

Gefördert von:

Erkenntnisse zu Studienerfolg und -abbruch aus digitalen Studiengängen

DiSEA - Digitale Studiengänge: Analyse von Erfolgs-
und Abbruchfaktoren

Überblick

1. Projektbeschreibung
2. Analyse der akademischen Daten
 1. Datenexploration
 2. Vorhersage von Erfolg/Abbruch
3. Analyse der Moodle-Daten
 1. Aufbau der Kurse
 2. Aktivitätsstreifen
4. Aktivitätsregelmäßigkeit
5. Ausblick

Eckdaten und Ziele

- <https://disea-projekt.de/>
- Das Projekt analysiert Verwaltungs- und Moodle-Daten aus den Online-Studiengängen der Virtuellen Fachhochschule (VFH)
- Entwicklung eines Frühwarnsystems im Lernmanagementsystem Moodle, das persönliches Feedback für die Studierenden bereitstellt
- Ableitung von Handlungsempfehlungen für weitere Stakeholders

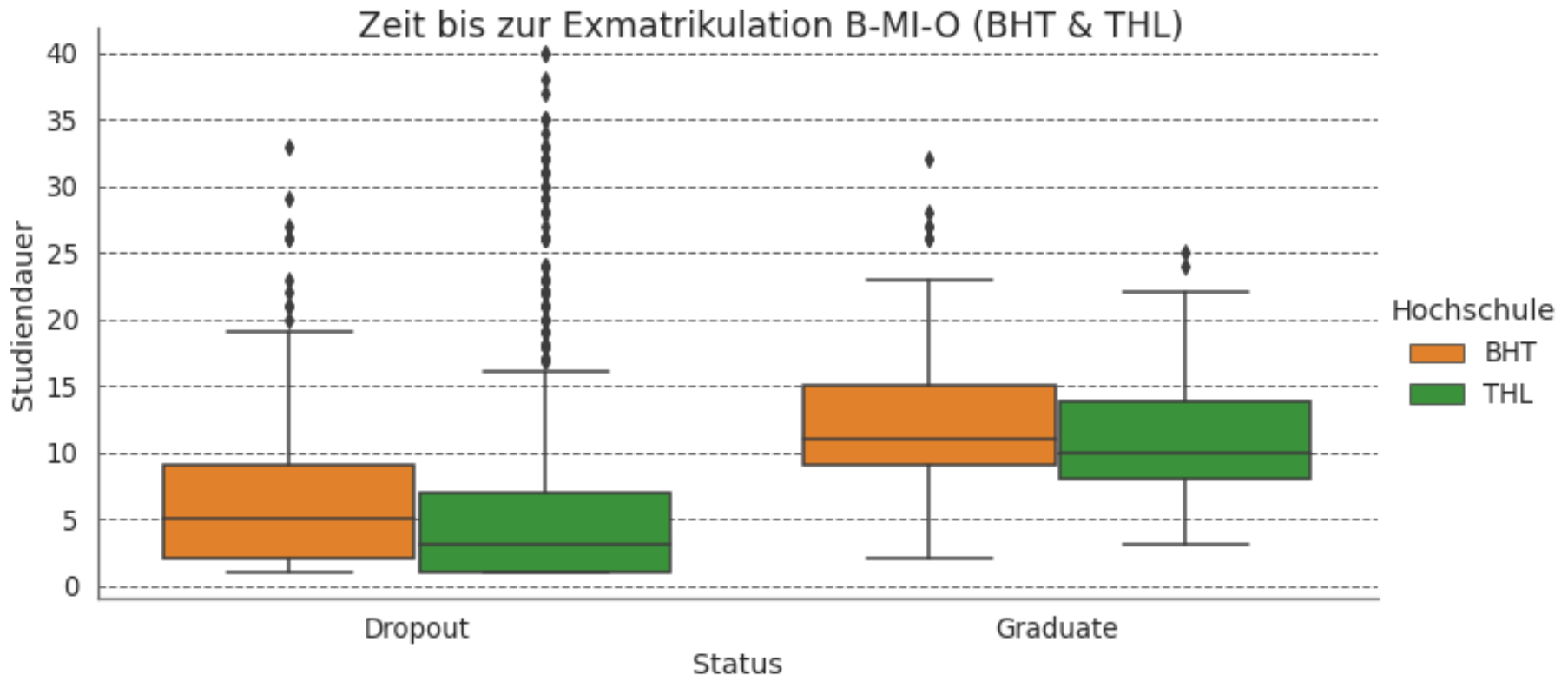
Methodik

- Datenexploration der Verwaltungs- und Moodle-Daten
- Modelle des maschinellen Lernens für die Entwicklung von Vorhersagemodellen
- Entwicklung eines Dashboards für Moodle, das den Studierenden Feedback gibt

Überblick

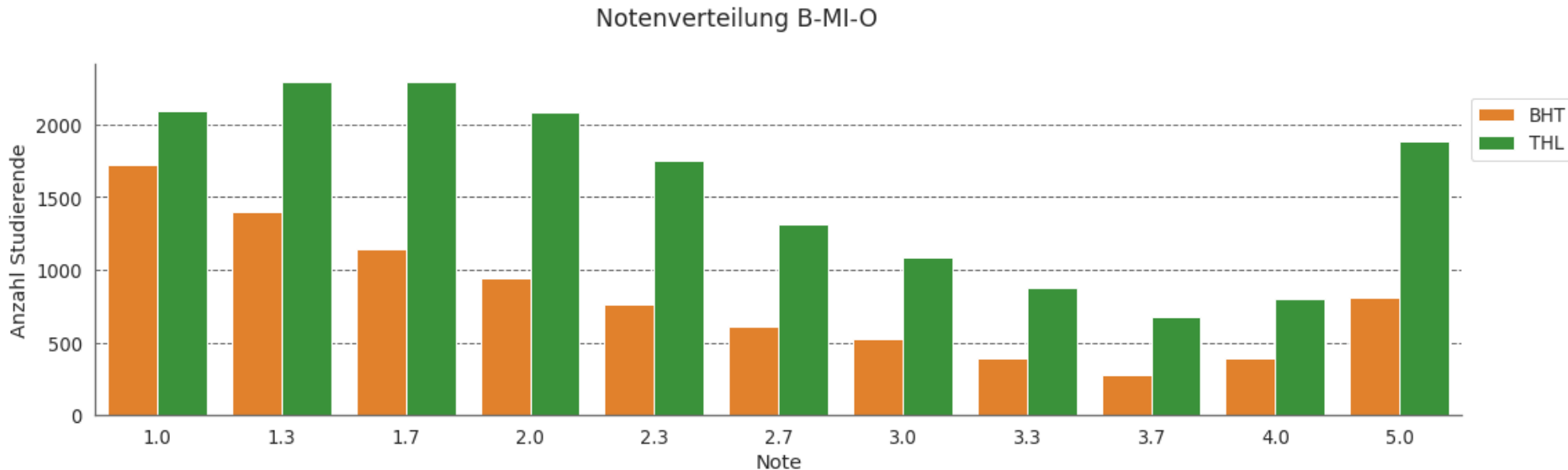
1. Projektbeschreibung
2. Analyse der akademischen Daten
 1. Datenexploration
 2. Vorhersage von Erfolg/Abbruch
3. Analyse der Moodle-Daten
 1. Aufbau der Kurse
 2. Aktivitätsstreifen
4. Aktivitätsregelmäßigkeit
5. Ausblick

Abschluss und Abbruch: Vergleich zwischen Hochschulen



- Betrachtete Hochschulen:
 - Berliner Hochschule für Technik (BHT)
 - Technische Hochschule Lübeck (THL)

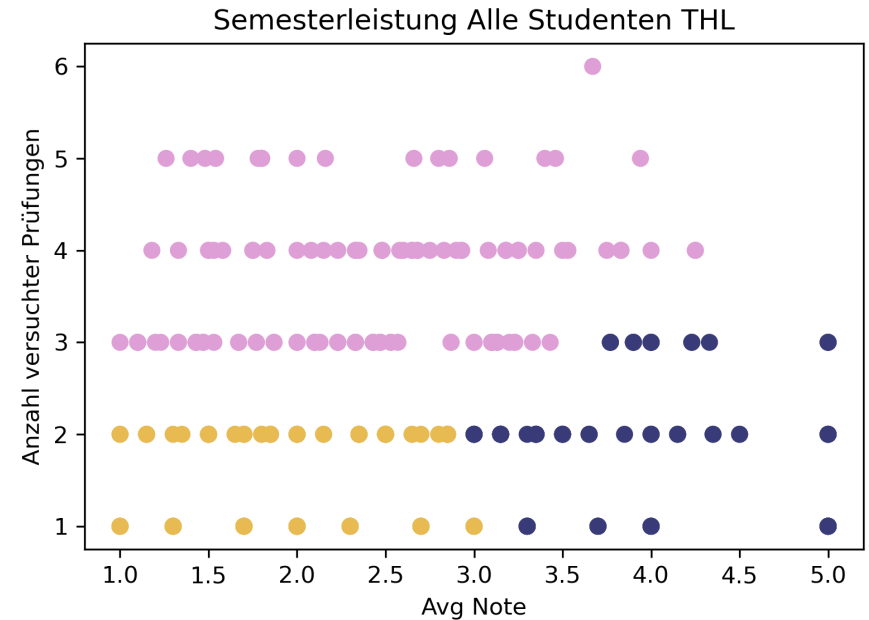
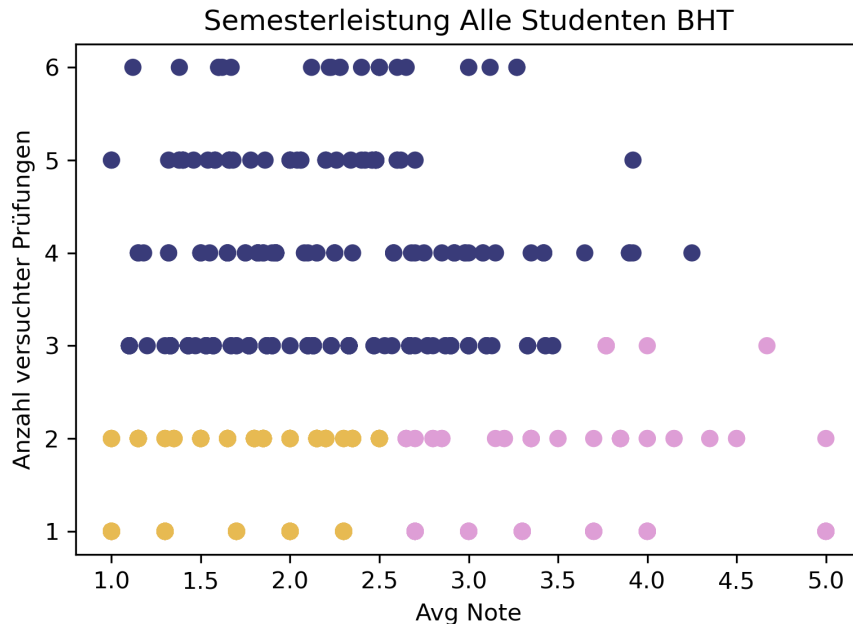
Notenverteilung



Die Noten sind nicht normalverteilt: Online-Studierende schließen ihre Prüfungen überwiegend mit sehr guten bis guten Noten ab.

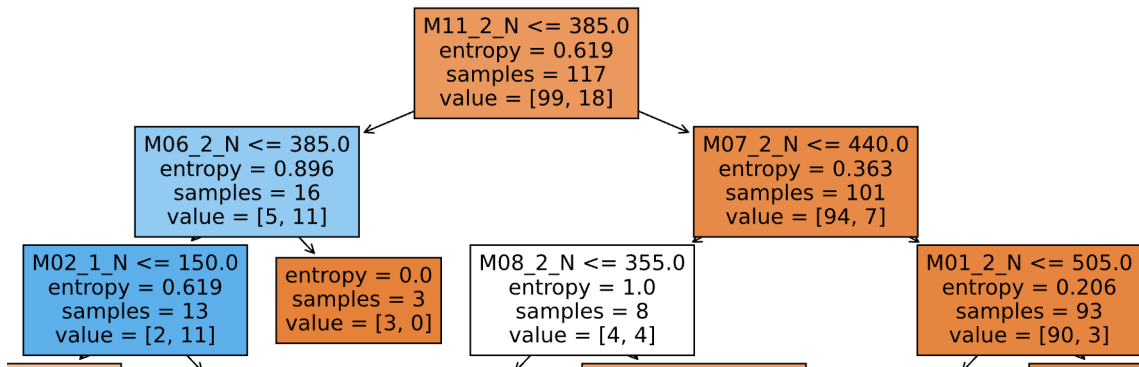
Die zwei Hochschulen unterscheiden sich.

Clustering von Studierenden anhand von Verwaltungsdaten



Die Mehrheit der Studierenden schreiben drei oder mehr Prüfungen im Semester. Die anderen zwei Clustern bestehen aus Studierenden, die unter dieser Grenze liegen. Der Unterschied hier besteht in der durchschnittliche erhaltene Note.

Entscheidungsbaum BHT, 2. Semester (B-MI-O)



Vorhergesagt 0 Vorhergesagt 1

Studienabbrecher (0)

31

2

Absolventen (1)

2

2

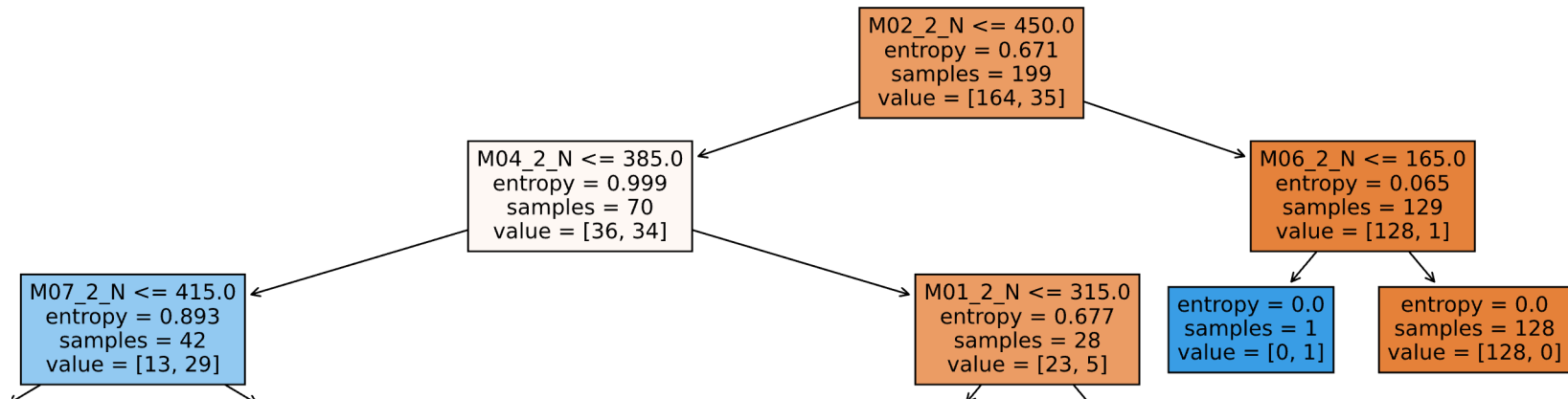
Avg. balanced accuracy: 62%

Avg. validation accuracy: 80%

Kurse

M11	Relationen und Funktionen	M02	Grundlagen der Programmierung 1
M06	Computerarchitektur und Betriebssysteme	M08	Mediendesign 2
M07	Grundlagen der Programmierung 2	M01	Einführung in die Informatik

Entscheidungsbaum THL, 2. Semester (B-MI-O)



Vorhergesagt 0 Vorhergesagt 1

Studienabbrecher (0)

46

6

Absolventen (1)

2

8

Avg. Balanced accuracy: 75%
Avg. Validation Accuracy: 85%

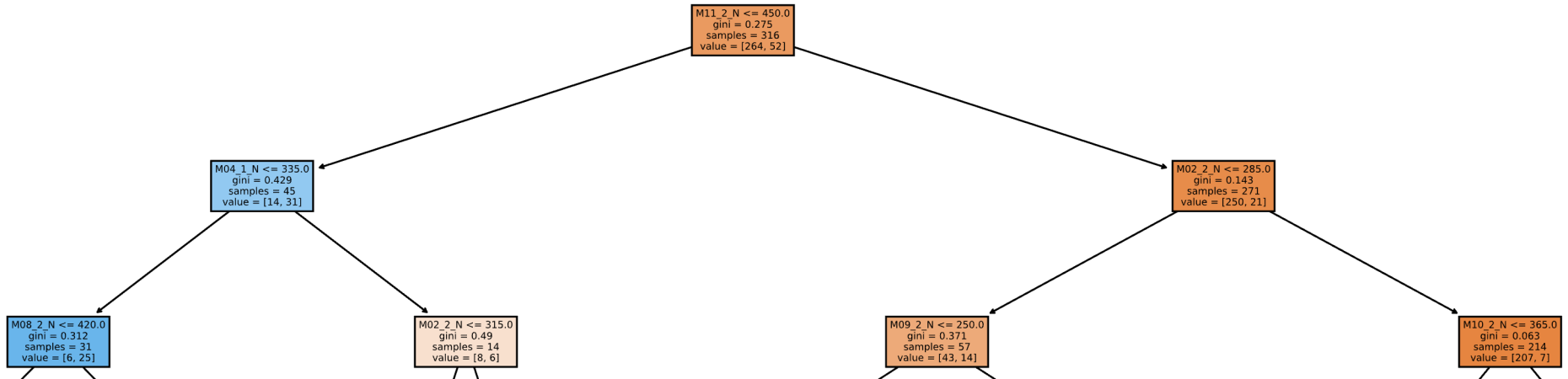
Kurse

M02	Grundlagen der Programmierung 1	M07	Grundlagen der Programmierung 2
M04	Grundlagen der Mathematik	M01	Einführung in die Informatik
M06	Computerarchitektur und Betriebssysteme		

Entscheidungsbäume Vergleich

- Viele Gemeinsamkeiten:
 - Für beide Standorte sind die folgenden Kurse wichtig: M01 Einführung in die Informatik, M02 Grundlagen der Programmierung 1, M06 Computerarchitektur und Betriebssysteme, M07 Grundlagen der Programmierung 2
 - für BHT ist der folgende Kurs M11 Relationen und Funktionen wichtig, für THL M04 Grundlagen der Mathematik, beide Mathe Kurse

Entscheidungsbaum BHT & THL, 2. Semester (B-MI-O)



Vorhergesagt 0 Vorhergesagt 1

Studienabbrecher (0)

72

10

Avg. balanced accuracy: 73%

Avg. validation accuracy: 85%

Absolventen (1)

9

8

Kurse

M11	Relationen und Funktionen	M08	Mediendesign 2
M04	Grundlagen der Mathematik	M09	Mensch-Computer-Kommunikation
M02	Grundlagen der Programmierung 1	M10	Theoretische Informatik

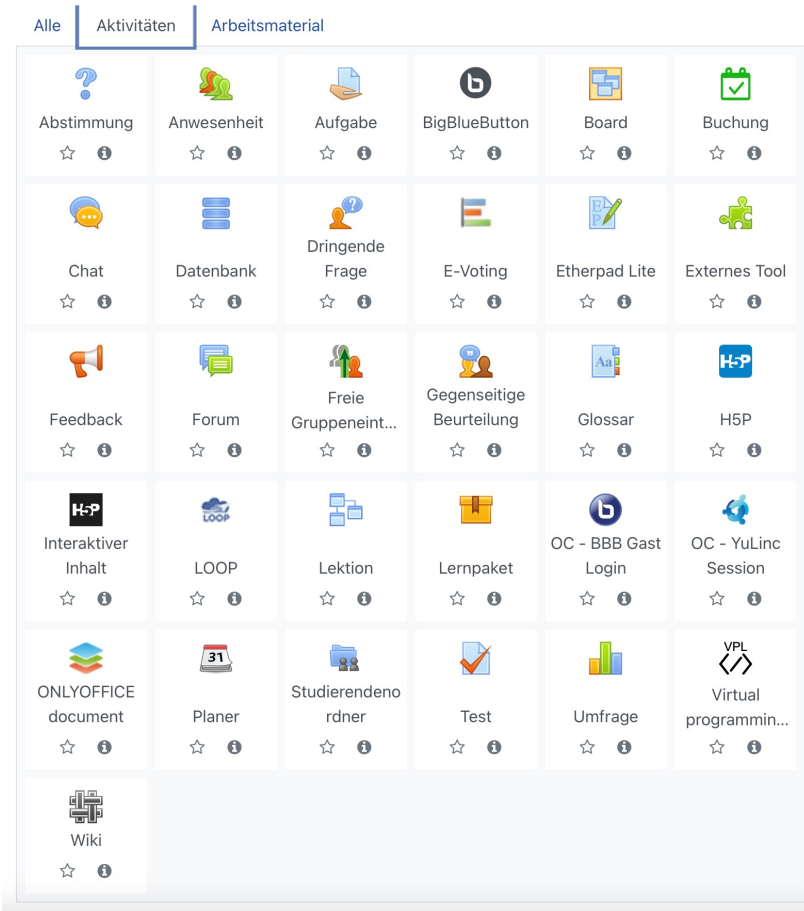
Herausforderungen bei der Vorhersage

- Kleiner und unausgeglichener Datensatz
- Die mehrfachen Änderungen der Studienordnung führt dazu, dass nicht der gesamte Datensatz für die Vorhersage geeignet ist (Daten ab WS 2014)
- Grund: Module wandeln in anderen Plansemestern oder werden durch inhaltlich nicht äquivalente Module ersetzt
- *Balanced accuracy* und *accuracy* weniger gut als in Präsenz-Studiengängen

Überblick

1. Projektbeschreibung
2. Analyse der akademischen Daten
 1. Datenexploration
 2. Vorhersage von Erfolg/Abbruch
3. Analyse der Moodle-Daten – Einwilligung der Studierende
 1. Aufbau der Kurse
 2. Aktivitätsstreifen
4. Aktivitätsregelmäßigkeit
5. Ausblick

Inhaltliche Dimension der Moodle-Kurse



userid	relateduserid	component	target	event_name
405	405	assignsubmission_file	submission	submission_created
554	-2	mod_bigbluebuttonbn	activity	activity_viewed
1309	-2	core	user_list	user_list_viewed
915	915	mod_assign	remove_submission_form	remove_submission_form_viewed
1534	-2	core	course	course_viewed

Ausschnitt aus Moodle-Log

- Im Moodle-Log kann eine Vielzahl von Aktivitäten detailliert nachvollzogen werden
- Aggregation der Moodle-Ereignisse nötig

Auswahl relevanter Aktivitäten für Kurserfolg in Moodle

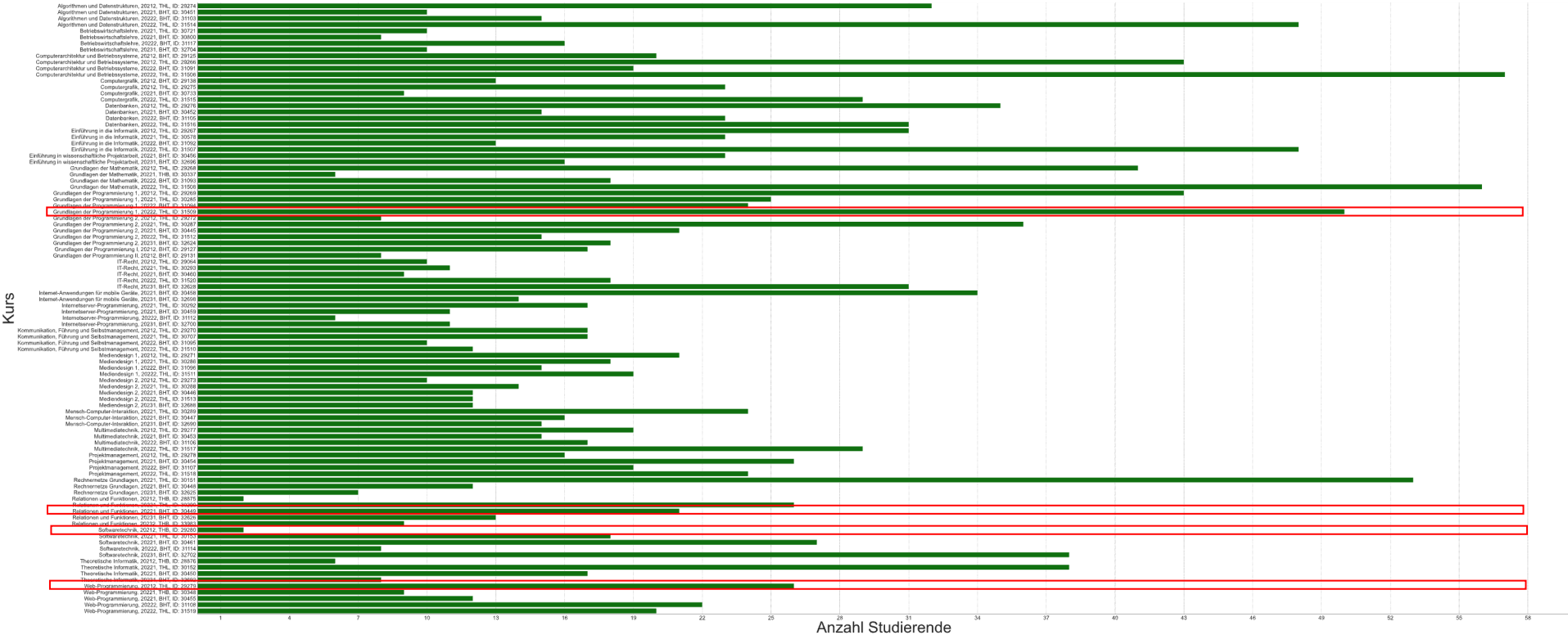
- Zusammenfassung von einzelnen Einträgen aus dem Moodle-Log zu Tasks: sinnvollen, zusammenhängenden Lernaktivitäten mit einer Dauer und einem eigenständigen Inhalt (z.B. Forum, Quiz, Einsendeaufgabe) (vgl. Rotelli & Monreale 2022, 2023)
- Analyse von Zusammenhängen zwischen Moodle-Aktivitäten und Kurserfolg
- Darstellung im Dashboard

3. Analyse der Moodle-Daten: Aufbau der Kurse



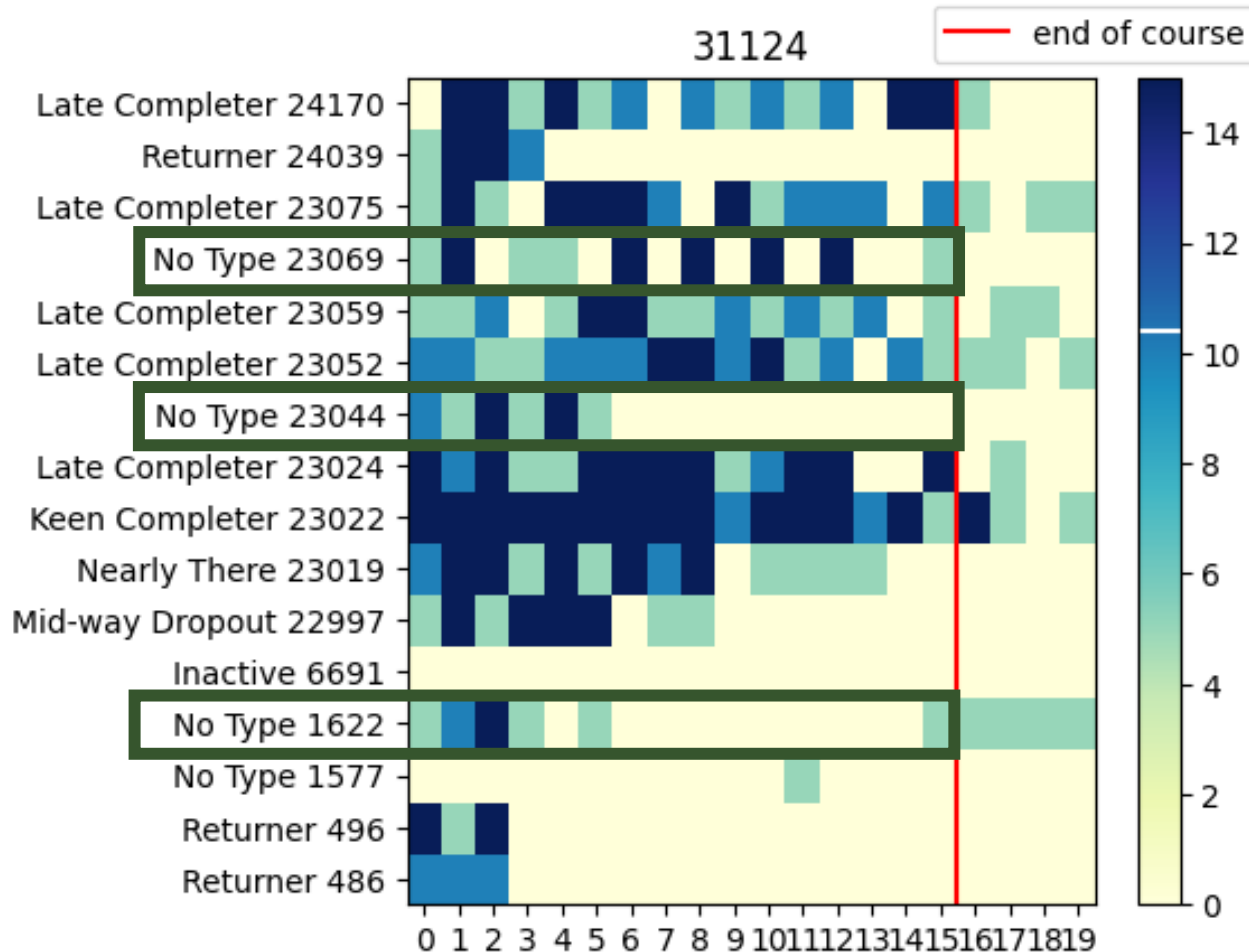
Gesamtanzahl der Studierenden in Kursen (B-MI-O)

Anzahl Studierender in den Kursen der Übersicht



Die Anzahl der Studierenden in den einzelnen Modulen variiert stark.

Analyse des Lernverhaltens in Moodle (Ferguson & Clow)



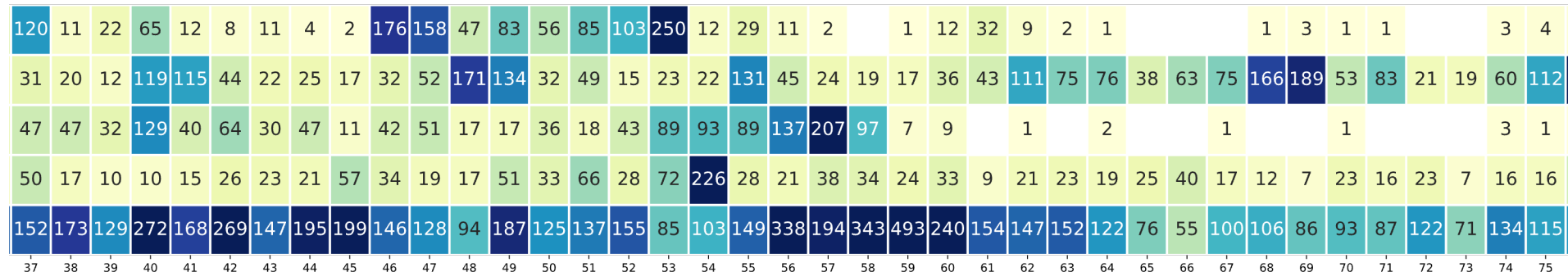
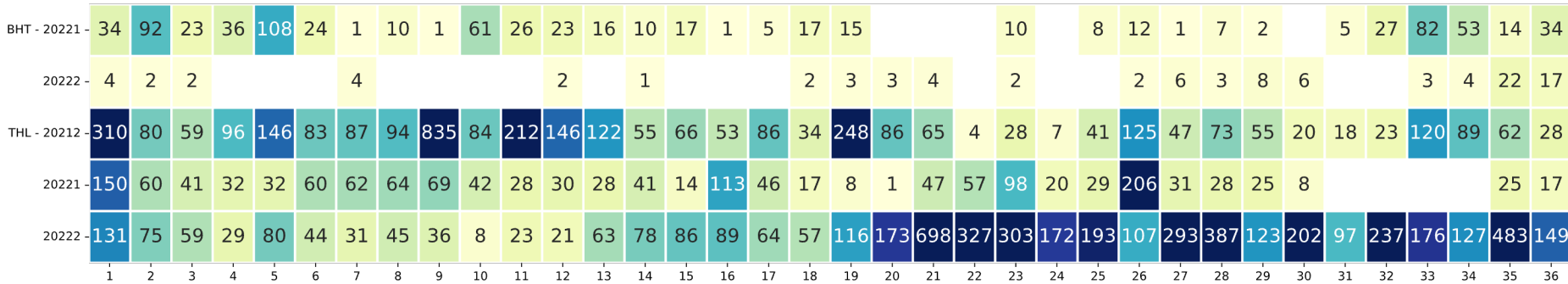
- Codierung der Aktivitäten
- Aktivität der Studierende wochenweise

Betrachteter Kurs: „English for Computer Scientists“
(BHT, B-WINF-O, WiSe 2022/23)

3. Analyse der Moodle-Daten: Aktivitätsstreifen

Aktivität im Kurs "Grundlagen der Programmierung 1"

Uni & Sem



Die Nutzung eines einzelnen Kurses kann von Semester zu Semester variieren.

Überblick

1. Projektbeschreibung
2. Analyse der akademischen Daten
 1. Datenexploration
 2. Vorhersage von Erfolg/Abbruch
3. Analyse der Moodle-Daten
 1. Aufbau der Kurse
 2. Aktivitätsstreifen
4. **Aktivitätsregelmäßigkeit**
5. Ausblick

Aktivitätsregelmäßigkeit

- Analyse von Zusammenhängen zwischen Moodle-Aktivitäten und Kurserfolg
- Die Aktivitätsregelmäßigkeit (*regularity*) in Moodle misst, ob die Studierenden über die Wochen des Kurses hinweg an ähnlichen Wochentagen arbeiten (vgl. Shirvani Boroujeni et al., 2016)
- *regularity* lässt sich durch eine Zahl zwischen 0 und 1 ausdrücken
- Der Maximalwert 1 ist erreicht, wenn der Studierende jede Woche an genau denselben Tagen aktiv ist

4. Aktivitätsregelmäßigkeit

Regularity im Kurs „Grundlagen der Programmierung 1“

student	regularity	bestanden	median_activity_week	Activity_All
1617	0.3563	ja	33	900
24182	0.2545	ja	25	467
24174	0.2534	ja	18	309
24189	0.2215	ja	8	384
24187	0.1795	nein	29	419
1621	0.1762	ja	15	431
24173	0.1742	nein	19	360
24196	0.1372	ja	18	298
26445	0.0771	ja	6	252

Grundlagen der Programmierung 1 (THL, WiSe 2022/23)

Nur ein kleiner Teil der Studierenden, die sich im Kurs anmelden, nehmen an der Prüfung teil.

student	regularity	bestanden	median_activity_week	Activity_All
24176	0.3356	k.A.	20	408
1395	0.2929	k.A.	74	1267
24182	0.2545	ja	25	467
24174	0.2534	ja	18	309
24188	0.2526	k.A.	31	726
29949	0.2421	k.A.	20	380
24175	0.2343	k.A.	26	542
24189	0.2215	ja	8	384
23056	0.2202	k.A.	14	292
24172	0.2077	k.A.	38	590
24201	0.2060	k.A.	19	419
24195	0.1861	k.A.	14	407
26443	0.1827	k.A.	19	385

Grundlagen der Programmierung 1 (THL, WiSe 2022/23)

(Ausschnitt)

4. Aktivitätsregelmäßigkeit

Regularity im Kurs „Relationen und Funktionen“

student	regularity	bestanden	median_activity_week	Activity_All
336	0.1728	ja	8	144
719	0.1337	ja	4	79
653	0.1155	ja	10	168
632	0.1046	ja	5	110
613	0.0000	ja	1	2
623	0.0000	ja	1	2
611	0.0000	ja	3	3
558	0.0000	ja	2	3
680	0.0000	ja	1	1

Relationen und Funktionen (BHT, SoSe 2022)

student	regularity	bestanden	median_activity_week	Activity_All
1514	0.2190	k.A.	11	274
336	0.1728	ja	8	144
719	0.1337	ja	4	79
653	0.1155	ja	10	168
632	0.1046	ja	5	110
685	0.1008	k.A.	3	62
658	0.0593	k.A.	6	120
584	0.0507	k.A.	8	106
369	0.0435	k.A.	4	59
621	0.0343	k.A.	4	54
1589	0.0119	k.A.	3	52
362	0.0119	k.A.	1	10
634	0.0000	k.A.	7	7
681	0.0000	k.A.	8	15
613	0.0000	ja	1	2
623	0.0000	ja	1	2
617	0.0000	k.A.	1	1
611	0.0000	ja	3	3
558	0.0000	ja	2	3
680	0.0000	ja	1	1

Relationen und Funktionen (BHT, SoSe 2022)

Weitere Schritte

- Erstellung eines *activity history* für manche Gruppen von Studierenden
- Vergleich der *regularity*-Wert mit den Werten in vorigen Semestern für Studierende, die eine niedrige *regularity*-Zahl (auch 0) bekommen haben, aber trotzdem bestanden haben
- Betrachtung der höchsten *regularity*-Werte für Studierenden, die keine Prüfung geschrieben haben (Vermerk: k. A.). Gibt es "Nachfolger"-Kurse, in denen es einen Prüfungsversuch gibt?

Ausblick

- Erste Ergebnisse wurden mit dem Modell Entscheidungsbaum erzielt
- Der Einsatz von weiteren Algorithmen ist geplant
- Identifikation der Kurse, die für den Studienabschluss maßgebend sind
- Die Entscheidungsbäume der zwei Hochschule unterscheiden sich in der Struktur, wenig in den Kursen.
- Die Moodle-Kurse sind sehr unterschiedlich, es werden signifikante Merkmale zum Vergleich benötigt, z.B. Moodle-Quiz statt Quizzes direkt im Kurs-Material

- Freiheit der Lehre: Keine Vereinheitlichung der Kurse empfohlen
- Eine Korrelation zwischen der Aktivitätsregelmäßigkeit in Moodle und Leistung wird noch untersucht
- Verfolgung von Studierenden-Aktivität über mehr als drei Semester
- *Sparsity* der Moodle-Logdaten
- Das Dashboard kann Feedback über Lernfortschritt oder Lernaktivitäten zeigen, kann aber keine Prognose über den Erfolg im Kurs berechnen und anzeigen (nicht immer erwünscht)

1. Ferguson, Rebecca & Clow, Doug (2015). Examining engagement: Analysing learner subpopulations in massive open online courses (moocs). In Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, LAK '15, page 51–58, New York, NY, USA, 2015. Association for Computing Machinery. doi:10.1145/2723576.2723606.
2. Shirvani Boroujeni, Mina & Sharma, Kshitij & Kidziński, Łukasz & Lucignano, Lorenzo & Dillenbourg, Pierre. (2016). How to Quantify Student's Regularity? In Proceedings of the 11th European Conference on Technology Enhanced Learning EC-TEL, pages 41-52, Cham, Springer, 9891. 277-291. 10.1007/978-3-319-45153-4_21.
3. Rotelli, Daniela & Monreale, Anna. (2022). Time-on-Task Estimation by data-driven Outlier Detection based on Learning Activities. 336-346. 10.1145/3506860.3506913.
4. Rotelli, Daniela & Monreale, Anna. (2023). Processing and Understanding Moodle Log Data and Their Temporal Dimension. Journal of Learning Analytics. 1-23. 10.18608/jla.2023.7867.

Gefördert von:

VDI|VDE|IT



Bundesministerium
für Bildung
und Forschung

Das Projekt DiSEA (Förderkennzeichen 01PX21001A / 01PX21001B) wird gefördert aus Mitteln des Bundesministeriums für Bildung und Forschung im BMBF-Förderschwerpunkt Wissenschafts- und Hochschulforschung, Förderbekanntmachung Studienerfolg/Studienabbruch II
<https://disea-projekt.de/>

Vielen Dank für Ihre Aufmerksamkeit!