Paper ID #44721

A Preliminary Study on the Impact of Lower-Division Mathematics Courses on Student Success in Engineering

Dr. Umut Can Cabuk, San Diego State University

Umut Can Cabuk received his B.Sc. degree in electronics engineering from Bursa Uludag University (Turkey) in 2012, his M.Sc. degree in information technology engineering from Aarhus University (Denmark) in 2015, and his Ph.D. degree at the International Computer Institute of Ege University (Turkey), where he also worked as a research assistant. He is currently continuing his post-doc studies at San Diego State University (USA). His research interests include drones, mobile and wireless networks, the Internet of Things, computer security, and graph theory. He has co-authored over 35 scholarly publications and has 4 patent applications.

Dr. Yusuf Ozturk, San Diego State University

A Preliminary Study on the Impact of Lower-Division Mathematics Courses on Student Success in Engineering

Umut Can Cabuk and Yusuf Ozturk
ucabuk@sdsu.edu
Department of Electrical and Computer Engineering
San Diego State University

Abstract

Students' success in engineering programs – in terms of retention and graduation rates – is significantly influenced by various cognitive and non-cognitive factors, as revealed in the literature. While earlier studies provided comprehensive analyses of many of those factors, we see that the impact of success in individual courses is often overlooked. Yet among the courses offered, lower-division mathematics classes play a critical role in student retention and graduation rates. Also, considering that not all students take the same set of courses in their first year due to transfers from other institutions, change of majors, and necessary preparatory programs, such an analysis based on individual courses becomes an important indicator. This study aims to address this oversight by examining the impact of early mathematics performance on students' long-term academic success in engineering disciplines. Our research employs regressions and advanced machine learning techniques to analyze the relationship between students' achievements in initial mathematics courses and their retention and subsequent success in higher-division engineering courses. The analysis demonstrates a strong correlation, as our predictive models forecast student success in upper-level courses with approx. 70% accuracy. This finding underscores the importance of early math education in shaping the future academic trajectory of students, especially in mechanical engineering and electrical and computer engineering. Moreover, the study offers significant insights for educational institutions. The results can inform strategic curriculum development, the design of more effective student support systems, and better resource allocation to address the specific needs of students in these fields. Beyond the academic literature on student success in engineering, this research also provides a practical framework for educational institutions to make informed decisions in curriculum design and planning in the evolving field of engineering education.

1 Introduction

Engineering curricula are intentionally designed to ensure contingency, consistency, and integrity, recognizing that later courses build progressively on the knowledge acquired in earlier ones. This is especially evident with mathematics-related courses, including but not limited to Calculus, Differential Equations, Trigonometry, Linear Algebra, and others[1]. The significance of mathematical prerequisites in achieving success in upper-division engineering courses is pivotal. Though not all courses are a continuation of an earlier one, multiple interweaved threads of courses are identifiable. In engineering curricula, such progression from basic to advanced courses is not only sequential but also interdependent, with each course designed to complement and reinforce the concepts of its predecessors.

Our research specifically examines the impact of performance in mathematics prerequisites on subsequent success in advanced engineering courses. We hypothesize that a strong performance in these foundational math courses is a reliable predictor of academic success in later, more complex engineering courses, as we work on the data from our College of Engineering as a case study. Identifying students who struggle in math prerequisites early can lead to timely and effective interventions, such as specialized tutoring or additional coursework tailored to strengthen their mathematical foundation.

This study leverages educational data mining and machine learning (ML) to develop a predictive model that assesses the influence of mathematical preparedness on future academic outcomes in the engineering disciplines. The model utilizes algorithmic and statistical techniques to identify patterns linking mathematical proficiency with success in upper-division engineering courses. Our novel contributions to the literature are centered on our student and grade database being very large, and newer ML methods have not yet been applied to a similar case.

2 Literature

The literature on engineering education does not lack research on predicting student success and retention, utilizing a variety of cognitive and non-cognitive metrics[2,3,4,5]. This section highlights key works related to this subject. Alshanqiti et al.[6] introduced a pioneering approach to predicting student performance. Their method integrated collaborative filtering, fuzzy set rules, and Lasso linear regression, focusing on dynamic weighting to enhance prediction accuracy. Nabil et al.[7] furthered this field by applying deep learning to educational data mining, using models like deep neural networks, decision trees, and random forests. Their work, centered on first-year course grades from a public university, tackled the common issue of imbalanced and incomplete datasets in educational data mining. These foundational studies have significantly influenced modern methodologies in this domain.

Similarly, Costa et al.[8] employed four different algorithms, including the Support Vector Machine method, to identify students at high risk of failure. Their research used weekly assessments and two exams in two courses to predict student success during a semester. Complementing these findings, Adnan et al.[9] utilized the Open University Learning Analytics Dataset (OULAD) for developing models that predict student success at various coursework stages. Employing six different ML algorithms, they segmented the course into six periods for predictive modeling, categorizing students into withdraw, pass, fail, and distinction. Their research underlined the importance of feature engineering, achieving high precision and recall rates of around 90%. These studies play a crucial role in establishing contemporary approaches within our study.

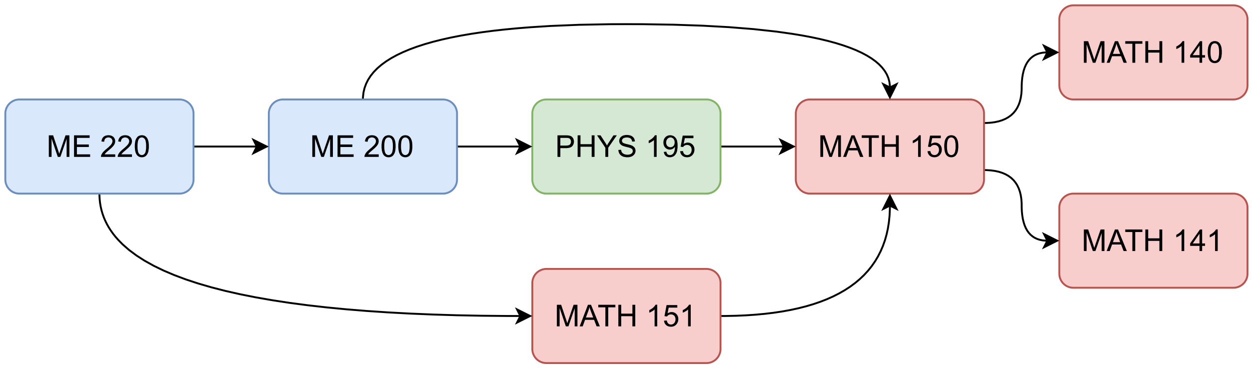


Figure 1: Prerequisite threads for ME200 Statics and ME220 Dynamics.

3 Methods

While some prerequisites prove more critical than others in forecasting student performance, understanding these relationships is key to our model's development. However, it is also important to recognize that correlation does not imply causation. For instance, a strong correlation between grades in *Data Structures* and *Digital Electronics* might stem from their shared difficulty level. Therefore, we limit our scope regarding prerequisite courses to math classes, which are almost always foundational in engineering education.

3.1 Dataset

Our dataset comprises comprehensive grade records for all engineering students enrolled in major and preparatory courses in the College of Engineering at San Diego State University since 2006. As of Fall 2021, the College of Engineering had an enrollment of 3,200 students across various undergraduate programs, including Aerospace Engineering, Civil Engineering, Construction Engineering, Computer Engineering, Electrical Engineering, and Mechanical Engineering. The dataset includes demographic features and academic performance metrics, though this study focuses exclusively on grades from prerequisite courses to estimate student success and, thus, identify at-risk students.

3.2 Data Preprocessing

Figure 1 illustrates a feature table considering the causality in prerequisite courses. It includes all math-related courses in the prerequisite chain for a target course. For instance, the prerequisite chain for the mechanical engineering courses ME200 involves immediate prerequisites like PHYS195 and MATH150, as well as earlier courses. Notably, some students completed major preparation courses at community colleges, leaving their grades absent from our dataset.

Our method for building the feature table for the target course involves identifying students who enrolled in the target course first and then compiling a list of their prerequisite courses and grades. If a prerequisite was not completed prior to the target course, it's excluded from consideration. Only the most recent attempt at a prerequisite is included, regardless of whether earlier attempts resulted in higher grades. Prerequisites with limited student enrollment and students who bypassed most prerequisites are omitted from the training data as outliers. To address the impact of noise in the dataset (e.g., outliers and missing parts), which can significantly impair classification accuracy and prediction quality, we excluded data older than seven years. This decision considers various evolving factors affecting academic performance, such as curriculum changes, teaching methodologies, policy shifts, and extraordinary circumstances like the COVID-19 pandemic. Aberrations in student records, like unexpectedly low grades from students with otherwise high prerequisite GPAs, are considered noise. These anomalies, possibly due to personal or medical reasons, are treated as outliers for data integrity.

3.3 Feature Engineering

Given the differences among courses in terms of weekly hours and credits, a weighted average ”grade points” calculation is typically used to gauge student success across courses (Eq. 1). However, feature scaling is crucial for machine learning applications like the k-Nearest Neighbor classifier (kNN)[10] we utilized. Without it, features with larger numerical values could disproportionately influence the distance metric and, consequently, the classifier’s output. Therefore, rescaling grades is necessary for more uniform and interoperable comparisons.

$$gradepoints \;=\; grade \;\times\; numberofunits \tag{1}$$

While rescaling grade points alters their initial meaning, it’s essential to account for courses where the grade range may not fully extend from 0 to 4.0. To ensure consistency across different courses, we applied MinMax scaling (Eq. 2) to standardize grades within a 0 to 1 scale uniformly, thus preserving the crucial information necessary for enhancing model accuracy.

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2}$$

3.4 Completing Missing Data

Missing data typically falls into three categories: ”missing completely at random (MCAR)”, ”missing at random (MAR)”, or ”missing not at random (MNAR)”[11]. Multiple imputation methods are commonly used for data classified as MAR or MCAR, offering the advantage of dealing with missing entries in a statistically robust and unbiased manner. It is important to distinguish between those: MCAR is characterized by a consistent distribution across different features, unlike MAR. In our study, the absent grades are predominantly MAR, not MCAR. This typically involves transfer students who completed equivalent prerequisite courses at community colleges or cases where higher-level courses fulfill lower-level course (1xx) requirements. The regression filling method is a refined approach that solves the problem by calculating a regression polynomial and uses values from the constructed polynomial, which means it is not impacted by outliers for cases with many prerequisites[12,13]. Hence, we opted to fill the gaps in our dataset by regression. Our regression is, accuracy-wise, comparable to using a row-median filling method.

3.5 Training The kNN Model

We decided to develop our model based on the kNN method as it suits our dataset well. kNN, as a non-parametric ML classifier, classifies students by evaluating their k nearest neighbors. So that the distance between each student is calculated instead of relying solely on model parameters. The distance between two students is often calculated by either the Euclidean distance given in Eq. 3 or the Manhattan distance[14] given in Eq. 4, depending on the case.

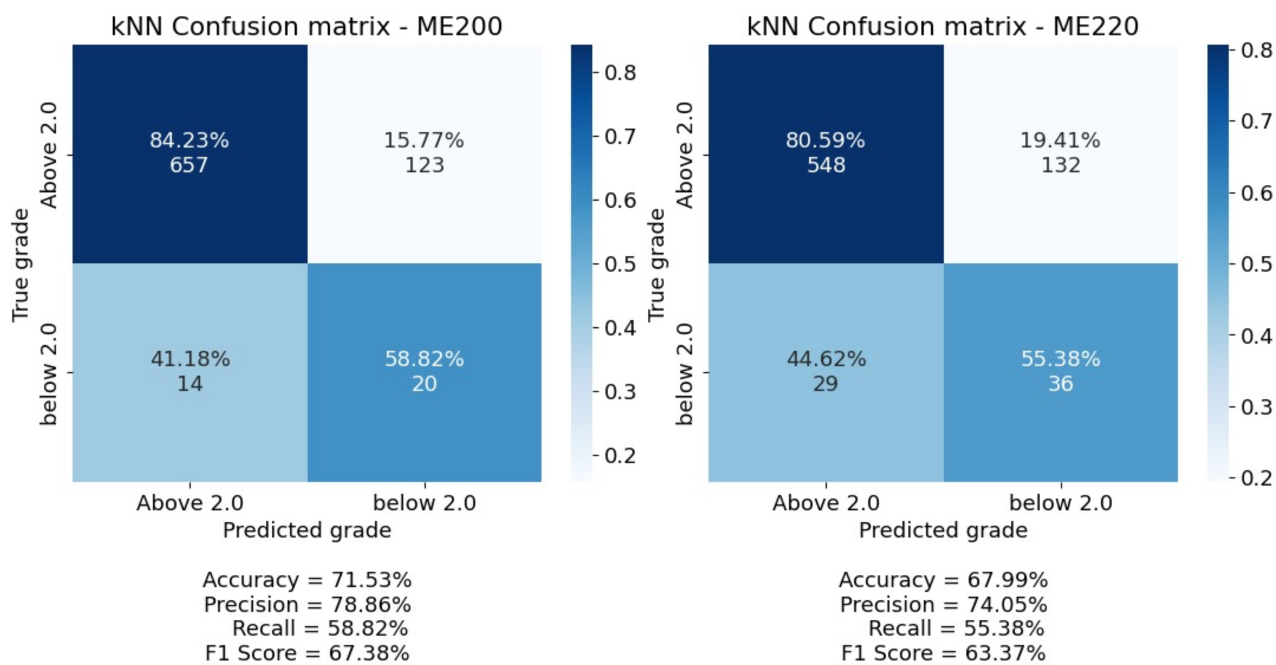
$$d(x,y) = \sqrt{\sum_{i=1}^{n} (y_i - \ x_i)^2} \tag{3}$$

$$d(x,y) = \sum_{i=1}^{D} |x_i - y_i| \tag{4}$$

4 Results

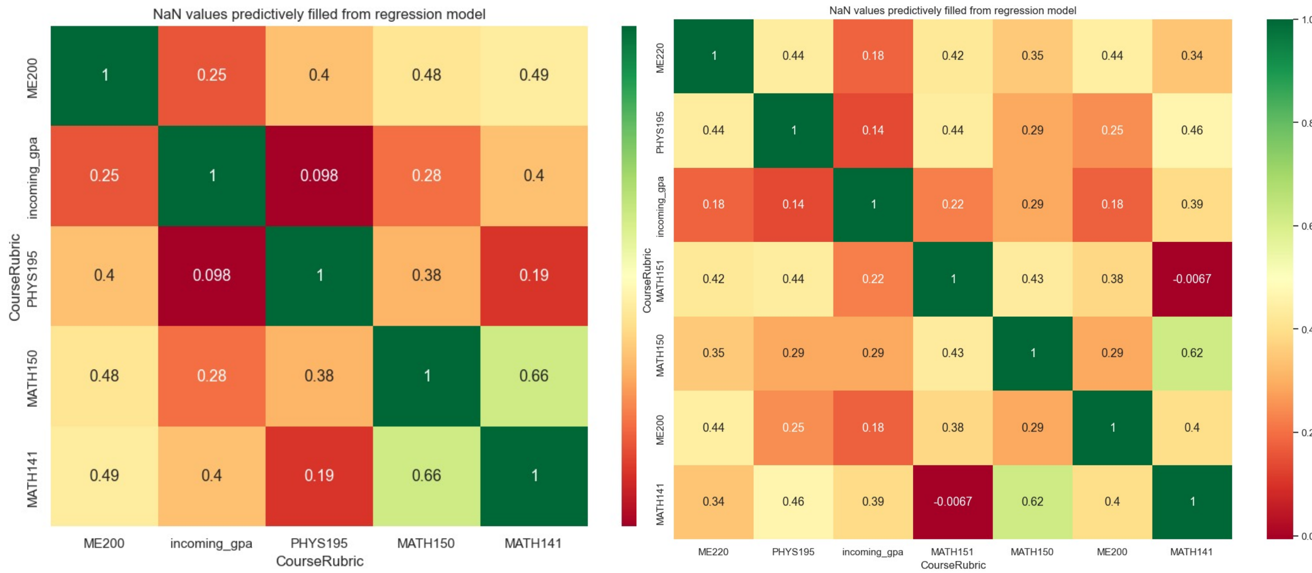
In this section, without commenting, we present our findings regarding the accuracy of our kNN-based success estimation model, which was built to estimate a student's expected grade in a target course by considering his/her earlier grades in all courses in the chosen target course's prerequisite chain, as well as that student's incoming GPA (i.e., high school or transfer).

Figure 2 shows the confusion matrices for the two courses that were selected as exemplary subjects. The confusion matrices were based on the students' grade estimation with a designated pass/fail threshold of 2.0 points. In those matrices, the top-left quadrants give the true positive rates as 84.23% and 80.59% for ME200 and ME220, respectively, whereas false positives, shown in the top-right quadrants, are 15.77% and 19.41% for ME200 and ME220. Likewise, the



(a) ME200 Statics. (b) ME220 Dynamics.

Figure 2: Confusion matrices showing model accuracy.



(a) ME200 Statics.

(b) ME220 Dynamics.

Figure 3: Correlation heatmaps showing the impact of prerequisites.

bottom-left quadrants show the false negatives as 41.18% and 44.62% for ME200 and ME220, whereas the bottom-right quadrants give the true negatives as 58.82% and 55.38% for ME200 and ME220, respectively. On the other hand, the overall accuracy of the model is calculated as 71.53% for ME200 and 67.99% for ME220.

Figure 3 shows the correlation heatmaps for both courses analyzed within our experiments, namely ME200 and ME220. The designated heatmaps are indeed visualized correlation matrices that show the calculated weight (i.e., impact) of success in each prerequisite course on success in the target course. While 1 indicates a direct positive correlation, 0 indicates no observable correlation, and -1 (which does not exist at all) would indicate a negative correlation. It is especially useful in evaluating target courses' dependency on their individual prerequisite courses.

5 Discussion

In the discussion of the results from the confusion matrices for courses ME200 and ME220, several key insights emerge regarding the effectiveness and limitations of our grade estimation model, particularly in its ability to discern pass/fail outcomes based on the threshold. Shown in Figure 2, the true positive rates (TPR) are commendably high, with ME200 at 84.23% and ME220 at 80.59%. This suggests that the model is effective in correctly identifying students who have (and would have) passed, with ME200 showing a slightly better performance in this regard, pointing out fluctuations among different courses. However, the false positive rates (FPR) stand at 15.77% for ME200 and 19.41% for ME220, indicating a notable proportion of students who were predicted to pass but actually failed. The higher FPR for ME220 suggests this course's predictions are less reliable in identifying students who shouldn't pass. Hence, we can say that the model's reliability is not uniform across all courses.

Nevertheless, the false negative rates (FNR) are concerning, being over 40% for both courses. This means a significant number of students who passed were predicted to fail. Such high FNRs are critical areas for improvement, as they suggest many students who are capable of passing might be underestimated by the model. Conversely, the true negative rates (TNR) are 58.82% for ME200 and 55.38% for ME220, showing moderate effectiveness in correctly identifying students who have (or would have) failed. The overall accuracy of the model stands at 71.53% for ME200 and 67.99% for ME220. While these figures are decent, they indicate that there is room for improvement. A promising way to further improve our model is to incorporate more features into the model-building and training phases. Those features may include demographic information and success in courses outside the target course' prerequisite chain.

Analyzing the correlations from the heatmaps, we see that ME200 and ME220 (likely other courses as well) interact differently with various academic predictors. In ME200, the correlation with a student's incoming GPA is at 0.25, implying a tangible, albeit not strong, relationship between a student's past academic performance and their success in ME200. The strongest linkage for ME200 is with MATH141 Precalculus at 0.49, signifying that mathematical skills are quite predictive of performance in this course, just like MATH150 Calculus I at 0.48. Yet PHYS195 shows a somewhat lower correlation at 0.40, suggesting that while related, it is not as significant a predictor of ME200 performance as MATH141 and MATH150.

In the case of ME220, the correlation with incoming GPA drops to 0.18, hinting that ME220, a more advanced course, might be less influenced by overall academic history and more by the competencies developed during the course. The correlation with PHYS195 at 0.44 is the most pronounced for ME220, reinforcing the idea that a solid understanding of physics concepts may be crucial for excelling in ME220. Interestingly, ME220 shows a relatively lower yet notable correlation with MATH150 at 0.35, which indicates that while mathematical skills are important, they do not predict success in ME220 as strongly as they do in ME200.

Comparatively, the correlation between ME200 and ME220 themselves is 0.44, revealing that performance in one is remarkably associated with performance in the other. This suggests some overlapping skills or knowledge bases are required for both courses but also points to distinct course-specific factors that influence student success. While prerequisite courses are indeed tied to each other in terms of success rates, the aforementioned distinctness, however, could be more evident in less related course pairs. Therefore, curriculum design efforts and targeted academic support should consider these ties and differences among various courses.

References

[1] Fathiah Zakaria, Siti Aishah Che Kar, Rina Abdullah, Syila Izawana Ismail, and Nur Idawati Md Enzai. A study on correlation of subjects on electrical engineering course using artificial neural network (ann). *Asian Journal of University Education*, 17(2):144–155, 2021.

[2] Dalia Abdulkareem Shafiq, Mohsen Marjani, Riyaz Ahamed Ariyaluran Habeeb, and David Asirvatham. Student retention using educational data mining and predictive analytics: a systematic literature review. *IEEE Access*, 2022.

[3] Xuansheng Lu, Yanmin Zhu, Yanan Xu, and Jiadi Yu. Learning from multiple dynamic graphs of student and course interactions for student grade predictions. *Neurocomputing*, 431:23–33, 2021.

[4] Yahia Baashar, Gamal Alkawsi, Nor'ashikin Ali, Hitham Alhussian, and Hussein T Bahbouh. Predicting student's performance using machine learning methods: A systematic literature review. In *2021 International Conference on Computer Information Sciences (ICCOINS)*, pages 357–362, 2021. doi: 10.1109/ICCOINS49721.2021.9497185.

[5] Boran Sekeroglu, Rahib Abiyev, Ahmet Ilhan, Murat Arslan, and John Bush Idoko. Systematic literature review on machine learning and student performance prediction: Critical gaps and possible remedies. *Applied Sciences*, 11(22), 2021. ISSN 2076-3417. doi: 10.3390/app112210907. URL https://www.mdpi.com/2076-3417/11/22/10907.

[6] Abdullah Alshanqiti and Abdallah Namoun. Predicting student performance and its influential factors using hybrid regression and multi-label classification. *IEEE Access*, 8:203827–203844, 2020.

[7] Aya Nabil, Mohammed Seyam, and Ahmed Abou-Elfetouh. Prediction of students' academic performance based on courses' grades using deep neural networks. *IEEE Access*, 9:140731–140746, 2021.

[8] Evandro B. Costa, Baldoino Fonseca, Marcelo Almeida Santana, Fabrsia Ferreira de Arajo, and Joilson Rego. Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*, 73:247–256, 2017. ISSN 0747-5632. doi: https://doi.org/10.1016/j.chb.2017.01.047. URL https://www.sciencedirect.com/science/article/pii/S0747563217300596.

[9] Muhammad Adnan, Asad Habib, Jawad Ashraf, Shafaq Mussadiq, Arsalan Ali Raza, Muhammad Abid, Maryam Bashir, and Sana Ullah Khan. Predicting at-risk students at different percentages of course length for early intervention using machine learning models. *Ieee Access*, 9:7519–7539, 2021.

[10] Kilian Q Weinberger, John Blitzer, and Lawrence Saul. Distance metric learning for large margin nearest neighbor classification. *Advances in neural information processing systems*, 18, 2005.

[11] Maisarah Zorkeflee, Aniza Mohamed Din, and Ku Ruhana Ku-Mahamud. Fuzzy and smote resampling technique for imbalanced data sets. 2015.

[12] Jiang Li, Xiaowei S Yan, Durgesh Chaudhary, Venkatesh Avula, Satish Mudiganti, Hannah Husby, Shima Shahjouei, Ardavan Afshar, Walter F Stewart, Mohammed Yeasin, et al. Imputation of missing values for electronic health record laboratory data. *NPJ digital medicine*, 4(1):147, 2021.

[13] Abdullah Z Alruhaymi and Charles J Kim. Why can multiple imputations and how (mice) algorithm work? *Open Journal of Statistics*, 11(5):759–777, 2021.

[14] Charu C. Aggarwal, Alexander Hinneburg, and Daniel A. Keim. On the surprising behavior of distance metrics in high dimensional space. In Jan Van den Bussche and Victor Vianu, editors, *Database Theory — ICDT 2001*, pages 420–434, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.