Can We Predict Your Performance? Assessing the Relationship of Admissions Data to Academic Performance in Gross Anatomy of First-Year Medical Students

Ashley N. Walker
Phuong Huynh
Kyle Rarey
Nancy Adams
University of Florida

Abstract

Educational data mining and predictive analytics in medical education have been justified to assist admissions committees and to help identify students struggling academically for purposeful interventions. This study's purpose is to see if medical school entry metrics could predict first semester anatomy performance. Block entry multiple regression analysis and logistic regression analysis were used with pre-admissions data from one cohort of 133 students on their anatomy lab practical scores. The results showed that Cumulative Science GPA, MCAT scores, and first-generation status are significant predictors of academic performance on the lab practicals. The long-term goal is to utilize the formulated regression model to encourage practitioners within medical education to consider programs and activities that assist in student development of at-risk students.

Keywords: Pre-admissions, anatomy, medical education

Background

Anatomy Performance

The medical school curriculum in the United States of America is traditionally divided into two years of "basic science" and two years of clinical clerkships or "rotations" (Finnerty et al., 2010). Basic sciences typically include physiology, biochemistry, histology, genetics, and gross anatomy. Gross anatomy is one of the first experiences in the medical education curriculum and is generally a challenging journey for a new medical student (Brooks et al., 2015). The demanding nature of learning gross anatomy has encouraged anatomy educators to consider how prior performance might predict current academic performance in gross anatomy. Medical educators have also used pre-admissions data to identify incoming medical students who might potentially struggle academically (Li et al., 2019; Yates & James, 2006).

Medical schools utilize pre-admissions data such as grade point averages (GPA), Medical College Admissions Test (MCAT) scores, and prior science grades to rank applicants; however, there are few studies exploring pre-admissions data related to performance in a particular basic science course such as gross anatomy. For example, Fredieu and Snyder (2015) reported that medical students possessing a prior master's degree in anatomy performed better on the United States Medical Licensing Examination (USMLE) Step 1, the national exam medical students take before starting the clerkship years. Also, Gauer et al. (2016) acknowledged that MCAT scores were significantly correlated with the USMLE Step 1 and Step 2 scores and are predictive of student performance on the USMLE.

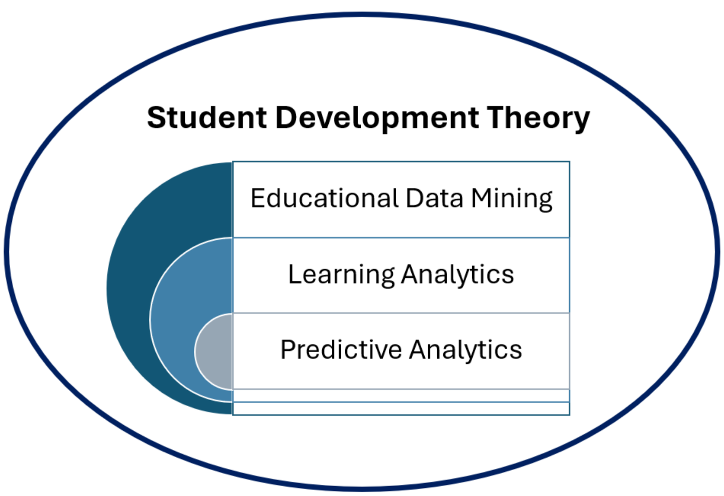
Similar studies exploring the relationship between MCAT scores and USMLE performance abound (Basco et al., 2002; Gliatto et al., 2016; Zhao et al., 2010), as well as studies examining relationships between other admissions data including race, sex, and GPA with USMLE performance (Veloski et al., 2000). However, there is little evidence describing the relationship of the pre-admissions data to specific coursework within the medical school curriculum.

Data Mining and Predictive Analytics as a Framework

Educational Data Mining (EDM) has been used to influence decisions about educational practice and techniques that influence student learning and development. EDM has been widely used by admissions committees to manage enrollment and student retention (Zhao et al., 2020). Considering the many uses of EDM involving student progression and development, Lei et al. (2017) proposed a framework for educational data mining for decision-making based on student development theory. This framework highlights that input variables, such as pre-admissions data, affect students' outcomes in a specific academic environment and that EDM and data-driven decision-makers can develop educational interventions. This framework explores the use of educational data to inform decisions being made that impact student learning and development.

Under the umbrella of EDM is Learning Analytics (LA). LA is also a data-driven technique to improve educational practices but includes predictive analytics, which is the use of data to foresee patterns. The use of predictive analytics in medical education has been justified to (a) assist admissions committees with choosing students that will be most likely to matriculate and perform well (Almarabheh, 2022) and (b) assist student support personnel with identifying students for the purpose of initiating interventions (Bird, 2021; Qahmash et al., 2023). The independent variables used in the literature are usually MCAT scores, various calculations of GPAs, and demographics. The outcomes of concern are usually completion and nationally standardized tests like USLME Steps 1 and 2. For example, Ferguson, et al. (2002) conducted a systematic review of literature reporting various independent variables on overall medical school performance. They found that 23% of student's performance was explained by previous academic performance measures, such as GPA and MCAT scores. They also summarized reports on personality tests, learning styles, interview results, personal statements, and references (Ferguson et al., 2002).

Figure 1. *Study Framework*



Note. The framework utilized for this study arose from Lei and colleagues (2017), who proposed educational data mining for decision-making based on student development theory and predictive analytics.

Purpose

Admission into medical schools in the United States typically requires an undergraduate degree; this is not the case in many other countries where medical students are admitted directly upon completion of secondary education. Other metrics like Medical College Admission Test (MCAT) scores and science GPAs are used to indicate preparedness for medical school. The purpose of this cross-sectional, observational study was to explore whether the medical school entry metrics, specifically cumulative science GPA, MCAT score, and first-generation status, could predict anatomy performance in a first-semester integrated course, utilizing Lei et al. (2017) educational data mining for student development and predictive analytics as a framework (Figure 1). This study has implications for utilizing a predictive model to inform medical school decision-makers about the best practices for incoming students and anatomy curricula.

The researchers specifically addressed the following research questions:

1. How do the pre-admission MCAT scores, cumulative science GPAs, and first-generation status relate to the continuous measure of average gross anatomy lab practical exam score?
2. How do the pre-admission MCAT scores, cumulative science GPAs, and first-generation status relate to the categorical measure of passing all three lab exams?

For each research question, the researchers developed both null and alternative hypotheses:

H_{01}: $\beta 1 = \beta 2 = ... = \beta p = 0$ (There is no relationship between pre-admissions data of interest and average lab practical scores.)
H_{a1}: At least one $\beta j \neq 0$ (There is a relationship between pre-admissions data of interest and average lab practical scores.)
H_{02}: $\beta 1 = \beta 2 = 0$ (There is no relationship between pre-admissions data of interest and passing all three anatomy lab practicals.)
H_{a2}: At least one $\beta j \neq 0$ (There is a relationship between the pre-admissions data of interest and passing all three anatomy lab practicals.)

Materials and Method

Data Sources

De-identified pre-admissions data and first-semester gross anatomy lab practical scores from one medical school cohort (*n* =133) were obtained via the institution's Office of Program Evaluation and Student Assessment. This cross-sectional data included continuous predictors, MCAT scores, and cumulative science GPA, and one categorical predictor, first-generation status (Table 1). The three anatomy lab practical scores were averaged together to form one continuous outcome (dependent variable). The three exams were also coded into a binary categorical outcome as "pass all exams" or "fail at least one."

Table 1. *Data Received from the Office of Program Evaluation and Student Assessment.*

Attribute	Type	Range	Description
MCAT score	Numerical; continuous	504–524	The min/max for the MCAT is 472–528
Cumulative science GPA	Numerical; continuous	3.03–4.00	Includes only undergraduate science courses
First-generation status	Categorical; binary	Yes = first-generation student; No = not	Defined as first generation to attend college
Average lab practical exam score	Numerical; discrete	68–99.33	Three lab exams averaged together
Lab exams passed	Categorical; binary	Yes = passed all three exams; No = failed one or more exams	"Passing" means a score of 75 or higher

Data Analysis

Descriptive statistics were conducted for cumulative science GPA, MCAT score, and first-generation status. Then, a block-entry (hierarchical) multiple regression analysis of the predictors on the average anatomy lab practical score was performed. Three regression models were created with one, two, and three of the predictors via the block entry method. To analyze the differences between the generated regression models, an ANOVA test was conducted to test the incremental R^2. The final regression model used in the study contained the three predictors: cumulative science GPA, MCAT score, and first-generation status. Missing data was eliminated via a listwise deletion, and model diagnostics were performed to explore the computed regression model.

To address the second research question, a logistic regression model was developed to explore the relationship between the categorical outcome (pass-all exams versus failing one or more) and the predictor variables: cumulative science GPA, MCAT score and first-generation status. Statistical analysis of the data was performed using R and RStudio statistical software, version 3.4.3 (R Foundation for Statistical Computing, Vienna, Austria).

Ethics Statement

This study was deemed exempt for educational purposes by the Institutional Review Board (Protocol number: 02-202301456).

Results

The aim of this study was to assess regressed anatomy lab practical performance on several pre-admissions predictors. Before running the block-entry regression analysis, descriptive predictors were run on the independent and dependent variables in the study. Table 2 and Table 3 present the descriptive statistics for the variables.

Table 2. *Descriptive Statistics for Predictors of Interest*

Measure	*n*	*M*	*SD*	*Skew*	*Kurtosis*
Continuous predictors					
MCAT score	126	514.22	4.61	0.04	–0.53
Cumulative science GPA	133	3.80	0.21	–1.18	0.02
Categorical predictor					
First-generation status	133*				

Note. *First-generation status was coded as yes (1) and no (0); yes = 22; no = 111.

Table 3. *Descriptive Statistics for Outcomes of Interest*

Measure	*n*	*M*	*SD*	*Skew*	*Kurtosis*
Continuous outcome					
Average lab exam score	133	86.80	5.65	–0.44	–0.07
Categorical outcome					
Pass/fail lab exams	133*				

Note. *Pass/fail lab exams was coded as pass all (1) and did not pass all (0): pass all = 104; did not pass all = 29.

Research Question 1: Multiple Linear Regression Model

The first model tested with only one predictor-cumulative science GPA-had an R^2 value reported as 0.097. After adding an additional predictor, the R^2 showed that the addition of MCAT score resulted in an improvement in the R^2 (R^2= 0.171) compared to the previous model. The final model included cumulative science GPA, MCAT score, and the first generation in college status improved the R^2 value (R^2 = 0.20) over the other two models.

The model diagnostics reported below relate to the final model containing all three predictors. The Cook's test was used to expose any outliers present in the data resulting in two observations, 77 and 121, presenting them as influential outliers. The DFBETAS measure showed that eliminating Observation 77 increases the MCAT coefficient by 0.14, the cumulative science GPA by 0.86, and the first-generation coefficient by 0.13. Meanwhile, Observation 121 increases the MCAT coefficient by 0.05, decreases the Cumulative Science GPA by 0.09, and decreases the first-generation coefficient by 0.16. The researchers acknowledged and discussed the potential influences of these cases but chose not to eliminate them in this observational study. To address multicollinearity (i.e., the degree to which the independent variables are correlated) the variance inflation factors (VIF) were assessed and presented in Table 4. Each VIF is between 1 and 2, indicating little to moderate correlation between the variables. PP and QQ normality plots were generated to assess residuals, and the results support the normality of the data. The final model shows no result of near colinear columns.

Finally, component residual plots were generated to examine the linear relationship between the predictors and lab practical averages. The results indicated linearity between the lab practical average outcome and MCAT score and first-generation status and subtle linearity with cumulative science GPA.

Table 4. *Variance Inflation Factors for MLR Models*

	Cumulative Science GPA	MCAT Score	First Generation Status
Model 2	1.06	1.06	
Final Model	1.07	1.09	1.02

Note. Model 1 only included the predictor cumulative science GPA.

For the continuous predictors, the final model predicts that cumulative science GPA and MCAT scores are statistically significant predictors of average lab practical scores. For the categorical predictors analyzed, first-generation status is a statistically significant predictor of average lab practical scores (Table 5).

Table 5. *Results of Multiple Regression*

			95% CI		
	β	*SE*	*LL*	*UL*	*p*
Intercept	0.599	50.604	–171.364	28.956	.162
Continuous					
Cumulative science GPA	0.286	2.161	3.214	11.771	.001***
MCAT score	0.210	0.101	0.053	0.454	.014 *
Categorical					
First generation status	–0.223	1.185	–5.604	–0.913	.007 **

Note. ***$p < .001$, **$p < .01$, *$p < .05$. The result of regressed academic performance as the average of three lab practicals by 1st-year medical students on cumulative science GPA, MCAT score, and first-generation college student. CI = confidence interval; LL = lower limit; UL = upper limit.

A one-point increase in cumulative science GPA is predicted to result in a 0.279-point increase in average lab practical score if all other predictors remain constant. A one-point increase in MCAT score predicts a 0.207-point increase in average lab practical score if all other predictors remain constant. Lastly, having first-generation status predicts a decrease in average lab practical score of 0.215; however, the standardized regression coefficient has a positive value of 0.078.

Research Question 2: Logistic Regression Model

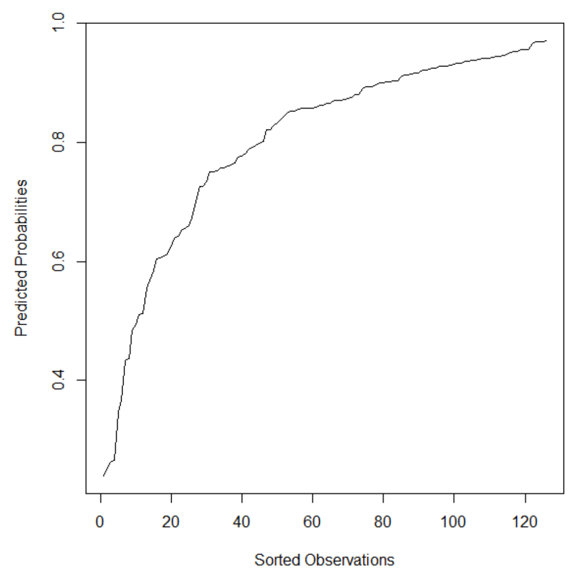
The multiple regression analysis for Research Question 1 allowed the researchers to evaluate predictors to a continuously distributed outcome—average lab practical scores. A logistic regression was used next to evaluate the relationship between the predictors and a categorical outcome- passing all three lab practical exams. The model, including all three predictors, resulted in the following partial slopes (Table 7).

Table 7. *Results of Logistic Regression*

Predictor Variable	β	*SE*	*z ratio*	*p*
Intercept	–64.381	28.609	–2.250	.024
Cumulative science GPA	2.510	1.109	2.263	.024
MCAT Score	0.111	0.056	1.957	.050
First-generation status	–1.593	0.557	–2.860	.004

The predicted probabilities of passing the lab practicals for each observation in the data set. The *x*-axis of the plot is the observations sorted by the predicted probabilities; the *y*-axis is the predicted probabilities of passing. Most observations are predicted to have high probabilities of passing all three lab exams (Figure 2).

Figure 2. *Predicted Probabilities of Passing of Each Observation*



As the *y* value of the inflection point in Figure 2 is close to 0.82, the present study sets $p = .82$ as the critical value (the observation with probability higher than 0.82 is predicted to be passing) to evaluate the model fit (Table 8.) The analysis shows both low sensitivity (true positive) and specificity (true negative) as well as a high false-positive rate (claiming passing when it doesn't exist) and a high false-negative rate (claiming not passing when passing occurs). Furthermore, the Hosmer-Lemeshow Goodness-of-Fit Test resulted in a test statistic of 8.492 with df = 8 and a *p*-value of .387, which affirms the poor fit of the logistic regression.

Table 8. *Result of Model Fit Evaluation*

	Sensitivity (%)	Specificity (%)	False Positive (%)	False Negative (%)
Pass/Fail	29.70	36.00	64.00	70.30

Discussion and Conclusions

The findings of this study showed that cumulative science GPA and MCAT scores, along with first-generation status, were statistically significant predictors of the continuous outcome of average lab practical scores in gross anatomy performance. However, the overall multiple regression model demonstrated a weak predictive power. Prior studies have explored the link between pre-admissions data and the basic sciences, but there are few reports of strong, definitive relationships. Wiley et al. (1996) concluded that there is a relationship between MCAT sub-section scores and medical school grades obtained during the basic science courses (an average GPA including all the basic sciences). However, the medical school grades were a combination of multiple basic science courses in the form of a cumulative GPA, thus emphasizing the point that there are few studies that examine the relationship between pre-admission data and a specific basic science course like gross anatomy.

A major limitation of this type of study is that there is no data collected for students who did not get accepted to the medical school cohort. For example, it remains unknown as to whether those potential students would have performed well in the gross anatomy course. Another limitation of the study is the restriction of range of the data points; students got accepted to medical school because of high GPAs and high MCAT scores. Finally, this pilot study only explored one cohort which can potentially explain how a weak association gets reported with statistical significance. For example, seven observations from this cohort were missing MCAT scores, which could be due to clerical errors, pipeline programs that forego MCAT scores, or other special considerations unbeknownst to the researchers.

While future studies incorporating multiple cohorts may assist in defining the relationship between specific predictors and performance in gross anatomy more thoroughly, adding a student’s prior anatomy coursework might be a rich avenue of exploration. Robertson (2020) concluded that there was no relationship between prior anatomy experience and anatomy grades; however, Forester (2002) found that prior anatomy coursework resulted in an increase of points accumulated in the anatomy course, particularly if the prior course involved cadaveric prosections. Peterson (2005) found that prior work in anatomy increased the students’ performances on the final exam and improved the student’s rank in the anatomy course. It should be noted that both Robertson (2020) and Forester (2002) used self-reported pre-admission anatomy experiences. To avoid that, the information about previous exposure to anatomy could be manually extracted from individual transcripts.

The focus of this current study was to explore the relationship between admissions data and gross anatomy performance. The long-term vision of this study aims to encourage educational policy and practice makers within medical education to utilize the formulated linear regression model and future predictive models to consider student progression and development. Lei et al., (2017) proposed a framework for decision-making based on student development theory and predictive analytics provide a foundation for the consideration of classroom data, pre-admissions data, and other metrics to inform student development more broadly and not just limited to an individual course. For example, using “in-class assessment data,” Gullo shows that performance early in medical school is a predictor of future performance in medical school (Gullo, 2016). Is this correlative or causative? Is this a place for qualitative methods to complement the readily collected quantitative data? Future research should consider the nature of these relationships and their implications for academic performance as a component of student development.

In conclusion, this sample of first-year medical students exhibited modest associations between pre-admission parameters and lab performance outcomes in gross anatomy at the medical school level. These models could inform supportive measures for medical students to improve performance and encourage successful progress.

References

Corresponding Author: Ashley N. Walker
Author Contact Information: ashleynwalker@ufl.edu

Almarabheh, A., Shehata, M. H., Ismaeel, A., Atwa, H., & Jaradat, A. (2022). Predictive validity of admission criteria in predicting academic performance of medical students: A retrospective cohort study. *Frontiers in Medicine*, *9*. https://doi.org/10.3389/fmed.2022.971926

Basco Jr, W. T., Way, D. P., Gilbert, G. E., & Hudson, A. (2002). Undergraduate institutional MCAT scores as predictors of USMLE step 1 performance. *Academic Medicine*, *77*(10), S13–S16. https://doi.org/10.1097/00001888-200210001-00005

Bird, K. A., Castleman, B. L., Mabel, Z., & Song, Y. (2021). Bringing transparency to predictive analytics: A systematic comparison of predictive modeling methods in higher education. *AERA Open*, *7*. https://doi.org/10.1177/23328584211037630

Brooks, W. S., Woodley, K. T. P., Jackson, J. R., & Hoesley, C. J. (2015). Integration of gross anatomy in an organ system-based medical curriculum: Strategies and challenges. *Anatomical Sciences Education*, *8*(3), 266–274. https://doi.org/10.1002/ase.1483

Ferguson, E., James, D., & Madeley, L. (2002). Factors associated with success in medical school: Systematic review of the literature. *BMJ: British Medical Journal*, *324*(7343), 952–957. https://doi.org/10.1136/bmj.324.7343.952

Finnerty, E. P., Chauvin, S., Bonaminio, G., Andrews, M., Carroll, R. G., & Pangaro, L. N. (2010). Flexner revisited: The role and value of the basic sciences in medical education. *Academic Medicine*, *85*(2), 349–355. https://doi.org/10.1097/ACM.0b013e3181c88b09

Forester, J. P., McWhorter, D. L., & Cole, M. S. (2002). The relationship between premedical coursework in gross anatomy and histology and medical school performance in gross anatomy and histology. *Clinical Anatomy*, *15*(2), 160–164. https://doi.org/10.1002/ca.1114

Fredieu, J. R., & Snyder, C. W. (2015). Positive impact of a master of science in applied anatomy program on USMLE Step 1 performance. *Anatomical Sciences Education*, *8*(1), 31–36. https://doi.org/10.1002/ase.1455

Gauer, J. L., Wolff, J. M., & Jackson, J. B. (2016). Do MCAT scores predict USMLE scores? An analysis on 5 years of medical student data. *Medical Education Online*, *21*(1). https://doi.org/10.3402/meo.v21.31795

Gliatto, P., Leitman, I. M., & Muller, D. (2016). Scylla and Charybdis: The MCAT, USMLE, and degrees of freedom in undergraduate medical education. *Academic Medicine*, *91*(11), 1498–1500. https://doi.org/10.1097/ACM.0000000000001247

Gullo C. A. (2016). The future is in the numbers: The power of predictive analysis in the biomedical educational environment. *Medical Education Online*, *21*. https://doi.org/10.3402/meo.v21.32516

Lei, X.-F., Yang, M., & Cai, Y. (2017). Educational data mining for decision-making: A framework based on student development theory. *Advances in Engineering Research*, 117, 628–641. https://doi.org/10.2991/eeeis-16.2017.76

Li, J., Thompson, R., & Shulruf, B. (2019). Struggling with strugglers: Using data from selection tools for early identification of medical students at risk of failure. *BMC Medical Education*, *19*(1), 415. https://doi.org/10.1186/s12909-019-1860-z

Little, R. J. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*, 90(431), 1112–1121. https://doi.org/10.1080/01621459.1995.10476615

Olsen, A. A., McLaughlin, J. E., & Harpe, S. E. (2020). Using multiple linear regression in pharmacy education scholarship. *Currents in Pharmacy Teaching and Learning*, *12*(10), 1258–1268. https://doi.org/10.1016/j.cptl.2020.05.017

Peterson, C. A., & Tucker, R. P. (2005). Undergraduate coursework in anatomy as a predictor of performance: Comparison between students taking a medical gross anatomy course of average length and a course shortened by curriculum reform. *Clinical Anatomy*, *18*(7), 540–547. https://doi.org/10.1002/ca.20154

Qahmash, A., Ahmad, N., & Algarni, A. (2023). Investigating students' pre-university admission requirements and their correlation with academic performance for medical students: An educational data mining approach. *Brain Sciences*, *13*(3), 456. https://doi.org/10.3390/brainsci13030456

Robertson, E. M., Thompson, K. L., & Notebaert, A. J. (2020). Perceived benefits of anatomy coursework prior to medical and dental school. *Anatomical Sciences Education*, *13*(2), 168–181. https://doi.org/10.1002/ase.1882

RStudio Team (2020). *RStudio: Integrated development for R*. http://www.rstudio.com/

Tucker R. P. (2008). Performance in a prematriculation gross anatomy course as a predictor of performance in medical school. *Anatomical Sciences Education*, *1*(5), 224–227. https://doi.org/10.1002/ase.48

Veloski, J. J., Callahan, C. A., Xu, G., Hojat, M., & Nash, D. B. (2000). Prediction of students' performances on licensing examinations using age, race, sex, undergraduate GPAs, and MCAT scores. *Academic Medicine*, *75*(10), S28–S30. https://doi.org/10.1097/00001888-200010001-00009

Wiley A, Koenig JA. (1996). The validity of the MCAT for predicting performance in the first two years of medical school. *Academic Medicine*, *71*, S83–S85. https://doi.org/10.1097/00001888-199610000-00052

Yates, J., & James, D. (2006). Predicting the "strugglers": A case-control study of students at Nottingham University Medical School. *British Medical Journal (Clinical Research Edition)*, *332*(7548), 1009–1013. https://doi.org/10.1136/bmj.38730.678310.63

Zhao, X., Oppler, S., Dunleavy, D., & Kroopnick, M. (2010). Validity of four approaches of using repeaters' MCAT scores in medical school admissions to predict USMLE Step 1 total scores. *Academic Medicine*, *85*(10), S64–S67. https://doi.org/10.1097/ACM.0b013e3181ed38fc

Zhao, Y., Qiangwen, X., Ming, C., & Gary, M. W. (2020). Predicting student performance in a Master of data science program using admissions data. *Proceedings of the 13th International Conference on Educational Data Mining*, 325–333.