

# Modelling Mathematics Performance of Ghanaian Junior High School Students Using Supervised Machine Learning

Author: Dominic Obeng Koranteng – MSc Mathematical Sciences, AIMS Ghana (June 2025)

## Introduction & Motivation

Low mathematics performance remains a key educational challenge in Ghana. Prior studies often used limited datasets or ignored socio-economic and environmental factors. This study applies supervised machine learning models to predict mathematics performance and identify the most significant influencing factors.

## Data & Methods

Dataset: 364 students, 26 variables (demographics, socio-economic background, study habits, school environment). Algorithms: Linear Regression, Random Forest Regression, Decision Tree, and Support Vector Machine. Evaluation: 70/30 train-test split, 5-fold cross-validation; metrics include  $R^2$ , MAE, RMSE, accuracy, precision, recall, and F1-score.

## Key Results

Regression: Random Forest ( $R^2 = 0.51$ , MAE = 10.71) outperformed Linear Regression ( $R^2 = 0.16$ , MAE = 16.27). Classification: Decision Tree achieved 80% accuracy (strong for passing students, weaker for failing students). SVM achieved 63% accuracy and failed to detect failing students. Feature Importance: Class size, distance to school, teacher's rating, study hours, and teacher quality were top predictors. Socio-economic and environmental factors (pocket money, family support, nutrition, classroom conditions) also played key roles.

## Discussion & Implications

Findings highlight the multi-faceted drivers of math performance beyond student ability. They demonstrate the value of interpretable ML for education stakeholders and suggest targeted interventions such as reducing class sizes, improving teacher quality, and supporting low-income students. Ethical considerations stress careful use of predictive models to avoid bias and stigmatization.

## Conclusion

Machine learning can effectively model student performance with interpretable results. Random Forest and Decision Tree models are promising for educational applications in Ghana. Future work should integrate ML tools into policy-making and adaptive learning platforms for broader impact.