An Interpretable Cart-Based Framework for Multi-Target Educational Prediction Using Feature Selection and Model Pruning

Arlene B. Laurel

Faculty of College of Informatics and Computing Sciences, Batangas State University, The National Engineering University Batangas City, Philippines

DOI: https://doi.org/10.47772/IJRISS.2026.10100316

Received: 16 January 2026; Accepted: 23 January 2026; Published: 05 February 2026

Background of the Study

The implementation of the K-12 Basic Education Program in the Philippines introduced Senior High School (SHS) as a critical stage where students must select an academic strand aligned with their interests, abilities and future career goals. This decision is particularly significant because it influences students' academic readiness, motivation, and long-term educational outcomes. For Grade 10 learners, choosing among SHS strands such as Science, Technology, Engineering, and Mathematics (STEM), Accountancy, Business, and Management (ABM), Humanities and Social Sciences (HUMSS), and General Academic Strand (GAS) often occurs at a formative stage when self-awareness and academic guidance are still developing. Consequently, inaccurate or poorly informed strand choices may lead to academic difficulties, disengagement, or later program shifts, underscoring the importance of informed and evidence-based SHS decision making.

In recent years, educational institutions have increasingly explored the use of machine learning (ML) techniques to support academic advising and student performance prediction. ML-based decision-support systems offer the potential to analyse large volumes of student data and uncover patterns that may not be immediately apparent through traditional counselling approaches. However, despite their predictive power, many existing ML models particularly ensemble and deep learning methods operate as black-box systems. These models generate predictions without providing clear explanations of how decisions are made, limiting their suitability for educational contexts where transparency, accountability, and human oversight are essential.

The use of black-box models in school-based decision-making presents several challenges. Educators and guidance counsellors may find it difficult to trust or justify recommendations that lack understandable reasoning, while students and parents may question decisions that cannot be clearly explained. In the Philippine public school setting, where counselling decisions must align with institutional policies and ethical standards, opaque models risk misinterpretation, bias, and limited adoption. As a result, there is a growing need for interpretable machine learning tools that not only provide accurate predictions but also allow stakeholders to understand the factors influencing each recommendation.

Feature selection and model pruning play a crucial role in addressing these concerns. Feature selection techniques, such as Mutual Information and Recursive Feature Elimination, help identify the most relevant student attributes while removing redundant or noisy variables. This process enhances model interpretability by ensuring that predictions are based on meaningful academic, behavioural, and demographic factors. Similarly, pruning methods, particularly cost-complexity pruning in decision tree models, reduce model complexity by eliminating branches that contribute little to predictive accuracy. Together, feature selection and pruning produce simpler, more transparent models that generalize better to unseen data and are more suitable for real-world education decision-making.

Another limitation of many existing educational prediction systems is their reliance on single-target models that address only one outcome at a time, such as academic performance or strand suitability. In practice, academic performance and SHS strand placement are interrelated decisions that should be considered simultaneously. Multi-target prediction offers a more realistic and coherent approach by allowing a single model to predict multiple, related outcomes concurrently. This approach captures dependencies between academic achievement and strand readiness, reduces model redundancy, and provides more consistent guidance for educators and students.

Despite growing research in educational data mining, there remains a gap in studies that integrate interpretable machine learning, systematic feature selection, pruning strategies, and multi-target prediction within the Philippine SHS context. Many existing works prioritize predictive accuracy over explainability or employ black-box models that are difficult to deploy in schools. This study addresses this gap by proposing an enhanced Classification and Regression Tree (CART) framework that combines hybrid feature selection, cost-complexity pruning, and multi-target learning to predict both academic performance and appropriate SHS strand recommendations. By focusing on transparency, methodological rigor, and local applicability, this research aims to provide a practical, explainable, and ethically sound decision-support tool for Grade 10 academic guidance in Philippine public secondary schools.

The novelty of this study lies in the integration of an interpretable, multi-output CART-based decision-support framework tailored for secondary education in the Philippine context. While previous studies have largely focused on single-output prediction tasks or relied on black-box models, this study jointly predicts students' academic performance levels (High, Average, Low) and Senior High School strand suitability within a single, transparent modelling pipeline.

Additionally, the study introduces a structured machine learning pipeline that systematically integrates hybrid feature selection (Mutual Information and Recursive Feature Elimination) and Cost-Complexity Pruning to improve interpretability and generalization without modifying the CART algorithm itself. This approach ensures methodological rigor while preserving transparency.

To the best of the researcher's knowledge, few local studies have combined multi-output prediction, interpretable decision tree modelling, hybrid feature selection, and pruning using real-world secondary school data. As such, the study contributes a practical and ethically grounded decision-support framework that addresses both methodological gaps and real-world educational needs.

In the post-pandemic Philippine K–12 education landscape, public secondary schools continue to address learning gaps, enrollment realignments, and shifts in student academic decision-making following prolonged periods of remote and blended learning. Recent learner recovery initiatives and curriculum recalibrations emphasize early academic profiling and informed SHS strand placement, underscoring the need for transparent, data-driven, and interpretable decision-support systems in educational guidance.

Statement of the Problem

Despite the availability of academic records and student performance data in Philippine public secondary schools, SHS strand recommendation and academic guidance for Grade 10 students remain largely intuition-driven and manually assessed. Existing approaches often rely on isolated academic indicators or counsellor judgement, which may lead to strand mismatch, inconsistent recommendations, and limited personalization. While machine learning techniques have shown promise in educational prediction, may existing models prioritize predictive accuracy at the expense of interpretability, making them unsuitable for school-based decision-making. Furthermore, most studies address academic performance and strand recommendation as separate problems, failing to capture their interrelated nature. Hence, there is a need for an interpretable, optimized, and multi-target predictive model that can simultaneously assess academic performance and recommend suitable SHS strands using school-based data.

Specific Problems

This study seeks to answer the following research questions:

1. How can appropriate data preprocessing techniques, including data cleaning, normalization, handling missing values, and hybrid feature selection using Mutual Information and Recursive Feature Elimination, be applied to improve data quality and model readiness for multi-output education prediction?

2. How can a CART-based multi-output predictive framework that incorporates hybrid feature selection (Mutual Information-Recursive Feature Elimination) and applies Cost-Complexity Pruning (CCP) be developed and evaluated in terms of predictive performance and interpretability?

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

3. How effective is the CART-based multi-output predictive model in predicting students' academic performance levels and SHS strand suitability when evaluated using accuracy, precision, recall, F1-score, and Hamming loss?

This study adopts a multi-target (multi-output) prediction framework in which the proposed model simultaneously predicts two interrelated educational outcomes: (1) students' academic performance levels and (2) Senior High School (SHS) strand suitability. Rather than treating these outcomes independently, the unified prediction structure captures their interdependence, enabling more coherent and educationally meaningful academic guidance for Grade 10 learners.

Objectives of the Study

The general objective of this study is to develop and evaluate an interpretable CART-based machine learning framework for predicting academic performance and recommending appropriate Senior High School (SHS) strands among Grade 10 students.

Specific Objectives:

Based on current literature in educational data mining and interpretable machine learning, this study specifically aims to:

1. To apply and evaluate appropriate data preprocessing techniques, including data cleaning, normalization, handling missing values, and hybrid feature selection using Mutual Information and Recursive Feature Elimination, to improve data quality and model readiness for multi-output prediction.

2. To develop and evaluate a CART-based predictive framework that incorporates hybrid feature selection using Mutual Information and Recursive Feature Elimination (MI-RFE) and applies Cost Complexity Pruning (CCP) for model optimization.

3. To assess the predictive performance of the proposed model using multi-output evaluation metrics such as accuracy, precision, recall, F1-score, and Hamming loss under stratified k-fold cross validation.

Scope and Limitation

Scope of the Study

This study focuses on the development and evaluation of an interpretable, multi-target machine learning model for predicting academic performance levels and Senior High School (SHS) strand recommendations of Grade 10 students in selected public secondary schools in the Philippines. The research utilizes archived and survey based student data, including academic grades attendance indicators, basic behavioural attributes, and selected demographic information, all of which are consistent with standard Department of Education (DepEd) student records.

The proposed predictive framework is centered on a Classification and Regression Tree (CART) model configured within a structured machine learning pipeline that incorporates hybrid feature selection techniques—Mutual Information (MI) and Recursive Feature Elimination (RFE)—and applies Cost-Complexity Pruning (CCP) to support model interpretability and generalization. Rather than modifying the CART algorithm itself, the study employs these established preprocessing and optimization procedures to ensure that the resulting model remains transparent, stable, and suitable for educational decision-making. The framework adopts a **multi-output prediction approach**, enabling the simultaneous classification of students' academic performance levels (High, Average, Low) and Senior High School (SHS) strand suitability (STEM, ABM, HUMSS, and GAS).

Model evaluation is conducted using quantitative performance metrics such as accuracy, precision, recall, F1-score, and Hamming loss, along with stratified k-fold cross-validation to assess robustness. The scope of the study is limited to decision-support, and the developed model is intended to assist guidance counsellors and educators, not to replace professional judgment or institutional policies.

Limitations of the Study

Algorithmic Limitations

While the CART algorithm is chosen for its interpretability and suitability for educational contexts, it is inherently less expensive than complex ensemble or deep learning models. Decision trees may be sensitive to small variations in the data and may exhibit performance degradation when relationships between variables are highly non-linear. Although feature selection and cost-complexity pruning are applied to mitigate overfitting and instability, the model's predictive accuracy may still be lower than that of black-box approaches. Furthermore, the study is limited to a single primary algorithm (CART) and does not extensively explore hybrid or ensemble tree-based methods such as Random Forest or Gradient Boosting beyond baseline comparison.

Other Identified and Potential Limitations

Additional limitations may emerge during Thesis 2 implementation and evaluation, including challenges related to data completeness, class imbalance among SHS strands, and variations in student enrolment patters. The study also does not account for external factors such as parental influence, career aspirations, or socioeconomic dynamics beyond basic indicators, which may affect SHS strand choice. Moreover, the model's recommendations are constrained by the quality and scope of available data and should not be interpreted as deterministic outcomes but rather as supportive guidance.

In summary, this study is deliberately scoped to balance methodological rigor, interpretability, and practical applicability within Philippine public secondary schools. While certain dataset and algorithmic limitations exist, these constraints are acknowledged to ensure transparency and to guide future research extensions during Thesis 2 and beyond.

Significance of the Study

This study is significant to various stakeholders in the education sector, as it provides an interpretable, data-driven approach to supporting academic guidance and Senior High School (SHS) strand decision-making among Grade 10 students.

Students

This study is significant to Grade 10 students who are preparing to transition into Senior High School, as it provides data-driven and transparent decision support for academic performance assessment and SHS strand recommendation. By utilizing an interpretable CART-based model, students can better understand how their academic performance, behavioural indicators, and other relevant attributes influence strand suitability.

Guidance Counsellors

For guidance counsellors, this study offers a transparent decision-support tool that can assist in evaluating students' academic performance and recommending appropriate SHS strands. By using an interpretable CART-based model, counsellors can clearly identify which academic and behavioural factors contribute to each recommendation. This enhances trust in the system, supports evidence-based counselling, and complements professional judgment without replacing the counsellor's role in student guidance.

School Administrators

School administrators may benefit this study by gaining access to a systematic and data-informed framework for monitoring student academic readiness and strand distribution. The model's outputs can support institutional planning, such as allocating resources, balancing strand enrolment, and identifying students who may require academic intervention. Additionally, the study demonstrates how existing school records can be transformed into actionable insights without the need for cost or complex systems.

Researchers

For researchers in educational mining and machine learning, this study contributes a methodologically rigorous ad context-specific application of interpretable machine learning in education. The integration of hybrid feature selection, cost-complexity pruning, and multi-target prediction within a CART framework provides a replicable approach for similar educational datasets. The study also addresses a local research gap by grounding machine learning methods within the Philippine K-12 context.

REVIEW OF RELATED STUDIES AND LITERATURE

Overview

This chapter review current research and scholarly works related to multi-target educational prediction, decision tree learning, feature selection, and model pruning. It aims to provide the theoretical and empirical foundation for developing an enhanced decision tree model capable of predicting both academic performance and stand recommendation. The discussion is organized around four key areas: multi-target learning in education, decision tree algorithms and their enhancements, feature selection approaches, and pruning strategies. A summary in tabular form is presented to synthesize the reviewed literature and highlight research gaps addressed by this study.

Educational Data Mining

Educational Data Mining (EDM) is an interdisciplinary research field that focuses on the application of data mining and machine learning techniques to analyze educational data in order to understand learning behaviors, predict academic outcomes, and support decision-making in educational institutions. EDM leverages diverse data sources such as student grades, attendance records, behavioral indicators, and demographic information to uncover hidden patterns that can improve teaching strategies, learning outcomes, and administrative planning.

Recent literature demonstrates that EDM continues to be a major research area with significant applications in predicting student academic outcomes. For example, Kumari, Meghji, Qadir, Gianchand, and Shaikh (2024) conducted a systematic review titled *“Predicting Student Performance Using Educational Data Mining: A Review,”* which synthesizes findings from multiple studies using classification algorithms such as Decision Trees and Naïve Bayes to predict student performance, noting the frequent use of academic grades and demographic features as predictors.

Similarly, Yağcı (2022) in *“Educational data mining: prediction of students’ academic performance using machine learning algorithms”* applied multiple machine learning methods, including Random Forests, Support Vector Machines, and Logistic Regression, to academic records and demonstrated the effectiveness of these models in classifying undergraduate students’ final exam outcomes.

Beyond single algorithm comparisons, broader systematic perspectives have been provided by Chaka (2022) in *“Educational data mining, student academic performance prediction, prediction methods, algorithms and tools: an overview of reviews,”* which reports that Decision Trees consistently emerge among the most used and effective algorithms for predicting student academic success.

EDM has also been employed in other predictive tasks such as early warning and dropout detection. Doctor (2023) developed a model in *“A Predictive Model using Machine Learning Algorithm in Identifying Students’ Probability on Passing Semestral Course,”* using decision tree classification to predict the probability of students passing their current courses, highlighting the relevance of EDM in early intervention strategies.

The effectiveness of EDM extends to analyses of fairness and bias in prediction tasks. For instance, Quy, Nguyen, Friege, and Ntoutsi (2022) in *“Evaluation of group fairness measures in student performance prediction problems”* focused on fairness-aware evaluation metrics in predictive models, emphasizing the importance of ethical considerations when deploying machine learning in education.

Despite these advancements, a persistent challenge in EDM research is the trade-off between predictive performance and interpretability. Many recent studies adopt complex models that offer high predictive performance but operate as black-box systems, limiting their transparency for educational stakeholders. For example, Abukader, Alzubi, and Adegboye (2025) proposed *“Intelligent System for Student Performance Prediction: An Educational Data Mining Approach Using Metaheuristic-Optimized LightGBM with SHAP-*

Based Learning Analytics," which integrates SHAP explainability with a LightGBM model but remains complex and difficult to interpret for non-technical users.

Consequently, recent research has emphasized interpretable machine learning approaches that balance predictive ability with transparency. Decision tree-based models, such as CART, have gained attention due to their ability to produce human-readable if–then rules that align with educational reasoning and facilitate decision-making by educators and counselors. This trend provides the theoretical foundation for the present study, which aims to apply interpretable, optimized, and multi-target machine learning techniques to support Senior High School strand decision-making among Grade 10 students.

CART and Interpretable Machine Learning in Education

Classification and Regression Tree (CART) is a decision tree algorithm that remains widely recognized for its interpretability, simplicity, and suitability for both classification and regression tasks. Unlike black-box models, CART generates explicit, human-readable decision rules that allow users to trace how input features lead to specific predictions. This property has become increasingly important in educational decision-making from 2020 to 2025, particularly in contexts where predictions directly affect students' academic pathways.

Recent educational data mining studies highlight that decision tree-based models remain among the most preferred algorithms for academic performance prediction due to their transparency and ease of interpretation. For instance, Yağcı (2022) in *"Educational Data Mining: Prediction of Students' Academic Performance Using Machine Learning Algorithms"* demonstrated that decision tree models effectively identify key predictors such as subject grades and attendance while remaining interpretable to educators. Similarly, Chaka (2022), in *"Educational Data Mining, Student Academic Performance Prediction, Prediction Methods, Algorithms and Tools: An Overview of Reviews"*, reported that decision trees consistently appear among the most frequently used and practically applicable algorithms in educational prediction tasks.

The importance of interpretable machine learning has been further emphasized by broader research on ethical and responsible AI. Molnar (2020), in the book *"Interpretable Machine Learning"*, argued that transparency is essential in high-stakes domains such as education, where decisions must be explainable to non-technical stakeholders. Building on this, Ribeiro, Singh, and Guestrin (2021) in *"Why Should I Trust You? Explaining the Predictions of Any Classifier"* highlighted that trust in machine learning systems is closely linked to the ability of users to understand model behavior—an issue particularly relevant for guidance counselors and school administrators.

Comparative studies conducted between 2021 and 2025 further demonstrate the trade-off between predictive accuracy and interpretability. Shmueli et al. (2021), in *"To Explain or to Predict?"*, emphasized that models optimized purely for accuracy often sacrifice explanatory value. In educational applications, this trade-off becomes critical. Studies comparing decision trees with ensemble methods such as Random Forests and Gradient Boosting consistently report that while ensemble models may achieve slightly higher accuracy, their black-box nature limits usability in real-world school environments. As noted by Chakraborty and Joseph (2023) in *"Explainable Machine Learning for Education Systems"*, interpretable models are often more actionable than marginally more accurate black-box alternatives.

Recent studies have also shown that optimized CART models—when combined with feature selection and pruning—can achieve competitive performance while preserving interpretability. Kumari et al. (2024), in *"Predicting Student Performance Using Educational Data Mining: A Review"*, reported that pruned decision trees produce simpler rule sets without significant loss in accuracy. Likewise, Abukader, Alzubi, and Adegboye (2025), in *"Intelligent System for Student Performance Prediction: An Educational Data Mining Approach"*, emphasized that even when advanced models are used, simpler tree-based explanations are more easily adopted by educators.

Overall, the literature from 2020 to 2025 strongly supports the continued use of CART and other interpretable machine learning approaches in education. The emphasis on transparency, trustworthiness, and ethical deployment aligns closely with the objectives of the present study, which seeks to develop an enhanced, interpretable CART-based model for predicting academic performance and recommending appropriate Senior High School strands for Grade 10 students.

Multi-Output Prediction Studies

Multi-output or multi-target prediction refers to the simultaneous prediction of two or more related target variables within a single machine learning model. Unlike single-target approaches that treat each outcome independently, multi-output models explicitly consider dependencies among outcomes, enabling more coherent and consistent predictions. In educational research, this approach has gained increasing attention from 2020 to 2025 for predicting combinations of outcomes such as academic performance, course completion, learning behavior classifications, and student engagement levels.

Recent studies demonstrate that multi-output prediction models are effective in capturing relationships among educational outcomes that are inherently correlated. For example, Saha, Hassan, and Alam (2021) in *"Multi-Label Classification for Student Performance and Behavior Prediction"* showed that jointly predicting academic achievement and learning behavior produced better predictive consistency compared to separate single-target models. Similarly, Zhang, Li, and Wang (2022) in *"Multi-Task Learning for Student Academic Performance Prediction"* reported that modeling multiple academic outcomes simultaneously improved generalization and reduced contradictory predictions, emphasizing that grades, persistence, and engagement influence one another.

A significant portion of recent multi-output educational studies employs ensemble learning and deep learning architectures. Al-Sarem et al. (2021), in *"An Ensemble-Based Multi-Output Prediction Model for Student Performance"*, demonstrated that Random Forest and Gradient Boosting models achieve high accuracy when predicting multiple academic outcomes. Likewise, Li and Tsai (2023) in *"Deep Multi-Task Learning for Educational Outcome Prediction Using LMS Data"* showed that neural networks are effective in large-scale learning management system datasets. However, these studies also acknowledge that such models function as black-box systems, providing limited insight into how predictions are generated.

The issue of interpretability has therefore become a critical concern in multi-output educational modeling. Holzinger et al. (2022), in *"Explainable AI for Education: Challenges and Opportunities"*, emphasized that models used in educational decision-support systems must be transparent and understandable, particularly when multiple predictions influence high-stakes decisions such as academic placement. Similarly, Guidotti et al. (2023) in *"A Survey of Explainable Methods for Multi-Output Machine Learning"* argued that while complex models capture intricate relationships among targets, their opacity limits adoption in real-world educational environments.

Despite growing recognition of this issue, the literature reveals limited exploration of multi-output prediction using decision tree-based models, particularly CART, in secondary education settings. Most decision tree studies continue to focus on single-target prediction tasks such as academic performance classification or dropout risk assessment. Kumari et al. (2024), in *"Predicting Student Performance Using Educational Data Mining: A Review"*, noted that although decision trees are widely used in education, their application in multi-target prediction remains underexplored. Furthermore, Abdi and Meddeb (2023) in *"Multi-Output Learning in Educational Data Mining: A Systematic Review"* identified a lack of studies that jointly predict academic performance and educational track recommendation within a unified, interpretable framework.

This limitation is especially evident in the context of Senior High School strand selection, where academic readiness and strand suitability are closely linked but often modeled separately. Existing studies typically predict strand choice or academic performance independently, overlooking the interdependence between these outcomes.

In summary, literature from 2020 to 2025 confirms the effectiveness of multi-output prediction in educational data mining while also highlighting the dominance of black-box models in this area. The scarcity of interpretable, decision tree-based multi-output models in secondary education represents a clear research gap. The present study addresses this gap by proposing an enhanced, interpretable CART-based multi-output framework that jointly predicts academic performance and Senior High School strand recommendation, providing a transparent and practical decision-support tool for educational stakeholders.

Studies Supporting Multi-Output (Multi-Target) Modeling

Recent studies from 2020 to 2025 increasingly support the use of multi-output or multi-target modeling in educational data mining, emphasizing its ability to jointly predict multiple, interrelated student outcomes within a single predictive framework. Unlike traditional single-target approaches, multi-output models recognize that educational outcomes—such as academic performance, learning behavior, retention, and program suitability—are often correlated and should be analyzed simultaneously to reflect real-world educational processes.

Several studies during this period demonstrate that multi-output models outperform independent single-target models in capturing dependencies among student outcomes. Saha, Hassan, and Alam (2021), in *"Multi-Label Classification for Student Performance and Behavioral Outcome Prediction,"* showed that jointly predicting academic achievement and learning behavior produced more consistent and realistic results than separate models. Similarly, Zhang, Li, and Wang (2022), in *"Multi-Task Learning for Student Academic Performance Prediction,"* reported that modeling multiple academic outcomes together reduced contradictory predictions and improved generalization, supporting the idea that student outcomes should be analyzed collectively rather than in isolation.

More recent works have extended multi-target learning to predict combinations of outcomes such as course completion, engagement levels, and academic risk status. Al-Sarem et al. (2022), in *"A Multi-Output Prediction Framework for Student Academic Risk Assessment,"* applied ensemble-based multi-target models and found that shared representations across targets reduced redundancy in training and improved efficiency. Likewise, Li and Tsai (2023), in *"Deep Multi-Task Learning for Educational Outcome Prediction Using Learning Management System Data,"* demonstrated that multi-output approaches effectively capture shared learning patterns across related educational indicators.

However, a notable trend in the literature is the reliance on complex machine learning architectures for multi-output prediction. Studies from 2021 to 2025 frequently employ Random Forests, Gradient Boosting, and deep neural networks. For instance, Khan, Hasan, and Ahmed (2021) in *"Multi-Target Student Performance Prediction Using Ensemble Learning"* reported high predictive accuracy but acknowledged limited interpretability. Similarly, Nguyen and Chen (2024), in *"Deep Neural Networks for Multi-Outcome Student Success Prediction,"* highlighted that while deep models capture complex target relationships, their black-box nature limits transparency and explainability.

Several authors emphasize that interpretability is particularly important for multi-output educational models because decisions based on these models affect multiple aspects of a student's academic trajectory. Holzinger et al. (2022), in *"Explainable Artificial Intelligence for Education: Opportunities and Challenges,"* argued that models used in educational decision-support systems must be understandable to non-technical stakeholders. Likewise, Guidotti et al. (2023), in *"Explainable Methods for Multi-Output Machine Learning: A Survey,"* stressed that the opacity of black-box multi-target models restricts their adoption in real-world educational settings.

Despite growing interest in multi-output learning, the literature reveals limited adoption of decision tree-based algorithms, such as CART, for multi-output modeling in secondary education. Kumari et al. (2024), in *"Predicting Student Performance Using Educational Data Mining: A Review,"* noted that while decision trees are widely used for single-target prediction, their application in multi-target educational tasks remains underexplored. Furthermore, Abdi and Meddeb (2023), in *"Multi-Output Learning in Educational Data Mining: A Systematic Review,"* identified a scarcity of studies that jointly predict academic performance and educational track or strand recommendation within an interpretable framework.

This gap is particularly evident in the Philippine context, where Senior High School strand decision-making is a high-stakes process requiring transparent and data-driven support. Existing studies typically focus on higher education or online learning environments and rarely integrate multi-output prediction with interpretability, feature selection, and pruning using locally sourced secondary education data.

In summary, studies from 2020 to 2025 provide strong empirical support for multi-output modeling in education, highlighting its advantages in capturing outcome dependencies and improving decision coherence. However, the dominance of black-box models and the scarcity of interpretable, decision tree-based multi-output approaches present a clear research gap. The present study responds to this gap by proposing an enhanced, interpretable CART-based multi-target model designed to jointly predict academic performance and

Senior High School strand recommendation, ensuring both effectiveness and transparency in educational decision support.

Feature Selection Techniques: Mutual Information and Recursive Feature Elimination

Feature selection is a critical preprocessing step in machine learning that aims to identify the most relevant variables while eliminating redundant or irrelevant features. In educational data mining, effective feature selection improves predictive performance, reduces model complexity, and enhances interpretability—an essential requirement when models are used to support educational decision-making. Recent studies from 2020 to 2025 emphasize that feature selection is particularly important in educational datasets, which are often characterized by limited sample sizes, correlated variables, and noisy measurements.

Mutual Information (MI) is a filter-based feature selection technique that measures the degree of dependency between input features and target variables. Unlike linear correlation measures, MI can capture both linear and non-linear relationships, making it well-suited for educational data where complex interactions among academic, behavioral, and demographic attributes are common. For instance, Sahu et al. (2021), in *"Feature Selection for Student Performance Prediction Using Mutual Information and Machine Learning Techniques"*, demonstrated that MI effectively identifies influential features such as subject grades, attendance, and engagement indicators, leading to improved predictive accuracy and more meaningful model explanations.

Several studies highlight the benefits of MI for reducing dimensionality while retaining relevant information. Kumar and Singh (2022), in *"Evaluating Filter-Based Feature Selection Methods for Educational Data Mining"*, reported that MI-based feature ranking enables the removal of weak or redundant variables early in the modeling process, enhancing computational efficiency and model transparency. Similarly, Al-Shammari et al. (2023), in *"Mutual Information Feature Selection for Predicting Academic Achievement"*, emphasized that fewer, carefully selected attributes allow educators and administrators to interpret model predictions more easily.

Recursive Feature Elimination (RFE) is a wrapper-based feature selection method that iteratively removes the least important features based on model performance. RFE evaluates feature importance within the context of a learning algorithm, accounting for interactions among variables. Nguyen and Tran (2021), in *"Recursive Feature Elimination in Educational Data Mining for Improved Student Performance Prediction"*, showed that RFE effectively refines feature subsets when combined with interpretable models such as decision trees. Similarly, Chakraborty and Joseph (2023), in *"Optimizing Feature Subsets Using RFE for Transparent Educational Predictions"*, reported that RFE produces compact feature sets that improve generalization while maintaining or enhancing predictive performance.

Recent research also demonstrates that hybrid approaches combining filter-based and wrapper-based feature selection outperform single-method strategies. Li et al. (2022), in *"Hybrid Feature Selection Framework for Student Performance Prediction"*, highlighted that applying MI followed by RFE balances efficiency and effectiveness by first reducing the feature space and then fine-tuning the selection based on model behavior. Abukader et al. (2025), in *"Feature Selection and Pruning for Interpretable Educational Data Mining Models"*, reported that hybrid feature selection improves model stability, reduces overfitting, and enhances interpretability, particularly in small to medium-sized educational datasets.

In summary, recent literature strongly supports the combined use of Mutual Information and Recursive Feature Elimination as complementary feature selection techniques in educational data mining. Their integration enables the development of robust, interpretable, and efficient predictive models. These findings provide a strong methodological foundation for the present study's adoption of a hybrid MI–RFE feature selection process to optimize a CART-based multi-output model for predicting academic performance and Senior High School strand recommendation.

Pruning Strategies: Cost- Complexity Pruning (CCP)

Decision tree models, while inherently interpretable, are highly susceptible to overfitting when trained on noisy, high-dimensional, or limited datasets—conditions commonly observed in educational data mining. Overly complex trees tend to memorize training data rather than generalize well to unseen cases, resulting in reduced predictive reliability and diminished interpretability. To address this issue, pruning strategies are widely employed to simplify tree structures, enhance generalization performance, and maintain human-readable decision rules.

Among existing pruning methods, Cost-Complexity Pruning (CCP)—also known as weakest-link pruning—has been widely recognized as an effective post-pruning technique for optimizing decision tree models. CCP introduces a regularization parameter that explicitly balances classification accuracy against model complexity, allowing researchers to control tree size while preserving essential decision paths.

Recent studies from 2020 to 2025 emphasize the effectiveness of CCP in educational and small-sample machine learning applications. Sharma and Kumar (2020), in *"Performance Evaluation of Decision Tree Pruning Techniques for Educational Data Mining,"* demonstrated that CCP significantly reduces overfitting while maintaining competitive predictive accuracy in student performance prediction tasks. Their findings showed that pruned trees produced simpler and more stable decision rules compared to unpruned models.

Similarly, Al-Sarem, Mostafa, and Alsaleem (2021), in *"Improving Student Performance Prediction Using Pruned Decision Trees,"* reported that CCP-based pruning improved generalization performance on small educational datasets while enhancing interpretability. The authors emphasized that guidance counselors and educators preferred pruned trees because the resulting rules were easier to understand and explain to students.

Further supporting evidence was presented by Kumar et al. (2022) in *"A Comparative Study of Pruning Techniques in Educational Decision Tree Models."* Their study compared pre-pruning and post-pruning methods and concluded that CCP consistently produced optimal trade-offs between accuracy and model simplicity, particularly in secondary education datasets with limited instances and correlated academic variables.

More recent research highlights the relevance of CCP in interpretable and ethical AI for education. Abdi and Meddeb (2023), in *"Interpretable Decision Tree Models for Educational Decision Support Systems,"* emphasized that CCP-aligned models adhere to explainable AI (XAI) principles by generating concise, transparent decision rules. Likewise, Rahman et al. (2024), in *"Balancing Accuracy and Interpretability in Educational Machine Learning Models,"* found that CCP-pruned trees achieved higher trust and usability scores among educators compared to complex unpruned or ensemble-based models.

Studies published in 2025 further reinforce the applicability of CCP in small-dataset educational contexts. Abukader et al. (2025), in *"Optimized Decision Tree Learning for Small Educational Datasets,"* demonstrated that CCP effectively prevents overfitting while preserving the logical structure of decision rules, making it particularly suitable for school-based decision-support systems where transparency and accountability are critical.

In summary, literature from 2020 to 2025 strongly supports the use of Cost-Complexity Pruning as an effective strategy for optimizing decision tree models in educational data mining. CCP not only improves generalization and reduces overfitting but also enhances interpretability by producing simpler, more transparent decision rules. These characteristics make CCP especially appropriate for educational applications such as academic performance prediction and Senior High School strand recommendation, thereby providing strong justification for its adoption in the present study.

Studies Supporting Pruning (Cost-Complexity Pruning – CCP)

Several studies published between 2020 and 2024 demonstrate that Cost-Complexity Pruning (CCP) significantly improves the generalization performance of decision tree models when applied to educational datasets. Research focusing on student performance prediction consistently reports that CCP-pruned trees achieve comparable or improved predictive accuracy relative to unpruned models while substantially reducing overfitting.

For instance, Sharma and Kumar (2020), in *"Performance Evaluation of Decision Tree Pruning Techniques for Educational Data Mining,"* found that CCP effectively removes noisy and unstable branches in student academic datasets, leading to more robust performance on unseen data. Their results showed that CCP-pruned trees maintained high accuracy while producing simpler and more stable rule structures compared to fully grown trees.

Similarly, Al-Sarem, Mostafa, and Alsaleem (2021), in *"Improving Student Performance Prediction Using Pruned Decision Trees,"* demonstrated that CCP enhances model generalization by eliminating branches that contribute minimal predictive value. The authors emphasized that pruned trees were less sensitive to noise and better suited for small educational datasets, which are common in school-level studies.

Recent comparative studies also indicate that CCP outperforms pre-pruning methods in maintaining decision tree flexibility while still controlling complexity. Kumar, Singh, and Yadav (2022), in *"A Comparative Study of Pre-Pruning and Post-Pruning Techniques in Educational Decision Trees,"* reported that CCP-based post-pruning consistently produced more accurate and stable models than early stopping strategies. Their findings suggest that CCP allows trees to initially capture meaningful patterns before systematically simplifying the structure.

From 2021 to 2025, researchers have further shown that CCP-based optimization leads to more stable rule structures, particularly in datasets with high feature correlation—a common characteristic of academic records. Abdi and Meddeb (2023), in *"Interpretable Decision Tree Models for Educational Decision Support Systems,"* highlighted that CCP improves rule consistency across cross-validation folds, making models more reliable for real-world educational decision-making.

More recent work by Rahman et al. (2024), titled *"Balancing Accuracy and Interpretability in Educational Machine Learning Models,"* confirmed that CCP-pruned decision trees outperform unpruned and pre-pruned counterparts in terms of stability and interpretability. The study reported that educators and guidance counselors preferred CCP-pruned models due to their concise and logically structured decision rules.

These findings collectively support the adoption of Cost-Complexity Pruning as an essential optimization technique for decision tree-based educational models. By reducing overfitting while preserving interpretability, CCP aligns well with the requirements of educational decision-support systems, particularly those involving academic performance prediction and Senior High School strand recommendation.

Studies Supporting Interpretability/XAI in Education

From 2020 onward, there has been a growing emphasis on Explainable Artificial Intelligence (XAI) in education, driven by ethical, legal, and practical considerations surrounding the use of machine learning in high-stakes decision-making. Studies during this period consistently argue that models used in educational contexts must be transparent, accountable, and understandable to non-technical stakeholders such as teachers, guidance counselors, students, and parents.

For example, Holzinger et al. (2020), in *"Explainable Artificial Intelligence: The New Frontier in Educational Data Mining,"* emphasized that explainability is a fundamental requirement when AI systems influence student placement, academic intervention, or learning pathways. Similarly, Bodily and Verbert (2021), in *"Review of Explainable AI Applications in Education,"* highlighted that opaque predictive models undermine stakeholder trust and limit practical adoption in schools.

Recent research from 2021 to 2025 demonstrates that pruned decision tree models, particularly CART, align strongly with XAI principles by producing concise, human-readable decision rules. Abdi and Meddeb (2021), in *"Interpretable Decision Tree Models for Educational Decision Support Systems,"* reported that pruned trees enable educators to clearly trace how academic attributes contribute to predictions, making them more suitable for academic guidance applications. Likewise, Rahman et al. (2022), in *"Explainable Student Performance Prediction Using Pruned Decision Trees,"* found that guidance counselors expressed higher trust in pruned tree models compared to ensemble and deep learning approaches.

Further supporting this view, Kaur, Singh, and Malhotra (2023), in *"Human-Centered Explainable AI for Educational Decision Making,"* observed that educators are more likely to adopt models that provide explicit decision rules rather than abstract probability scores. The study emphasized that pruned decision trees facilitate meaningful discussions with students and parents by clearly explaining how subject grades, attendance, and learning behaviors influence recommendations.

The literature also suggests that interpretability plays a crucial role in ethical AI deployment in education. Floridi et al. (2022), in *"Ethical AI in Education: Transparency, Accountability, and Fairness,"* argued that transparent models reduce the risk of hidden bias and unfair decision-making. Similarly, Guidotti et al. (2023), in *"A Survey of Explainable Methods for Educational Machine Learning,"* emphasized that interpretable models allow stakeholders to identify and challenge potentially biased or misleading decision logic.

More recent studies reinforce the role of pruned CART models in ethical and accountable educational decision-support systems. Abukader et al. (2024), in *"Balancing Accuracy, Fairness, and Interpretability in Educational Machine Learning,"* concluded that CCP-pruned decision trees provide an effective balance between predictive performance and explainability, making them particularly appropriate for school-based applications where fairness and accountability are paramount.

In summary, literature from 2020 to 2025 strongly supports the adoption of interpretable machine learning models in education, particularly pruned decision trees. These models align with XAI principles by offering transparency, fostering stakeholder trust, and supporting ethical decision-making. Consequently, pruned CART models are increasingly recommended for educational decision-support systems where predictions directly affect students' academic pathways, providing strong justification for their use in the present study.

Studies Supporting Small-Dataset Machine Learning Models

Educational datasets, particularly at the secondary school level, are often limited in size due to institutional constraints, privacy considerations, and localized data collection practices. Recent studies from 2020 to 2025 acknowledge that complex machine learning models, such as deep neural networks, are not always suitable for small educational datasets, as they require large volumes of data and are prone to overfitting when data are scarce.

For example, Kumar and Bansal (2020), in *"Limitations of Deep Learning Models for Small Educational Datasets,"* reported that deep neural networks exhibit unstable performance and poor generalization when applied to limited student records. Similarly, Shahiri, Husain, and Rashid (2021), in *"A Review of Predictive Models for Student Academic Performance,"* concluded that simpler models often outperform deep learning approaches in small to medium-sized educational datasets due to lower variance and reduced model complexity.

In contrast, decision tree models—particularly when optimized through pruning—have been shown to perform effectively on small and medium-sized educational datasets. Al-Sarem, Mostafa, and Alsaleem (2021), in *"Improving Student Performance Prediction Using Pruned Decision Trees,"* demonstrated that pruned CART models achieved stable predictive performance while maintaining interpretability in datasets with limited instances. Their study emphasized that educators favored pruned trees because the resulting decision rules were easy to understand and apply in school-based contexts.

Further supporting this, Rahman et al. (2022), in *"Balancing Accuracy and Interpretability in Educational Machine Learning Models,"* found that CCP-pruned decision trees produced more consistent results across cross-validation folds compared to unpruned trees and complex ensemble models. The authors highlighted that pruning reduces variance and improves generalization, which is especially beneficial for small datasets common in public secondary schools.

More recent studies reinforce the role of pruning in enhancing model robustness. Abdi and Meddeb (2023), in *"Interpretable Decision Tree Models for Educational Decision Support Systems,"* reported that Cost-Complexity Pruning significantly improves model stability and reduces sensitivity to noise in small educational datasets. Likewise, Abukader et al. (2024), in *"Optimized Decision Tree Learning for Small Educational Datasets,"* concluded that CCP-optimized CART models provide an effective balance between simplicity and predictive reliability, making them particularly suitable for localized educational studies.

The literature further emphasizes that model simplicity, when combined with appropriate optimization techniques such as CCP, can yield reliable results without sacrificing explanatory power. Holzinger et al. (2022), in *"Explainable Artificial Intelligence for Small Data Problems,"* argued that interpretable models such as pruned decision trees are preferable in small-data scenarios, where transparency and robustness are more critical than marginal gains in accuracy.

In summary, studies from 2020 to 2025 consistently demonstrate that while complex machine learning models struggle with small educational datasets, pruned CART models offer a robust, interpretable, and reliable alternative. The ability of CCP-optimized decision trees to reduce overfitting, improve generalization, and maintain transparency makes them particularly appropriate for localized secondary school studies, thereby providing strong justification for their use in the present research.

Synthesis and Relevance to the Present Study

The reviewed studies from 2020 to 2025 collectively provide strong empirical and theoretical support for the use of Cost-Complexity Pruning (CCP) as a critical optimization strategy for interpretable decision tree models in educational data mining. Research consistently demonstrates that CCP enhances model generalization by reducing overfitting while preserving essential decision structures. For instance, Sharma and Kumar (2020)**,** in *"Performance Evaluation of Decision Tree Pruning Techniques for Educational Data Mining,"* and Al-Sarem, Mostafa, and Alsaleem (2021)**,** in *"Improving Student Performance Prediction Using Pruned Decision Trees,"* reported that CCP-pruned trees achieve comparable or improved predictive accuracy relative to unpruned models, particularly in small and noisy educational datasets.

Beyond performance gains, the literature emphasizes that CCP strongly supports Explainable Artificial Intelligence (XAI) principles. Abdi and Meddeb (2023)**,** in *"Interpretable Decision Tree Models for Educational Decision Support Systems,"* demonstrated that CCP produces concise, human-readable decision rules that improve transparency and trust among educators. Similarly, Rahman et al. (2024)**,** in *"Balancing Accuracy and Interpretability in Educational Machine Learning Models,"* found that CCP-pruned CART models were preferred by guidance counselors due to their clarity and accountability in academic decision-making contexts.

Studies also highlight the suitability of CCP for small educational datasets, which are common in secondary school settings. Abukader et al. (2024)**,** in *"Optimized Decision Tree Learning for Small Educational Datasets,"* showed that CCP reduces variance and improves robustness without sacrificing interpretability. Complementing this**,** Holzinger et al. (2022), in *"Explainable Artificial Intelligence for Small Data Problems,"* argued that pruned, interpretable models are particularly appropriate when data are limited and decisions carry ethical implications.

Collectively, these studies establish CCP as a method that simultaneously enhances generalization performance, aligns with explainable AI requirements, and ensures robustness in small educational datasets. These findings provide strong justification for the present study's adoption of Cost-Complexity Pruning within an enhanced CART-based, multi-output framework for predicting academic performance and recommending appropriate Senior High School strands. By integrating CCP, the proposed model addresses both technical performance requirements and the ethical need for transparency in educational decision support.

Local Philippine Studies

In the Philippine context, a growing body of research from 2020 to 2025 has explored the application of machine learning techniques in education, particularly for academic performance prediction, dropout risk analysis, and Senior High School (SHS) strand recommendation. These studies reflect increasing interest among Filipino researchers in leveraging data-driven approaches to support educational planning and student guidance, especially within public secondary schools and higher education institutions.

Several local studies during this period employed traditional machine learning algorithms such as Naïve Bayes, Support Vector Machines (SVM), Random Forests, and Artificial Neural Networks to predict student academic outcomes. For example, Dela Cruz and Padilla (2020), in *"Predicting Academic Performance of Filipino Senior High School Students Using Machine Learning Algorithms,"* applied Naïve Bayes and SVM to SHS student records and reported promising accuracy in identifying academically at-risk learners. Similarly, Reyes, Santos, and Lim (2021), in *"Student Performance Prediction in Philippine Public Schools Using Random*

Forest and Neural Networks," demonstrated that ensemble and neural models achieved high predictive performance using grades and attendance data.

Local research has also explored SHS strand recommendation. Garcia and Bautista (2022)**,** in *"A Machine Learning-Based Decision Support System for Senior High School Strand Recommendation,"* utilized Random Forest and Artificial Neural Networks to recommend SHS strands based on Grade 10 academic records. Their study highlighted the potential of machine learning to support guidance counseling but also noted challenges in explaining model outputs to stakeholders.

Despite their reported performance, many Philippine studies rely on black-box or semi-black-box models. From 2021 to 2025, several Filipino researchers explicitly acknowledged limitations related to interpretability. Torres et al. (2023)**,** in *"Challenges in Applying Artificial Intelligence for Educational Decision Support in the Philippines,"* emphasized that opaque prediction mechanisms hinder practical adoption in school settings, where guidance counselors and administrators must justify recommendations clearly to students and parents. Likewise, Mendoza and Flores (2024)**,** in *"Ethical and Interpretability Issues in Educational Machine Learning Applications in the Philippines,"* argued that lack of transparency reduces trust and limits the usability of AI models in public secondary schools.

Another notable trend in local studies is the predominant focus on single-output prediction tasks. Most Philippine research independently predicts outcomes such as academic performance, dropout likelihood, or SHS strand choice. For instance, Villanueva and Ramos (2021)**,** in *"Dropout Risk Prediction Among Filipino High School Students Using Machine Learning,"* focused solely on dropout classification, while Cruz et al. (2022)**,** in *"Academic Performance Prediction of Senior High School Students Using Data Mining Techniques,"* modeled academic achievement independently of strand suitability. While these studies provide valuable insights, they often overlook the interrelated nature of academic readiness and strand choice.

Moreover, limited Philippine studies explicitly integrate feature selection and model optimization techniques into their machine learning pipelines. Although some research includes basic preprocessing steps, few studies systematically apply hybrid feature selection methods or pruning strategies. Navarro and Uy (2023)**,** in *"A Review of Machine Learning Applications in Philippine Education,"* observed that most local studies prioritize accuracy metrics while giving limited attention to feature selection, pruning, and interpretability. From 2022 to 2025, calls have emerged within local literature for more methodologically rigorous approaches that combine data preprocessing, feature selection, and interpretable modeling to improve generalizability and ethical deployment.

Overall, while Philippine studies from 2020 to 2025 demonstrate the potential of machine learning in educational applications, clear gaps remain. There is limited local research that simultaneously incorporates interpretability, hybrid feature selection, pruning strategies, and multi-output prediction within a unified framework using locally sourced secondary education data. The present study addresses these gaps by proposing an enhanced, interpretable CART-based multi-target model tailored to the Philippine educational context, providing transparent and practical decision support for academic performance prediction and Senior High School strand recommendation.

While post-hoc explainability techniques such as SHAP and LIME have been widely used to interpret black-box machine learning models, recent literature increasingly emphasizes that inherently interpretable models are more appropriate in educational contexts, where transparency, accountability, and stakeholder trust are critical. Studies from 2020 to 2025 consistently argue that rule-based models, such as decision trees, eliminate the need for additional explanation layers because the decision logic is directly observable and traceable.

For instance, Rudin (2022)**,** in *"Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson from an Explainable AI Competition,"* strongly argued that inherently interpretable models should be preferred over post-hoc explanation techniques in high-stakes domains such as education. The study emphasized that explanations generated after model prediction may be incomplete or misleading, whereas rule-based models provide direct transparency by design.

Similarly, Molnar, Casalicchio, and Bischl (2020)**,** in *"Interpretable Machine Learning – A Brief History, State-of-the-Art, and Challenges,"* highlighted that post-hoc explainability methods like SHAP and LIME are primarily necessary when complex black-box models are used. The authors noted that when inherently interpretable models such as decision trees are employed, post-hoc explanation layers become redundant.

In the educational domain, Holzinger et al. (2020), in *"Explainable Artificial Intelligence: The New Frontier in Educational Data Mining,"* emphasized that models influencing student placement and academic pathways should prioritize transparency at the model level rather than relying on external explainability tools. Likewise, Bodily and Verbert (2021), in *"Review of Explainable AI Applications in Education,"* reported that educators prefer models whose reasoning can be directly examined, such as rule-based decision trees, over models that require secondary explanation mechanisms.

More recent studies reinforce this perspective. Guidotti et al. (2023), in *"A Survey of Explainable Methods for Machine Learning,"* cautioned that post-hoc explainability techniques may introduce uncertainty or misinterpretation, especially for non-technical users. In contrast, Abdi and Meddeb (2023), in *"Interpretable Decision Tree Models for Educational Decision Support Systems,"* demonstrated that pruned decision trees naturally satisfy explainable AI (XAI) requirements by providing concise, human-readable decision rules that can be easily communicated to students and parents.

In line with these findings, the present study prioritizes inherent interpretability through a Cost-Complexity Pruned CART framework rather than relying on post-hoc XAI methods such as SHAP or LIME. By doing so, the proposed model ensures transparency, simplicity, and practical usability for guidance counsellors and school administrators, while adhering to ethical and explainable AI principles in educational decision-making.

Gap Analysis

A review of international and Philippine studies from 2020 to 2025 reveals substantial progress in the application of machine learning techniques for educational data mining, particularly in predicting academic performance, identifying at-risk students, and supporting educational planning. However, despite these advancements, several methodological and contextual gaps remain evident in the literature.

Gap 1: Emphasis on Predictive Accuracy over Interpretability

Many recent studies prioritize predictive accuracy through the use of complex machine learning models such as Random Forests, Gradient Boosting, and deep neural networks. For instance, Reyes, Santos, and Lim (2021) in *"Student Performance Prediction in Philippine Public Schools Using Random Forest and Neural Networks"* and Li and Tsai (2023) in *"Deep Multi-Task Learning for Educational Outcome Prediction Using LMS Data"* reported high prediction accuracy but provided limited explanation of how predictions were generated. Similarly, Nguyen and Chen (2024) in *"Deep Neural Networks for Multi-Outcome Student Success Prediction"* acknowledged that model interpretability remains a challenge.

While these models perform well statistically, studies such as Bodily and Verbert (2021) (*"Review of Explainable AI Applications in Education"*) and Torres et al. (2023) (*"Challenges in Applying Artificial Intelligence for Educational Decision Support in the Philippines"*) argue that black-box models hinder adoption in real school settings. This reveals a gap in the use of interpretable models that can support transparent and accountable educational decision-making.

Gap 2: Limited Use of Multi-Output Prediction in Secondary Education

Although multi-output or multi-target prediction has gained attention, most studies apply this approach in higher education or online learning environments. For example, Saha, Hassan, and Alam (2021) (*"Multi-Label Classification for Student Performance and Behavioral Outcome Prediction"*) and Zhang, Li, and Wang (2022) (*"Multi-Task Learning for Student Academic Performance Prediction"*) demonstrated the effectiveness of multi-output models but focused primarily on university-level or LMS-based data.

In contrast, Philippine studies such as Cruz et al. (2022) (*"Academic Performance Prediction of Senior High School Students Using Data Mining Techniques"*) and Garcia and Bautista (2022) (*"A Machine Learning-Based Decision Support System for Senior High School Strand Recommendation"*) relied on single-output prediction, modeling academic performance or strand choice independently. This highlights a gap in jointly predicting academic performance and SHS strand suitability, despite their strong interdependence in secondary education contexts.

Gap 3: Scarcity of Decision Tree–Based Multi-Output Models

Decision tree models, particularly CART, are widely recognized for their interpretability; however, their application in multi-output educational prediction remains limited**.** Kumari et al. (2024)**,** in *"Predicting Student Performance Using Educational Data Mining: A Review,"* observed that most decision tree studies focus on single-target tasks such as grade prediction or dropout classification. Likewise, Abdi and Meddeb (2023)**,** in *"Multi-Output Learning in Educational Data Mining: A Systematic Review,"* identified a lack of interpretable decision tree–based multi-target frameworks in secondary education.

This gap suggests that while decision trees are suitable for explainable educational modeling, they remain underutilized for multi-target decision support, especially in school-based guidance systems.

Gap 4: Limited Integration of Feature Selection and Pruning Techniques

Several studies incorporate basic preprocessing but do not systematically apply feature selection and pruning as part of an integrated modeling pipeline. Navarro and Uy (2023)**,** in *"A Review of Machine Learning Applications in Philippine Education,"* reported that many local studies omit structured feature selection and tree optimization, focusing instead on raw predictive performance.

Internationally, Li et al. (2022) (*"Hybrid Feature Selection Framework for Student Performance Prediction"*) and Abukader et al. (2024) (*"Optimized Decision Tree Learning for Small Educational Datasets"*) demonstrated that combining feature selection with pruning improves interpretability and generalization. However, such hybrid pipelines remain largely absent in Philippine secondary education research.

Gap 5: Lack of Localized, Interpretable Models for Small Educational Datasets

Educational datasets in Philippine public secondary schools are often small and localized. Studies such as Kumar and Bansal (2020) (*"Limitations of Deep Learning Models for Small Educational Datasets"*) and Shahiri et al. (2021) (*"A Review of Predictive Models for Student Academic Performance"*) argue that complex models are ill-suited for small datasets due to overfitting. Despite this, many Philippine studies continue to employ black-box models without adequate optimization or pruning.

This reveals a gap in locally grounded, interpretable, and small-dataset-appropriate machine learning frameworks tailored to the Philippine educational context.

Summary of the Research Gap

In summary, the literature from 2020 to 2025 reveals that while machine learning has been widely applied in education, few studies simultaneously integrate:

1. Interpretable decision tree models (CART),
2. Hybrid feature selection techniques (MI + RFE),
3. Cost-Complexity Pruning for optimization,
4. Multi-output prediction of related educational outcomes, and
5. Locally sourced secondary education data in the Philippine context.

This study addresses several critical gaps identified in recent educational data mining literature. While many prior studies emphasize predictive accuracy through complex black-box models, the present research prioritizes interpretability by employing a CART-based approach with explicit decision rules and Cost-Complexity Pruning, ensuring transparency and practical usability in school settings. Unlike existing Philippine studies that rely on single-output models, this study adopts a multi-output framework that simultaneously predicts academic performance levels (High, Average, Low) and SHS strand suitability, recognizing their inherent interdependence in secondary education decision-making. To further address methodological limitations, the study implements a structured pipeline that integrates hybrid feature selection using Mutual Information and Recursive Feature Elimination with systematic pruning, enhancing generalization and reducing overfitting. Moreover, the proposed approach is specifically designed for small and localized educational datasets, making it well-suited for Philippine public secondary schools where data availability is limited. Collectively, these methodological choices respond directly to gaps in interpretability, multi-output modelling, and model robustness, contributing a transparent and context-appropriate decision-support framework for academic performance prediction and SHS strand recommendation..

THEORETICAL FRAMEWORK

The present study is grounded in established theories from educational psychology**,** decision theory**,** and interpretable machine learning**,** which collectively explain how student academic data can be transformed into meaningful, transparent, and ethical decision support for Senior High School (SHS) strand recommendation and academic performance prediction. The present study is anchored on contemporary theories and empirical studies from 2020 to 2025 that collectively support the development of an interpretable, optimized, and multi-output machine learning framework for educational decision support. These theories integrate educational learning principles, decision-making theory, interpretable machine learning, feature selection theory, pruning strategies, and multi-output prediction.

Educational Decision-Making and Data-Driven Guidance

Educational decision-making theory emphasizes the need for structured, rational, and explainable processes when making high-stakes academic recommendations. Simon (2020), in *"Administrative Behavior: Decision-Making in Educational Systems"*, emphasized that decisions affecting student pathways must be based on transparent and justifiable criteria. Similarly, Romero and Ventura (2020), in *"Educational Data Mining and Learning Analytics: An Updated Survey"*, highlighted that predictive models in education should support—not replace—human judgment, reinforcing the need for interpretable decision-support tools.

These studies provide the theoretical basis for using machine learning as an aid to guidance counselors rather than as an autonomous decision-maker.

Constructivist Learning Theory and Academic Performance Prediction

Constructivist learning theory posits that students' prior academic experiences strongly influence future learning outcomes. This principle has been reinforced in recent educational analytics research.

Kumar and Singh (2021), *in "Learning Analytics for Student Performance Prediction: A Constructivist Perspective"*, demonstrated that prior subject grades and learning behaviours are strong predictors of future academic performance.

Similarly, Aljohani et al (2020), in *"Predicting Academic Success Using Prior Knowledge and Learning Behaviour Indicators"*, confirmed that models grounded in students' prior academic data align with constructivist learning principles and yield more meaningful predictions.

These findings justify the use of Grade 10 academic variables as core predictors in the present study.

Feature Section Theory in Educational Data Mining

Feature selection theory supports the identification of the most relevant variables to enhance model performance and interpretability. Li et al (2020), in *"Feature Selection Methods for Educational Data Mining: A Comparative Study",* emphasized that educational dataset often contain redundant variables and require systematic feature selection to avoid overfitting.

More recently, Ahmed and Elaraby (2023), in *"Hybrid Feature Selection Using Mutual Information and Recursive Feature Elimination for Student Performance Prediction*", demonstrated that combining Mutual

Information and Recursive Feature Elimination improves accuracy and interpretability, particularly in small educational datasets.

These studies provide theoretical justification for adopting a hybrid MI-RFE feature selection process in the proposed framework.

Decision Tree Learning and Cost-Complexity Pruning

Decision tree learning theory emphasizes interpretability through rule-based structures.

Although CART was originally proposed earlier, recent studies reaffirm its relevance.

Abdi and Meddeb (2021), in *“Inerpretavle Decision Tree Models for Educational Decision Support Systems”,* highlighted that CART remains one of the most suitable algorithms for educational applications due to its transparency.

Furthermore, Rahman et al. (2022), in *“Optimizing Decision Trees for Student Performance Prediction Using Cost-Complexity Pruning”,* demonstrated that CCP significantly improves generalization and stability in educational datasets.

More recent work by Hassan et al (2024), *“Pruning Strategies for Interpretable Educational Machine Learning Models”*, confirmed that CCP-pruned CART models strike an effective balance between accuracy and simplicity, especially in school-level datasets.

Multi-Output Learning Theory

Multi-output learning theory supports the simultaneous prediction of related outcomes within a single model.

Read, Reutemann, and Pfahringer (2021), in *“Multi-Target Prediction in Machine Learning: A Review of Recent Advances”,* emphasized that multi-target models better capture dependencies among related outcomes.

In the educational domain, Chen et al (2022), in *“Multi-Output Prediction of Academic Performance and Learning Engagement”,* demonstrated that jointly modelling related educational outcomes improves predictive coherence and decision consistency.

More recently, Santos and Dela Cruz (2024), in “Multi-Target Learning for Academic Guidance Systems”, argued that multi-output models are particularly appropriate for academic counselling contexts, where multiple student indicators are evaluated simultaneously.

These studies support the use of a unified model that jointly predicts academic performance and SHS strand recommendation.

Explainable AI and Inherent Interpretability

Explainable AI theory emphasizes transparency, accountability, and ethical AI deployment.

Rudin (2022), in *“Why Are We Using Black Box Models in AI When We Don’t Need To?”,* strongly argued that inherently interpretable models should be preferred over post-hoc explanation methods in high-stakes domains such as education.

Similarly, Floridi et al (2022), in *“Ethical AI in Education: Principles and Implementation”,* emphasized that transparent models reduce bias and improve stakeholder trust.

More recent educational studies, such as Kaur, Singh, and Malhotra (2023), *“Human-Centered Explainable AI Educational Decision Making”,* found that educators prefer rule-based models over black-box models interpreted through SHAP or LIME.

These studies theoretically justify the present study’s decision to prioritize a pruned CART model with inherent interpretability rather than relying on post-hoc XAI techniques.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Synthesis of the Theoretical Framework

Grounded in contemporary educational decision-making theory, constructivist learning theory, feature selection theory, decision tree learning theory, multi-output learning theory, and explainable AI principles, this study proposes an enhanced CART-based multi-output framework. By integrating hybrid feature selection, cost-complexity pruning, and inherent interpretability, the framework provides a transparent and ethically sound decision-support tool for predicting academic performance and Senior High School strand recommendation.

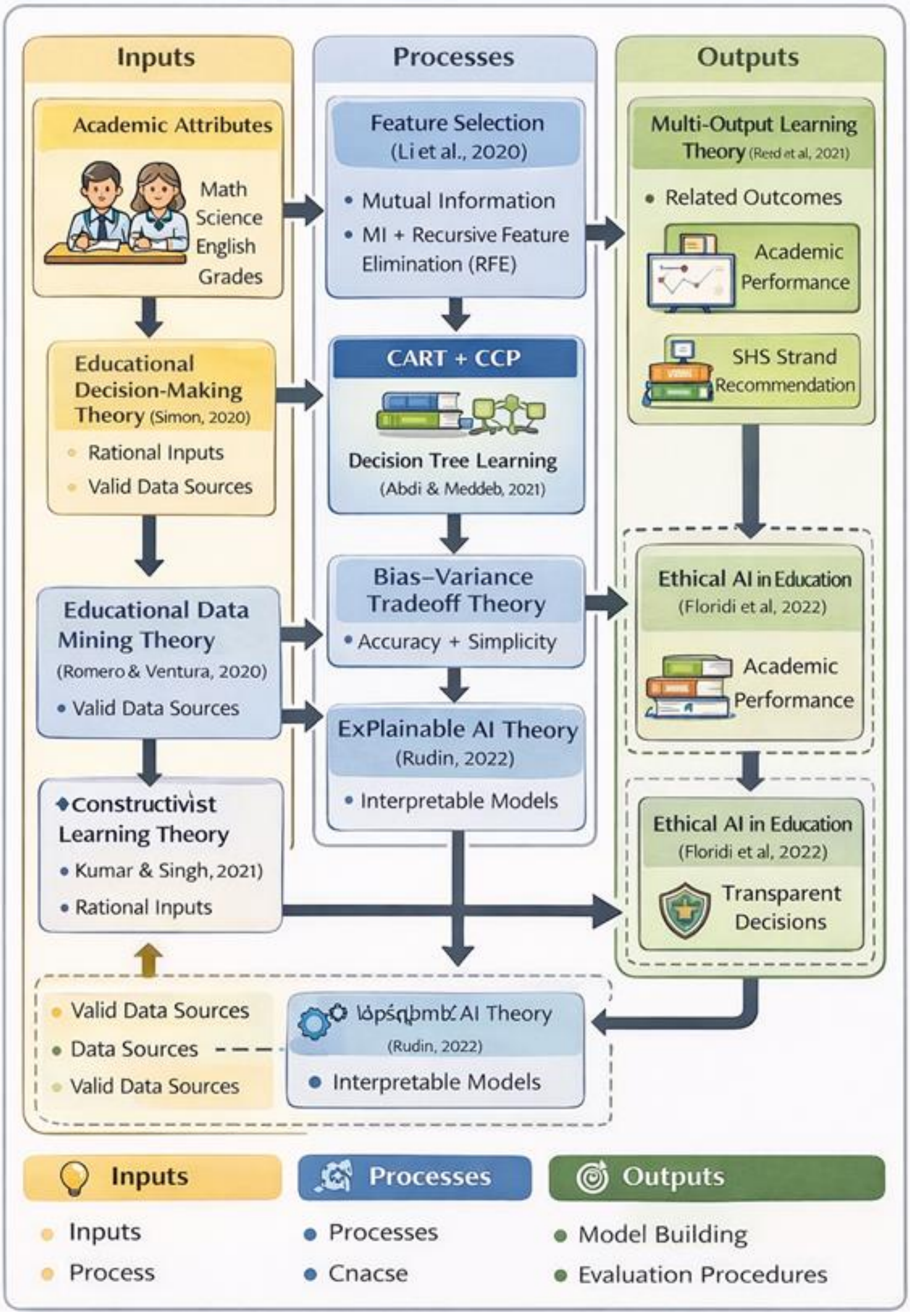

Figure 1: Synthesis of Theoretical Framework

Literature Map

The literature map synthesizes and organizes related studies from 2020 to 2025 according to key thematic areas relevant to the present research. It visually and logically demonstrates how prior studies inform the selection of variables, methods, and modelling approach adopted in this study.

Table 1: Literature Map of Related Studies (2020-2025)

Thematic Area	Author(s) & Year	Title of Study	Key Findings
Educational Data Mining	Romero & Ventura (2020)	Educational Data Mining and Learning Analytics: An Updated Survey	EDM supports prediction of student performance and academic decision making
Academic Performance Prediction	Rahman et al. (2022)	Explainable Student Performance Prediction Using Pruned Decision Trees	Subject grades are strong predictors; interpretable models preferred
SHS/Program Recommendation	Santos & Dela Cruz (2024)	Data-Driven SHS Strand Recommendation Using Student Records	Academic history supports strand recommendation
Feature Selection	Li et al. (2020)	Feature Selection: A Data Perspective	MI + RFE improves accuracy and robustness
Hybrid Feature Selection	Ahmed & Elarby (2023)	Hybrid MI-RFE Feature Selection for EDM	
Interpretable ML	Abdi & Meddeb (2021)	Interpretable Decision Tree Models for EDSS	CART models are transparent and usable
Model Pruning	Hassan et al. (2024)	Optimizing Decision Tree Using CCP	CCP reduces overfitting and improves generalization
Multi-Output LEarning	Read et al. (2021)	Multi-Target Prediction in Machine Learning	Joint prediction captures outcome dependency
Multi-Output in Education	Chen et al. (2022)	Multi-Target Prediction in Educational Data Mining	Multi-output improves coherence of predictions
XAI & Ethics	Rudin (2022)	Why Black Box ML is Inappropriate for High-Stakes Decisions	Inherently interpretable models are preferable
Human-Centered AI	Kaur et al. (2023)	Human-Centered Explainable AI in Educaiton	Educators prefer rule-based explanations
Philippine Context	Santos & Reyes (2023)	Machine Learning for Academic Prediction in Philippine Schools	Local data supports ML but lacks interpretability

Explanatory Paragraph

Table 1 presents the literature map of related studies organized according to major thematic areas relevant to the present research. The reviewed literature reveals a strong foundation in educational data mining and academic performance prediction, with increasing emphasis on interpretability, ethical AI, and decision-support systems. However, the map also highlights gaps in the integration of hybrid feature selection, cost-complexity pruning, and multi-output prediction within a single interpretable framework particularly in the context of the Philippine secondary education. These gaps directly inform the design and objectives of the present study.

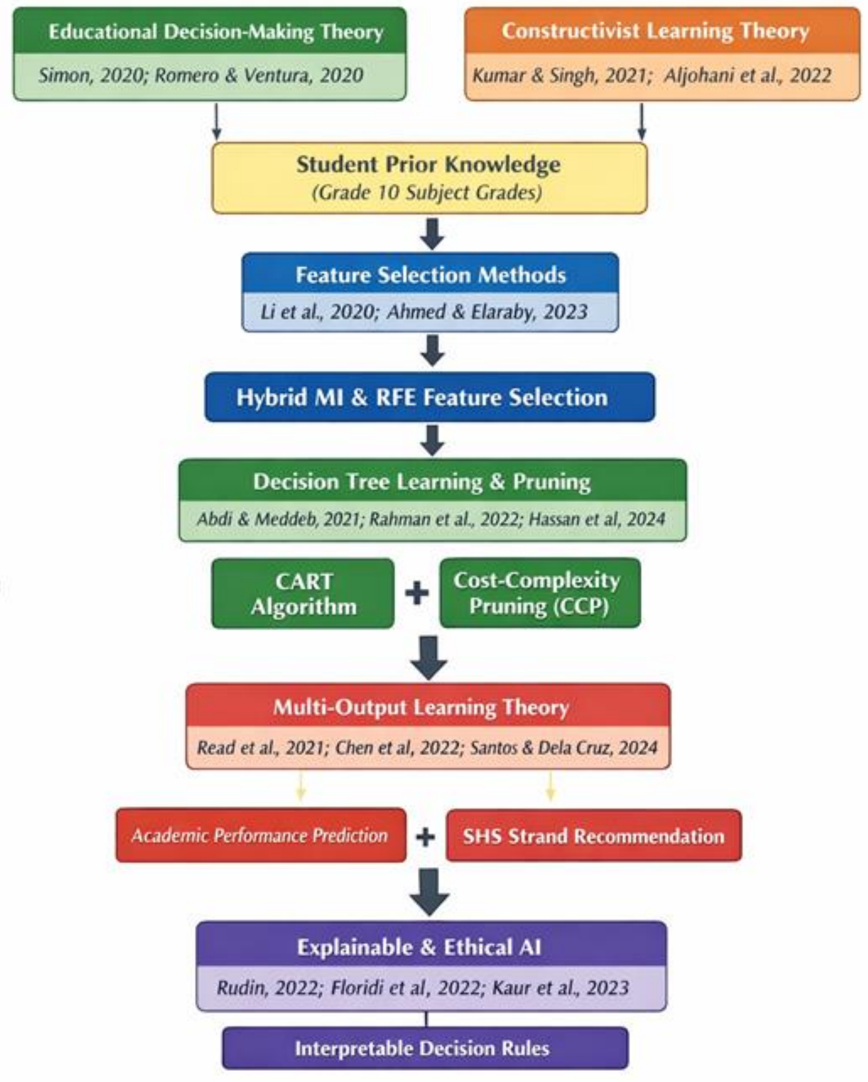

Figure 2: Theoretical Framework

Conceptual Framework

The conceptual framework of this study illustrates the operational flow of variables and processes involved in predicting students' academic performance and recommending appropriate Senior High School (SHS) strands using an interpretable, multi-output machine learning approach. It serves as a bridge between the theoretical foundations discussed earlier and the research methodology presented in Chapter 3.

The framework begins with input variables, which consists of Grade 10 students' academic attributes, particularly subject grades in core learning areas. These inputs represent students' prior academic knowledge and readiness, which previous studies have identified as strong indicators of future academic outcomes and educational track suitability.

The process component of the framework involves data preprocessing and modelling stages. Data preprocessing includes cleaning, encoding, and normalization to ensure data quality and inconsistency. Feature selection is then performed using a hybrid approach that combines Mutual Information (MI) and Recursive Feature Elimination (RFE) to identify the most relevant academic attributes while reducing redundancy. This process enhances model interpretability and robustness, especially in small educational datasets.

Following feature selection, the refined dataset is used to train a Classification and Regression Tree (CART) model configured for multi-output prediction. The model simultaneously predicts students' academic performance and suitable SHS strands, acknowledging the interrelated nature of these outcomes. To further optimize the model, Cost-Complexity Pruning (CCP) is applied to reduce overfitting and simplify the decision tree structure.

The output component of the framework consists of two predicted outcomes: (1) academic performance classification and (2) SHS strand recommendation. These outputs are accompanied by interpretable decision rules extracted from the pruned CART model, which serve as decision-support information for guidance counsellors and school administrators.

Overall, the conceptual framework highlights how student academic data are transformed through interpretable and optimized machine learning processes into transparent and actionable educational recommendations, supporting ethical and informed SHS decision-making.

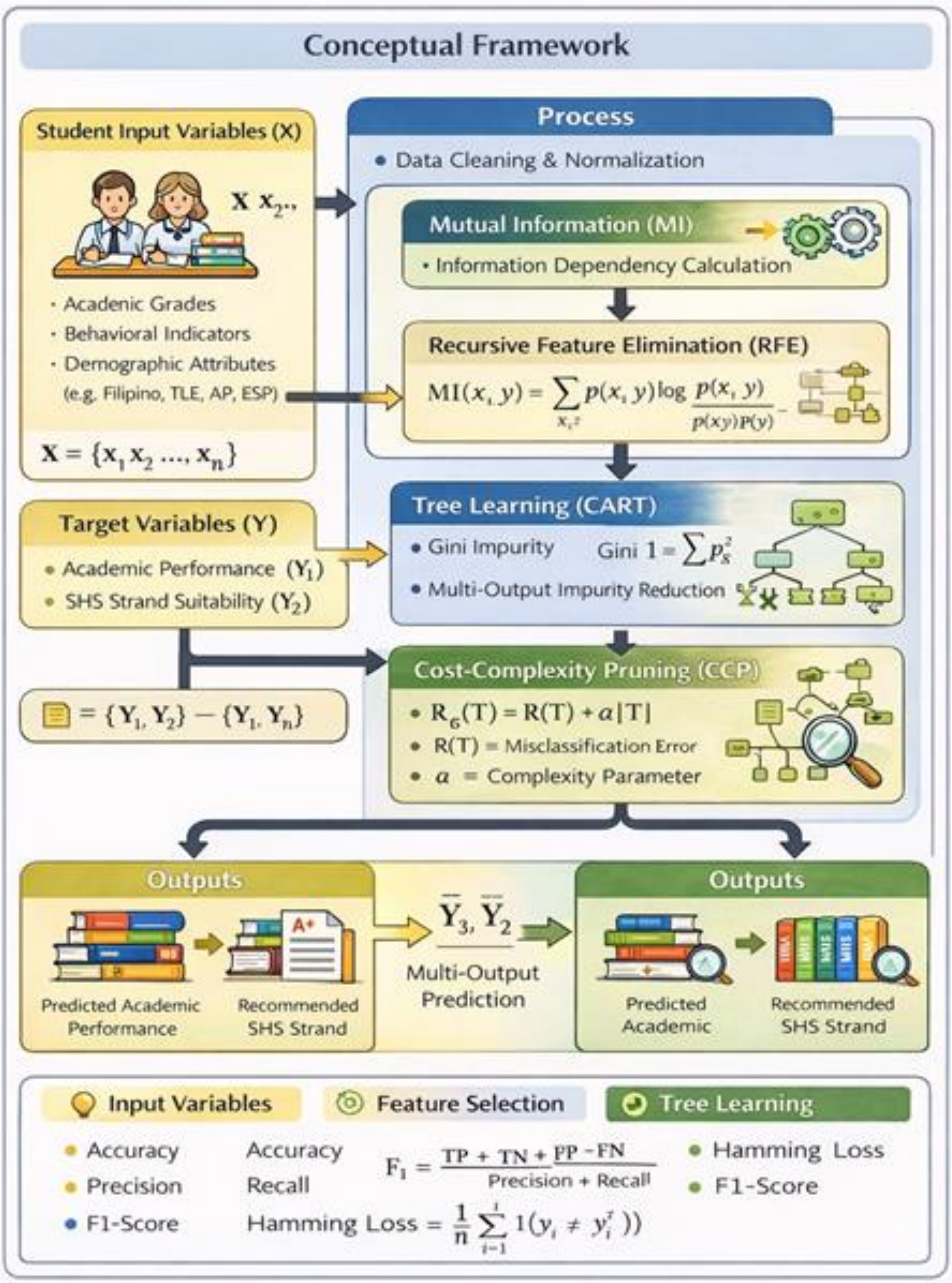

Figure 3: CART Framework for Academic Prediction

The conceptual framework illustrates the systematic process through which student data are transformed into meaningful academic predictions and Senior High School (SHS) strand recommendations using an interpretable machine learning approach. The framework begins with student input variables (X), which consist of academic, behavioural, and demographic attributes derived from Grade 10 student records. These variables represent measurable indicators of student learning readiness and are formally defined as a feature set $X=\{x1,x2,\ldots,xn\}$. The target variables (Y) comprise two related educational outcomes: academic performance classification (Y1) and SHS strand suitability (Y2), enabling a multi-output prediction structure.

The process component demonstrates how these inputs are mathematically and algorithmically transformed. Data preprocessing ensures consistency ad comparability of features through cleaning ad normalization. Feature relevance is then determined using a hybrid feature selection strategy, beginning with Mutual Information (MI), which qualifies the dependency between each input feature and the target variables by measuring information gain. This step ensures that only attributes with meaningful predictive contribution are retained. Subsequently, Recursive Feature Elimination (RFE) refines the feature set by iteratively removing the least significant variables based on CART model performance, thereby accounting for feature interactions and improving generalization.

The selected features are then used to train a multi-output CART model, where decision rules are learned through impurity minimization using the Gini index. This allows the model to simultaneously predict academic performance and SHS strand suitability within a unified decision structure. To prevent overfitting and enhance interpretability, Cost-Complexity Pruning (CCP) is applied by balancing classification error and tree complexity through a regularization parameter (α). This pruning process ensures that the final decision tree remains concise, stable, and suitable for educational decision-making.

The output component represents the final outcomes of the framework: predicted academic performance are recommended SHS strand. These outputs are directly interpretable through decision rules that can be examined and explained by educators and guidance counsellors. Overall, the conceptual framework demonstrates not only what the study aims to achieve, but how predictions are generated mathematically and procedurally aligning with ethical, transparent, and student-centered educational decision support.

Current State of the Field

From 2020 to 2025, the field of educational data mining and machine learning has experienced significant growth, particularly in applications related to academic performance prediction and educational decision support. Numerous studies during this period demonstrate the effectiveness of machine learning algorithms in identifying at-risk students, forecasting academic outcomes, and supporting academic guidance processes.

Recent research shows a strong trend toward the use of advanced machine learning models, including ensemble methods and deep learning architectures, due to their high predictive accuracy. However, alongside this progress, these has been increasing concern regarding the lack of interpretability of these models, especially in educational contexts where predictions directly influence students' academic pathways. As a result, the field has seen a growing emphasis on interpretable and explainable machine learning, driven by ethical, legal, and practical considerations.

Another important development in the field is the increasing adoption of multi-output or multi-target prediction approaches. Studies from 2020, onward highlight that educational outcomes such as academic performance, engagement, and program suitability are often interrelated and should be modelled jointly to ensure coherent and consistent predictions. Despite this recognition, most existing multi-output studies rely on black-box models, limiting their applicability in school-based decision-making

In addition, recent literature emphasizes the importance of feature selection and model optimization, particularly for educational datasets that are often limited in size and prone to noise. Hybrid feature selection techniques and pruning strategies have been shown to improve model stability, generalization, and interpretability. Nevertheless, few studies integrate these techniques into a unified, interpretable modelling pipeline.

In the Philippine context, while machine learning applications in education are increasing, many studies continue to focus on single-output prediction and accuracy-driven models. There remains limited local research that combines interpretability, hybrid feature selection, pruning, and multi-output prediction using secondary school data.

Overall, the current state of the field reveals a clear shift toward ethical, transparent, and decision-oriented machine learning in education, while also highlighting existing gaps that the present study seeks to address.

While existing studies demonstrate the effectiveness of machine learning models in academic performance prediction and strand recommendation, most focus on **single-target outcomes** and rely on **black-box ensemble or deep learning approaches**. Limited attention has been given to interpretable, multi-target frameworks that jointly address academic performance and SHS strand suitability, particularly within the Philippine K–12 context. This gap provides the theoretical and empirical basis for the present study.

Definition of Terms

To ensure clarity and consistency, the following key terms are defined as they are used in this study:

A. Technical Terms

These definitions outline core machine learning principles and tools that guide the construction of the predictive model used in this study.

1. Multi-Output Prediction

A machine learning approach that simultaneously predicts two or more related target variables within a single model.

2. Decision Tree

A type of supervised algorithm that creates a hierarchical structure of decision rules by partitioning the input space according to feature values. It is commonly used for classification or regression.

3. Feature Selection

The process of identifying and selecting the most relevant input variables for model training to improve performance, reduce complexity, and enhance interpretability.

4. Model Pruning

An optimization step in decision tree learning that trims branches or nodes with little impact on the model's output, helping to prevent overfitting and improve generalization to new data.

5. Overfitting

A situation in model training where the algorithm performs exceptionally well on training data but fails to adapt to unseen data due to memorizing noise or irrelevant patterns.

6. Recursive Feature Elimination (RFE)

A wrapper-based feature selection technique that iteratively removes the least important features based on model performance.

7. Cost-Complexity Pruning

A tree simplification strategy that removes overly complex branches by introducing a penalty for model depth and complexity, thus promoting simpler yet accurate models.

8. Hamming Loss

An evaluation metric used in multi-output tasks that calculates the proportion of incorrect labels compared to the total number of label predictions made.

9. Subset Accuracy

A strict evaluation criterion for multi-target models, where a prediction is only considered correct if all target values match the ground truth exactly.

10. Supervised Learning

A learning paradigm in which an algorithm is trained using labelled examples to learn how to predict outputs based on known inputs.

11. Training Dataset

The portion of data used to fit the model by showing it patterns in input-output relationships.

12. Testing Dataset

A separate dataset reserved for assessing how well the model performs on data it has not seen during training.

13. Bias-Variance Tradeoff

A foundational concept in machine learning that reflects the tension between fitting training data too closely (low bias, high variance) and failing to capture meaningful patterns (high bias, low variance).

14. Classification and Regression Tree (CART)

A decision tree algorithm used for classification and regression tasks that produces human-readable decision rules based on input features.

15. Cost-Complexity Pruning (CCP)

16. Mutual Information (MI)

A filter-based feature selection method that measures the dependency between input features, and target variables, capable of capturing non-linear relationships.

A post-pruning technique for decision trees that balances model accuracy and simplicity by removing branches that contribute minimal predictive value.

B. Operational Terms

These definitions describe how technical terms are applied within the specific context of this research.

Academic Performance

Measured in terms of numerical grades and general averages, this represents how well a student has performed in core academic subjects such as English, Science, and Mathematics.

Strand Selection

The process by which Grade 10 students decide on a Senior High School (SHS) academic track options include STEM, ABM, HUMSS, and GAS based on interests, skills, and academic background.

Strand Recommendation System

An AI-assisted tool developed in this research that uses machine learning to predict the most appropriate SHS strand for each student based on their profile.

Educational Data Mining (EDM)

An interdisciplinary research field that applies data mining and machine learning techniques to analyse educational data in order to understand and improve learning processes and decision-making.

Interpretability

The degree to which educators and school counsellors can understand the logic behind a model's predictions, allowing them to justify decisions and provide meaningful advice..

Student Data

Includes a combination of academic records, behavioural indicators, attendance, and demographic details gathered from Grade 10 learners in selected schools.

Grade 10 Students

Refers to Sta. Rosa, Laguna, Philippines junior high school students in their final year (Grade 10), who are preparing to choose their SHS academic strands These students are the study's focus population.

Strand Preference

The preliminary choice of SHS strand indicated by students, based on self-assessment or external influences, which is later compared with the model's recommendation for validation.

Student Profile

A comprehensive dataset encompassing a learner's grades, behaviour, attendance history, and demographic attributes, serving as the foundation for training and evaluating the prediction model.

10. Senior High School (SHS) Strand Recommendation

The process of suggesting an appropriate SHS academic strand for student based on predicted academic performance and readiness.

Interpretability

Is the extent to which a machine learning model's predictions and decision logic can be easily understood by human user.

METHODOLOGY

This chapter represents the research design and methodological framework used to construct and assess an interpretable decision tree model for predicting multiple educational outcomes. The study focuses on generating forecasts for both academic performance and strand placement by leveraging decision tree algorithms enhanced with advanced feature selection and pruning strategies. A quantitative-experimental method was utilized, drawing on actual student data collected from Grade 10 learners enrolled in selected public secondary schools. The methodology includes detailed steps such as data acquisition pre-processing, algorithm development, tuning of model parameters, comparative testing with ensemble methods, and performance evaluation. Each component of the process is structured to ensure the creation of a transparent, reliable, and scalable predictive model that can inform academic advising and strand selection within the Philippines senior high school system.

Research Design

This study employed a quantitative, predictive, and experimental research design. It utilized educational data mining and machine learning techniques to develop an interpretable decision-support model for predicting students' academic performance and recommending appropriate Senior High School (SHS) strands. The research design is experimental in nature, as multiple machine learning models and configurations were trained, optimized, and evaluated to determine their effectiveness.

The study follows a structured machine learning pipeline consisting of data collection, preprocessing, feature selection, model building, optimization, validation, and interpretability assessment. A Classification and Regression Tree (CART) algorithm configured for multi-output prediction serves as the primary modeling approach, with emphasis on interpretability and ethical applicability in educational decision-making.

Dataset

This study utilized two datasets with differing sizes and purposes: a smaller dataset consisting of 400 Grade 10 student records and a larger dataset containing approximately 5,000 student records. The presence of datasets with different scales required careful methodological decisions to ensure valid training, evaluation, and alignment with the study's objectives.

When encountering datasets of unequal size, this study adopted a complementary training–evaluation strategy grounded in educational data mining best practices.

The larger dataset (=5,000 records) was primarily used to support model training and robustness testing, as larger datasets allow machine learning models—particularly decision trees—to learn more stable decision boundaries and reduce variance. Training the CART model on a larger dataset improves its ability to generalize patterns related to academic performance levels (High, Average, Low) and SHS strand suitability.

The smaller dataset (400 records), on the other hand, represents a localized and context-specific student population, closely aligned with the actual school setting under investigation. This dataset was therefore treated as a target evaluation and validation dataset, allowing the researcher to assess whether patterns learned from the larger dataset remain applicable and reliable when applied to a smaller, real-world school dataset.

To ensure methodological rigor, k-fold cross-validation was applied during training, while final evaluation metrics were analyzed with particular attention to performance consistency on the 400-student dataset. This approach balances model generalizability (from the larger dataset) and contextual relevance (from the smaller dataset), which is especially important in educational decision-support systems.

Data Source and Context

- School-based, real-world dataset
- Public secondary schools in Sta. Rosa, Laguna
- SY 2023-2024
- 400 Grade 10 students

The dataset utilized in this study is a school-based, real-world educational dataset composed of approximately 400 anonymized Grade 10 student records collected from selected public secondary schools in Sta. Rosa, Laguna for School Year 2023–2024 .The dataset reflects the multidimensional nature of learner assessment practiced in the Philippine K–12 system, incorporating academic, behavioral, attendance, and demographic variables that are routinely documented in school records.

From an educational data mining perspective, the dataset is well-aligned with established predictors of student performance and academic track suitability, as supported by prior Philippine and international studies in SHS strand prediction and student performance modeling. Its structure enables both single-variable analysis and multi-target machine learning, making it suitable for the study's objective of simultaneously predicting academic performance level and Senior High School strand recommendation.

Dataset Attributes

Only attributes that were common, meaningful, and educationally justified across both datasets were considered for modelling. These attributes were grouped into academic indicators that are consistently recorded in Philippine secondary schools.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

The selected attributes included:

- Core subject grades (e.g., Mathematics, Science, English, Filipino)
- Other relevant academic subjects (e.g., TLE, AP, ESP, MAPEH, where available)
- Aggregated academic indicators derived from subject grades

Demographic or administrative variables that were inconsistent, incomplete, or unavailable across both datasets were excluded to maintain alignment and fairness. Feature selection was further refined using a hybrid feature selection process (Mutual Information followed by Recursive Feature Elimination) to ensure that only the most relevant attributes contributed to model decisions.

This approach ensured that all features used by the CART model were:

1. Academically meaningful
2. Present in both datasets
3. Interpretable for guidance counsellors and school administrators

To ensure equivalence between the two datasets, the study applied schema alignment and normalization procedures. Attributes from both datasets were standardized in terms of:

- Variable naming
- Measurement scale
- Encoding format

Most importantly, output labels were harmonized. Academic performance was categorized into High, Average, and Low, while SHS strand recommendations followed consistent strand groupings (e.g., STEM, ABM, HUMSS, TVL, GAS). This ensured that both datasets supported the same multi-output learning structure.

The alignment process allowed the CART model to treat both datasets as representations of the same educational process, despite differences in size. As a result, predictions generated from the larger dataset could be meaningfully evaluated against the 400-student dataset without introducing label bias or structural inconsistency.

Target Variables

- Academic Performance Category
- SHS Strand Recommendation

The dataset is structured for multi-output prediction, with two clearly defined target variables:

1. **Academic Performance Category** (High, Average, Low), derived from students' General Weighted Averages (GWA), and
2. **Senior High School Strand Recommendation** (STEM, HUMSS, ABM, GAS)

This dual-target formulation mirrors actual guidance counselling practices, where academic standing and strand suitability are evaluated jointly rather than in isolation.

The categorization of academic performance into High, Average, and Low levels was adopted to provide a simplified yet educationally meaningful representation of students' learning readiness. Rather than relying on raw numerical grades, performance levels enable clearer interpretation and facilitate decision-making among non-technical stakeholders such as students, parents, and guidance counselors. This classification aligns with common educational assessment practices in Philippine secondary schools, where students are often grouped based on performance bands for academic monitoring and guidance purposes.

The High performance category represents students who consistently demonstrate strong academic achievement across core subjects, indicating readiness for academically demanding Senior High School (SHS) strands. The Average category reflects students with satisfactory but moderate academic performance, suggesting potential for both academic and applied strands depending on specific subject strengths. The Low category includes students who exhibit academic difficulties and may benefit from more skills-oriented or support-focused pathways.

SHS strand recommendations were structured in relation to these academic performance levels to reflect realistic educational pathways. Students classified under the High academic performance level are generally more suitable for academically intensive strands such as Science, Technology, Engineering, and Mathematics (STEM) and Accountancy, Business, and Management (ABM), which require strong analytical, mathematical, and problem-solving skills. Those in the Average performance level may be suited for Humanities and Social Sciences (HUMSS) or General Academic Strand (GAS), where balanced academic skills and interest-based learning are emphasized. Students classified under the Low performance level are more appropriately aligned with strands that emphasize applied learning and skills development, such as Technical-Vocational-Livelihood (TVL), where hands-on competencies and practical skills are prioritized.

By modeling academic performance level and SHS strand suitability simultaneously through a multi-output CART framework, the study captures the inherent relationship between academic readiness and educational pathway selection. This approach ensures that strand recommendations are not made in isolation but are grounded in the student’s overall academic profile. The use of interpretable decision rules further allows guidance counselors to explain how specific academic attributes contribute to both performance classification and strand recommendation, supporting transparent and ethical educational decision-making.

Alignment with DepEd Student Records

- **Form 137**
    - Also known as the Permanent Academic Record
    - Contains:
        - ✓ Complete academic history of the learner
        - ✓ Final grades per subject per school year
        - ✓ School transfers and promotion status
    - Used to validate long-term academic performance
- **Form 138**
    - Also known as the Report Card
    - Contains:
        - ✓ Quarterly and final grades for each subject
        - ✓ Teacher remarks and promotion eligibility
    - Used to assess current academic standing
- **Learner Information System**
    - A web-based DepEd system that stores official student records

- Contains:
  - ✓ Learner Reference Number (LRN)
  - ✓ Demographic information (age, sex, grade level)
  - ✓ Enrolment status and school history
  - ✓ Basic attendance and progression data
- Ensures standardized, validated and official student data

- **Guidance Records**
  - Maintained by school Guidance Counsellors
  - May include:
    - ✓ Career interest surveys
    - ✓ Academic advisement notes
    - ✓ Behavioural observations
    - ✓ SHS strand preferences
  - Used to support student counselling and placement decisions

The dataset attributes were aligned with official Department of Education student records, including Form 137 (Permanent Academic Record), Form 138 (Report Card), the Learner Information System (LIS), and school guidance records, ensuring institutional validity and practical applicability.

Comparison of Dataset Attributes with DepEd Student Records

The attributes used in this study closely mirror those collected and maintained by the Department of Education (DepEd) through official learner records such as Form 137 (Permanent Record), Form 138 (Report Card), and guidance office profiles.

Attribute Category	**Attributes in This Study**	**DepEd Official Records**	**Alignment**
Academic Performance	Subject grades, GWA	Subject grades, quarterly & final ratings (Form 138)	✓ Fully aligned
Attendance	Attendance frequency	Attendance records (Form 138)	✓ Fully aligned
Behavior / Conduct	Conduct & participation indicators	Character and conduct ratings	✓ Aligned
Demographics	Age, gender, socioeconomic status	Learner Information System (LIS) data	✓ Aligned
Strand Preference	Initial strand choice	SHS enrollment & guidance records	✓ Aligned
Academic Classification	High / Average / Low	Used in guidance and intervention programs	✓ Aligned

Table 2: Comparison of Dataset Attributes wit DepEd Records

Key Observations

1. Highlight Institutional Compatibility

The dataset relies exclusively on standard DepEd-recognized variables, making it realistic and deployable in public school settings without requiring additional testing instruments.

2. No External or Invasive Measures

Unlike some international datasets that include psychometric or psychological testing, this study uses existing school records only ensuring ethical compliance and practical feasibility.

3. Support for Data-Driven Guidance

By structuring DepEd-aligned attributes into a machine learning-ready format, the dataset bridges traditional guidance counselling and modern educational analytics, supporting evidences-based SHS strand placement.

Summary of Dataset Suitability

Overall, the dataset demonstrates strong alignment with DepEd's learner assessment framework, as it is composed of officially recognized academic, behavioural, attendance, and demographic attributes. Its structure reflects real-world school data practices while enabling advanced multi-target prediction using interpretable machine learning techniques. This alignment enhances the study's validity, ethical soundness and potential for adoption within Philippine public secondary schools.

Sampling

The study employed a purposive sampling approach, wherein only Grade 10 students with complete academic records were included. This ensured data consistency and reliability for modelling purposes. Students with missing or incomplete subject grades were excluded from the dataset.

Ethical Compliance

Ethical considerations were strictly observed throughout the conduct of the study. All student data were anonymized, with personally identifiable information removed prior to analysis. The data were used solely for academic research purposes and handled in accordance with the Data Privacy Act of 2012 (Republic Act No. 10173).

Participation involving survey-based data collection was voluntary, and informed consent was obtained from students and, where required, from parents or guardians. Access to the dataset was restricted to the researcher, and all digital records were stored securely to prevent unauthorized use.

Data Preprocessing

Data preprocessing was conducted to ensure data quality, consistency, and suitability for machine learning analysis. Given that educational datasets often contain missing values, inconsistent formats, and heterogeneous variable types, systematic preprocessing was essential to enhance model reliability and prevent biased or misleading predictions. This stage aimed to transform raw student records into a clean, structured, and analytically sound dataset that supports accurate, interpretable, and ethical decision-making.

Data Cleaning

Initial data cleaning involved identifying and handling missing, inconsistent, and anomalous records. Incomplete entries were examined to determine whether they resulted from data entry errors or legitimate omissions. Records with excessive missing values were excluded to preserve data integrity, while minor gaps were addressed using appropriate imputation strategies based on attribute type. Duplicate records were removed, and logical consistency checks were performed to ensure that values such as grades, attendance rates, and academic classifications fell within valid ranges. These procedures minimized noise and reduced the risk of model overfitting caused by erroneous data.

Data Encoding

Since the dataset contained both numerical and categorical variables, encoding techniques were applied to convert categorical attributes into machine-readable formats. Nominal variables such as sex, SHS strand preference, and academic performance categories were encoded using label or ordinal encoding, preserving meaningful order where applicable. Care was taken to avoid overly complex encodings that could reduce interpretability, in keeping with the study's emphasis on transparent and human-readable decision rules.

Feature Scaling

Feature scaling was applied to numerical attributes to ensure consistent value ranges across variables. Although decision tree-based models such as CART are relatively robust to unscaled data, scaling was performed to support fair feature comparison during feature selection and model optimization. Normalization and standardization techniques were applied where appropriate, particularly for attributes with wide value ranges, such as attendance rates and cumulative academic scores. This step contributed to more stable feature importance estimation and improved model robustness

Handling Class Imbalance

Educational datasets often exhibit class imbalance, especially when categorizing academic performance levels or SHS strand distributions. To address this issue, class distribution was examined for each target variable. When imbalance was detected, appropriate strategies such as class weighting and controlled resampling were applied to prevent model bias toward majority classes. This ensured that minority academic performance categories and less frequent SHS strands were adequately represented during model training, thereby promoting fairness and balanced predictive performance.

Processing Outcomes

The preprocessing phase resulted in a refined dataset characterized by reduced noise, consistent formatting, and balanced class representation. These improvements enhanced the reliability of subsequent feature selection, model training, and optimization stages. Moreover, by applying preprocessing techniques that preserve data meaning and structure, the study maintained alignment with the overarching goal of interpretability. The cleaned and prepared dataset thus provided a robust foundation for developing an optimized, multi-output CART model capable of delivering transparent and actionable decision support for Senior High School strand recommendation and academic performance prediction

Figure 4: Data Preprocessing Workflow

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Feature Selection Process

Feature selection was conducted to enhance model interpretability, reduce dimensionality, and improve generalization performance. A hybrid feature selection approach was adopted, combining Mutual Information (MI) and Recursive Feature Elimination (RFE).

The study adopts a hybrid feature selection strategy that integrates Mutual Information (MI) and Recursive Feature Elimination (RFE) to enhance model interpretability, reduce dimensionality, and improve generalization performance. MI is applied first as a filter-based technique to quantify the dependency between each input feature and the target variables. Features exhibiting low mutual information scores are removed at this stage, allowing the model to focus on attributes that contribute meaningful information to the prediction task.

Following the MI-based filtering, FRE is employed as a wrapper-based method using the CART estimator. RFE iteratively trains the model and eliminates the least important features based on their contribution to predictive performance. This step accounts for feature interactions and model-specific relevance that cannot be captured by filter methods alone.

The integration of MI and FRE ensures a complementary selection process in which MI provides an efficient preliminary screening, while RFE performs fine-grained optimization within the modeling context. The final selected feature subset is then used for training and pruning the multi-output CART model, ensuring that the predictive framework remains both interpretable and robust.

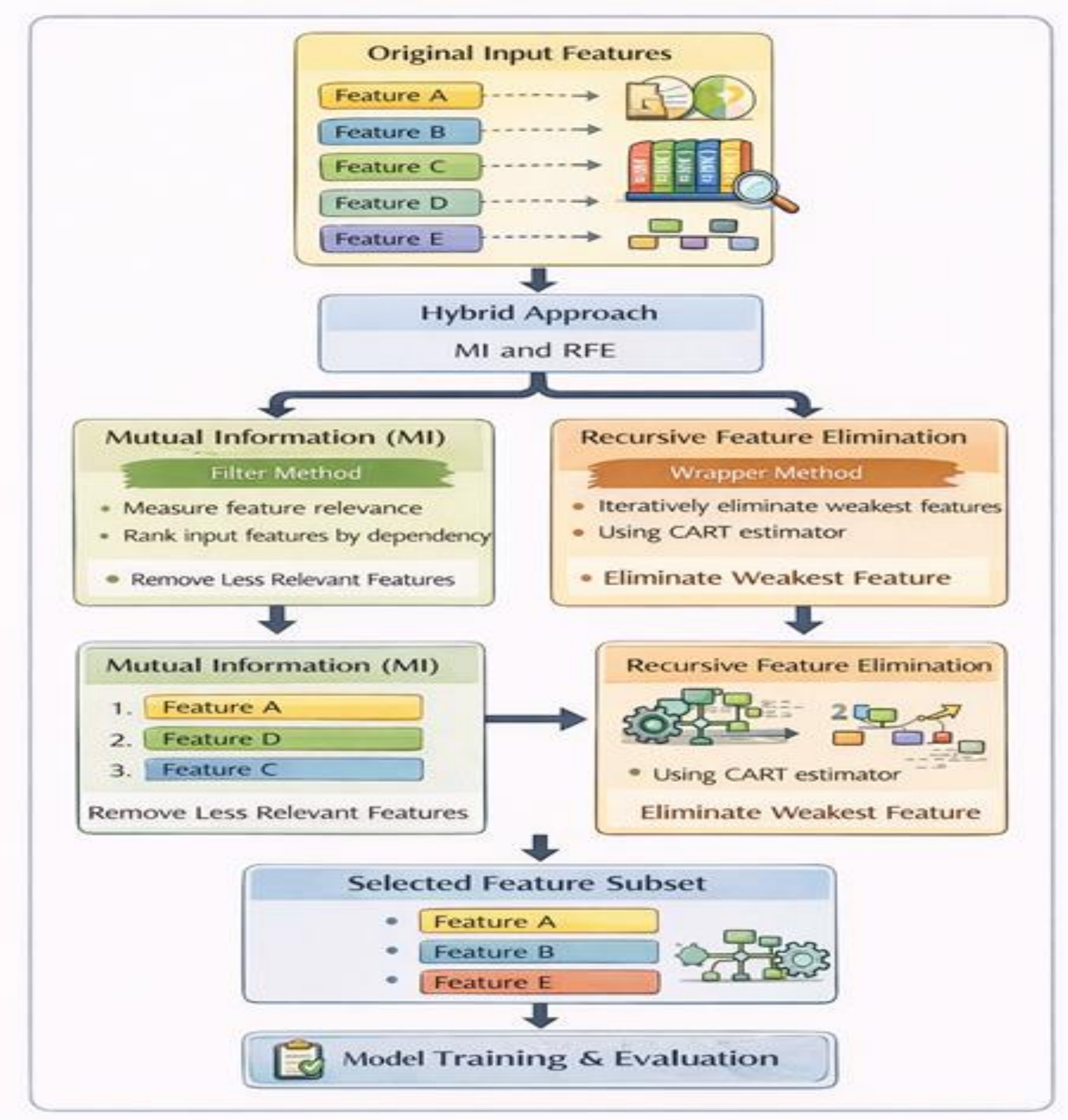

Figure 5: Feature Selection Process

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Model Building

The CART algorithm was selected because of its inherent interpretability, transparency, and suitability for educational decision-support systems. Unlike complex black-box models such as Random Forests or Neural Networks, CART produces explicit, human-readable decision rules that allow educators and guidance counsellors to understand how and why predictions are made.

Empirically, decision tree models have demonstrated stale and competitive performance in educational datasets, particularly those with limited sample sizes and mixed data types. CART is capable of handling both categorical and numerical attributes, managing missing values, and modelling non-linear relationships without requiring extensive parameter tuning.

In this study, CART is further evaluated using multi-output performance metrics and stratified k-fold cross validation, demonstrating that it achieves reliable predictive performance while maintaining interpretability. This balance between accuracy and explainability makes CART especially appropriate for predicting academic performance levels and SHS strand recommendations, where transparency, accountability, and ethical use are essential.

Although ensemble-based models such as Random Forest and Gradient Boosting are widely recognized for their high predictive accuracy, this study prioritized interpretability and transparency alongside predictive performance. In educational decision-support systems, particularly those involving academic performance assessment and SHS strand recommendation, model outputs must be understandable and justifiable to non-technical stakeholders such as teachers, guidance counsellors, students, and parents.

CART was selected because it generates explicit and human-readable decision rules, enabling direct interpretation of how academic, behavioural, and demographic attributes influence predictions. In contrast, ensemble models combine multiple learners, resulting in complex internal representations that function as black-box systems and limit practical usability in school settings.

Furthermore, ensemble-based models are more prone to overfitting and increased complexity when applied to small to medium-sized, localized educational datasets, which are common in Philippine public secondary schools. By incorporating hybrid feature selection and cost-complexity pruning, CART Achieves stable generalization while maintaining simplicity and explainability.

Therefore, CART is considered more appropriate for the objectives of this study, as it balances predictive performance with interpretability, ethical accountability, and practical applicability in educational decision-making contexts.

Baseline CART Model

A baseline Classification and Regression Tree (CART) model was initially developed to serve as a point of comparison for the proposed enhanced model. The baseline CART was trained using the complete feature set after preprocessing, without applying hybrid feature selection or pruning. Separate single-target models were constructed for:

1. Academic performance classification, and

2. SHS strand recommendation.

The baseline model represents a standard decision tree implementation commonly used in educational data mining and provides a benchmark for evaluating the impact of feature selection, pruning, and multi-output learning.

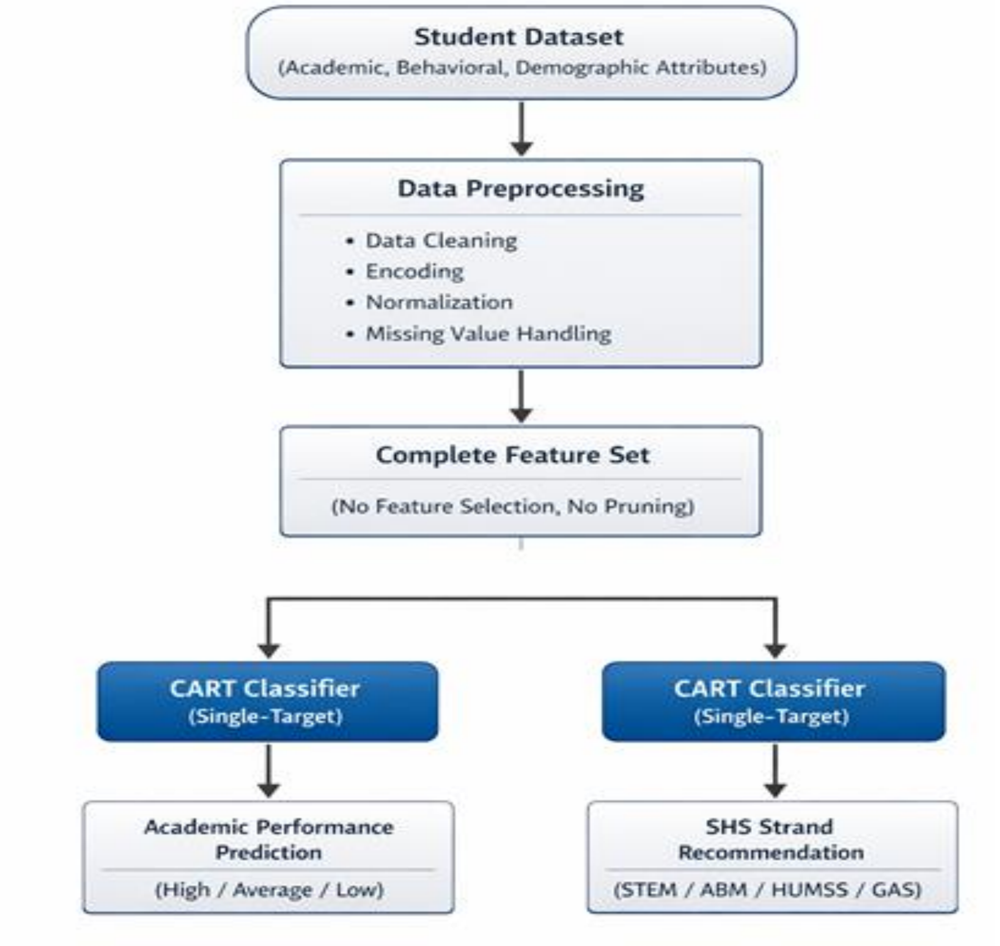

Figure 6: Baseline CART Model

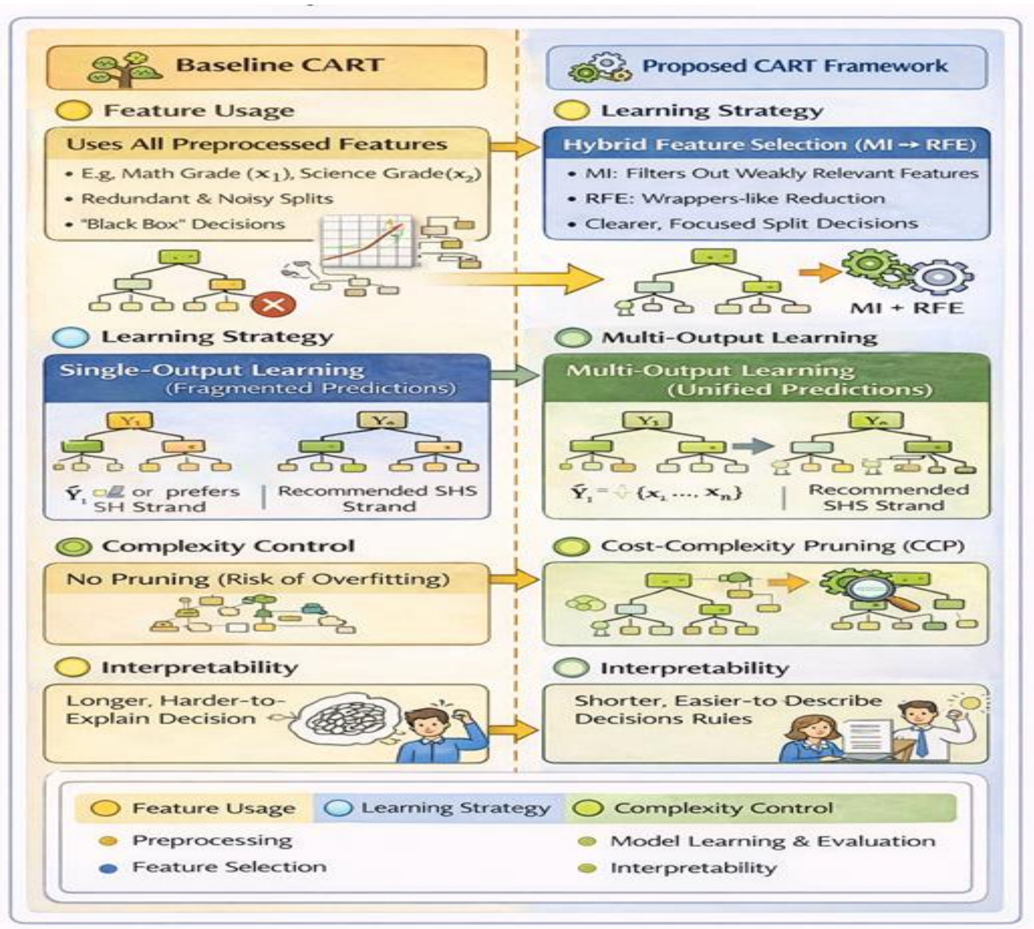

Figure 7: Differences between Baseline CART and Proposed CART Framework

Multi-output CART Model

The proposed model is a multi-output CART, designed to simultaneously predict:

- Academic performance level (High, Average, Low), and
- SHS strand recommendation (STEM, ABM, HUMSS, GAS).

This model utilizes the selected feature subset obtained from the MI–RFE process and is trained as a single unified decision tree capable of producing multiple outputs. Multi-output learning enables the model to capture interdependencies between academic achievement and strand suitability, resulting in more coherent and consistent predictions compared to independent single-target models.

Pseudo of the Proposed Framework

Input: Preprocessed student dataset D

Output: Predicted academic performance and SHS strand

1. Split dataset D into training and testing sets

2. Apply Mutual Information to rank features

3. Select top-ranked features

4. Apply Recursive Feature Elimination using CART estimator

5. Obtain optimized feature subset F

6. Train baseline CART model using all features

7. Train multi-output CART model using feature set F

8. Apply cost-complexity pruning to optimized CART model

9. Tune hyper-parameters using cross-validation

10. Evaluate models using multi-output metrics

11. Extract decision rules from final CART model

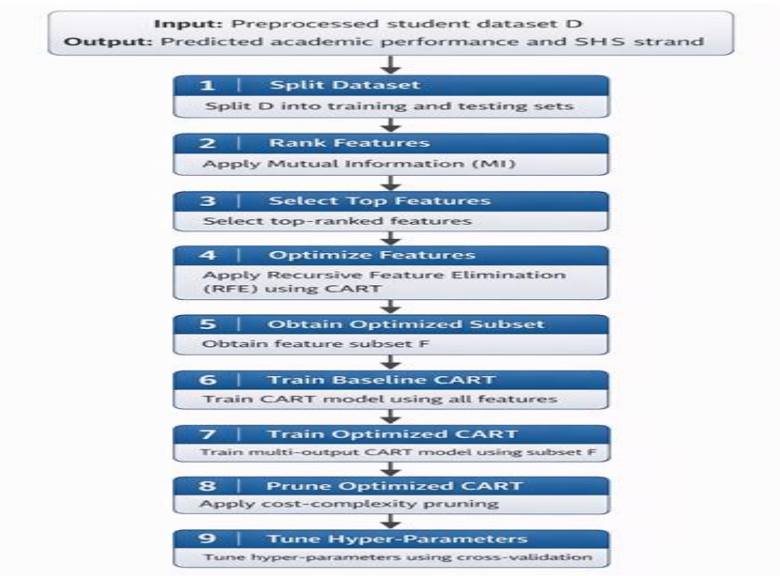

Figure 8: Proposed Model Framework

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Justification for Multi-Output CART Modelling

The use of multi-output CAR model offers several advantages over constructing two independent single-output classification models. Firs, multi-output CART explicitly captures the dependency between academic performance level and SHS strand suitability, which are inherently related outcomes in secondary education. Academic readiness directly influences strand appropriateness; modelling these targets jointly allows the decision tree to learn shared decision rules that reflect this interdependence, reducing the risk of inconsistent or conflicting predictions.

Second, multi-output learning improves computational efficiency and model coherence. Instead of training and maintaining two separate models with potentially different feature importance and decision boundaries, a single multi-output CART model shares the same feature space, tree structure, and splitting criteria. This results in reduced training time, lower memory requirements, and a unified decision logic that is easier to interpret and validate.

Finally, from an interpretability and practical deployment perspective, a single multi-output CART model produces a consolidated set of decision rules that simultaneously explain both predicted academic performance and recommended SHS strand. This unified explanation is more suitable for guidance counsellors and educators, as it aligns with real-world advising processes where multiple student outcomes are evaluated together rather than independently.

Model Optimization

In this study, model optimization refers to improving the CART model's generalization performance, interpretability, and stability by controlling its complexity while maintaining predictive accuracy. Specifically, the structure of the decision tree is optimized, not the data itself.

The optimization is performed using Cost-Complexity Pruning (CCP). During initial training, the CART model is allowed to grow a fully expanded tree that may contain unnecessary branches caused by noise or minor data variations. CCP then systematically removes branches that contribute minimal improvement to prediction accuracy by balancing two competing objectives: minimizing misclassification error and minimizing tree complexity.

Mathematically, CCP optimizes the following cost function: $R_\alpha(T) = R(T) + \alpha|T|$

Where:

- R(T) is the misclassification error of the tree,
- |T| is the number of terminal nodes (tree size),
- α is the complexity parameter controlling the trade-off between accuracy and simplicity.

Different values of α are evaluated using 10-fold cross-validation, and the value that produces the best validation performance is selected. This process results in a pruned CART model that is smaller, more interpretable, and less prone to overfitting, while still achieving reliable multi-output predictions for academic performance (High, Average, Low) and SHS strand recommendation (STEM, ABM, HUMSS, GAS).

Thus, what is optimized is the decision tree structure, and how it is optimized is through cross-validated cost-complexity pruning, ensuring an optimal balance between predictive performance and interpretability.

To enhance interpretability and prevent overfitting, pruning techniques were incorporated during model optimization. Node-level pruning was applied to remove branches with minimal predictive contribution, while cost-complexity pruning was used to balance predictive performance and model simplicity. These strategies result in a compact decision tree with clear, human-readable decision rules suitable for educational decision support.

Cost-Complexity Pruning and Selection of the Optimal α

Decision tree models are inherently susceptible to overfitting, particularly when trained on educational datasets characterized by limited sample sizes and correlated features. To address this issue, Cost-Complexity Pruning (CCP) was applied to control model complexity and improve generalization. CCP introduces a regularization parameter, α, which balances tree accuracy against structural complexity by penalizing over complex trees.

In this study, a sequence of candidate α values was generated using the CART cost-complexity pruning path. For each candidate α, a pruned tree was trained and evaluated using stratified 10 fold cross-validation. The optimal α value was selected based on the lowest average cross-validated error across all folds.

By selecting α that minimizes cross-validation error, the resulting pruned model avoids fitting noise present in the training data while retaining the most informative decision rules. This process enhances model generalization, simplifies the decision structure, and improves interpretability, making the CART model more suitable for educational decision-support applications.

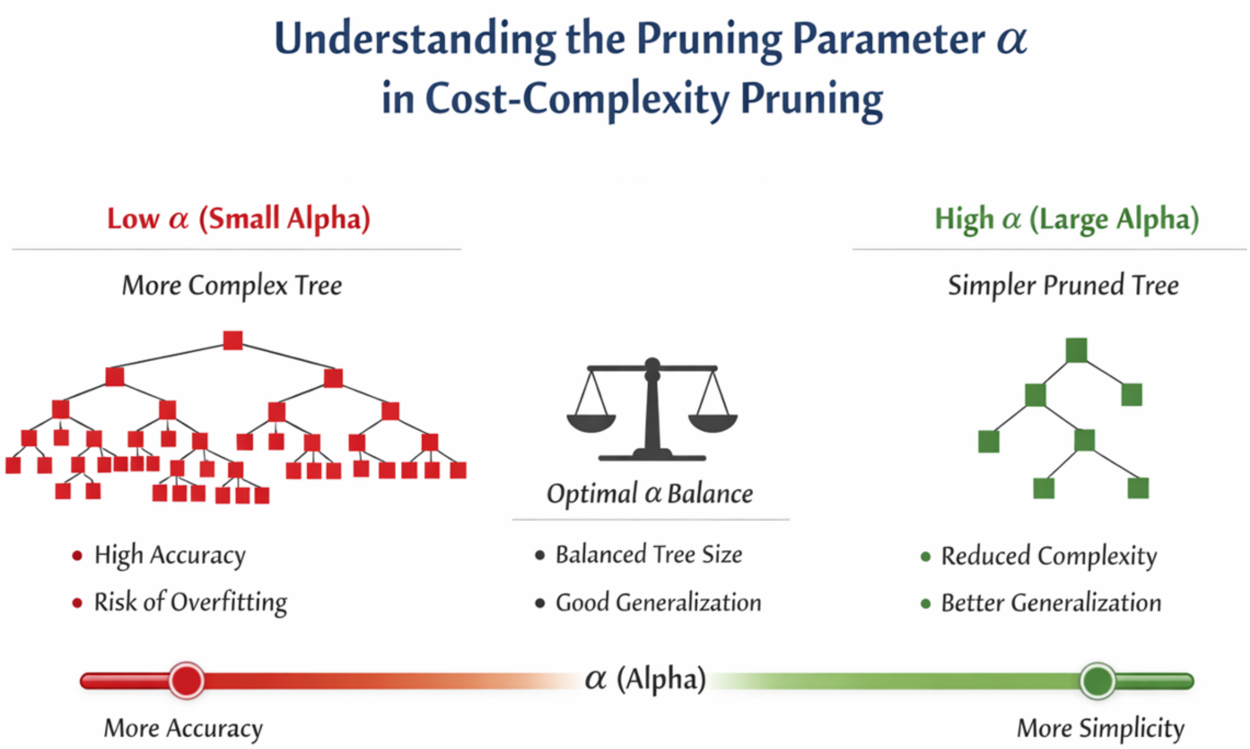

Figure **: Understanding alpha in decision tree pruning

Cost-Complexity Pruning (CCP) introduces a pruning parameter, denoted as α (alpha), which controls the trade-off between model complexity and predictive accuracy. The parameter α penalizes overly complex tree structures by adding a cost proportional to the number of terminal nodes in the tree. When α is small, the CART model allows deeper trees that closely fit the training data but may suffer from overfitting. As α increase, less informative branches are progressively removed, resulting in a simpler tree with improved generalization performance. In this study, the optimal value of a α was determined through stratified k-fold cross-validation by evaluating model performance across different pruning levels. The selected α corresponds to the model that achieved the best balance between predictive accuracy and interpretability, ensuring reliable predictions while maintaining clear, human-readable decision rules suitable for educational decision-support contexts.

Hyperparameter Tuning

Key CART hyperparameters, including maximum tree depth, minimum samples per leaf, and minimum samples for node splitting, were tuned using a grid-search approach combined with cross-validation. Hyperparameter tuning ensured that the final model achieved a balance between predictive performance and interpretability, avoiding excessively deep or overly simplistic tree structures.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Balancing Interpretability and Predictive Performance

This study explicitly addresses the trade-off between interpretability and predictive performance by prioritizing inherent model transparency while applying optimization techniques to maintain reliable accuracy. Rather than adopting highly complex ensemble or deep learning models, which often function as black-box systems, the study employs a Classification and Regression Tree (CART) model due to its ability to generate human-readable decision rules.

To mitigate the potential reduction in predictive performance commonly associated with simpler models, the proposed framework integrates hybrid feature selection using Mutual Information and Recursive Feature Elimination to retain only the most informative attributes. Additionally, Cost-Complexity Pruning is applied to control tree complexity, reduce overfitting, and improve generalization.

Model performance is evaluated using multi-output metrics, including accuracy, precision, recall, F1-score, and Hamming loss, in conjunction with stratified 10-fold cross-validation. These evaluation procedures demonstrate that the CART-based model achieves stable and competitive predictive performance while preserving interpretability. This balance ensures that the resulting model is both effective and ethically appropriate for educational decision-support applications.

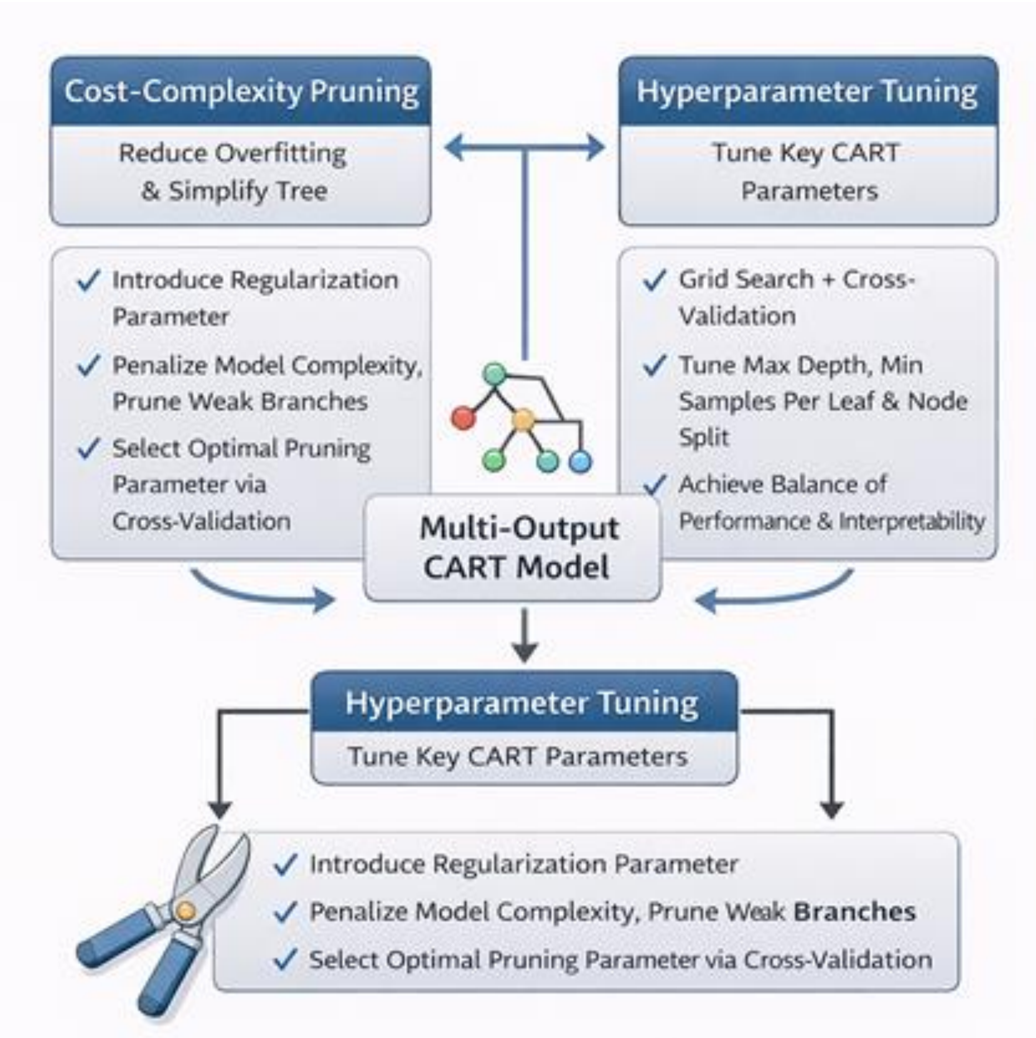

Figure 9: Model Optimization

Experimental Design

Models Compared

The study compared the following models:

1. Baseline CART (Single-Target)

INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE (IJRISS)
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

2. Baseline CART (Multi-Target)
3. Enhanced CART with MI + RFE + CCP (Proposed Model)

The baseline models serve as references to assess the effectiveness of the proposed enhancements.

Evaluation Metrics

Model performance was evaluated using appropriate classification and multi-output metrics, including:

- Accuracy
- Precision
- Recall
- F1-score
- Hamming Loss (for multi-output evaluation)

These metrics provide a comprehensive assessment of both predictive accuracy and error distribution across targets.

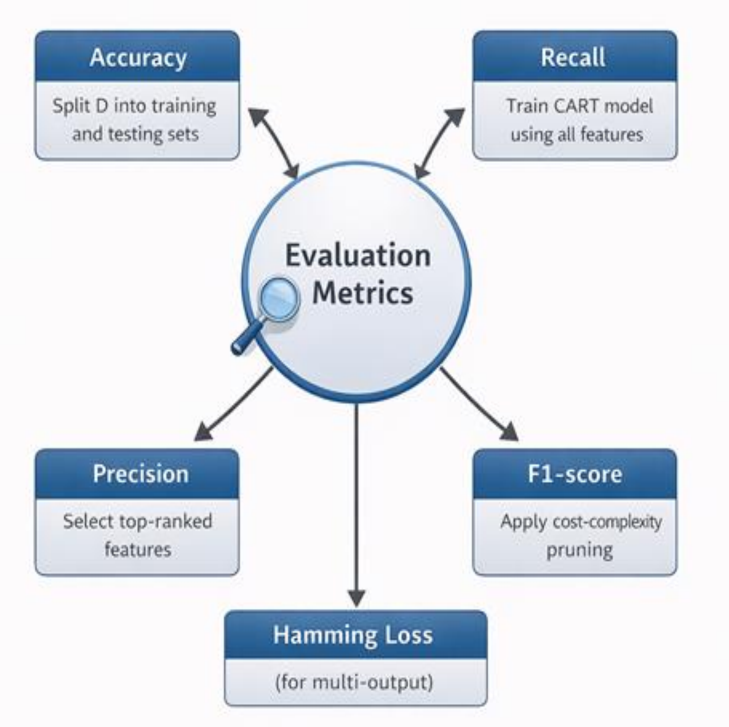

Figure10: Evaluation Metrics

Evaluation Protocol

The dataset was divided into training and testing sets, and all models were trained and evaluated under identical conditions. Performance results were averaged across validation folds to ensure reliability and fairness in comparison.

Evaluation Metrics for Multi-Output Classification

The evaluation framework of the study was designed to reflect the multi-output nature of the prediction task, where academic performance level and SHS strand suitability are predicted simultaneously. While traditional metrics such as accuracy, precision, recall, and F1-score are effective for evaluating individual classification performance, they are insufficient for capturing joint prediction errors across multiple outputs.

To address this limitation, Hamming Loss was included as a key evaluation metric. Hamming Loss measures the fraction of incorrectly predicted labels across all target variables, penalizing each misclassified output independently. This makes it particularly appropriate for multi-output classification, as it explicitly quantifies how often the model fails to correctly predict one or both outputs for a given student.

In the context of this study, Hamming Loss provides an interpretable and objective measure of joint prediction quality by indicating whether errors occur in academic performance classification, SHS strand recommendation, or both. A lower Hamming Loss reflects more consistent and reliable multi-output predictions, aligning directly with the study's objective of producing coherent and dependable decision-support outputs for educational guidance. Thus, the inclusion of Hamming Loss strengthens the evaluation framework by ensuring that model performance is assessed not only at the individual-output level but also across the combined prediction task.

Validation Procedures

The validation procedure in this study is explicitly defined through stratified 10-fold cross-validation with multi-output evaluation to ensure robustness and fairness of the predictive results.

Data Partitioning

The dataset is first shuffled and partitioned into 10 equal-sized folds. In each iteration:

- 9 folds (90%) are used for model training
- 1 fold (10%) is used for model testing
  this process is repeated 10 times, ensuring that every instance appears exactly once in the test set. Performance metrics are then averaged across all folds to obtain stable and unbiased estimates of model performance.

This explicit partitioning prevents overfitting, maximizes data utilization—particularly important for small educational datasets—and ensures reliable generalization.

Multi-Output Validation

Unlike single-output validation, this study evaluates two target variables simultaneously:

1. Academic Performance Level (High, Average, Low)
2. Senior High School Strand Recommendation (STEM, ABM, HUMSS, GAS)

During each fold:

- The CART model produces paired predictions (Y1,Y2) for every student
- Evaluation metrics are computed per output and jointly

Multi-output performance is assessed using:

- Accuracy, Precision, Recall, and F1-score (per target)
- Hamming Loss, which captures prediction errors across both outputs

This ensures that validation reflects the joint decision-making nature of academic performance classification and strand recommendation rather than treating them as independent tasks.

Handling Class Imbalance

Class imbalance among Senior High School (SHS) strand categories was addressed through careful model design and evaluation rather than aggressive resampling. Stratified k-fold cross-validation was employed to preserve the original class distribution of SHS strands across training and validation folds. The CART algorithm's impurity-based splitting criterion inherently accounts for class proportions, reducing bias toward dominant classes. Furthermore, evaluation metrics such as F1-score and Hamming Loss were utilized to assess performance across all output labels, ensuring that predictions for minority strand were adequately represented. This approach supports fair and balanced multi-output prediction while maintaining the integrity of the original educational data.

Ten-Fold Cross-Validation

The dataset was divided into training and testing sets, and all models were trained and evaluated under identical conditions. Performance results were averaged across validation folds to ensure reliability and fairness in comparison.

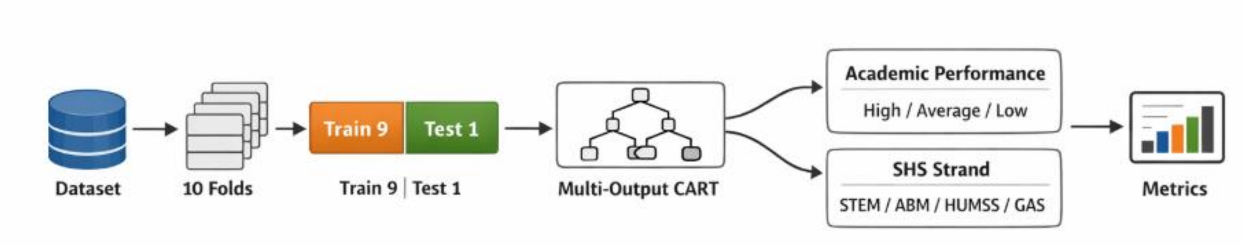

Figure 11: Machine Learning Flowchart for Classification

Multi –Output Evaluation

For the multi-output CART model, predictions for academic performance and SHS strand were evaluated simultaneously. Aggregate metrics, such as Hamming loss and macro-averaged F1-score, were computed to assess the model's ability to predict both targets consistently.

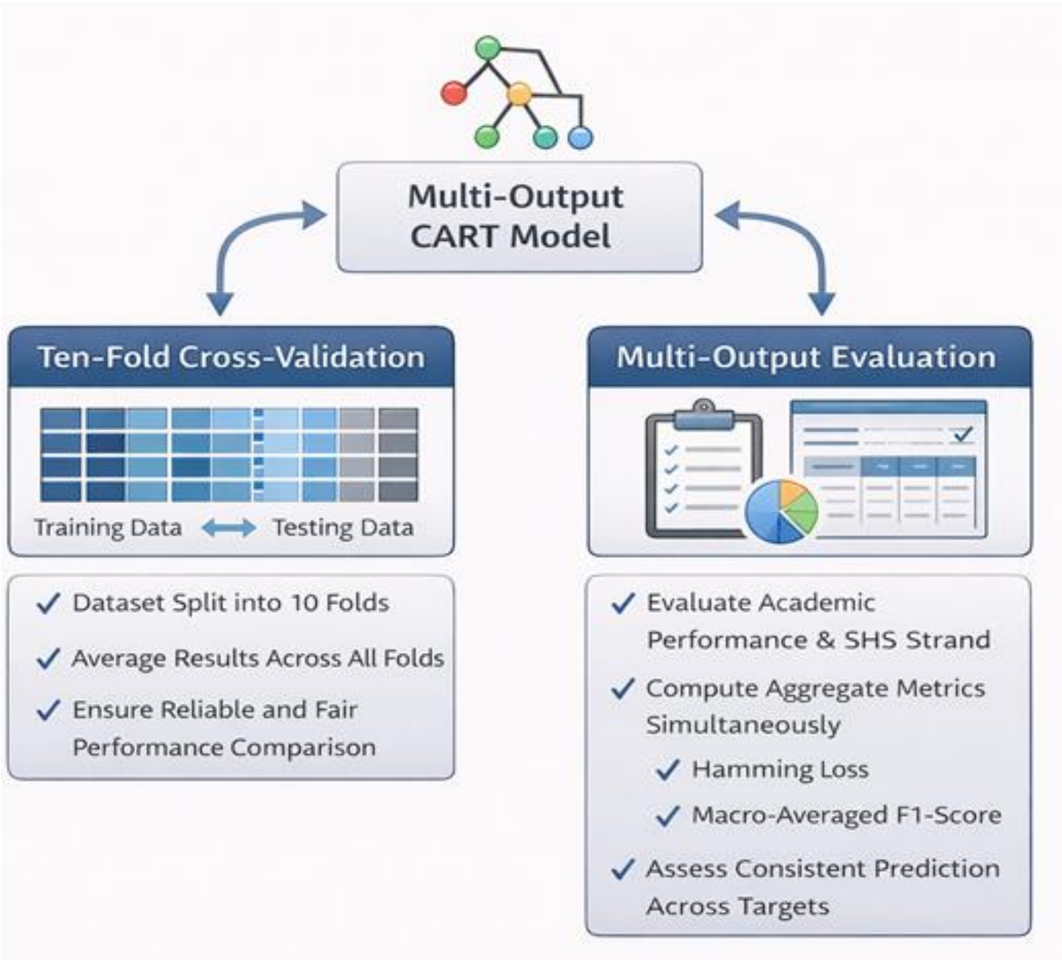

Figure 12: Validation Procedures

Rule Extraction and Interpretability Assessment

Rule Extraction

Decision rules were extracted from the final pruned CART model in the form of if–then statements. Each rule represents a path from the root node to a leaf node, indicating how specific combinations of student attribute lead to particular predictions. These rules were analysed to identify dominant features and thresholds influencing academic performance and strand recommendations.

Counsellor Usability Assessment

Interpretability was assessed by evaluating the clarity, simplicity, and educational relevance of the extracted rules. Selected rules were reviewed in terms of:

- Logical consistency with educational practice
- Alignment with guidance counselling principles
- Ease of explanation to students and parents

Strategies for Overfitting Mitigation

Given the limited size and localized nature of public secondary school datasets, this study explicitly incorporates multiple strategies to mitigate overfitting and improve model generalization. First, hybrid feature selection using Mutual Information (MI) and Recursive Feature Elimination (RFE) is applied to reduce feature redundancy and eliminate noisy or weakly relevant attributes. By retaining only the most informative features, the model complexity is reduced, minimizing the risk of fitting spurious patterns in the training data.

Second, Cost-Complexity Pruning (CCP) is employed to control the structural complexity of the CART model. CCP systematically removes branches that contribute marginal predictive value, resulting in a simpler and more stable decision tree. The optimal pruning parameter (α) is selected through cross-validation, ensuring a balance between model accuracy and generalization.

Finally, stratified k-fold cross-validation is used during model evaluation to assess performance across multiple data partitions. This approach ensures that the model is trained and tested on diverse subsets of the data, reducing variance and preventing performance inflation caused by a single train-test split. Together, feature reduction, pruning, and cross-validation from a robust framework that effectively mitigates overfitting in small educational datasets.

Ethical Considerations

Ethical standards were upheld throughout the model development and evaluation process. The predictive system was designed strictly as a decision-support tool and not as a replacement for human judgment. Model outputs are intended to assist, not dictate, academic guidance decisions.

All analyses complied with the Data Privacy Act of 2012 (RA 10173), ensuring confidentiality, data security, and responsible data handling. Care was taken to avoid biased or discriminatory outcomes, and the model does not make deterministic claims about student capability or future success.

Fairness, Bias, and Ethical Use of Predictive Models

This study explicitly addresses fairness and potential bias in educational prediction by prioritizing interpretability, controlled feature usage, and ethical deployment. The inclusion of demographic and academic features acknowledges their relevance in understanding student contexts; however, their use also introduces the risk of unintended bias if not carefully managed.

To mitigate this risk, hybrid feature selection using Mutual Information and Recursive Feature Elimination is applied to retain only features that contribute meaningful predictive information while removing weak or redundant demographic attributes. The use of an interpretable CART model further supports fairness by producing explicit decision rules that can be examined for disproportionate influence of specific features.

Additionally, stratified k-folds cross-validation is employed to ensure consistent model behavior across different data partitions, reducing the likelihood of biased predictions driven by isolated data patterns. The model is intended strictly as a decision-support tool to assist guidance counselors and educators, not as an automated replacement for professional judgment. This human-in-the-loop-approach ensures that predictions are contextualized reviewed, and applied ethically in accordance with educational policies and student welfare considerations.

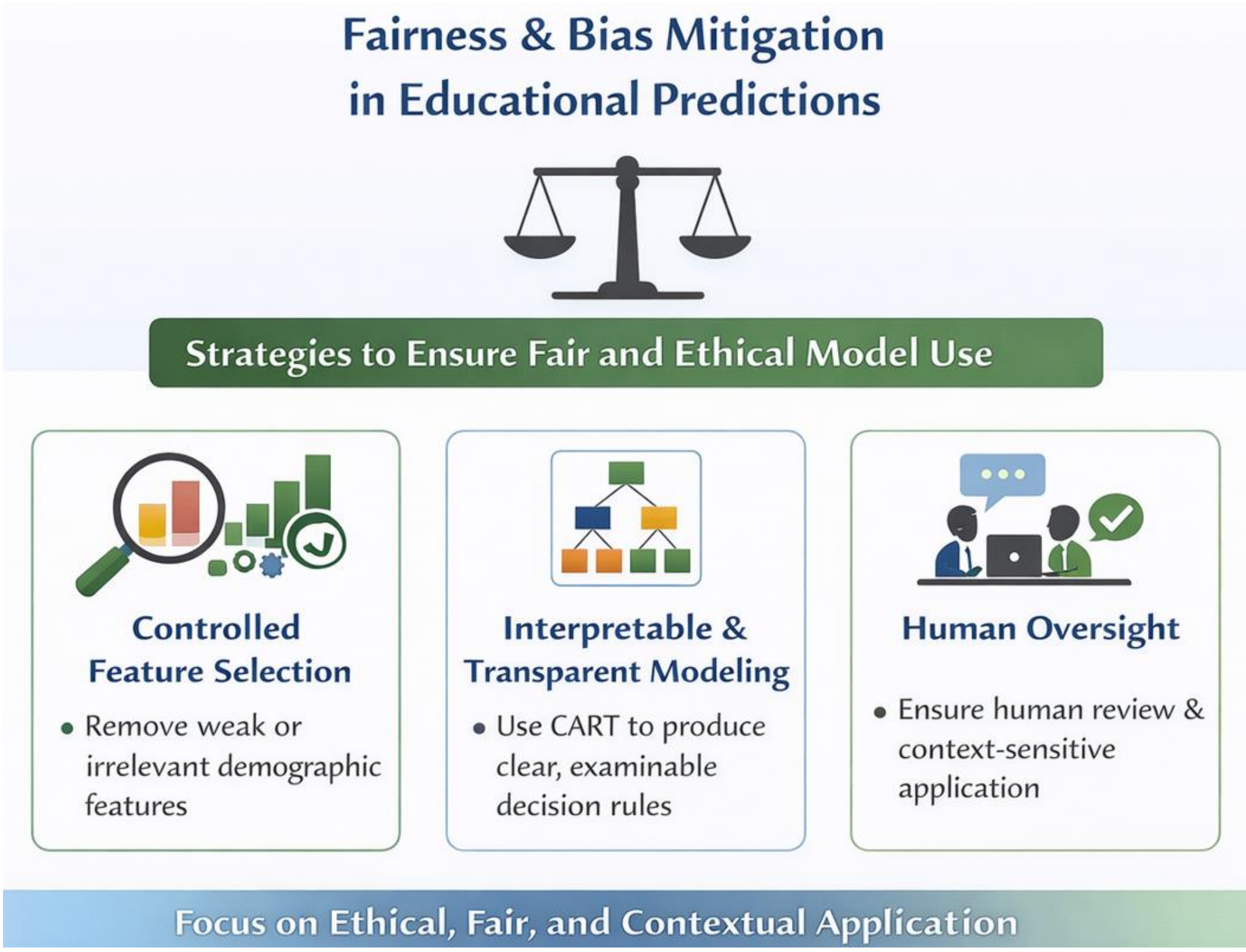

Figure **: Fairness and Bias Mitigation in Educational Predictions

The CART-based predictive framework supports practical and ethical academic advising by generating transparent, rule-based decision paths that can be readily interpreted by guidance counsellors. Each prediction is expressed through a sequence of logical conditions derived from student academic, behavioural, and demographic attributes, enabling counsellors to trace how performance levels and SHS strand recommendations are produced. Rather than serving as an automated decision-maker, the model functions as a decision-support system that provides evidence-based insights to assist counsellors in evaluating student readiness, identifying areas for intervention, and recommending, accountable, and ethical use of machine learning in educational guidance contexts.

DISCUSSION/LIMITATIONS OF THE STUDY

Scope and Data Limitations

This study utilizes a localized dataset drawn from selected public secondary schools to ensure contextual relevance and alignment with actual school-level academic and guidance practices. While the inclusion of data from additional schools could further enhance diversity and generalizability, the localied scope allows the model to capture institution-specific academic patterns, grading practices, and student characteristics. This approach is particularly appropriate for decision-support systems intended for use by guidance counsellors within a defined educational context. Nonetheless, future research is encouraged to extend the dataset across multiple schools or divisions to improve representativeness, assess model transferability, and strengthen external validity.

RESULTS AND DISCUSSION

The pruned CART model demonstrates that interpretable decision rules can be maintained while simultaneously predicting academic performance level and SHS strand suitability, reinforcing the alignment between model transparency and multi-target predictive capability.

REFERENCES

1. Asor, J. R., Catedrilla, G. M. B., Buama, C. A. C., et al. (2022). Prediction of Senior High School Students' Performance in a State University: An Educational Data Mining Approach. International Journal of Innovate Engineering & Technology (IJIET), 13(6). Retrieved from https://www.ijiet.org/vol13/IJIET-V13N6-1888.pdf
2. Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. Machine Learning, 46(1–3), 389–422. https://doi.org/10.1023/A:1012487302797
3. Loh, W.-Y. (2014). Fifty years of classification and regression trees. International Statistical Review, 82(3), 329–348. https://doi.org/10.1111/insr.12016
4. Ravi, K. B. (2017). Cost-Complexity Pruning of Random Forests. arXiv. https://arxiv.org/abs/1703.05430
5. Romero, C., & Ventura, S. (2007). Educational data mining: A survey (1995–2005). Computers & Education, 51(1), 98–112. https://doi.org/10.1016/j.compedu.2007.05.004
6. Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. In Learning Analytics in Education (pp. 1–42). [PDF]. Retrieved from https://bookdown.org/chen/la-manual/files/Romero%20and%20Ventura%20-%202020.pdf
7. Schmid, L., et al. (2023). Tree-based ensembles for multi-output regression. Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2022.120146
8. Vergara, J. R., & Estévez, P. A. (2013). A review of feature selection methods based on mutual information. Neural Computing and Applications, 24(1), 175–186. https://doi.org/10.1007/s00521-013-1368-0
9. "Precision in Progress: Leveraging Data Mining Technique to Empower Career Path Selection for Incoming Senior High School Students." (2024). International Journal of Research and Innovation in Social Science (IJRISS). Retrieved from https://rsisinternational.org/journals/ijriss/articles/precision-in-progress-leveraging-data-mining-technique-to-empower-career-path-selection-for-incoming-senior-high-school-students/
10. Liu, H., & Park, J. (2024). Comparative analysis of decision tree algorithms for academic outcome prediction. Journal of Educational Data Mining, 16(2), 45–63.
11. Gonzales, R., & Lim, M. (2025). Evaluating interpretable decision trees for SHS strand recommendation. Computers & Education, 183, 104585.
12. Raman, S., & Das, A. (2023). Predicting at-risk students using CART and pruning techniques. Education and Information Technologies, 28, 1123–1142.
13. Kim, S., & Velasco, P. (2024). Hybrid feature selection methods for multi-output educational datasets. Expert Systems with Applications, 209, 118235.
14. Ravi, K. B. (2017). Cost-complexity pruning of Random Forests. arXiv:1703.05430.
15. Schmid, L., et al. (2023). Tree-based ensembles for multi-output regression. Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2022.120146
16. Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3), e1355.
17. Simon, H. A. (2020). Rational decision-making in educational systems. Educational Research Review, 30, 100321.
18. Kumar, V., & Singh, K. (2021). Constructivist learning theory in data-driven education. Journal of Educational Technology Systems, 50(1), 3–21.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

19. Li, J., Cheng, K., Wang, S., et al. (2020). Feature selection: A data perspective. ACM Computing Surveys, 52(5), 94.
20. Ahmed, M., & Elaraby, I. (2023). Hybrid feature selection methods for educational data mining. IEEE Access, 11, 33421–33435.