzial, das Studierende von sich aus mitbringen.
- Die strukturelle Ebene wird den Studierenden auferlegt; sie umfasst z. B. den Inhalt der Curricula, Zugangsbeschränkungen zu Studien und Lehrveranstaltungen sowie die generelle Infrastruktur, die zur Abwicklung der Lehre verwendet wird (BURCK & GRENDEL, 2011).

Der vorliegende Ansatz geht von diesem Verständnis der Studierbarkeit aus und versucht über Machine-Learning-Methoden Studienerfolg als Ergebnisindikator

von Studierbarkeit zu prognostizieren. Dabei sollen folgende Forschungsfragen beantwortet werden:

1. Wie treffsicher lässt sich Studienerfolg mithilfe unterschiedlicher Machine-Learning-Verfahren prognostizieren?
2. Welche Variablen haben Einfluss auf die Prognose des Studienerfolgs?

Die Beantwortung dieser Forschungsfragen gliedert sich in die folgenden Abschnitte: Zunächst wird in Kapitel 2 auf das Konzept der Studierbarkeit und auf Überschneidungen zu klassischen Studienerfolgstheorien eingegangen. Anschließend wird der Stand der Forschung hinsichtlich der Erklärung, Prognose und Simulation von Studienerfolg überblicksartig dargestellt. Die Beantwortung der Forschungsfragen erfolgt in Kapitel 6 („Ergebnisse“) unter Verwendung der in Kapitel 4 und 5 beschriebenen Analysemethoden und Datengrundlagen. Abschließend wird in Kapitel 7 die Möglichkeit einer konzeptionellen Erweiterung durch Integration des Prognoseansatzes in eine Simulation vorgestellt.

Die hier dargelegten Ansätze wurden im Rahmen des vom Österreichischen BMBWF unterstützten Projekts „Predictive Analytics Services für Studienerfolgsmanagement“ (PASSt) der TU Wien, WU Wien und JKU Linz praktisch erprobt. Das Projekt ist eingebettet in das thematische Cluster „Digitale und soziale Transformation in der Hochschulbildung“, welcher auch das Projekt „LA – Studierende im Fokus“ der Universität Wien, TU Graz und KFU Graz beinhaltet. PASSt ermöglicht potenziell Anwendungsfelder für verschiedene Zielgruppen – unter anderem für das Hochschulmanagement, wofür der Aufbau eines Managementinformationssystems (beispielsweise generisch adaptierbare Cockpits wie z. B. eine Dashboard- oder Reporting-Lösung) zur Visualisierung der Ergebnisse mit einfacher Handhabbarkeit und hoher Skalierbarkeit angestrebt wird. Darüber hinaus wird der rechtliche Rahmen in allen Phasen des Projekts geprüft und Rahmenbedingungen werden abgeklärt. Die Ergebnisse sollen zusammengefasst anderen Bildungseinrichtungen zur Verfügung gestellt werden.

Die vorliegende Arbeit fokussiert auf die methodischen Grundlagen des Projekts. Es diskutiert die Wahl des Prognoseverfahrens zur individuellen Studierbarkeit sowie Potenziale der Integration der Prognose in die Simulation.

2 Theorie

Obwohl immer häufiger im Rahmen von Akkreditierungs- oder Leistungsvereinbarungen von Hochschulen die Gewährleistung von Studierbarkeit gefordert wird, existiert bislang weder eine einheitliche Begriffsdefinition noch ein etabliertes Instrumentarium, mit dem Phänomenbereich Studierbarkeit an Hochschulen umzugehen. RICHTER (2000) definiert Studierbarkeit sehr allgemein als die „Abwesenheit von Faktoren, die das Studium behindern“ (RICHTER, 2000, S. 161f.). Häufig wird auf eine individuelle und eine strukturelle Komponente von Studierbarkeit verwiesen (AQ AUSTRIA, 2018; ZUCHA, ZAUSSINGER & UNGER, 2020).

Strukturelle Studierbarkeit kann „als institutionelles Arrangement verstanden werden, das Handlungen der Studierenden, z. B. das überschneidungsfreie Besuchen von Lehrveranstaltungen oder die reibungslose Organisation eines Auslandsaufenthaltes, ermöglicht oder behindert“ (BURCK & GREDEL, 2011, S. 102). Nach ZUCHA, ZAUSSINGER & UNGER (2020) kann strukturelle Studierbarkeit beispielsweise durch gelungene Studienplangestaltung und -organisation, ausreichend Platzangebot in Lehrveranstaltungen, Unterstützungs- und Infrastrukturangebote sowie Qualität und Digitalisierung der Lehre gewährleistet werden.

Individuelle, studierendenbezogene Aspekte von Studierbarkeit beziehen sich auf Merkmale der Studierenden und können deren Lebenssituation und Herkunftsbedingungen, Studienengagement sowie soziale und akademische Integration an der Hochschule sein (ZUCHA, ZAUSSINGER & UNGER, 2020).

Sowohl individuelle, studierendenbezogene als auch strukturelle Aspekte finden sich in theoretischen Ansätzen zur Erklärung von Studienerfolg wieder, welche sich bereits seit den 1970er-Jahren mit der Erklärung von Studienerfolg beschäftigen. So wird auf individuelle Studierbarkeit in Form von Herkunftsbedingungen beispielsweise bei BOURDIEU & PASSERON (1971) fokussiert, auf soziale und akademische Integration bei TINTO (1975). Das individuelle Studienengagement wird beispielsweise bei BEAN & EATON (2000) berücksichtigt. Auf strukturelle Aspekte wird insbesondere bei BEAN (1980) eingegangen, welcher strukturelle Faktoren der Hochschule mit individuellen Merkmalen der Studierenden in Bezug setzt.

In aktuellen empirischen multiperspektivischen Studienerfolgsmodellen (beispielsweise HEUBLEIN et al., 2017 oder UNGER et al., 2009) wird ebenfalls betont, dass

sich individuelle und institutionelle Merkmale wechselseitig beeinflussen. Für eine gelungene Passung zwischen Studierenden und Studium wären Adaptionsleistungen sowohl vonseiten der Studierenden als auch vonseiten der Hochschule nötig. Universitäre Maßnahmen zur Studierbarkeit (strukturelle Dimension) sollten sich an der Diversität und den Lebensbedingungen der Studierenden (individuelle Dimension) orientieren (BUSS, 2019). Auch in der aktuellen österreichischen Studierenden-Sozialerhebung (ZUCHA, ZAUSSINGER & UNGER, 2020) wird hervorgehoben, dass die erste Säule der Studierbarkeit, also Maßnahmen zur Verbesserung der Studierbarkeit, an die zweite Säule, individuelle Merkmale der spezifischen Studierendenschaft einer Universität, angepasst werden sollten. Studierbarkeit wäre in diesem Sinne „all das [...], was eine Hochschule an Rahmenbedingungen und Infrastruktur bereitstellt, um den Studierenden ein möglichst reibungsloses und flexibles Studieren zu ermöglichen. Dies kann bzw. sollte unter Berücksichtigung der Diversität der Studierenden erfolgen.“ (ZUCHA, ZAUSSINGER & UNGER, 2020, S. 14).

Sichtbar und messbar wird Studierbarkeit durch die Analyse von Ergebnisindikatoren der Studierbarkeit. BUSS (2019) unterscheidet zwischen objektiven und subjektiven Ergebnisindikatoren. Als objektive Ergebnisindikatoren führt BUSS beispielsweise Studienabschluss/Studienabbruch, Studiengeschwindigkeit oder Noten an. Ein subjektiver Ergebnisindikator könnte die Studienzufriedenheit sein. Einflussfaktoren auf subjektive und objektive Ergebnisindikatoren wären das Studierverhalten (Besuch von Lehrveranstaltungen, Lernzeiten, Ablegen von Prüfungen) sowie weitere Einflussfaktoren, die sich in der Studiererfolgsforschung als relevant erwiesen haben (BUSS, 2019).

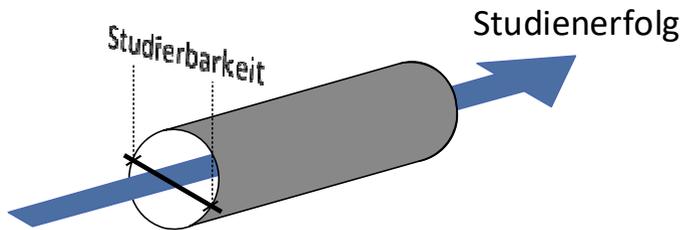


Abb. 1: Studierbarkeit vs. Studienerfolg

Studierbarkeit kann somit zusammenfassend als das teils individuelle, teils institutionelle Potenzial verstanden werden, während der Studienerfolg dessen Realisierung darstellt (siehe Abbildung 1).

3 Stand der Forschung

Sowohl Studierbarkeit, als auch Studierfähigkeit sind vielschichtige und mehrdimensionale Phänomene. Um die komplexe Realität dieser Phänomene analytisch nutzbar zu machen bzw. zu vereinfachen gibt es unterschiedliche Ansätze. Solche Modelle können entlang ihres (Wissenschafts-)Ziels in Beschreibungs-, Erklärungs-, Prognose- und Entscheidungsmodelle eingeteilt werden (STACHOWIAK, 1973).

In den letzten Jahren haben sich eine Reihe von multidisziplinären Studien mithilfe von **Beschreibungs- und Erklärungsmodellen** der Erklärung von individueller Studierbarkeit bzw. Studienerfolg gewidmet (vgl. beispielsweise KUH et al., 2006; HEUBLEIN et al., 2017; UNGER et al., 2009). HEUBLEIN et al. (2017) entwerfen ein theoretisches Modell des Studienabbruchprozesses, in welchem sich soziale Herkunft und Bildungsherkunft sowie die Persönlichkeit der Studierenden auf Bildungssozialisation und Studienfachentscheidungen, aber auch auf den individuellen Studienprozess auswirken. Der individuelle Studienprozess wiederum wird durch das Studierverhalten, die Studienmotivation, psychische und physische Ressourcen sowie durch die Studienleistungen beeinflusst. Auf diese wirken wiederum externe Faktoren wie Studienbedingungen, Lebensbedingungen, Informationen und Alternativen. Multivariate Modelle zur Erklärung von Studienerfolg berücksichtigen die Komplexität und Vielschichtigkeit von Studienerfolg und -abbruch und beziehen unterschiedliche Einflussfaktoren entlang der von HEUBLEIN et al. (2017) beschriebenen Dimensionen mit ein. In solchen Modellen weisen Vorleistungen (z. B. Schulnoten, Assessment-Tests, erste Studienleistungen) die stärksten Effekte auf (BARON-BOLDT, 1989; ROBBINS et al., 2004; TRAPMANN et al., 2007).

Prognosemodelle auf Basis von Machine-Learning-Ansätzen wurden im deutschsprachigen Hochschulraum beispielsweise von BERENS et al. (2019) und WALKMANN (2020) mit dem Ziel aufgestellt, Erfolgsfaktoren von Studierbarkeit früh auf individueller Ebene vorhersagen zu können. Meist wird hier auf Einflussfaktoren zurückgegriffen, welche als Administrativdaten vorliegen und sich innerhalb von

Studienerfolgs-Erklärungsmodellen als signifikant erwiesen haben. BERENS et al. (2019) verwenden hierzu ein Set an Administrativdaten zu Studienbeginn, das sie im Laufe des Studiums um Prüfungsdaten ergänzen. Dieser Ansatz verwendet zur Modellierung den AdaBoost Algorithmus und erreicht mit der Studienabbruchprognose zu Studienbeginn etwa 80% und im Verlauf des Studiums etwa 90% Prognosegenauigkeit (BERENS et al., 2019). WALKMANN (2020) kombiniert Administrativdaten mit Paneldaten und erreicht mit einem Logit-Modell nach Kreuzvalidierung sogar über 96% Treffsicherheit. In beiden Publikationen sind Vorleistungen (bisher durchschnittlich erreichte ECTS bzw. Durchschnittsnote) der einflussreichste Faktor im Hinblick auf die Prognosegüte.

Entscheidungsmodelle wie beispielsweise Simulationen (siehe auch „Ausblick“ in Abschnitt 7) bauen auf solchen Modellen auf und sollen Entscheidungsfindungen am Modell unterstützen. Simulationsmodelle sind im Kontext der Hochschulen bisher nur wenig zum Einsatz gekommen. Eine Ausnahme bildete das MoreSpace-Projekt (EMRICH et al., 2009), welches an der TU Wien in Zusammenarbeit mit dwh simulation services entwickelt wurde, um die Nutzung von Lehrräumen zu optimieren. In diesem Projekt wurden universitäre Abläufe und individuelle Entscheidungen auf Ebene der Studierenden abgebildet; im Zentrum der Beobachtung stand jedoch nicht der Studienerfolg, sondern die Präsenz der Studierenden am Universitätscampus und der daraus resultierende Raumbedarf.

4 Methode

Für die Prognose von Bildungserfolgen eignen sich neben klassischen Ansätzen wie (linearen) Regressionen auch Verfahren aus dem Bereich des überwachten maschinellen Lernens (BAKER & YACEF, 2009; KERSTING, 2019). Im Rahmen des Ansatzes werden (Boosted) Logistic Regressions, Gradient Boosting Machine-Modelle (GBM), Random Forest, Support Vector Machine und Neuronale Netze berechnet und hernach verglichen.

Boosting bedeutet ganz allgemein, aus schwachen Klassifikatoren stufenweise einen guten Klassifikator zu bilden, indem falsch klassifizierte Datenpunkte identifiziert und im Weiteren stärker gewichtet werden, um einen guten Fit zu erzielen (SCHAPIRE, 1990). Für Boosted Logistic Regressions wurde der LogitBoost-Al-

gorithmus von FRIEDMANN, HASTIE & TIBSHIRANI (2000) verwendet, zusätzlich wurden GBM-Modelle gerechnet.

Bei GBM-Modellen werden mehrere Entscheidungsbäume (Decision Trees) trainiert, wobei diese nicht voneinander unabhängig sind, sondern sequenziell unter Verwendung von Informationen aus den vorherigen Bäumen aufeinander aufgebaut werden. Random Forest (BREIMAN, 2001) basiert ebenfalls auf Entscheidungsbäumen, welche im Gegensatz zu GBM jedoch nicht korreliert sind und parallel mit einem Algorithmus trainiert werden. Support Vector Machine (SVM) versucht Objekte mithilfe von Trennungslinien oder -ebenen zu separieren. Zur Optimierung wird der Abstand zwischen Support Vektoren (Margins) maximiert. Durch die Anwendung des Kernel-Tricks lässt sich die Methode auch bei nicht-linearen Entscheidungsgrenzen einsetzen: Hierfür werden die Trennungsvektoren in eine zusätzliche Dimension transformiert (vgl. KARATZOGLOU, SMOLA & HORNIK, 2016). Neuronale Netze schließlich sind Algorithmen, die Muster aus den Daten selbst erkennen und lineare und nicht-lineare Wechselbeziehungen zwischen den unabhängigen Variablen datengetrieben berücksichtigen. Sie bestehen aus einer Eingabeschicht („Input Layer“ bzw. unabhängige Variablen), einer verdeckten Schicht („Hidden Layer“) und einer Ausgabeschicht („Output-Layer“ bzw. Vorhersage der abhängigen Variable) (REY & WENDER, 2018). Zusätzlich wurden klassische logistische Regressionen (Maximum-Likelihood-Lösung) gerechnet, um vergleichen zu können, ob anspruchsvollere Modelle in diesem Anwendungsfall eine höhere Vorhersagequalität aufweisen.

Zur Bewertung der Prognosequalität werden Gesamttrefferquote (Accuracy), Sensitivität und Spezifität sowie Receiver Operating Characteristics-Kurven (ROC-Kurven) berechnet (vgl. JAMES et al., 2013). Die Gesamttrefferquote kann als Anteil der korrekten Vorhersagen an allen getroffenen Prognosen verstanden werden. Im vorliegenden Fall ist das der prozentuelle Anteil der korrekt als prüfungsaktiv/-inaktiv klassifizierten Studierenden. Sensitivität wäre hier der Anteil korrekt klassifizierter prüfungsaktiver Studierender unter den tatsächlich Prüfungsaktiven und Spezifität umgekehrt der Anteil korrekt klassifizierter Nicht-Prüfungsaktiver unter den tatsächlich Nicht-Prüfungsaktiven. Zusätzlich wird die Prognosegüte mittels ROC-Kurve dargestellt. Hierfür werden in einem Diagramm die Falsch-Positiv-Rate (1-Spezifität) auf der x-Achse und die Richtig-Positiv-Rate (Sensitivität) auf der y-Achse abgetragen. Je weiter sich die Kurve eines Prognosemodells an die obere

linke Ecke der Grafik annähert, desto besser ist somit die Trefferquote. Die Fläche unter der ROC-Kurve wird als AUC (Area under the Curve) bezeichnet und kann maximal 1 sein – wobei ein Wert unter 0,5 bedeuten würde, dass eine Zufallszuweisung bessere Vorhersagen erzielen würde.

5 Definitionen & Daten

Ein Spezifikum der öffentlichen Finanzierung von Universitäten in Österreich ist, dass die Finanzierung von Universitäten mit einem Erfolgsindikator von Studierbarkeit (im weitesten Sinne) wesentlich in Beziehung steht. Dieser Erfolgsindikator wird als „Prüfungsaktivität“ bezeichnet. Prüfungsaktive Studien in der Definition der österreichischen Wissensbilanz (BGBl. II Nr. 97/2016, §2.A.69) werden als jene Bachelor-, Diplom- und Masterstudien definiert, in denen im Studienjahr mindestens 16 ECTS-Punkte oder positiv beurteilte Studienleistungen im Umfang von acht Semesterstunden erbracht werden. Dies ist vor allem für die Zuteilung des Unibudgets (siehe auch die österreichische Universitätsfinanzungsverordnung (BGBl. II Nr. 202/2018)) relevant. In den folgenden Prognosemodellen wird daher entlang dieser Definition die Prüfungsaktivität von Studierenden eines Bachelorstudiums im Studienjahr 2019/20 als Maßgröße für Studienerfolg und als Ergebnisindikator für Studierbarkeit vorhergesagt.

Als unabhängige Variablen wurden in der Studienerfolgsvorschung als relevant betrachtete Einflussfaktoren bzw. als individuelle Merkmale der Studierbarkeit geltende Faktoren einbezogen: Erreichte ECTS-Punkte im Vorjahr und Lerntage auf der Lernplattform (als Indikatoren des Studierverhaltens) sowie Alter, Doppelstudien, Geschlecht, Mobilitätserfahrung und Staatsbürgerschaft (als individuelle Einflussfaktoren).

Das Verfahren wurde zusätzlich an 15 unterschiedlichen Programmen validiert. Alle statistischen Auswertungen wurden auf anonymen Datensätzen gerechnet. Die unabhängigen Variablen für diesen exemplarischen Anwendungsfall zeigt Tabelle 1.

Aufgrund der in Tabelle 1 ersichtlichen unterschiedlichen Skalendimensionen wurden die Daten im Zuge der Modellformulierung skaliert und standardisiert. Die Berechnung wurde in R mittels caret (KUHN, 2008) durchgeführt. Hierfür wurde

der Gesamtdatensatz zunächst randomisiert in einen Trainingsdatensatz (n=5010), einen Validierungsdatensatz (n=1074) und einen Testdatensatz (n=1074) aufgeteilt. Die Modelle wurden am selben Trainingsdatensatz mit denselben Resampling-Parametern (10-fache Kreuzvalidierung) trainiert. Die Modellperformance wurde anschließend mit dem Validierungsdatensatz verglichen und die Vorhersagekraft des besten Modells am Testdatensatz evaluiert.

Tab. 1: Unabhängige Variablen im Prognosemodell

Unabhängige Variablen	Definition	Werte
Alter	Alter zu Studienbeginn	Jahre
Doppelstudien	zu mehreren Studien zugelassen	Ja(0)/Nein(1)
ECTS-Punkte im Vorjahr	Erreichte positive ECTS-Punkte im Vorjahr	Anzahl ECTS
Geschlecht	Geschlecht	m(0)/w(1)
Lerntage	Lerntag = mind. 1 Click/Tag auf einer Lernplattform	Tage
Mobilitätserfahrung	Auslandssemester (im Rahmen des Studiums) absolviert	Ja (0)/nein(1)
Staatsbürgerschaft	Inländische (Österreich) vs. Ausländische Staatsbürgerschaft	Inländer*in(0)/ Ausländer*in(1)

6 Ergebnisse

Tabelle 2 zeigt, dass rund 86% aller getroffenen Vorhersagen korrekt waren, womit die Gesamttrefferquote allgemein sehr hoch ist. Die höchste Accuracy erzielt Random Forest mit 87,1%, wobei sich die anderen Modelle davon nur um maximal 1,3 Prozentpunkte unterscheiden. Die schlechteste Accuracy weisen logistische Regression und SVM auf.

Hinsichtlich der Spezifität erreichen GLM und SVM mit rund 92% korrekt als inaktiv vorhergesagten Studierenden sehr gute Vorhersagewerte; Boosted Logistic Regression die schlechtesten. Random Forest und Boosted Logistic Regression weisen bei der Sensitivität mit 80% richtig vorhergesagten aktiven Studierenden die besten Werte auf. Die Sensitivität ist bei SVM mit 0,698 am geringsten.

Tab. 2: Modellvergleich (Validierungsdatensatz)

	Accuracy	Spezifität	Sensitivität	AUC
Random Forest	0,871	0,898	0,804	0,851
Logistische Regression (GLM)	0,858	0,923	0,698	0,810
Boosted Logistische Regression	0,862	0,855	0,807	0,846
Neuronale Netze	0,866	0,894	0,797	0,846
SVM	0,858	0,924	0,698	0,811
GBM	0,860	0,899	0,788	0,839

Die ROC-Kurven (siehe Abbildung 2) illustrieren darüber hinaus einerseits, dass die Vorhersagekraft des Modells eine zufällige Zuweisung übertrifft und sich alle Kurven der linken oberen Ecke annähern. Zum anderen wird mittels AUC erneut die schlechtere Vorhersagegüte von Support Factor Machine und logistischen Regressionen im Vergleich zu den anderen Modellen bestätigt. Random Forest weist mit 0,851 den beste AUC-Wert auf.

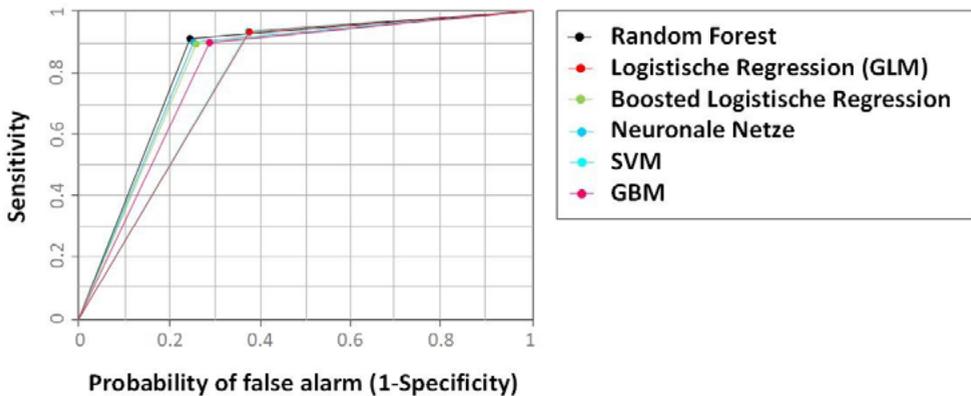


Abb. 2: ROC-Kurvenvergleich

Die Validität des besten Modells (basierend auf dem Validierungsdatensatz, siehe Tabelle 2) wurde nochmals anhand des Testdatensatzes validiert und erreicht hier eine Treffsicherheit von 87% (Random Forest: Accuracy=0,870; Spezifität=0,907; Sensitivität=0,787; AUC=0,847).

In Abbildung 3 werden die Wichtigkeitsmaße des Random Forest Modells dargestellt, um zu zeigen, welche Merkmale eine entscheidende Rolle bei der Vorhersage gespielt haben. Hierfür wird bei Klassifikationsproblemen (nominalskalierte abhängige Variable) der Gini Index als Maß für die Homogenität in den einzelnen Blattregionen (Node Impurity) verwendet und analysiert, wie sich diese bei Weglassen einer der unabhängigen Variablen verändern würde (vgl. DANIYA, GEETHA & KUMAR, 2020).

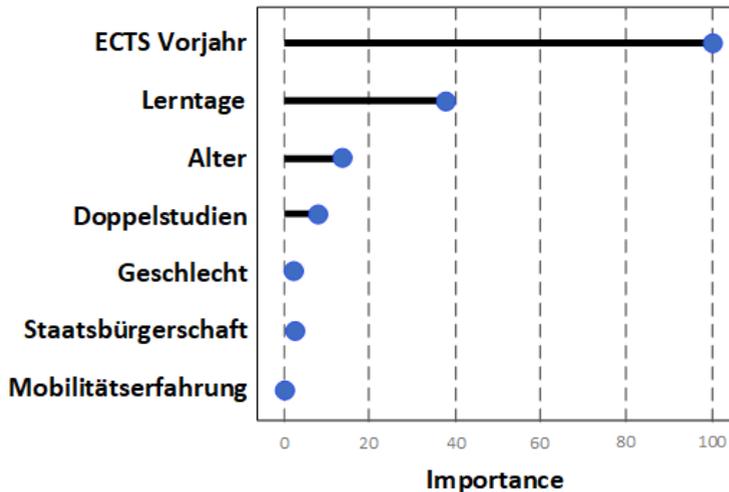


Abb. 3: Wichtigkeitsmaße des Random Forest

Deutlich wird, dass Vorleistungen wie erwartet ein wichtiger Faktor zur Prognose von Prüfungsaktivität und somit des Studienerfolgs sind. Sowohl die Zeit, die für das Studium investiert wird (Lerntage, Doppelstudien) als auch das Alter der Studierenden (und damit verbunden möglicherweise Kontextfaktoren wie Erwerbstätigkeit oder Betreuungspflichten) sind relevant. Kaum einen Einfluss auf die Vorhersage haben hingegen Geschlecht, Staatsbürgerschaft oder Mobilitätserfahrungen der Studierenden.

7 Fazit und Ausblick

Im vorliegenden Beitrag wurde gezeigt, dass Machine-Learning-Modelle mit hoher Treffsicherheit für die Analyse und Prognose von Ergebnisindikatoren von Studierbarkeit (z. B. potenziell inaktive Studierende) herangezogen werden können. Es wurde gezeigt, dass Machine-Learning-Modelle in diesem Anwendungsfall eine bessere Vorhersagegüte ermöglichen als klassische logistische Regressionen. Darüber hinaus scheinen Support Vector Machines, welche weniger gut funktionieren, wenn es keine klaren Margins zwischen den Klassen gibt und sich die Gruppen nicht klar voneinander abgrenzen lassen, für diese Fragestellung schlechter geeignet. Der wichtigste Faktor für die Prognose des Studienerfolgs sind wie erwartet Vorleistungen (in diesem Fall die im Vorjahr erreichten ECTS).

Geplant ist darüber hinaus, das Modell mit Umfragedaten anzureichern, um Faktoren aus der individuellen Lebenssituation (beispielsweise Erwerbstätigkeit), Studienmotivationen und -verhalten, aber auch zur (sozialen) Herkunft der Studierenden einbeziehen zu können. Zur weiteren Verbesserung der Vorhersagequalität könnten darüber hinaus Schulnoten oder die Ergebnisse von Aufnahmeverfahren (beispielsweise Zulassungstests) herangezogen werden.

Die beschriebenen Methoden nutzen Parameter auf Personenebene, um den individuellen Studienerfolg zu prognostizieren. Angedacht ist – noch einen Schritt weitergehend – eine Integration mit einer **Simulation zur Prognose der strukturellen Studierbarkeit** durchzuführen, wie in Abbildung 4 dargestellt und im Folgenden ausgeführt wird.

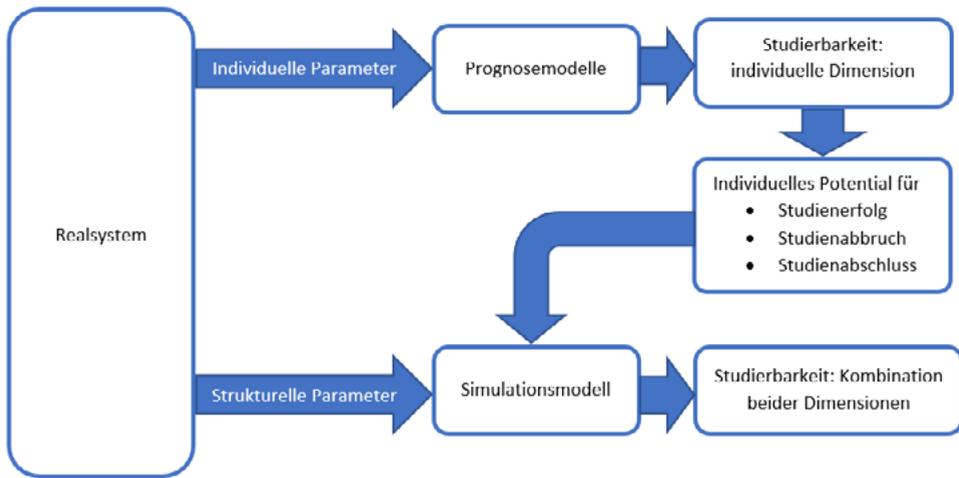


Abb. 4: Verzahnung der Modelle

Von der Ebene der einzelnen Studierenden (Agenten einer Simulation; vgl. BONABEAU, 2002; MACAL & NORTH, 2005) wird auf die Ebene des Studienprogramms übergegangen, um das Zusammenwirken individueller und struktureller Faktoren abzubilden. Genauer gesagt werden Lehrveranstaltungen modelliert, welche als kapazitätsbeschränkte Ressourcen der Simulation fungieren und von Studierenden in einem bestimmten Semester genutzt werden. Individuelle Parameter der Studierenden, wie beispielsweise das Potenzial an ECTS pro Semester, werden über die vorhin beschriebenen Prognosemodelle errechnet und gehen damit direkt in die Simulation ein. Engpässe, die sich aus der Struktur des Curriculums (Reihenfolgen, Voraussetzungsketten) oder den Lehrveranstaltungen (Anzahl Prüfungsplätze, Abhaltefrequenz) ergeben, können so aufgezeigt werden. Mittels Simulationsexperimenten können zudem Aussagen über die prognostizierte Studierbarkeit nach Veränderung des Status quo erlangt werden (Szenarienrechnung): Änderungen betreffend Kapazität oder Lage von Lehrveranstaltungen innerhalb des Curriculums können dann entweder manuell oder vollautomatisch (mittels Parametervariation) errechnet und hernach mit Bezug zu einer Zielfunktion gerankt werden.

Die Ergebnisse der Simulation und des Prognosemodells können zielgruppenorientiert ausgewertet und präsentiert werden: Beispielsweise kann eine Prognose von Risikostudierenden dabei helfen, frühzeitig mehr über diese Gruppe zu lernen (Wie viele Studierende sind betroffen? Welche Gruppen bzw. Cluster an Studierenden sind betroffen? Welche Parameter sind für Inaktivität verantwortlich?), um zielgerichtete Maßnahmen zu entwickeln. Die Implementierung dieser Maßnahmen kann dann wieder auf Ebene der einzelnen Studierenden, bei den Lehrenden oder in den einzelnen Curricula vorgenommen werden. Sowohl die Ergebnisse der Prognosemodelle als auch jene der Simulation können zur Entscheidungsunterstützung im Zuge der Planung herangezogen werden.

Das Abschätzen des Einflusses struktureller Faktoren auf den Studienerfolg ist maßgeblich in der Entscheidungsfindung bei der strategischen Planung in der Hochschulentwicklung. Bewirken diese Entscheidungen Änderungen im Originalsystem, so muss das Simulationssystem entsprechend angepasst werden, ebenso wie die davorliegenden Prognosemodelle.

8 Literaturverzeichnis

AQ Austria (2018). *Qualitätssicherung an österreichischen Hochschulen – Studierbarkeit*, Bericht gemäß §28 HS-QSG. Wien: Facultas.

Baker, R. & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3–17.

Baron-Boldt, J. (1989). *Die Validität von Schulabschlußnoten für die Prognose von Ausbildungs- und Studienerfolg. Eine Metaanalyse nach dem Prinzip der Validitätsgeneralisierung*. Frankfurt am Main: Peter Lang.

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, 12(2), 155–187.

Bean, J. P. & Eaton, S. (2000). A psychological model of college student retention. In J. M. Braxton (Hrsg.), *Reworking the departure puzzle: New theory and research on college student retention*. Nashville: University of Vanberbilt Press.

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(3), 1–14.

Bonabeau E. (2002). Agent-based modeling: methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99 Suppl 3, 7280–7287.

Bourdieu, P. & Passeron, J.-C. (1971). *Die Illusion der Chancengleichheit. Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs*. Stuttgart: Klett.

Breiman L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

Burck, K. & Grendel, T. (2011). Studierbarkeit – ein institutionelles Arrangement? *Zeitschrift für Hochschulentwicklung*, 6(2), 99–105.

Buß I. (2019). *Flexibel studieren – Vereinbarkeit ermöglichen. Studienstrukturen für eine diverse Studierendenschaft*. Wiesbaden: Springer VS.

Czanderle, B. (2017). Studierbarkeit. Wesen und Bedeutung für die Akteure im berufs begleitenden Fernstudium. *Hochschule und Weiterbildung*, 2017(1), 39–44.

Daniya, T., Geetha, M. & Kumar, K. S. (2020). Classification and Regression Trees with Gini Index. *Advances in Mathematics: Scientific Journal*, 9(10), 8237–8247.

Emrich, S., Tauböck, S., Wiegand, D., Breitenecker, F., Popper, N. & Zerlauth, S. (2009). *MORESPACE – A Hybrid Dynamic Approach for Modelling Lecture Room Management*, ARGESIM Report No. 35.

Friedman, J., Hastie, T. & Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, 28(2), 337–407.

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.

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. New York: Springer VS.

Karatzoglou, A., Smola, A. & Hornik, K. (2016). R-Paket kernlab. Kernel-Based Machine Learning Lab. In CRAN. *The R Foundation*

Kersting, K. (2019). *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*. Berlin: Springer.

Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, B. K. & Hayek, J. C. (2006). *What Matters to Student Success: A Review of the Literature. Commissioned report for the National Symposium on Postsecondary Student Success: Spearheading a Dialog on Student Success*. Washington, DC: National Postsecondary Education Cooperative.

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, 28(5), 1–26.

Macal, C. M. & North, M. J. (2005). *Tutorial on agent-based modeling and simulation*. Proceedings of the 37th conference on Winter simulation (WSC '05). Winter Simulation Conference, 2–15.

Österreichische Wissensbilanzverordnung (2016). BGBl. II Nr. 97/2016, §2.A.6.

Rey, G. D. & Wender, K. F. (2018). *Neuronale Netze. Eine Einführung in die Grundlagen, Anwendungen und Datenauswertung*. Bern: Hogrefe AG.

Richter, R. (2000). Studierbarkeit des Studiums herstellen – Ein Auftrag für die Studienreform. *Das Hochschulwesen*, 48(5), 158–162.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130(2), 261–288.

Schapire, R. (1990). The strength of weak learnability. *Machine Learning*, 5(2), 197–227.

Stachowiak H. (1973). *Allgemeine Modelltheorie*. Wien: Springer VS.

Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, 45(1), 89–125.

Trapmann, S., Hell, B., Weigand, S. & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs – eine Metaanalyse. *Zeitschrift für pädagogische Psychologie*, 21, 11–27.

Unger, M., Wroblewski, A., Latcheva, R., Zaussinger, S., Hofmann, J. & Musik, C. (2009). *Frühe Studienabbrüche an Universitäten in Österreich*. Endbericht. Studie im Auftrag des Bundesministeriums für Wissenschaft und Forschung BMWF.

Walkmann, R. (2020). Prognosemodelle zur Früherkennung potenzieller Studienabbrüche. Chancen und Grenzen. In E. Deuer & T. Meyer (Hrsg.), *Studienverlauf und Studienerfolg im Kontext des dualen Studiums. Ergebnisse einer Längsschnittstudie* (S. 57–72). Bielefeld: wbv.

Zucha, V., Zaussinger, S. & Unger, M. (2020). *Studierbarkeit und Studienzufriedenheit*. Zusatzbericht der Studierenden-Sozialerhebung 2019. Institut für Höhere Studien – Institute for Advanced Studies (IHS).

Autor*innen



Julia SPÖRK, MA || WU Wien, Evaluierung & Qualitätsentwicklung || Welthandelsplatz 1, A-1220 Wien

www.wu.ac.at

julia.spoerk@wu.ac.at



Dr. Karl LEDERMÜLLER || WU Wien, Evaluierung & Qualitätsentwicklung || Welthandelsplatz 1, A-1220 Wien

www.wu.ac.at

karl.ledermueller@wu.ac.at



Mag. Robert KRIKAWA || WU Wien, Akademisches Controlling || Welthandelsplatz 1, A-1220 Wien

www.wu.ac.at

robert.krikawa@wu.ac.at



Priv.-Doz. Dipl.-Ing. Dr. Gabriel WURZER || TU Wien, Zentrum für strategische Lehrentwicklung || Karlsplatz 13, A-1040 Wien

www.tuwien.ac.at

gabriel.wurzer@tuwien.ac.at



Dipl.-Ing. Dr. Shabnam TAUBÖCK || TU Wien, Zentrum für strategische Lehrentwicklung || Karlsplatz 13, A-1040 Wien

www.tuwien.ac.at

shabnam.tauboeck@tuwien.ac.at