

# [Titel der Ausarbeitung]

[Untertitel der Ausarbeitung]

**Ausarbeitung**

im Studiengang Medieninformatik

Fachbereich Medien  
Hochschule Düsseldorf

Direnc Timur  
Matrikel-Nr.: 123456  
Datum: Juni 2022

Till Pilarczyk  
Matrikel-Nr.: 765335  
Datum: Juni 2022

Prüfer  
Prof. Dr.-Ing. Thomas Rakow

# Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Ausarbeitung selbständig und ohne unzulässige fremde Hilfe angefertigt habe. Die verwendeten Quellen sind vollständig zitiert. Diese Arbeit wurde weder in gleicher noch ähnlicher Form einem anderen Prüfungsamt vorgelegt oder veröffentlicht. Ich erkläre mich ausdrücklich damit einverstanden, dass diese Arbeit mittels eines Dienstes zur Erkennung von Plagiaten überprüft wird.

*Ort, Datum*

*Direnc Timur*

*Ort, Datum*

*Till Pilarczyk*

## **Kontaktinformationen**

Direnc Timur

direnc.timur@study.hs-duesseldorf.de

—

Till Pilarczyk

till.pilarczyk@study.hs-duesseldorf.de

# Zusammenfassung

[Titel der Ausarbeitung]

Direnc Timur

Dies ist die Zusammenfassung Ihrer Arbeit ...

# Abstract

[Titel der Ausarbeitung]

Direnc Timur

Diese möchten Sie natürlich auch auf Englisch bereitstellen . . .

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Aktuelle Forschung</b>	<b>3</b>
2.1	Whitepaper künstliche Intelligenz in der Hochschullehre . . . . .	3
2.1.1	Lehren und Lernen mit KI in der Hochschulausbildung . . . .	3
2.1.2	Lehren und Lernen über KI in der Hochschulausbildung . . . .	5
2.1.3	KI und Ethik in der Hochschulbildung . . . . .	6
2.2	AIED . . . . .	6
2.3	Pass/Fail Prediction with Moodle . . . . .	9
2.3.1	Vorgehen . . . . .	9
2.3.2	Ergebnis . . . . .	10
2.3.3	Fazit . . . . .	11
<b>3</b>	<b>Fazit</b>	<b>12</b>
<b>4</b>	<b>Ausblick</b>	<b>13</b>

# Abbildungsverzeichnis

1.1	Anzahl der Pressemitteilungen von Hochschulen und weiteren Wissenschafts-Institutionen, die das Schlagwort „Künstliche Intelligenz“ enthalten [3, S. 9]	1
-----	---	---

# Tabellenverzeichnis

# Abkürzungsverzeichnis

**AIED** Artificial Intelligence in Education

**KI** Künstliche Intelligenz

**LMS** Lernmanagementsystem



# Kapitel 1

## Einleitung

Das Thema Künstliche Intelligenz (KI) erlangt zunehmend an Relevanz und durchdringt viele Felder. Auch die Hochschullehre kann vom Einsatz künstlicher Intelligenzen profitieren. Es existieren erste Versuche intelligente Systeme in die Lehre zu integrieren, wie etwa Chatbots in eingesetzten Lernmanagementsystemen, qualitative Auswertungen von Lernendendaten durch Learning Analytics und weitere [4, S. 18, S. 14ff.].

Abbildung 1.1 veranschaulicht die Entwicklung der Verbreitung von KI generell und in der Kategorie „Studium und Lehre“. Es ist eine Tendenz dahin erkennbar, dass generell künstliche Intelligenz zunehmend vorkommt. Zudem ist KI in der Rubrik Studium und Lehre zwar relativ seltener vertreten, aber die Tendenz steigt.

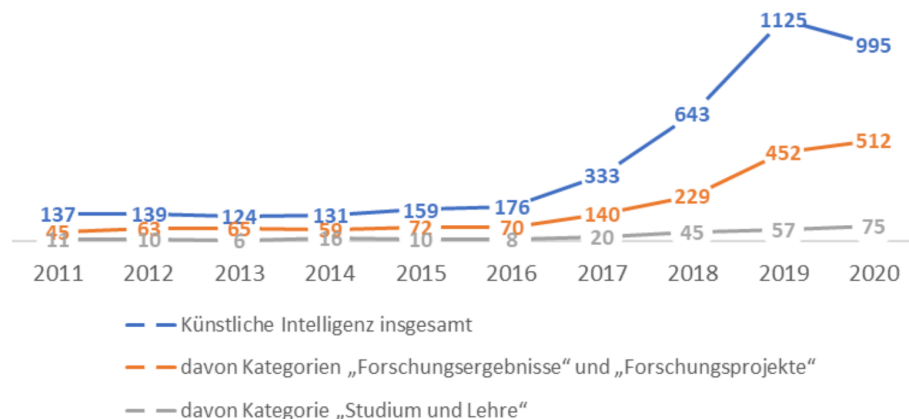


Abbildung 1.1: Anzahl der Pressemitteilungen von Hochschulen und weiteren Wissenschafts-Institutionen, die das Schlagwort „Künstliche Intelligenz“ enthalten [3, S. 9]

Zugriffsdaten sind Daten, die beim Zugriff auf (Web-)Ressourcen erzeugt werden. Sie können bspw. für das Reproduzieren und Beheben von Fehlern eingesetzt werden. Daher ist davon auszugehen, dass sie in Systemen in großen Mengen vorhanden sind. Sie können jedoch auch für Analysezwecke genutzt werden, um vom Zugriffsverhalten Rückschlüsse auf andere Erkenntnisse zu erlangen.

Für die Entwicklung von KIs sind Daten notwendig, welche für das Training einer der KI eingesetzt werden. Hochschulen setzen in großen Teilen bereits Systeme für die Lehre ein und besitzen demnach Daten, besonders Zugriffsdaten. Die zentrale Ar-

beitsfrage dieser Ausarbeitung ist, wie Zugriffsdaten für den Einsatz von künstlichen Intelligenzen im Kontext der Hochschullehre eingesetzt werden können.

Hierfür wird zunächst eine Einführung in das Themenfeld KI im Kontext der Hochschullehre gegeben. Die Paper „Künstliche Intelligenz in der Hochschullehre“ von de Witt, Rampelt und Pinkwart und „Another 25 Years of AIED? Challenges and Opportunities for Intelligent Educational Technologies of the Future“ von Pinkwart geben einen Überblick über aktuelle Entwicklungen in der Hochschullehre mit KIs. Hierfür werden die Erkenntnisse der Paper herausgestellt. Anschließend folgt eine Analyse einer KI, die im Lernmanagementsystem (LMS) Moodle die Zugriffsdaten des LMS auswertet. Die Ergebnisse werden im Fazit zusammengefasst und eine Bewertung zu den Entwicklungen gegeben.

# Kapitel 2

## Aktuelle Forschung

In diesem Kapitel werden einige Paper vorgestellt, in den sich mit dem Thema KI an Hochschulen oder KIs an Hochschulen, die mit Zugriffsdaten arbeiten vorgestellt.

### 2.1 Whitepaper künstliche Intelligenz in der Hochschullehre

Das Paper „Whitepaper künstliche Intelligenz in der Hochschule“ umfasst eine Sammlung von verschiedenen Werken zum Thema Künstliche Intelligenz in der Hochschullehre. Das Paper ist in die folgenden vier Kernthemen gegliedert:

1. Lehren und Lernen mit KI in der Hochschulausbildung
2. Lehren und Lernen über KI in der Hochschulausbildung
3. KI und Ethik in der Hochschulausbildung
4. Zukunftsperspektiven für KI in der Hochschulbildung

Nachfolgend folgt eine inhaltliche Zusammenfassung der vier Kernthemen. Dabei wird auf die wichtigsten Erkenntnisse und Aussagen eingegangen.

#### 2.1.1 Lehren und Lernen mit KI in der Hochschulausbildung

Beim Kapitel „Lehren und Lernen mit KI in der Hochschulausbildung“ geht es darum, wie künstliche Intelligenzen eingesetzt werden können, um die Lehre sowohl aufseiten der Lernenden, als auch für die Lehrenden und Bildungsorganisationen zu verbessern. Es werden mehrere Gründe genannt, warum KI eingesetzt werden sollte:

- Verwendung von Daten zur Unterstützung von Studierenden bei der Weiterentwicklung des eigenen Lernverhaltens und Lehrenden bei der Verbesserung ihrer Didaktik.
- Förderung von kritischem und kreativem Denken
- Entdeckung neuer Erkenntnisse durch Analyse von Daten
- Vorzeitige Erkennung von Risiko-Studierenden

[4, S. 11f.]

Die Gründe werden davon gestützt, dass eine Reihe von Herausforderungen in der aktuellen Lehre bestehen. Als Beispiel wird genannt, dass die Zeit von Lehrenden begrenzt ist, sodass sie Lernende nicht immer vollständig betreuen können. Als mögliche Maßnahme werden virtuelle intelligente Assistenten aufgeführt, die bspw. Lernende anstelle von Lehrenden betreuen und Hilfestellungen leisten. Auch die Auswertung von längeren Texteinreichungen wird als aktuelle Herausforderungen herausgestellt. Hierbei werden Automated Essay Scoring Anwendungen vorgeschlagen, die solche Texteinreichungen systematisch analysieren, bewerten und Lernenden Feedback geben können. [4, S. 12]

#### **Learning Analytics**

Ein aufgeführtes Konzept zur intelligenten Analyse von Leistungsdaten von Lernenden ist Learning Analytics. Learning Analytics kann als Prozess verstanden werden, welcher den Prozess des Lernens unterstützt. Der Prozess wird in drei Ebenen eingeteilt: Mikroebene, Mesoebene und Makroebene. Der Prozess kann je Ebene unterschiedliche Zielgruppen unterstützen, wie Lernende, Lehrende oder auch gesamte Organisationen. Die Ebenen beziehen sich dabei auf den Umfang und zeitlichen Kontext des Lernens. In der Mikroebene können Lernende bspw. im Rahmen einer Lernsitzung oder im Verlaufe eines Kurses unterstützt werden. In der Mesoebene können Lernende über ein gesamtes Semester begleitet und Daten kursübergreifend ausgewertet werden. Auf Grundlage dieser Daten kann Learning Analytics auch für die Erkennung von gefährdeten Lernenden eingesetzt werden. [4, S. 14ff.]

Auch bei Learning Analytics bestehen Herausforderungen, die es zu adressieren gilt. Allen voran ist zu beachten, dass Learning Analytics bei der Auswertung lediglich einen Ausschnitt des Lernprozesses berücksichtigen kann. Weiterhin sollten intelligente Systeme die persönlichen Eigenschaften sowie den sozialen und kulturellen Kontext der Lernenden berücksichtigen. [4, S. 16]

#### **Arten von Künstlichen Intelligenzen in der Hochschulbildung**

Das Whitepaper listet eine Reihe von Arten künstlicher Intelligenzen in Hinsicht auf die Hochschullehre auf.

**Personalisierte adaptive Lernumgebungen (ALU)** Unter Personalisierten adaptiven Lernumgebungen werden digitale Lernumgebungen für die autonome Lehre gut strukturierter Inhalte verstanden. Dabei werden drei Modelle unterschieden: Domänenmodell (Fach), Lernermodell und Didaktikmodell. Das Domänenmodell behandelt den Inhalt des jeweiligen Faches. Das Lernermodell betrifft den Zustand des Lernenden und welches Wissen er besitzt. Das Didaktikmodell schlägt auf Basis der Erkenntnisse des Lernermodells Lernstrategien und Handlungsempfehlungen für Lernende vor. [4, S. 17f.]

**Chatbots** Lernende können durch Chatbots moderiert und unterstützt werden, wobei sie gleichzeitig Lehrende entlasten. Als Beispiel wird ein Chatbot in einem Hochschulforum aufgeführt, welche Fragestellungen von Lernenden anstelle von Lehrenden beantwortet. [4, S. 18ff.]

Darüber hinaus können auch Empfehlungssysteme, Edu-Robots und KI-Schreibbots eingesetzt werden. [4, S. 18-21]

### **Veränderung der Betreuung**

Anstelle von klassischer Lehre fachlichen Wissens, sollen Lernende mehr zu kritischem Denken gefördert werden. Das schließt nicht aus, dass Fachkräfte für spezielle Fachrichtungen benötigt werden. Lernende sollen jedoch mehr dazu gebracht werden Wissen zu generieren, anstatt nur vermittelt zu bekommen. Dadurch wird das Risiko eingegangen, dass nicht zu einer „Allwissenheit“ ausgebildet wird. Es wird vermutet, dass sich die Betreuung von Lernenden durch Lehrende im Sinne eines Mentoring verändern wird. Beim Mentoring ist die Erkennung des Status des Lernenden wichtig, um darauf reaktiv handeln zu können. Das Mentoring wird in die Phasen Vorbereitung, Lernprozess und Nachbereitung aufgeteilt.

- Vorbereitung: Erhebung der Lerninhalte und Planung der Ausgestaltung des Lernprozesses.
- Lernprozess: Ausführung der geplanten individualisierten Aktivitäten zur An eignung der Lerninhalte.
- Nachbereitung: Reflexion des Lernprozesses über Lernerfolg.

[4, S. 22f.] Es wurde gezeigt, dass intelligente Tutor- und Lernsysteme zur Verbesserung der Lernergebnisse führen können. Es bestehen allerdings heute noch fehlende Tutoring- und Lernanwendungen mit KI. Bisher wurden nur Vorhersage von Erfolg und Misserfolg des Lernprozesses eingesetzt bzw. geprüft. [4, S. 24]

### **2.1.2 Lehren und Lernen über KI in der Hochschulausbildung**

Damit KI effektiv eingesetzt werden kann, muss auch Wissen über KI vermittelt werden. Dies beginnt bei der Verwendung von KI und geht über die Auswahl geeigneter KI bis hin zur Entwicklung neuer KI. Es wird erwartet, dass die Nachfrage an Fachkräften im Bereich KI stark wachsen wird. Daher sollte der Fokus mehr auf Lehre über KI gesetzt werden. [4, S. 26]

Damit die Lehre über KI stattfinden kann, müssen einige Herausforderungen beachtet und gelöst werden.

#### **Herausforderungen bei der Lehre über KI**

- Uneinheitliches Verständnis über KI.
- Aktuell am meisten Verwendet in Ingenieurwissenschaften, Tendenz interdisziplinär.
- Fehlende Standards digitaler Lernangebote.

[4, S. 27]

### **Digitale Lehre im Bereich KI**

Es wird gefordert, dass mehr qualifizierte digitale Lernangebote zu KI angeboten werden. Angebote sollten dabei für mehrere Zielgruppen zugänglich und verwendbar sein. Es sollen demnach auch Angebote vorhanden sein, die keine tiefen technischen Kenntnisse erfordern. Zudem sollten Lernangebote offen lizenziert und transparent sein. [4, S. 28]

In Deutschland existieren bereits digitale Plattformen, auf denen auch Lernangebote zum Thema KI zu finden sind. Darunter fallen bspw. oncampus von der TH Lübeck, OPEN VHB von der virtuellen Hochschule Bayern, sowie OER-Content.nrw vom Bundesland Nordrhein-Westfalen. Insgesamt existieren noch zu wenig Lernangebote auf den Plattformen. [4, S. 28ff.]

#### **2.1.3 KI und Ethik in der Hochschulbildung**

KIs als autonome Systeme interagieren mit Menschen und geben Handlungsempfehlungen auf Grundlage von Daten der Menschen. Aus diesem Grund besteht das Risiko, dass diese Systeme den Menschen Schaden zufügen können. Dieses Risiko ist ein wichtiger Faktor bei der Berücksichtigung der Ethik von KIs im Zusammenhang mit dem Menschen. [4, S. 38]

**Qualität der (Trainings-)Daten** Die Qualität der Daten beeinflusst unmittelbar die Ergebnisse der KI. Bestehen Tendenzen zu Diskriminierung, Unterrepräsentation, Rassismus oder ähnlicher negativer sozialer Einstellungen in den Trainingsdaten einer KI, greift die KI diese auf und integriert sie in Entscheidungen und Handlungsempfehlungen. Daher gilt es, die Qualität der Daten zu sichern. [4, S. 39]

**Sensibilisierung für KI** Eine Analyse der Daten von Menschen und anschließende Auswertungen können bei Menschen Misstrauen durch entstehende Gefühle von Überwachung und Kontrolle erzeugen. Weiterhin wird durch die Verwendung unterschiedlichster Daten zur Fremdbestimmung statt Selbstbestimmung tendiert. Daher müssen Sensibilisierungsmaßnahmen getroffen werden, damit die Akzeptanz und das Verständnis gegenüber KIs erhöht werden. Die kann bspw. durch frühzeitige Aufklärungsarbeiten über die Funktionsweise und Herausforderungen von KIs stattfinden. Ein notwendiges Mittel für die Transparenz von KIs ist die Erklärbarkeit, welches im Rahmen der Disziplin Explainable AI erforscht wird. [4, S. 39ff.]

## **2.2 AIED**

Professor Niels Pinkwart ist bei der Humbolt-Universität in Berlin als Leiter für „Didaktik der Informatik / Informatik und Gesellschaft“ tätig. Er hat über 250 Publikationen veröffentlicht. Darunter auch den Artikel „Another 25 Years of AIED? Challenges and Opportunities for Intelligent Educational Technologies of the Future“, der im Jahre 2016 im „International Journal of Artificial Intelligence in Education“ veröffentlicht wurde. Dieser Artikel wurde laut der Webseite Springer 45 mal von

anderen Autoren zitiert.<sup>1</sup>

In dem Artikel wird versucht die aktuellen Trends und Entwicklungen in der Informatik, der Lehre und der Lehre mit Technologien zu analysieren. Dies hilft dabei rückschlüsse auf die Weiterentwicklung von Artificial Intelligence in Education (AIED) (dt. Künstliche Intelligenz in der Lehre) in den nächsten 25 Jahren schließen zu können. Basierend auf diesen Trends werden zwei mögliche Zukunftsszenarien vorgestellt. Es gibt eine Utopie und eine Dystopie. Zuletzt werden mögliche Herausforderungen und Chancen für die Entwicklung von AIED Systeme identifiziert, die in den nächsten 25 Jahren entstehen können. Da es nicht möglich ist 25 Jahre in die Zukunft zu schauen, kann der Artikel in Zukunft benutzt werden um zu schauen wie sich AIED entwickelt hat im Vergleich zu dem was gedacht wurde. [1, S. 1f]

Professor Niels nennt sieben Herausforderungen, die sich aus der Untersuchung und den Szenarien hervorgehen, sind im folgenden genannt und erläutert.

1. intercultural and global dimensions (dt. Interkulturelle und Globale Dimensionen)

Lernsoftware kann in einem kulturellen Kontext sehr gut funktionieren in einem anderen nicht. Dies liegt an vielen verschiedenen Faktoren, der Sprache, den verschiedenen Inhalten der Lehrpläne, der Lehrkultur und die unterschiedlichen Art und Weisen, wie die Interaktion zwischen Lernenden und Lehrenden ist. Es kann vermutet werden, dass in 25 Jahren überall auf der Erde Internet verfügbar sein, und damit möglich jedem Lernenden AIED Systeme zur Verfügung gestellt werden kann. Es muss daher in dem Bereich Interkulturelle Lehre geforscht werden, damit AIED Systeme entwickelt werden können, die allen Lernenden hilft. [1, S. 9f]

2. practical impact (dt. praktische Auswirkungen)

Es gibt bereits einige AIED Systeme in Schulen eingesetzt werden, diese passiert oft mit kommerziellen Initiativen innerhalb des AIED-Bereichs. Diese Verbindungen zu Schulen sind zwar wichtig, um den Einfluss von AIED Systemen auf die Bildung weiter zu erhöhen. Dennoch werden in Forschungssystemen gewisse Interessensgruppen wie Lehrkräfte und Schulen oft nicht eingebunden. Besonders in Schulen ist die Frage, nach technischem Support, Fehlerbehebungen, Stabilität und kontinuierliche Verfügbarkeit sehr wichtig und wird häufiger nicht ausreichend beachtet. Daher besteht die Herausforderung darin, Wege zu finden die über Forschungssysteme hinausgehen und dauerhaft bestehen. [1, S. 10]

3. privacy (dt. Privatsphäre)

Dadurch das AIED Systeme als Begleiter fürs Lebenslange Lernen dienen sollen, werden viele Daten benötigt und gesammelt [1, S. 10]. Diese müssen geschützt werden, damit diese z. B. nicht genutzt werden um möglicherweise Kandidaten bei einer Bewerbung im Vorhinein auszusortieren, sollten die Daten von Firmen verkauft werden [1, S. 9f]. Dies kann passieren, wenn AIED Systeme von privaten Unternehmen entwickelt werden. Daher spielt der Datenschutz eine sehr wichtige Rolle. Die Regeln könnten sich in Zukunft verschärfen. Dies ist

---

<sup>1</sup><https://link.springer.com/article/10.1007/s40593-016-0099-7>, zuletzt aufgerufen am 10.07.2022

zwar für den Benutzer gut, stellt jedoch für die Unternehmen eine weitere Herausforderung da. Diese möchten Gewinne erzielen, diese werden häufiger oft mit Daten der Nutzer erzielt. Ist dies durch die strikteren Regeln nicht mehr möglich, muss eine andere Möglichkeit gesucht werden. Eine der Möglichkeiten wäre, Werbung schalten, dies würde aber den Lernfluss stören. Eine andere Möglichkeit wäre die Lizenzierung, dann könnte jedoch das Problem entstehen, dass sich einige Institutionen sich diese Software nicht leisten können. Sollte also mit den Daten der Lernenden Gewinn erzielt werden, werden klare und für den Nutzer transparente Regeln nötig. [1, S. 10f]

#### 4. Interaction methods (dt. Interaktionsmethoden)

Die vierte Herausforderung sind die Interaktionsmethoden. Durch die Entwicklung von Mensch-Computer-Interaktion ist die Interaktion der Lernenden und den AIED Systemen einfacher und intuitiver als früher. Daher ist es möglich, dass sich dieser Trend weiter fortsetzt. Es muss überlegt werden, wie die Lernenden in Zukunft mit AIED Systemen umgehen sollen. Es muss für die Lernenden möglich sein, leicht Informationen einzugeben und bequemes Feedback zurückzubekommen. Die Schwierigkeit besteht darin, die Eingaben, die über unterschiedliche Kanäle (z. B. Sprache oder Text) besonders gut und ohne Störungen zu empfangen, damit diese analysiert werden können. [1, S. 11]

#### 5. collaboration at scale (dt. Zusammenarbeit im großen Maßstab)

Es gibt viele globale Online-Kurse, die tausenden von Lernenden erreichen. Diese setzen jedoch nur auf einfache Bildungstechnologien wie Video, Texte, Multiple-Choice-Tests und Diskussionsforen. Möglichkeiten und einige Grundlagen für das globale gemeinsame Lernen mit intelligenten Techniken, wie eine personalisierte Anleitung und zum Zusammenarbeiten sind schon vorhanden. Damit Lehrkräften und Teilnehmern an großen Online-Kursen besser unterstützt und gefördert werden, muss an diesen Techniken noch weiter geforscht werden. [1, S. 11]

#### 6. effectiveness in multiple domains (dt. Effektivität in verschiedenen Bereichen)

Die sechste Herausforderung besteht darin, dass AIED Systeme auch in anderen Bereichen aktiv eingesetzt werden, die heutzutage nicht vertreten sind. Solche Bereiche sind die, die nicht genau definiert sind. Heutzutage gibt es viele AIED Systemen in Bereichen, wo die Domäne bekannt und gut modelliert werden kann z. B. Geometrie. Damit in Zukunft auch Ethikaufsätze oder ein Virtual-Reality-Chirurgietraining mithilfe von AIED Systemen abgedeckt werden können, wird weitere Forschung und Entwicklung in den wissenschaftlichen Bereichen wie der Technologie zur Verarbeitung natürlicher Sprache oder der Wissensdarstellung benötigt. [1, S. 11f]

#### 7. role of aAIED in educational technology (dt. Rolle von AIED in Bildungstechnologien)

In vielen Bereichen wie z. B. die Mensch-Computer-Interaktion gibt es bereits globale Konferenzen (CHI-Konferenz) und Veranstaltungen die einen gewissen Prestige vorweisen und die weitere Entwicklung des Bereichs stark verändern können. Im Bereich der Bildungstechnologien gibt es derzeit nichts vergleichbares. Es gibt jedoch eine große Bandbreite an Veranstaltungen und damit viele



Richtungen in welche die Forschung gehen kann. Dies ist erstmal nichts negatives, da sich so ein Bereich auch weiterentwickelt. Die Herausforderung besteht nun darin eine globale Gemeinschaft der Forschung im Bereich der Bildungstechnologie aufzubauen und so starke Fragmentierung zu vermeiden. Eine zu starke Fragmentierung kann dazu führen, dass sich ein Bereich nicht stark und gut weiterentwickelt und damit keine Gesamtwirkung nachweisen kann. [1, S. 12]

## 2.3 Pass/Fail Prediction with Moodle

Rory Quin und Professor Geraldine Gray sind an der „Technological University Dublin“ im Fachbereich für „Informatics and Engineering PhD “ angestellt. Sie haben in dem Artikel „Prediction of student academic performance using Moodle data from a Further Education setting“ welches im Irish Journal of Technology Enhanced Learning im Jahre 2019 veröffentlicht wurde und nach researchgate 14 mal Zitiert worden ist, untersucht ob es Möglich ist anhand von Zugriffsdaten des Lernmanagementsystems Moodle die an einer Hochschule gesammelt wurden, vorhersagen über die Zukünftigen Laufbahn eines Students innerhalb eines Kurses zu treffen. <sup>2</sup> Diese Art von Forschung wurde schon von einigen anderen Professoren und Wissenschaftlern untersucht, jedoch immer nur an Universitäten und nie mit Daten von einer Hochschule. [2, S. 5]

Rory Quin und Professor Geraldine Gray versuchen in dem Paper die Frage, ob es möglich ist, mit Aktivitätsdaten von Moodle, die während der Dauer eines Kurses gesammelt werden vorherzusagen, wie gut ein Student in dem Kurs abschneidet und damit auch ob dieser den Kurs besteht oder durchfällt zu beantworten. Ebenso versuchen Sie die Frage, ob es möglich ist, mit den selben Daten jedoch nur von den ersten sechs bzw. zehn Wochen vorherzusagen ob ein Student den Kurs bestehen oder durchfallen wird. [2, S. 5] Dies hat den Grund, dass Dozenten Präventionen einleiten können für Studenten die Gefahr laufen den Kurs nicht zu bestehen. Dies würde nicht nur den Studenten helfen sondern auch der Hochschule und dem Dozenten, denn diese würde langfristig den Zeit und Geld sparen. [2, S. 15]

### 2.3.1 Vorgehen

Um die Fragen zu beantworten werden Daten von insgesamt 29 Kursen von 9 verschiedenen Modulen im selben Fachbereich und unter der Leitung eines Dozenten genommen. Die Kurse fanden zwischen den Jahren 2011 und 2018 statt. Die Kurse hatten unterschiedlich lange Laufzeiten, der Kürzeste Kurs hat eine Länge von 11 Wochen und der Längste 33 Wochen. Insgesamt konnten so Daten von 690 Kursteilnehmern. Die 690 Kursteilnehmer setzten sich aus insgesamt 410 Studierenden zusammen. [2, S. 6]

Die erste Klassifizierungsaufgabe besteht darin den Notenbereich der Studierenden hervorzusagen und damit ob diese den Kurs bestehen oder durchfallen. Es gibt die Klassen, Early Exit (Abbruch), Fail (durchgefallen), Pass (bestanden mit 50% bis

---

<sup>2</sup>[https://www.researchgate.net/publication/340908692\\_Prediction\\_of\\_student\\_academic\\_performance\\_using\\_Moodle\\_data\\_from\\_a\\_Further\\_Education\\_setting](https://www.researchgate.net/publication/340908692_Prediction_of_student_academic_performance_using_Moodle_data_from_a_Further_Education_setting), zuletzt aufgerufen am 10.07.2022

64%), Merit (bestanden mit 65% bis 79%), Distinction (bestanden mit 80% bis 100%). [2, S. 6f]

Um die zweite Frage zu beantworten werden die ersten beiden Klassen Early Exit und Fail, sowie die anderen drei Klassen zusammengefasst. Es gibt jetzt nur noch die Klassen bestanden und durchgefallen. [2, S. 6]

Um die Fragen zu beantworten werden Zugriffsdaten verwendet. Es werden z. B. die Anzahl an Log einträgen, die Anzahl an Tagen an denen sich die Studierenden eingeloggt haben und die Anzahl wie oft sich die Studierenden die Aufgabe angeschaut haben genutzt um die Vorhersage zu treffen. Diese sind jedoch nur einige der Zugriffsdaten. Insgesamt wurden 21 solcher Daten verwendet. [2, S. 7f]

Es werden vier verschiedene Algorithmen getestet, um die bestmögliche Vorhersagen zu treffen. Der Random Forest, Gradient Boosting, k Nearest Neighbours und Linear Discriminant Analysis. Es werden sich für die Algorithmen getestet, da diese in anderen Forschungen bereits gute Ergebnisse erzielt. Jeder der Algorithmen wird mit 70% der Daten trainiert. Die restlichen 30% werden benutzt um die Genauigkeit zu überprüfen. Die Genauigkeit wird außerdem mit der „no-information rate“ verglichen. Diese gibt die Genauigkeit an, wenn immer die am häufigsten vorkommende Klasse vorhergesagt wird. [2, S. 8f]

### 2.3.2 Ergebnis

Das Trainieren und Testen mit allen Daten ergibt, dass alle Algorithmen eine ähnliche Genauigkeit haben, wenn es darum geht den Notenbereich vorherzusagen. Sie unterscheiden sich nur in wenigen Prozentpunkten. Der Random Forest Algorithmus trifft mit einer Genauigkeit von 60,5% die besten Vorhersagen. Wenn es nur darum geht vorherzusagen, ob ein Student den Kurs besteht oder durchfällt hat der Random Forest Algorithmus eine Genauigkeit von 92,2% erreicht. Auch hier schneidet der Algorithmus am besten ab. Mit den 92,2% ist der Algorithmus Signifikant besser als die „no information rate“. Diese liegt aufgrund der Tatsache, dass es mehr Studierende gab die Kurse bestanden haben, als die die nicht bestanden haben bei 73,5%. Der Random Forest Algorithmus kann insgesamt 121 von 152 Studierende die den Kurs bestanden haben richtig vorhersagen und 42 von 54 Studierende die den Kurs nicht bestanden haben. [2, S. 9f]

#### Frage 2

Das Trainieren und Testen mit Daten von sechs Wochen ergibt, dass kein Algorithmus bessere Voraussagen treffen kann, als wenn man die „no information rate“ benutzt, die bei 75,5% liegt. Die unterschiedliche „no information rate“, setzt sich daraus zusammen, dass noch nicht alle Studierenden innerhalb der ersten 6 Wochen des Kurses Moodle benutzen. [2, S. 10f]

Wurden hingegen die Daten der ersten zehn Wochen benutzt kann eine Signifikante Verbesserung gegenüber der „no information rate“ erreicht werden. Auch hier schneidet der Random Forest Algorithmus am besten ab. Er erreichte eine Genauigkeit von 82,18% wenn es darum ging zu bestimmen ob Studierende den Kurs bestehen oder nicht. [2, S. 10f] Die „no information rate“ liegt bei 74,8%.

Am meisten Einfluss auf die Vorhersage, wenn alle Daten oder die der ersten 10 Wochen benutzt wurden, hatte die Anzahl an Tagen an denen sich die Studierenden in den Kurs eingeloggt haben. [2, S. 15]

### 2.3.3 Fazit

Die Testergebnisse zeigen, dass es möglich ist mithilfe von Zugriffsdaten die von Moodle im Laufe eines Kurses gesammelt werden Signifikant bessere Vorhersagen über die Noten der Studierenden zu treffen als mit der „no-information rate“. Dabei haben alle Algorithmen gut abgeschlossen und sich nur innerhalb von maximal fünf Prozentpunkten unterschieden. Der Random Forest Algorithmus schließt dabei mit 60.5% am ab. Der Random Forest Algorithmus kann 97% der Studierenden die den Kurs bestehen und 78% der Studierenden die den Kurs nicht bestehen vorhersagen. [2, S. 14]

Es reicht jedoch nicht aus am Ende eines Kurses vorherzusagen zu können ob ein Studierender den Kurs besteht oder nicht. Ein frühzeitiger Hinweis welche Studierenden möglicherweise nicht bestehen wäre für die Lehrenden hilfreich, da sie so gezielt Hilfe leisten können. Das Testergebnis zeigt, dass es nicht möglich ist mit Daten von sechs Wochen festzustellen welche Studierenden möglicherweise Durchfallen. Mit den gesammelten Daten von zehn Wochen kann jedoch eine Signifikant bessere Vorhersage getroffen werden als mit der „no-information rate“. Trotzdem können nur weniger als die Hälfte der Studierenden ermittelt werden, die Durchfallen. Dies deutet darauf hin, dass die Zugriffsdaten von Moodle hilfreich für ein Frühwarnsystem sein können. Mehr Daten wären hilfreich, wenn es darum geht bessere Vorhersagen treffen zu können. Man könnte noch weitere Zugriffsdaten, Ergebnisse von Zwischenabgaben, oder Anzahl an Interaktionen mit anderen Studierenden einzubeziehen. [2, S. 14ff]

Da der Regelmäßige Login so eine hohe Relevanz bei den Vorhersagen hat, zeigt darauf hin, dass diese Student ein besseres Lern und Zeit Management haben. Dies ist ein wichtiger Faktor für den Erfolg. [2, S. 15f]

Abschließend kann gesagt werden, die Zugriffsdaten von Moodle können benutzt werden um ein Frühwarnsystem zu erstellen. Mehr und genauere Daten wären hilfreicher. Die Ergebnisse dieser Studie sollten jedoch noch mit einer größeren Datenmenge wiederholt werden um die Ergebnisse zu bestätigen. [2, S. 16f]

# Kapitel 3

## Fazit

# Kapitel 4

## Ausblick

# Literatur

- [1] Niels Pinkwart. „Another 25 Years of AIED? Challenges and Opportunities for Intelligent Educational Technologies of the Future“. In: *International Journal of Artificial Intelligence in Education* 26.2 (2016), S. 771–783. ISSN: 1560-4292. DOI: 10.1007/s40593-016-0099-7.
- [2] Rory Joseph Quinn und Geraldine Gray. „Prediction of student academic performance using Moodle data from a Further Education setting“. In: *Irish Journal of Technology Enhanced Learning* 5.1 (2020). DOI: 10.22554/ijtel.v5i1.57.
- [3] Klaus Wannemacher und Laura Bodmann. „Künstliche Intelligenz an Hochschulen: Potenziale und Herausforderungen in Forschung, Studium und Lehre sowie Curriculumentwicklung“. In: (2021), S. 0–66. URL: [https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD\\_AP\\_59\\_Kuenstliche\\_Intelligenz\\_Hochschulen\\_HIS-HE.pdf](https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_59_Kuenstliche_Intelligenz_Hochschulen_HIS-HE.pdf) (besucht am 10.07.2022).
- [4] Claudia de Witt, Florian Rampelt und Niels Pinkwart. *Whitepaper "Künstliche Intelligenz in der Hochschulbildung"*. 2020. DOI: 10.5281/zenodo.4063722.