

## Klausurprognose mit Hilfe von E-Assessment-Nutzerdaten

Till Massing<sup>1</sup>, Natalie Reckmann<sup>1</sup>, Benjamin Otto<sup>2</sup>, Kim J. Hermann<sup>1</sup>, Christoph Hanck<sup>1</sup>  
und Michael Goedicke<sup>2</sup>

**Abstract:** Elektronisches Assessment wird heutzutage verbreitet genutzt. Die dabei anfallenden Nutzerdaten geben Einblicke zum Lernverhalten der Studierenden. Wir analysieren die Daten mit Hilfe von Data Mining Methoden dahingehend einen eventuellen Erfolg in der Abschlussklausur vorzeitig vorherzusagen. Es zeigt sich, dass bereits während des Semesters eine hohe Prognosegenauigkeit vorliegt. Die entsprechenden Vorhersagen lassen sich für ein Frühwarnsystem nutzen, um die Studierenden in ihrer Lernbereitschaft zu motivieren.

**Keywords:** E-Assessment, Data Mining, Klausurprognose, Frühwarnsystem

### 1 Einleitung

Um großen Studierendenzahlen angemessen begegnen zu können, werden elektronische Übungssysteme in der Studieneingangsphase verbreitet eingesetzt. Online-Assessments ersetzen dabei den in vielen Studienfächern üblichen, klassischen Übungsbetrieb. In der Vorlesung „Induktive Statistik“ an der Universität Duisburg-Essen wird das selbst entwickelte E-Assessment-System JACK genutzt. Das Konzept beinhaltet formative und summative Assessments in Form von parametrisierten Lückentextaufgaben [Sc17] sowie Programmieraufgaben in der Programmiersprache R [Ot17]. Analysen haben gezeigt, dass die Note der Abschlussklausur mit dem Lernerfolg bei diesen Assessments positiv korreliert [Ma18].

Dieser Beitrag analysiert die Nutzerdaten der Studierenden und zeigt, dass schon während des laufenden Semesters das Bestehen der Abschlussklausur mit hoher Genauigkeit prognostiziert werden kann. Hierzu wird ein Vorhersagemodell mit Data Mining Methoden erstellt. Dieses kann in künftigen Versionen des Kurses genutzt werden, um Studierende in ihrem Lernverhalten zu beeinflussen und zu motivieren.

Der Rest dieses Artikels gliedert sich wie folgt: Zunächst werden in Kapitel 2 verwandte Arbeiten beschrieben. Kapitel 3 erläutert das Konzept der Vorlesung. In Kapitel 4 wird das Prognosemodell vorgestellt. Kapitel 5 gibt einen technischen Ausblick.

---

<sup>1</sup> Universität Duisburg-Essen, Wirtschaftswissenschaften, Universitätsstr. 12, 45117 Essen,  
{vorname.nachname}@vwl.uni-due.de

<sup>2</sup> Universität Duisburg-Essen, Paluno - The Ruhr Institute for Software Technology, Gerlingstraße 16, 45127  
Essen, {vorname.nachname}@paluno.uni-due.de

## 2 Verwandte Arbeiten

Educational Data Mining beschäftigt sich unter anderem mit der Prognose von Noten. Einen umfassenden Überblick über die jüngste Literatur bieten [Me16]. Häufig angewandte Educational Data Mining Methoden sind: Naive Bayes, Neural Networks, Support Vector Machines, Entscheidungsbäume [Co17] sowie multiple und logistische Regressionen [MD10]. Während viele Autoren sich auf eine möglichst genaue Prognose der Noten fokussieren (siehe unter anderem [BP11]), ist für Lehrpersonen eine möglichst frühe Prognose von größerer Bedeutung [Me16]. Diese Vorhersagen können genutzt werden, um ein Frühwarnsystem einzuführen, Risikostudierende frühzeitig zu identifizieren und die Durchfallquote durch angemessene Interventionen zu senken [Me16], [MD10]. [Co17] konnten bereits nach der ersten Hälfte eines Onlinekurses (5 von 10 Wochen) zur Einführung in die Programmierung mit einer Genauigkeit von 82 % mittels eines Entscheidungsbaums vorhersagen, ob Studierende den Kurs bestehen oder nicht. [Me16] entwickelten einen Algorithmus, der für alle Studierende individuell zum frühestmöglichen Zeitpunkt die Kursnote vorhersagt. Bereits nach vier Wochen konnte für 85 % der Studierenden eine Prognose mit einer Genauigkeit von 76 % gestellt werden. Beispiele für die Implementierung eines solchen Frühwarnsystems stellen [Bi13] vor.

## 3 Vorlesungskonzept

Im Sommersemester 2017 wurde JACK im Rahmen der Lehrveranstaltung §Induktive Statistik§ an der Universität Duisburg-Essen am Lehrstuhl für Ökonometrie eingesetzt. Die Hörschaft der Veranstaltung setzte sich zu Beginn zusammen aus insgesamt etwa 750 Studierenden aus dem 2. bis 4. Semester der Bachelorstudiengänge VWL, BWL, Wirtschaftsinformatik und Lehramt Wirtschaftswissenschaften.

Die Veranstaltung §Induktive Statistik§ besteht aus einer wöchentlich stattfindenden Vorlesung, in der die statistischen Verfahren gelehrt werden sowie einer ebenfalls wöchentlich stattfindenden Übung, in der die erlernten Methoden in praktischen Anwendungen vertieft werden. Um das aktive, selbständige Lernen der Studierenden zu fördern, wurden, wie bereits in vergangenen Semestern, Aufgaben zum selbständigen Bearbeiten auf der Lernplattform JACK bereitgestellt. Insgesamt wurden 173 Übungsaufgaben freigeschaltet, davon 125 Übungsaufgaben zu klassischem Vorlesungsstoff ohne Programmierinhalte. Die restlichen 48 Aufgaben wurden als R-Aufgaben konzipiert. Durch die Bereitstellung von Feedback zu häufig begangenen Fehlern beim Lösen einer Aufgabe sowie optional wählbarer Hinweise wird der individuelle Lernerfolg der Studierenden unterstützt.

Zur Unterstützung einer kontinuierlichen Arbeitsweise der Studierenden wurden vier Testate abgehalten, für die JACK ebenfalls genutzt wurde. Die Online-Testate hatten eine Dauer von 40 Minuten und fanden zu vorgegebenen Zeiten statt. Dies bot den Studierenden die Möglichkeit ihren persönlichen Wissensstand kontinuierlich über die

Vorlesungszeit hinweg zu überprüfen. Bonuspunkte bildeten den Anreiz zur Ablegung dieser Testate, wobei maximal 2,5 pro Testat im Falle einer am Semesterende bestandenen Klausur angerechnet wurden. Ein zusätzliches 5. Testat am Ende der Vorlesungszeit beinhaltete (im Gegensatz zu den vorherigen Testaten) R-Aufgaben, die ebenfalls zu einer festgelegten Zeit auf der Lernplattform JACK bereitgestellt wurden. Die Anrechnung von maximal 2 Bonuspunkten auf die am Semesterende abzulegende Klausur - auch im Falle des Nichtbestehens - bildete einen zusätzlichen Anreiz. Die abschließenden Klausuren hatten eine Dauer von 70 Minuten und wurden als PC-gestützte Klausuren auf der Plattform JACK abgelegt. 8 der 60 möglichen Punkte wurden in Form von R-Aufgaben abgefragt.

## 4 Bestehensprognose

Dieses Kapitel stellt ein Prognosemodell für den Klausurerfolg vor. Das Fach *Induktive Statistik* stellt eine klassische Einstiegshürde für Studierende dar. Tabelle 1 zeigt eine Übersicht über die Anzahlen an Studierenden im Sommersemester 2017 und die Einreichungen je Gruppe. Eine große Anzahl der Studierenden besuchte die Veranstaltung nur sporadisch und wurde im Laufe des Semesters abgehängt. Der Kurs (evtl. auch das Studium) wurde abgebrochen, bzw. die Klausur verschoben oder nicht bestanden.

Studierende, die	Anzahl	Einreichungen
am Kurs teilgenommen haben	752	163.444
sich zu einer Klausur angemeldet haben	438	152.232
an einer Klausur teilgenommen haben	378	147.868
eine Klausur bestanden haben	185	87.382

Tab. 1: Studierendenzahlen im SS2017

Mit Hilfe der gesammelten Nutzerdaten soll nun vorhergesagt werden, ob Studierende eine Abschlussklausur bestehen oder nicht. Die Daten umfassen alle Online-Einreichungen inklusive des Zeitpunkts und der erreichten Punktzahl von 0 bis 100. Zusätzlich stehen die erreichten Punkte je Testat gemessen in Prozent zur Verfügung.

Für das Prognosemodell wurde das Semester in insgesamt 5 Zeiträume unterteilt. Der 1. Zeitraum beginnt mit Vorlesungsbeginn und endet nach dem 1. Onlinetestat. Danach erstrecken sich die Zeiträume jeweils bis zum nächsten Onlinetestat. Der letzte Zeitraum endet mit dem abgehaltenen R-Testat. Für jeden Zeitraum wurde für jeden Studierenden die Anzahl aller JACK-Einreichungen von Übungsaufgaben, die Anzahl der korrekten

Einreichungen sowie der individuelle Score gespeichert. Der Score ist definiert als Summe aller Punkte der jeweils letzten Einreichung über alle verschiedenen Übungsaufgaben. Dieser ändert sich immer dann, wenn in einer Aufgabe eine andere Punktzahl als beim vorherigen Mal erzielt wurde oder diese zum ersten Mal gelöst wurde. Der Score kann bei allen Studierenden täglich gemessen werden und Werte von 0 bis 100 mal 173 (Anzahl an Übungsaufgaben) annehmen, siehe auch [Ma18].

Außerdem liegt für alle 4 Testate sowie das R-Testat die erreichte Punktzahl in Prozent vor. Die zu vorhersagende Variable nimmt den Wert 1 an, wenn eine der abschließenden Klausuren bestanden wurde. In allen anderen Fälle (nicht angemeldet, nicht erschienen oder durchgefallen) nimmt die Variable den Wert 0 an.

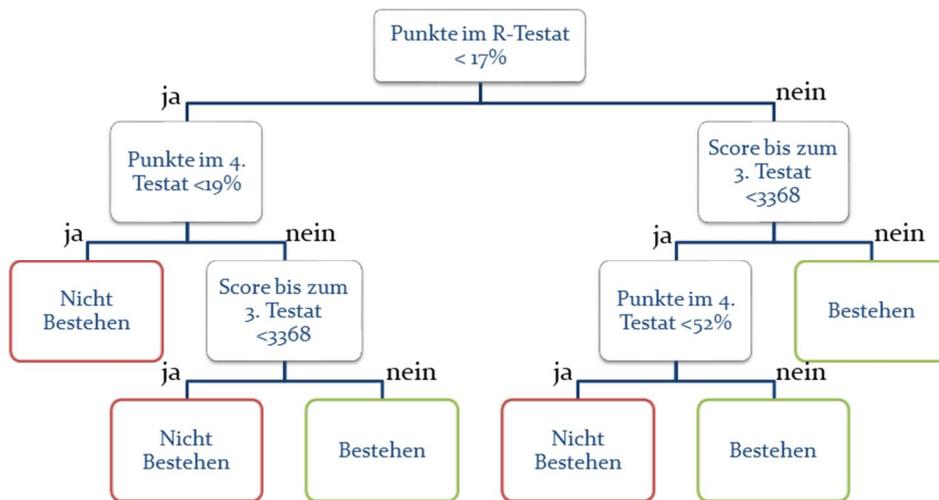


Abb. 1: Entscheidungsbaum zur Klausurprognose

Nutzt man die zuvor beschriebenen Variablen, um einen Entscheidungsbaum aufzustellen, zeigen sich interessante Ergebnisse. Der Entscheidungsbaum wird mit Hilfe des CART-Algorithmus [Br84] erstellt. Dabei werden die Daten zufällig in einen Trainingsdatensatz mit 480 Studierenden und einen Testdatensatz mit 272 Studierenden unterteilt. Die erste Verzweigung des Entscheidungsbaumes nutzt zunächst die erreichten Punkte im R-Testat, siehe Abb. 1. Dies zeigt, dass das Erlernen der Statistiksoftware einen hohen Einfluss auf das Bestehen der Klausur hat. Da der R-Anteil in der Klausur vergleichsweise gering ist, scheinen die R-Übungen einen über Programmierkenntnisse hinausgehenden Einfluss auf das allgemeine Statistikverständnis zu haben, wie es auch [Ot17] dargelegt haben. Hat man in diesem Testat weniger als 17 % erreicht, erfolgt die nächste Aufteilung zum 4. Testat. Wurden in diesem weniger als 19 % der Punkte erreicht, wird ein Nichtbestehen vorhergesagt. Im anderen Fall wird nach dem Score am Tag des dritten Testats entschieden. Ist dieser kleiner als 3368 Punkte, erfolgt eine negative, ansonsten eine positive Bestehensprognose. Verfolgt man den Entscheidungsbaum von der ersten

Verzweigung aus nach rechts, so sind auch hier die Entscheidungsvariablen der Score bis zum 3. Testat und die Punkte im 4. Testat.

Wendet man den Entscheidungsbaum auf die Testdaten an, so trifft man in 86 % der Fälle eine richtige Entscheidung. Ein Problem von Entscheidungsbäumen ist die hohe Varianz, weswegen man oft auf Random Forests zurückgreift. Diese stellen eine große Anzahl an verschiedenen Bäumen auf und mitteln abschließend aus allen Bäumen. In diesem Fall jedoch erzielt ein Random Forest aus insgesamt 1000 Bäumen eine Genauigkeit am Testdatensatz von 86 %, was der Genauigkeit des zuvor aufgestellten Baumes entspricht. Der Entscheidungsbaum sagt im Testdatensatz insgesamt 87 % der Studierenden, welche die Klausur nicht bestanden haben, richtig vorher. Der Anteil der richtig vorhergesagten Bestehenden beträgt 80 %.

## 5 Ausblick & Implementierung

Die Informationen aus diesem Prognosemodell können auf verschiedene Weise helfen, das formative Assessment in Zukunft zu verbessern. Dabei kann das errechnete Modell auch auf künftige Veranstaltungen angewandt werden. Im Moment wird JACK deshalb um Kurse erweitert, in denen einerseits Studierenden Feedback zu ihrem Lernverhalten gegeben und andererseits eine adaptive Auswahl von Aufgaben aufgrund des Modells angeboten wird.

So wird es beim Betreten eines solchen Kurses ein UI-Element geben, welches Studierenden, denen  $\hat{\omega}$  die Klausur nicht bestehen  $\hat{\omega}$  prognostiziert wird, eine entsprechende Nachricht anzeigt. Hierbei muss eine Formulierung gefunden werden, die nicht demotivierend wirkt. Ferner sollen die Studierenden explizit darauf hingewiesen werden, dass diese Aussage natürlich mit einem statistischen Unsicherheitsfaktor zu verstehen ist. Eine Forschungsfrage wird dann sein, ob solche Hinweise die schwächeren Studierenden genug motivieren, um eine relevante Verbesserung in der Bestehensquote zu erhalten. Weiterhin ist offen, ob auch Studierenden, die mit der jetzigen Leistung vermutlich eher bestehen werden, ein solcher Hinweis angezeigt werden soll. Hierbei könnte ein positiver Effekt dadurch zunichte gemacht werden, dass Studierende annehmen könnten, nicht mehr lernen zu müssen.

Ferner kann das Modell dazu genutzt werden, Kurse in JACK adaptiv zu gestalten. Bis jetzt unterscheidet das Modell zwischen Lernständen zu verschiedenen Zeitpunkten. In Zukunft soll es erweitert werden, um zusätzlich zwischen Aufgaben, die einen großen Effekt auf die Erfolgsprognose haben und solchen, die weniger geeignet sind, einen nützlichen Beitrag zur Prognose zu leisten, zu unterscheiden. Dies ließe sich sehr gut mit der im letzten Absatz erwähnten Warnung kombinieren, die dann nicht nur eine Aussage dazu treffen könnte, dass generell mehr gelernt werden muss, um die Klausur zu bestehen, sondern auch konkret wie dies am besten zu erreichen ist.

Durch die Kombination dieser beiden Features und damit besserem Feedback zum aktuellen Lernstand zu jeder Zeit im Semester, erhoffen wir uns eine wesentlich bessere Introspektion und Motivation insbesondere der schwächeren Studierenden und in Konsequenz eine weitere Verbesserung der Bestehensquote in der Statistikausbildung. Diese Hypothese werden wir nach erfolgreicher Implementierung in der nächsten Vorlesung „Induktive Statistik“ überprüfen.

## Danksagung

Diese Arbeit wurde gefördert durch das Projekt „Professionalisierung für Vielfalt“ der Qualitätsoffensive Lehrerbildung des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01 JA 1610.

## Literaturverzeichnis

- [Bi13] Bin Mat, U. et al.: An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention. IEEE 5th Conference on Engineering Education (ICEED) 2013.
- [BP11] Baradwaj, B.K.; Pal, S.: Mining Educational Data to Analyse Students' Performance. International Journal of Advanced Computer Science and Applications 2/06, S. 63-69, 2011.
- [Br84] Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J.: Classification and regression trees. Wadsworth International Group, Belmont, California, 1984.
- [Co17] Costa, E.B. et al.: Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming. Computers in Human Behavior 73/01, S. 247-256, 2017.
- [Ma18] Massing, T. et al.: E-Assessment Using Variable-Content Exercises in Mathematical Statistics. Wiedereingereicht beim Journal of Statistics Education.
- [MD10] Macfadyen, L.P.; Dawson, S.: Mining LMS data to develop an "early warning system" for educators: A proof of concept. Computers and Education 54/02, S. 588-599, 2010.
- [Me16] Meier, Y. et al.: Predicting Grades. IEEE Transactions on Signal Processing 64/04, S. 959-972, 2016.
- [Ot17] Otto, B. et al.: Evaluation einer Statistiklehrveranstaltung mit dem JACK R-Modul. Bildungsräume 2017.
- [Sc17] Schwinning, N. et al.: Towards digitalization of summative and formative assessment in academic teaching. Proceedings of LaTiCE 2017.