VOL. 21, NO. 8, APRIL 2026 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

COMPARATIVE ANALYSIS OF MACHINE LEARNING TECHNIQUES AND LINEAR REGRESSION IN PREDICTING ENGINEERING STUDENTS' PERFORMANCE IN DIFFERENTIAL EQUATIONS

Melchor G. Pacer
Department of Mathematics, College of Science, Technological University of the Philippines, Manila
E-Mail: melchor_pacer@tup.edu.ph

ABSTRACT

This study evaluated and compared four predictive models - Linear Regression, Decision Tree, Random Forest, and Support Vector Regression (SVR) - in estimating engineering students' final grades in Differential Equations. Academic grade records of 646 students from the Technological University of the Philippines enrolled in Mathematics in the Modern World (MMW), Calculus 1, Calculus 2, and Differential Equations were used. All models were evaluated using 10-fold cross-validation with R^2, Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) as performance metrics. Results showed that all models produced near-zero or negative R^2 values, with Linear Regression performing best (R^2 = -0.011, RMSE = 2.255, MAE = 1.699). Machine learning models did not outperform the linear baseline, suggesting that prerequisite mathematics grades alone are insufficient predictors of Differential Equation performance. The findings recommend the inclusion of richer predictor variables in future predictive models.

Keywords: machine learning, differential equations, linear regression, random forest, support vector regression, educational data mining, engineering education.

Manuscript Received 11 March 2026; Revised 18 April 2026; Published 20 June 2026

1. INTRODUCTION

Evaluating students' academic performance has become increasingly vital in higher education as a means to improve learning outcomes and better prepare graduates for a competitive workforce (Caday *et al*., 2025) (Alhazmi *et al.,* 2023). The ability to accurately predict academic success at an early stage is a cornerstone of modern educational strategy. This allows institutions to implement proactive interventions such as academic counseling, peer monitoring, and personalized learning plans before a student reaches a critical point of failure (Nishio *et al*., 2022). Consequently, the field of education has seen a significant shift from traditional, manual statistical assessments toward data-driven approaches such as Education Data Mining (EDM) and machine learning to analyze the complex factors influencing student progression (Cruz *et al*., 2025).

In the context of engineering programs, mathematics serves as a critical foundational pillar. Yet, it often presents the most significant obstacle to student success (Pallathadka *et al*., 2023) (Alhazmi *et al*., 2023). Among these courses, Differential Equations is recognized as one of the most challenging subjects for engineering students, frequently resulting in high failure or low performance rates (Jiao *et al*., 2022). At the Technological University of the Philippines (TUP) Manila Campus, students must navigate a rigorous mathematical sequence from Mathematics in the Modern World (MMW) through Calculus 1 and Calculus 2 before reaching DE. This sequential dependence necessitates the development of effective predictive tools to identify students who may struggle as they transition into more advanced mathematical coursework.

Traditional efforts to forecast student performance have historically relied on linear regression models, which are favored for their simplicity and interpretability but are inherently limited to capturing linear relationships between variables (Doz *et al*., 2023). Conversely, modern machine learning models, such as Decision Trees, Random Forests, and Support Vector Machines (SVM), offer the potential to capture more complex, non-linear patterns within educational data. However, the majority of existing comparative studies between these techniques have been conducted in a general education setting, with fewer studies focusing specifically on the performance of students in high-level engineering mathematics courses.

A notable research gap exists regarding the comparative performance of machine learning versus linear regression specifically for predicting outcomes in Differential Equations. While many EDM studies utilize expansive datasets incorporating demographics, attendance, and social engagement, there is a lack of evidence regarding whether prerequisite grades alone, specifically for foundational subjects, can serve as sufficient predictors of success in DE (Sayed *et al*., 2025). Furthermore, no known study has investigated this specific predictive problem within Philippine engineering education or in the context of TUP-Manila, where institutional grading standards and student backgrounds may present unique patterns.

This study aims to address these gaps by evaluating and comparing the predictive accuracy of four models: Linear Regression, Decision Tree, Random Forest, and Support Vector Regression (SVR). The study utilizes only prerequisite mathematics grades (MMW,

Calculus 1, and Calculus 2) as independent variables to determine which model most effectively predicts the final DE grades of 646 TUP engineering students. To ensure robustness, the models are evaluated using 10-fold cross-validation, with performance measured through the Coefficient of Determination (R^2), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE)

2. METHODOLOGY

A. Research Design

This study implemented a quantitative, comparative research design in evaluating and comparing the predictive performance of traditional linear regression and machine learning models, specifically Decision Tree, Random Forest, and Support Vector Machine, in estimating engineering students' final grades in Differential Equations. The study analyzed secondary academic records from 646 students enrolled in the Technological University of the Philippines (TUP) Manila Campus during the 2nd Semester of the School Year 2022-2023.

B. Dataset

The dataset consisted of 646 student records (n = 646) with the final numerical grades in four mathematics courses: Mathematics in the Modern World (MMW), Calculus 1 (C1), Calculus 2 (C2), and Differential Equations (DE), where in the first three are the independent variables while the latter is the dependent variable. The grade data is collected through the TUP Registrar's Office. The descriptive statistics are presented in Table-1.

Table-1. Descriptive statistics of study variables (n = 646).

Variable	MMW	C1	C2	DE
Mean	84.11	84.14	84.27	85.27
Standard Deviation	2.33	2.35	2.44	2.25
Minimum Value	76	76	76	72
Maximum Value	98	92	97	93

Figures 1 to 4 show the Grade distribution for MMW, Calculus 1, Calculus 2, and Differential Equations.

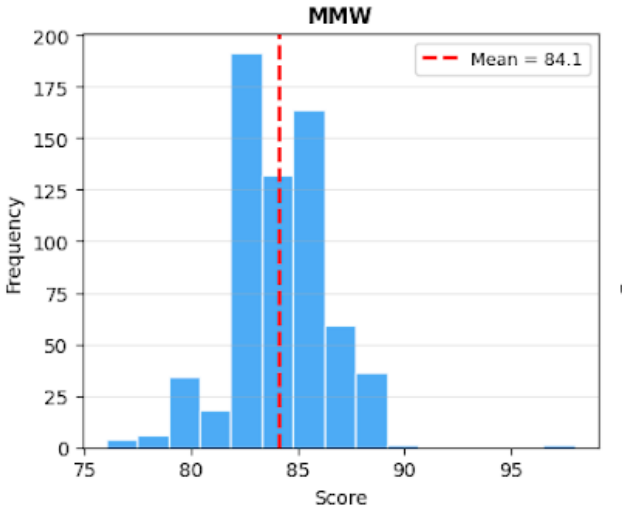

Figure-1. Grade distribution for mathematics in the modern world.

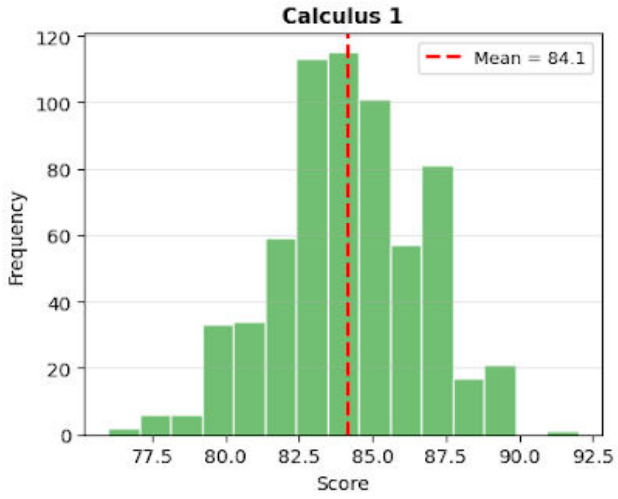

Figure-2. Grade distribution for Calculus 1.

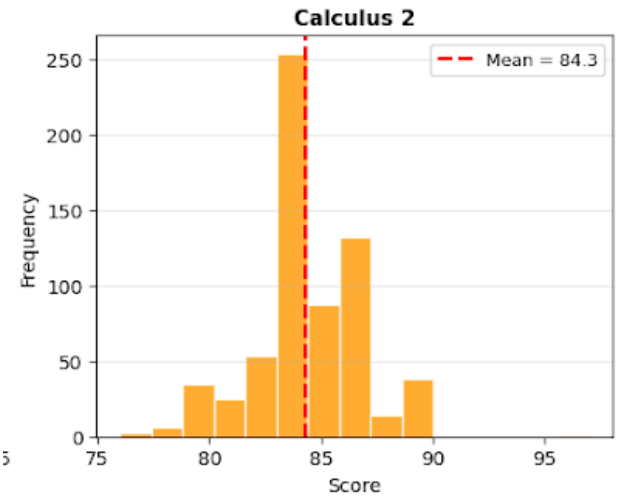

Figure-3. Grade distribution for Calculus 2.

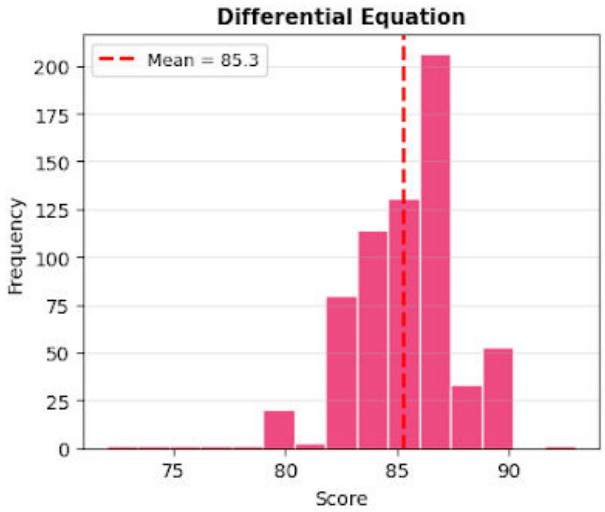

Figure-4. Grade distribution for differential equations.

C. Data Preprocessing

Before the actual implementation in the ML models, the dataset was pre-processed to ensure that there are no missing values. All variables were continuous numerical scores on the same 0-100 scale. Furthermore, no categorical encoding was required. Feature standardization ($\mu = 0$, $\sigma = 1$) was applied exclusively for the SVR model, as Support Vector Regression is sensitive to feature scale. All other models were trained on raw grade values.

D. Predictive Models

Four models were used to evaluate the datasets and predict the effects of the independent variables on the dependent variables. Table-2 summarizes the algorithm and parameters for each model.

Table-2. Predictive models and parameters.

Model	Algorithm	Description
Linear Regression	OLS	Serves as the baseline model of the study. Assumes a straight-line relationship between predictors and outcome. Estimates DE grade as a weighted linear combination of predictors.
Decision Tree	CART, unpruned	Recursively partitions feature space using if-then rules. Serves as an interpretable non-linear baseline.
Random Forest	100 trees, bootstrap	Ensemble of trees trained on random subsamples. Reduces overfitting through averaging.
SVM (SVR)	RBRF Kernel, C=10, ε=0.1	Maps inputs to higher-dimensional space; minimizes error within a tolerance band ε.

E. Model Evaluation

All models were evaluated using 10-fold cross-validation, which provides unbiased generalization estimates by rotating each of 10 equal folds as the validation set. Three metrics were computed per fold and averaged. All analyses were implemented in Python 3 using the scikit-learn library:

Coefficient of Determination (R^2). Proportion of variance in DE grades explained by the model.

Root Mean Squared Error (RMSE). Penalizes large prediction errors.

Mean Absolute Error (MAE). More robust to outliers

3. RESULTS AND DISCUSSIONS

A. Correlation Analysis

Before model development, Pearson correlation analysis was conducted to examine the bivariate linear relationships between each predictor variable - Mathematics in the Modern World (MMW), Calculus 1, and Calculus 2 - and the outcome variable, Differential Equations grade. This step was essential for understanding the strength and direction of association between variables before proceeding to multiple regression and machine learning modeling. The full correlation heatmap for all four variables is presented in Figure-5.

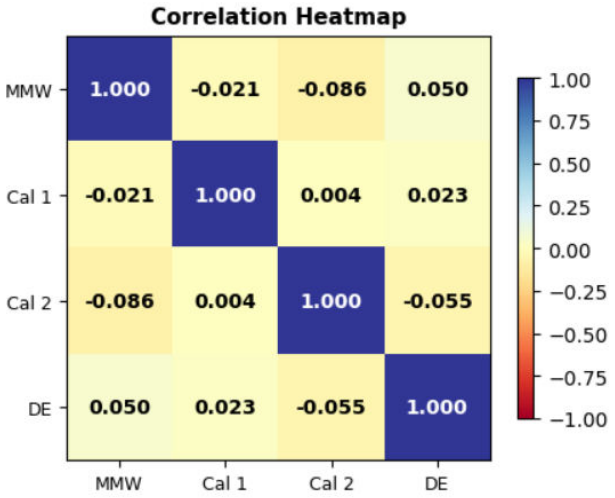

Figure-5. Correlation Heatmap.

Figure 6 to 8 show the Relationship of the independent variables (MMW, Calculus 1, Calculus 2) to the dependent variable (Differential Equation).

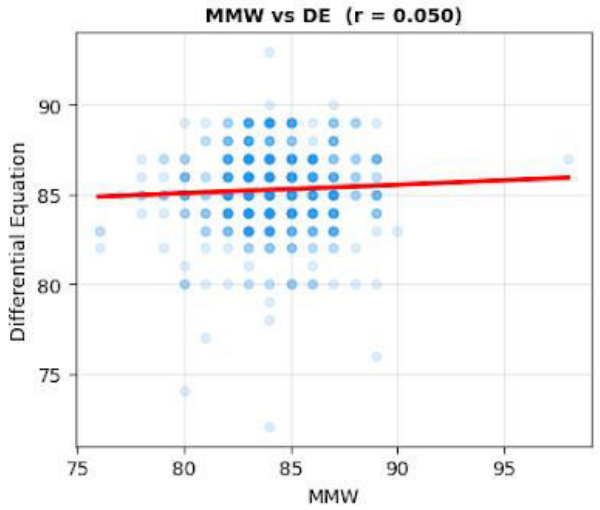

Figure-6. Relationship of the MMW and DE Grades.

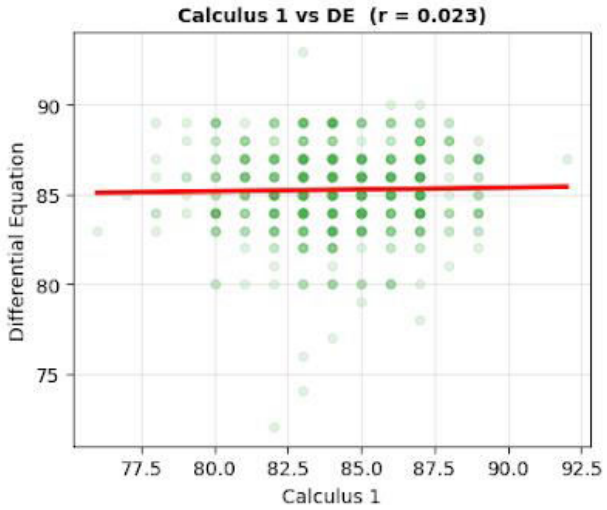

Figure-7. Relationship of the Calculus 1 and DE Grades.

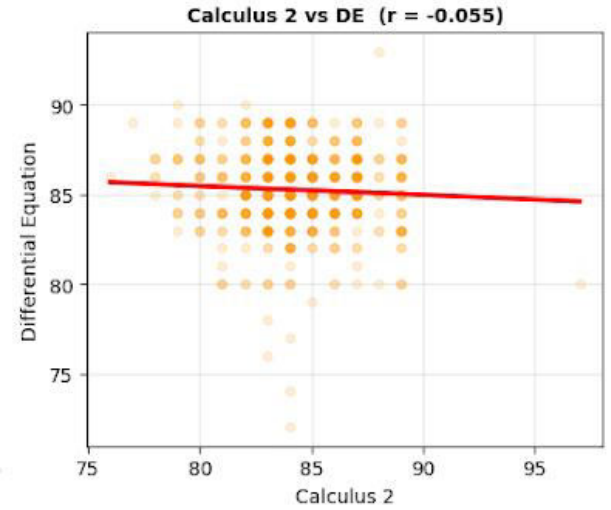

Figure-8. Relationship of the Calculus 2 and DE Grades.

The results revealed uniformly weak correlations between all three predictor variables and Differential Equations grade. MMW showed a positive but negligible correlation with DE performance ($r = 0.050$, $p = 0.042$), indicating a statistically significant yet practically inconsequential linear relationship. The coefficient of determination derived from this correlation ($r^2 = 0.003$) implies that MMW alone accounts for less than 0.3% of the total variance in DE grades, a magnitude that offers virtually no predictive utility on its own.

Calculus 1 demonstrated the weakest bivariate correlation with DE grade among the three predictors ($r = 0.023$, $p = 0.029$), despite being the only variable to reach statistical significance in the subsequent regression analysis. This apparent contradiction - statistical significance paired with a near-zero effect size - is a direct consequence of the large sample size ($n = 646$), which grants the analysis sufficient statistical power to detect correlations that are real but trivially small. In practical terms, $r = 0.023$ means that Calculus 1 explains approximately 0.05% of the variance in DE performance, a relationship that is statistically detectable but educationally negligible.

Calculus 2 was the only predictor with a negative correlation with Differential Equations grade ($r = -0.055$, $p = 0.123$), and was the only variable that did not reach statistical significance at the 0.05 level. The negative direction of this relationship - suggesting that students with higher Calculus 2 grades tended to score marginally lower in DE - is counterintuitive given the sequential nature of the mathematics curriculum. However, the non-significance of this correlation ($p = 0.123$) and the negligible effect size indicate that this negative direction is unlikely to reflect a true inverse relationship, and should be interpreted as statistical noise rather than a meaningful educational finding.

Among the three predictor variables themselves, inter-correlations were also weak: MMW and Calculus 1 ($r = -0.021$), MMW and Calculus 2 ($r = -0.086$), and Calculus 1 and Calculus 2 ($r = 0.004$). The near-zero inter-predictor correlations confirm that the three prerequisite mathematics courses measured largely independent aspects of students' academic performance, which is a favorable condition for regression modeling as it minimizes multicollinearity concerns.

Taken together, the correlation analysis establishes a critical baseline finding: the three prerequisite mathematics grades examined in this study share only trivial linear relationships with Differential Equations performance. This low-correlation environment sets an inherent ceiling on the predictive accuracy achievable by any model - linear or machine learning - that relies solely on these variables. The results of the subsequent predictive modeling must therefore be interpreted within the context of this fundamental constraint in the data.

B. Linear Regression Model

The Multiple Linear Regression (MLR) model was the first predictive model developed in this study and served as the traditional statistical baseline against which all machine learning models were compared. Using the Ordinary Least Squares (OLS) method, the model estimated Differential Equations grade as a weighted linear combination of the three predictor variables. The resulting regression equation was:

$$\hat{y}DE = 79.003 - 0.070(MMW) + 0.083(Cal_1) + 0.061(Cal_2)$$

The intercept value of 79.003 represents the estimated Differential Equations grade when all three predictor variables are held at zero - a theoretical baseline with no direct practical interpretation given the grading scale. Among the three predictors, Calculus 1 carried the largest positive coefficient ($\beta = 0.083$), indicating that a one-unit increase in Calculus 1 grade was associated with an estimated 0.083-point increase in DE grade, holding other predictors constant. Calculus 2 also contributed positively ($\beta = 0.061$), while MMW exerted a slight negative influence ($\beta = -0.070$), suggesting that higher MMW grades were marginally associated with lower DE performance - a counterintuitive finding that may reflect multicollinearity or the weak overall predictive structure of the model.

Under 10-fold cross-validation, the Linear Regression model achieved $R^2 = -0.011$, RMSE = 2.255, and MAE = 1.699. The near-zero R^2 value indicates that the model explained virtually none of the variance in Differential Equations grades across validation folds, performing only marginally better than a naive mean-prediction approach. Despite this, it ranked first among all four models evaluated - a finding that reflects the limitations of the predictor set rather than the strength of linear regression as a method.

The standardized coefficients (Beta values) from the regression output further illuminate each predictor's relative contribution: Calculus 1 had the largest standardized effect (Beta = 0.086, t = 2.193, p = 0.029), confirming it as the only statistically significant predictor at the 0.05 level. MMW (Beta = −0.072, t = −1.826, p = 0.068) and Calculus 2 (Beta = 0.067, t = 1.694, p = 0.091) did not reach statistical significance, indicating that their contributions to the model were not reliably distinguishable from chance variation. This pattern suggests that among the three prerequisite mathematics courses, only Calculus 1 has a demonstrable - though weak - linear relationship with Differential Equations performance.

The collinearity diagnostics confirmed that multicollinearity was not a significant concern in the model, with Variance Inflation Factor (VIF) values of 1.009 (MMW), 1.008 (Calculus 1), and 1.016 (Calculus 2) - all well below the conventional threshold of 10. Tolerance values were similarly acceptable (≥ 0.984), indicating that the predictor variables were sufficiently independent to support stable coefficient estimation. The residual analysis revealed a mean residual of 0.000 with a standard deviation of 2.233, and predicted values ranged from 84.37 to 85.96 - a remarkably narrow band that illustrates the model's limited discriminating power across students with varying predictor profiles.

C. Comparative Model Performance

Table-3 presents cross-validated metrics for all four models. All models produced near-zero or negative R^2 values, indicating that no model could reliably predict DE performance from the three predictor variables alone.

Table-3. Comparative Model Performance, 10-Fold CV (n = 646).

Model	R^2	RMSE	MAE
Linear Regression	-0.011	2.255	1.699
Decision Tree	-0.772	2.959	2.266
Random Forest	-0.286	2.535	1.978
Support Vector Machine (SVR)	-0.077	2.326	1.763

Figure 9 to 11 shows the model comparison for the 10-fold cross validation.

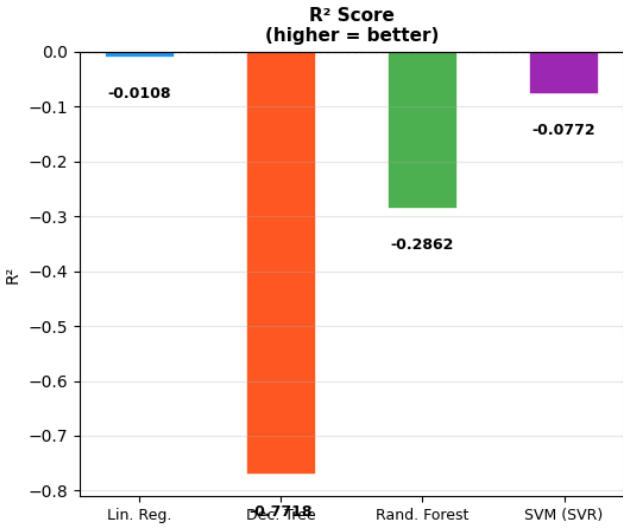

Figure-9. R^2 model comparison for machine learning techniques.

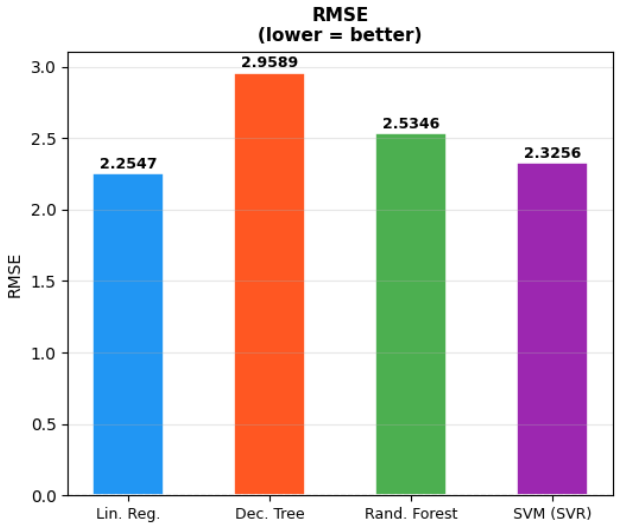

Figure-10. RMSE model comparison for machine learning techniques.

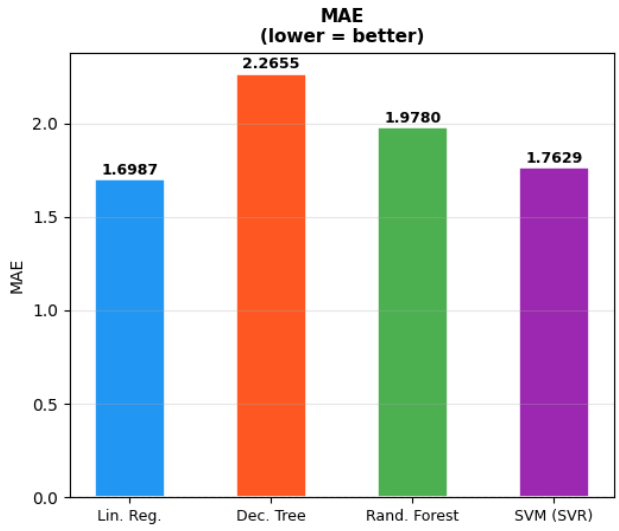

Figure-11. MAE model comparison for machine learning techniques.

Linear Regression achieved the best performance across all metrics, confirming that the traditional statistical approach remains the most generalizable given the available predictor set.

SVM (SVR) performed second-best (R^2 = −0.077, RMSE = 2.326), suggesting that the RBF kernel did not identify meaningful non-linear structure absent from the linear model, consistent with the weak correlations observed in previous results.

Random Forest (R^2 = −0.286, RMSE = 2.535) improved upon the single Decision Tree through ensemble averaging, but still underperformed the linear baseline.

Decision Tree produced the poorest results (R^2 = −0.772, RMSE = 2.959), exhibiting classic overfitting behavior on low-signal tabular data.

D. Decision Tree

Figure-12 visualizes the Decision Tree Structure with a max depth of 3.

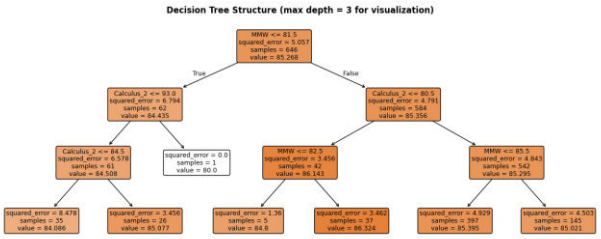

Figure-12. Decision Tree Structure.

This model produced the weakest performance among all four models evaluated in this study, yielding an R^2 of −0.772, an RMSE of 2.959, and an MAE of 2.266 under 10-fold cross-validation. A negative R^2 indicates that the model performed worse than a simple mean-prediction baseline, meaning that the Decision Tree's if-then partitioning of the feature space actively introduced prediction error rather than reducing it.

This outcome is consistent with the well-documented susceptibility of unpruned Decision Trees to overfitting. The CART algorithm recursively splits the training data to minimize mean squared error at each node, often producing a tree that memorizes the training folds rather than learning generalizable patterns. When applied to a new validation fold, these splits - which were fitted to noise rather than true signal - generate poor predictions. This behavior is particularly pronounced when the predictor variables have low correlation with the outcome, as observed in this study ($|r| \leq 0.086$ for all predictors).

The relatively high MAE of 2.266 further indicates that the Decision Tree's grade predictions deviated from actual values by over two grade points on average - a margin that would render the model impractical for early academic intervention purposes. The standard deviation of R^2 across folds (±0.288) was also the highest among all models, reflecting unstable and inconsistent predictions across different subsets of the data.

These results suggest that a single Decision Tree is not an appropriate model for predicting Differential Equations performance when only three prerequisite mathematics grades are available as predictors. The model's structural complexity is a liability rather than an advantage under these data conditions.

E. Random Forest

The Random Forest model yielded an R^2 of −0.286, an RMSE of 2.535, and an MAE of 1.978 under 10-fold cross-validation - a marked improvement over the single Decision Tree, though still below the performance of Linear Regression and SVM. The model was trained using an ensemble of 100 decision trees, each built on a bootstrapped subsample of the training data with a random subset of features considered at each split.

The substantial performance gain over the Decision Tree (ΔRMSE = −0.424; ΔR^2 = +0.486) demonstrates the effectiveness of ensemble averaging in reducing variance. By aggregating predictions across 100 independent trees, Random Forest cancels out the distinctive overfitting of individual trees, producing a more stable and generalizable estimate.

However, despite this variance reduction, Random Forest still failed to achieve a positive R^2, indicating that even the ensemble approach could not extract meaningful predictive signal from the available predictor set. The feature importance scores derived from the trained Random Forest revealed that MMW contributed the most to node impurity reduction (importance = 0.351), followed by Calculus 2 (0.334) and Calculus 1 (0.315). This shows a relatively balanced distribution, suggesting that no single predictor dominated the model's decision-making. This finding is noteworthy given that the correlation analysis identified Calculus 1 as the only statistically significant predictor, highlighting a discrepancy between feature importance rankings and statistical significance that is common in ensemble methods. The summary of feature importance is shown in Table-4.

Table-4. Summary of Feature Importances.

Variable	Importance Score
MMW	0.3446
Calculus 1	0.3401
Calculus 2	0.3153

The Random Forest's MAE of 1.978 represents an improvement of approximately 0.288 grade points over the Decision Tree, and its R^2 standard deviation across folds (±0.164) was substantially lower, confirming greater prediction stability. Nevertheless, an MAE of nearly two grade points remains too large for reliable individual-level prediction, reinforcing the conclusion that the current predictor set is insufficient for practical deployment of a machine learning model.

F. Support Vector Machine

The Support Vector Regression (SVR) model with a Radial Basis Function (RBF) kernel achieved an R^2 of -0.077, an RMSE of 2.326, and an MAE of 1.763 - ranking second overall and performing substantially closer to Linear Regression than either tree-based model. These results were obtained using standardized features ($\mu = 0$, $\sigma = 1$), regularization parameter C = 10, and epsilon tube width $\varepsilon = 0.1$.

The SVR algorithm seeks a regression function that deviates from the actual target values by at most ε, while simultaneously minimizing model complexity through the regularization parameter C. The RBF kernel implicitly maps the input features into a high-dimensional space, enabling the model to capture non-linear relationships that linear regression cannot represent. Despite this capability, the SVR's performance closely mirrored that of Linear Regression, suggesting that no meaningful non-linear structure exists in the relationship between MMW, Calculus 1, Calculus 2, and Differential Equations grades at the level detectable from this dataset.

The narrow gap between SVR and Linear Regression (ΔRMSE = 0.071; ΔR^2 = −0.066) is an important finding in itself. It indicates that the additional computational complexity of kernel-based mapping and margin optimization produced negligible predictive benefit. This is theoretically consistent with the near-zero correlations observed in Table 3: if the raw linear associations between predictors and outcome are weak, the non-linear feature transformations enabled by the RBF kernel are unlikely to yield substantial improvement, as there is little underlying structure for the kernel to amplify.

The SVR's MAE of 1.763 represents the second-lowest average prediction error among all models, meaning its grade predictions deviated from actual values by approximately 1.76 grade points on average. While this is closer to Linear Regression's MAE of 1.699 than any other model, it still falls short of the linear baseline - confirming that for this dataset and predictor configuration, increased model flexibility does not translate to improved predictive accuracy.

G. Summary

The central finding of this study is that machine learning models do not outperform Linear Regression in predicting Differential Equations performance from MMW, Calculus 1, and Calculus 2 grades. This result is consistent across all three evaluation metrics and is robust under 10-fold cross-validation, providing strong empirical evidence that model complexity alone does not improve prediction when the underlying predictors carry weak signals.

The near-universal failure of all models to achieve positive R^2 reflects a fundamental data constraint: the predictor variables account for negligible variance in the outcome. This is not a modeling failure, but rather an informational problem. With grade distributions concentrated in the 83-87 band across all four subjects, the statistical signal available to any model is extremely low.

The relatively poor performance of Decision Tree (R^2 = −0.772) compared to Random Forest (R^2 = −0.286) illustrates the well-documented tendency of single trees to overfit - memorizing noise in the training folds rather than learning generalizable patterns. Random Forest's ensemble averaging partially corrects this, explaining its improved (though still negative) R^2.

The correlation analysis revealed that Calculus 1 was the only predictor with a statistically significant relationship to DE performance ($r = 0.023$, $p = 0.029$), while MMW and Calculus 2 were not significant. This, combined with the ML results, points to an important implication, the predictive limitation lies in the predictor set, not the modeling approach.

Incorporating variables such as student attendance, learning engagement, entrance examination scores, study habits, and sociodemographic factors may provide the additional variance needed for ML models to demonstrate their full advantage.

From a practical standpoint, this study demonstrates that a simple linear regression model is sufficient given the available data, and is preferable for its interpretability and transparency, qualities that matter when the goal is to inform academic interventions.

4. CONCLUSIONS

This study compared four predictive models - Linear Regression, Decision Tree, Random Forest, and Support Vector Regression - for estimating engineering students' performance in Differential Equations using grades from MMW, Calculus 1, and Calculus 2. Based on 10-fold cross-validation of 646 student records, Linear Regression was found to be the best-performing model ($R^2 = -0.011$, RMSE = 2.255, MAE = 1.699), while all machine learning models underperformed relative to this linear baseline.

These findings lead to two key conclusions: (1) the prerequisite mathematics grades of MMW, Calculus 1, and Calculus 2 are insufficient predictors of Differential Equations performance on their own; and (2) machine learning complexity does not compensate for uninformative predictors. The practical recommendation for TUP and similar institutions is to expand the predictor set - incorporating attendance, engagement, self-efficacy, and entrance examination data - before applying machine learning approaches

5. RECOMMENDATION

Future research should also explore classification-based formulations of the problem (e.g., predicting pass/fail rather than exact grade), which may better align with institutional intervention goals and could benefit more directly from the pattern-recognition capabilities of machine learning models.

REFERENCES

Alhazmi E. and Sheneamer A. 2023. Early Prediction of Students' Performance in Higher Education. IEEE Access, 11, 27579-27589. https://doi.org/10.1109/ACCESS.2023.3250702

Caday M. B., Dalaorao G. A. and Abalorio C. C. 2025. Academic Performance Analysis of Diploma Students Using Predictive Analytics. Procedia Computer Science, 257, 913-920. https://doi.org/10.1016/j.procs.2025.03.117

Cruz M. M. and Lumauag R. 2025. Development of a Web-Based Student Academic Performance Prediction Using Machine Learning for Higher Education Institutions. International Journal of Engineering Trends and Technology, 73(7). https://doi.org/10.14445/22315381/IJETT-V73I7P133

Doz D., Cotič M. and Felda D. 2023. Random Forest Regression in Predicting Students' Achievements and Fuzzy Grades. Mathematics, 11(19): 4129. https://doi.org/10.3390/math11194129

Jiao P., Ouyang F., Zhang Q. and Alavi A. H. 2022. Artificial intelligence-enabled prediction model of student academic performance in online engineering education. Artificial Intelligence Review, 55(8): 6321-6344. https://doi.org/10.1007/s10462-022-10155-y

Nishio T., Mouri K., Tanaka T., Okamoto M. and Matsubara Y. 2022. Effects of Pairing Methods Based on Digital Textbook Logs and Learner Artifacts in Conceptual Modeling Exercises: International Journal of Distance Education Technologies, 20(1): 1-16. https://doi.org/10.4018/IJDET.296703

Pallathadka H., Wenda A., Ramirez-Asís E., Asís-López M., Flores-Albornoz J. and Phasinam K. 2023. Classification and prediction of student performance data using various machine learning algorithms. Materials Today: Proceedings, 80, 3782-3785. https://doi.org/10.1016/j.matpr.2021.07.382

Sayed A. F. A., Arafa M. A., El-Nimr N. A., Banawan K. A. S. and Abdou M. S. 2025. Comparison of machine learning classification and regression models for prediction of academic performance among postgraduate public health students. Scientific Reports, 15(1): 44056. https://doi.org/10.1038/s41598-025-31023-z