

Hausarbeit

Einflussfaktoren auf die Schulleistungen im Rahmen der Oberschule

MADS2100 Reporting und Visualisierung 23oB

Leon Henne

Köln, den 6. März 2024

Betreut durch Dr. Robert Stahlbock

Inhaltsverzeichnis

Tabellenverzeichnis	I
Abbildungsverzeichnis	II
1 Problemstellung	1
2 Zielsetzung	2
2.1 Untersucher Datensatz	2
2.2 Forschungsfragen	2
3 Grundlagen	4
3.1 how to Visualisierungen ?	4
3.2 how to analyse Visualisierungen ?	4
4 Fazit	5
Literatur	6

Tabellenverzeichnis

Abbildungsverzeichnis

1 Problemstellung

Der für jede Nation erstrebenswerte langfristige ökonomische Fortschritt wird unter anderem stark durch das vorherrschende Bildungsniveau beeinflusst. Cortez und Silva, 2008, S. 1 Um dieses sich ergebende Bildungsniveau stärker zu durchleuchten und final zu verbessern, wird zur Unterstützung der Schüler und Lehrkräfte die Modellierung von Schulleistungen eingesetzt. Cortez und Silva, 2008, S. 1 So kann eine zeitabhängige Vorhersage der Leistungen, lernschwächere Schüler detektieren und damit Lehrkräfte frühzeitig befähigen, mit entsprechenden Maßnahmen in den Lernprozess einzugreifen. Namoun und Alshanqiti, 2021, S. 2 Verstärkt wurde dieser Bedarf durch die in der Vergangenheit eingetretenen Covid-Pandemie, und den damit verbundenen Schulschließungen, welche für neue erhebliche Herausforderungen sorgten. Clark et al., 2021, S. 2 Durch Clark et al., 2021, S. 13 konnte hierzu aufgezeigt werden, welchen positiven Effekt digitaler Lernunterstützung auf die Schülergruppen ausmachte. Aus der Arbeit von Namoun und Alshanqiti, 2021, S. 9 geht jedoch hervor, dass bereits seit 2017 erneut das Interesse anstieg hinsichtlich der Modellierung von Lernergebnissen. Dabei wurde jedoch auch deutlich, dass seitdem besonders ein Fokus auf das Bildungsniveau von Bachelorstudiengängen besteht und die Untersuchung weiterführender Schulen lediglich ein Anteil von in etwa 12% ausmacht. Namoun und Alshanqiti, 2021, S. 11 Die in den letzten Jahren erforschte Modellierung von Studierendenergebnissen lässt jedoch häufig unbeachtet, wie einzelne Faktoren innerhalb der maschinellen Lernmethoden zu den Vorhersagen führen. Namoun und Alshanqiti, 2021, S. 19 Die Gesamtheit dieser aktuellen Gegebenheiten motiviert die nachfolgende Untersuchung des gewählten Datensatzes anhand der im Rahmen dieser Arbeit formulierten Forschungsfragen.

2 Zielsetzung

Inhalt dieses Kapitels ist zunächst die Erläuterung des gewählten Datensatzes. Zum Anderen werden daraus abgeleitete Forschungsfragen formuliert, welche im Rahmen dieser Arbeit durch die Entwicklung und Analyse von Visualisierungen untersucht werden.

2.1 Datensatz

Der in dieser Arbeit betrachtete Datensatz entstammt der Arbeit von Cortez und Silva, 2008. Motiviert wurde diese Forschung durch Statistiken, welche Portugal im europäischen Vergleich als deutlich unterdurchschnittlich klassifizierten, aufgrund von hohen Durchfallquoten. Cortez und Silva, 2008, S. 1 Daher wurde mit dieser Arbeit ein realer Datensatz erhoben. Hierfür wurden Schulleisten und schulbezogene Faktoren vom Berichtswesen gesammelt und demografische und soziale Faktoren durch Befragungen ermittelt. Cortez und Silva, 2008, S. 1 Die schulbezogenen Faktoren beziehen sich dabei auf die Leistungen in den Schulfächern Mathematik und Portugiesisch, da Inhalte dieser Fächer übergreifend in anderen Fächern zum Einsatz kommen. Cortez und Silva, 2008, S. 2 Zielgruppe der Untersuchung waren Schüler der dreijährigen zweiten Bildungsphase in Portugal, welche auf der ersten neunjährigen Phase aufbaut. Cortez und Silva, 2008, S. 2 Mittels der Berichte und Umfragen konnten schulbezogene-, demografische-, und soziale Faktoren von X Mathematik-Schülern und Y Portugiesisch-Schülern der beiden Schulen erhoben werden. Alle erhobenen Faktoren, welche im Rahmen der Analyse auch als Variablen oder Merkmale bezeichnet werden, können mit ihrer Kurzbeschreibung und ihren Ausprägungen der nachfolgenden Liste entnommen werden:

2.2 Forschungsfragen

Beschreibung	Datentyp	Ausprägungen
Schulbezeichnung	Binär	GP - Gabriel Pereira; MS - Mousinho da Silveira
Geschlecht	Binär	F - weiblich; M - männlich
Alter	Numerisch	15 bis 22
Wohngegend	Binär	U - urban; R - ländlich
Familiengröße	Binär	LE3 - kleiner oder gleich 3; GT3 - mehr als 3
Zusammenleben der Eltern	Binär	T - leben gemeinsam; A - leben auseinander
Mutters Bildungsgrad	Numerisch	0 - kein Bildungsgrad; 1 - Grundschulabschluss (4. Klasse); 2 - 5. bis 9. Klasse; 3 - Oberstufenabschluss; 4 - höherer Bildungsgrad
Vaters Bildungsgrad	Numerisch	0 - kein Bildungsgrad; 1 - Grundschulabschluss (4. Klasse); 2 - 5. bis 9. Klasse; 3 - Oberstufenabschluss; 4 - höherer Bildungsgrad
Mutters Arbeitsbereich	Nominal	Lehrerin; Gesundheitswesen; Sozialwesen (Verwaltung oder Polizei); zu Hause; anderer Bereich
Vaters Arbeitsbereich	Nominal	Lehrer; Gesundheitswesen; Sozialwesen (Verwaltung oder Polizei); zu Hause; anderer Bereich
Grund der Schulentscheidung	Nominal	Nahe dem Zuhause; Ruf der Schule; Kurspräferenz; anderer Grund
Erziehungsberechtigter	Nominal	Mutter; Vater; Anderer
Pendelzeit	Numerisch	1 - <15 min.; 2 - 15 bis 30 min.; 3 - 30 min. bis 1 Stunde; 4 - >1 Stunde
Wöchentliche Lernzeit	Numerisch	1 - <2 Stunden; 2 - 2 bis 5 Stunden; 3 - 5 bis 10 Stunden; 4 - >10 Stunden
Anzahl durchgefallener Kurse	Numerisch	1 bis 3; 4
externe Lernunterstützung	Binär	Ja; Nein
familiäre Lernunterstützung	Binär	Ja; Nein
Extra bezahlter Unterricht im befragungsfach	Binär	Ja; Nein
außerschulische Aktivitäten	Binär	Ja; Nein
Besuch der Vorschule	Binär	Ja; Nein
Absicht zur Weiterbildung	Binär	Ja; Nein
häuslicher Internetzugang	Binär	Ja; Nein
Partnerliche Beziehung	Binär	Ja; Nein
Qualität der Familienbeziehungen	Numerisch	von 1 - sehr schlecht bis 5 - exzellent
außerschulische Freizeit	Numerisch	von 1 - sehr schlecht bis 5 - sehr gut
soziale Aktivitäten	Numerisch	von 1 - sehr schlecht bis 5 - sehr gut
Alkoholkonsum unter der Woche	Numerisch	von 1 - sehr schlecht bis 5 - sehr gut
Alkoholkonsum am Wochenende	Numerisch	von 1 - sehr schlecht bis 5 - sehr gut
aktueller Gesundheitszustand	Numerisch	von 1 - sehr schlecht bis 5 - sehr gut
Anzahl der Fehltage	Numerisch	von 0 bis 93
erste Vorabnote	Numerisch	von 0 bis 20
zweite Vorabnote	Numerisch	von 0 bis 20
finale Note	Numerisch	von 0 bis 20

3 Grundlagen

3.1 how to Visualisierungen ?

3.2 how to analyse Visualisierungen ?

4 Fazit

Literatur

- Clark, A. E., Nong, H., Zhu, H., & Zhu, R. (2021). Compensating for academic loss: Online learning and student performance during the COVID-19 pandemic. *China Economic Review*, 68, 101629. <https://doi.org/10.1016/j.chieco.2021.101629>
- Cortez, P., & Silva, A. M. G. (2008). Using data mining to predict secondary school student performance. <https://api.semanticscholar.org/CorpusID:16621299>
- Namoun, A., & Alshanqiti, A. (2021). Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review. *Applied Sciences*, 11(1). <https://doi.org/10.3390/app11010237>

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit eigenständig und ohne fremde Hilfe angefertigt habe. Textpassagen, die wörtlich oder dem Sinn nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche kenntlich gemacht.

Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Köln, den 6. März 2024

Leon Henne