Exploring Student Performance Patterns Using Tree-Based Techniques

1st Felix Indra Kurniadi
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia,11530
felix.indra@binus.ac.id

2ndMeta Amalya Dewi
Information Systems Department
BINUS Online Learning
Bina Nusantara University
Jakarta, Indonesia,11480
meta.dewi@binus.edu

3rdDina Fitria Murad
Information Systems Department
BINUS Online Learning
Bina Nusantara University
Jakarta, Indonesia,11480
dmurad@binus.edu

4thSucianna Ghadati Rabiha
Information Systems Department
BINUS Online Learning
Bina Nusantara University
Jakarta, Indonesia,11480
sucianna.rabiha@binus.ac.id

5thAwanis Romli
Faculty of Computing
Universiti Malaysia Pahang
Pahang, Malaysia,26600
awanis@ump.edu.my

Abstract—Due to its direct impact on the development and progress of nations, predicting student performance has acquired considerable importance in modern society. The evaluation of student performance measures the effectiveness of educational institutions and their capacity to influence the next generation. As a result, enhancing the educational process has become a necessity, compelling governments and institutions to devote significant resources to its ongoing development. Based on the Student Grade Data obtained from the Binus Online Learning platform at Bina Nusantara University, this study analyzes and predicts student performance using tree-based methods, specifically Decision Tree and Random Forest. The dataset includes pupil information and variables pertaining to performance. By contrasting the performance of these tree-based models, it is possible to gain insight into their accuracy and efficacy in predicting student outcomes. The experimental results demonstrate that both the Decision Tree and Random Forest models predict student performance with high accuracy. These results demonstrate the potential of tree-based methods in educational data analysis and prediction, providing educators, administrators, and policymakers with valuable insights for identifying at-risk students and implementing timely interventions to enhance educational outcomes.

Keywords—student performance, tree-based methods, predictive modeling, educational data analysis, decision tree, random forest, academic success, performance prediction.

I. INTRODUCTION

There is a growing awareness of the significance of predicting students performance in today's society. This issue is crucial to the progress and development of nations because it has a direct impact on the educational system and the production of a capable generation that can assume leadership roles and contribute to the progress of the nation in various facets of life, such as science, the economy, society, and defense. The evaluation of student performance serves as a measure of the efficacy of educational institutions, which play a crucial role in shaping successive generations in accordance with the varying phases of life in each country. Consequently, prioritizing the improvement of the educational process has become a necessity, compelling governments and educational institutions to exert substantial and meticulous efforts towards the continuous and progressive development of education [1].

The value of predicting student performance extends beyond individual academic accomplishments. It indicates the overall caliber of education provided by institutions and their ability to prepare students for future challenges. Consequently, governments and educational institutions are obligated to prioritize the ongoing and accelerating development of the educational process. The objective is to ensure that educational systems adapt to the ever-changing needs of the contemporary world.

The issue of university dropouts is a particular difficulty that requires attention. Dropping out of college has far-reaching effects on education systems, academic and financial administrators, educators, and society as a whole, in addition to the students themselves. The economic and social impact of dropping out of college is significant, with implications for the self-perception, feelings of failure and frustration, and overall well-being of individuals [2].

Education stakeholders must take proactive measures to resolve the critical problem of university dropout. This includes identifying the challenges encountered by students at risk and implementing early intervention strategies. Early identification of students with academic performance issues enables educators to provide targeted assistance, such as supplementary courses, additional assignments, or individualized assistance. Education systems can mitigate the negative effects of university attrition and increase the likelihood of student success by intervening early [3].

The field of machine learning has grown as a promising method for predicting pupil performance in this context. Ed-