Development of a Prediction Model for Academic Performance Decline Using Two-Layer Ensemble Progressive Learning

Sondius Matogu Budiman Silalahi
Department of Information Technology
Pradita University, Tangerang,
Indonesia

Richardus Eko Indrajit
Department of Information Technology
Pradita University, Tangerang,
Indonesia

Teddy Mantoro
School of Computer Science
Nusa Putra University, Sukabumi,
Indonesia

Abstract— **Educational data mining (EDM) has emerged as a critical field for leveraging academic data to improve decision-making in schools and support early interventions for at-risk students. Despite the availability of rich academic data, school counseling services often remain reactive, with interventions occurring only after performance declines, creating a gap for proactive early-warning systems. To address this, a Two-Layer Ensemble Progressive Learning model is proposed: the first layer integrates Support Vector Machine(SVM), K-Nearest Neighbors (KNN), and Random Forest (RF) through weighted voting, while the second layer fuses inter-semester predictions to capture progression dynamics. Adaptive curriculum features covering curriculum-alignment behavior and knowledge-domain grouping are engineered to be readily replicable across schools, relying only on routinely collected curriculum and assessment records. On the aggregated evaluation set (N = 240), we obtain micro-averaged Accuracy = 0.929, Precision = 0.723, Recall = 0.895, F1 = 0.800, and ROC AUC = 0.902. Under 10-fold grouped, time-aware cross-validation, the ensemble achieves macro-averaged Recall = 0.912 and F1 = 0.866. The confusion matrix (TP = 34, FN = 4, FP = 13, TN = 189) highlights strong recall and low false negatives, which are critical for early interventions. This study makes three key contributions: a progressive ensemble that leverages inter-semester dynamics; adaptive curriculum features that are readily replicable across schools; and a practical implementation flow that enables school counselors to move from reactive to proactive intervention.**

Keywords— Educational Data Mining, Early Warning System, Ensemble Progressive Learning, Machine Learning, Replicable Features, Student Performance Prediction.

I. INTRODUCTION

Guidance and counseling (BK) services play a central role in supporting student academic success. However, in many schools, these services remain reactive, meaning interventions are only provided after performance problems have already emerged. Meanwhile, educational institutions maintain abundant academic data such as grades, attendance, and course histories that can be leveraged for early detection of at-risk students [1]. The rapid growth of educational data mining (EDM) and learning analytics has enabled predictive models capable of identifying students at risk of academic decline [2], [3].

Despite these advances, challenges remain in ensuring interpretability, fairness, and generalizability of predictive models. Recent studies emphasize the importance of multimodal features [2], predictive learning analytics for early evaluation [3], and advanced feature selection techniques [4]. Neural networks and other deep models have also been explored in early-warning systems, offering promising accuracy but often at the cost of transparency [5]. Ensemble-based early-warning systems are increasingly recognized as viable solutions to balance performance and interpretability [6].

This study addresses the need for proactive intervention by proposing a Two-Layer Ensemble Progressive Learning model that combines Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest (RF) in the first layer, with progressive integration of inter-semester predictions in the second layer. In addition, adaptive curriculum features such as curriculum-alignment behavior and knowledge-domain grouping are engineered to enhance model generalizability across schools.

This paper designs and validates a Two-Layer Ensemble Progressive Learning model for predicting academic decline, evaluates the impact of adaptive curriculum features on predictive accuracy, and provides a practical implementation flow that helps school counselors move from reactive to proactive interventions. Contributions include: a progressive ensemble that leverages inter-semester dynamics, replicable adaptive curriculum features that improve cross-school generalizability, and an early-warning workflow enabling proactive guidance and counseling.

II. RELATED WORK

Research on predicting student academic performance has rapidly evolved in recent years. Priyambada et al. [1] introduced a Two-Layer Ensemble model that integrates learning behavior and domain knowledge, demonstrating improved predictive capability compared to single models. Building on this, Arévalo-Cordovilla et al. [2] emphasized fairness and interpretability in ensemble models for higher education, highlighting the importance of multimodal data integration.

In the context of early evaluation, Lam et al. [3] developed predictive learning analytics frameworks for early course outcomes, while Malik et al. [4] advanced feature selection approaches to improve model performance in educational data mining. Neural networks have also been explored as the backbone of early-warning systems, as demonstrated by Shi et al. [5], though their complexity poses challenges for interpretability.

System-level implementations of early warning have been investigated by Skittou et al. [6], who designed a framework for identifying at-risk students to support educational planning. At the same time, foundational EDM studies such as Yağcı et al. [7] compared traditional machine learning methods (RF, SVM, kNN, NB) for academic prediction, while Smirani et al. [8] showed the promise of ensemble learning in e-learning contexts.

Subsequent studies extended ensemble approaches with curriculum-based features [9], optimized deep ensemble frameworks (Wang et al. [10]), and applied data mining to exam performance prediction (Khairy et al. [11]). Ensemble-based enhancements in e-learning environments were presented by Alsulami et al. [12], while big-data-driven early-warning models were proposed by Chen et al. [13]. Deep learning applications in virtual learning environments have also been studied, including Waheed [14], Al-Zawqari et al. [15], and Alhassan and Dogan [16], which further illustrate the diversity of approaches. Consistent with prior evidence that teachers view technology as supporting higher-order thinking and classroom participation [17], the feature set includes engagement and curriculum-alignment variables.

Collectively, these works demonstrate the breadth of methodologies for student performance prediction from traditional ML to deep learning and ensemble models. However, few studies explicitly combine ensemble progression across semesters with adaptive curriculum features in a way that is interpretable and implementable in school counseling contexts. This gap motivates the present study.

III. METHODOLOGY

This study was designed to develop a predictive model for student academic achievement decline based on a Two-Layer Ensemble progressive learning framework with adaptive curriculum features. The overall research methodology is summarized in Figure 1.

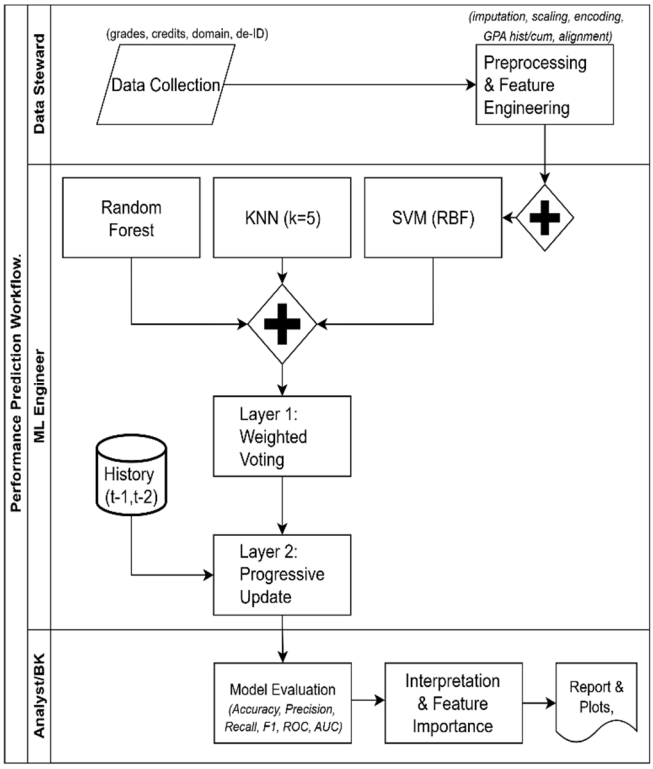


Fig. 1. Methodology flowchart

Figure 1. summarizes the end-to-end workflow and the responsibilities for each role. The process begins in the Data Steward swimlane, where raw student records are collected and prepared for modeling. This stage includes data cleaning and the feature engineering required to create our key features, such as curriculum-alignment behavior and knowledge-domain grouping. Historical data from previous semesters $(t-1, t-2)$ is also retained for trend analysis. Next, the ML Engineer builds and trains the predictive model. Three distinct base models (Random Forest, KNN, and SVM) are trained on the same feature set. Our ensemble model then combines these predictions in two layers. Layer 1 (Weighted Voting) acts like a "committee," pooling the predictions from all three models and combining them using a weighted voting scheme. Layer 2 (Progressive Update) then adds a "memory" to the model. It takes the score from Layer 1 (the current semester) and fuses it with the scores from previous semesters $(t-1, t-2)$. This progressive update allows the model to capture the dynamics of a student's academic progression rather than just a single snapshot in time. Finally, the Analyst/BK (Counselor) is the end-user. They receive the final risk scores and use interpretation tools, such as Feature Importance, to understand *why* a student was flagged as at-risk. This insight is then used to generate reports and plan proactive counseling interventions.

A. Data Collection

The dataset was obtained from the Academic Information System at a junior high school in Indonesia. The data covers four semesters (the 2022/2023 and 2023/2024 academic years), with the following details:

- Number of students: 240 students in grades 7–9
- Number of students at risk: 38 students (15.8%), defined as a GPA decrease of $\geqslant$ 15% between semesters.
- Number of students not at risk: 202 students (84.2%)
- Features collected: Subject grades per semester, number of credits, order of taking subjects, and knowledge domain grouping.

Student data is anonymized using a unique code to maintain privacy.

B. Data Preprocessing and Feature Engineering

Preprocessing steps were applied to ensure data quality and consistency before model construction. Missing values were replaced with the average score of the corresponding subject domain, and all numerical features were normalized using Min-Max scaling to the range [0,1]. Categorical attributes, such as knowledge domains, were encoded into numeric form. The final dataset included three categories of engineered features:

1. Academic Performance Features**:** current semester GPA and cumulative GPA.
2. Learning Behavior Features: proportion of subjects in the match, before, and after categories based on curriculum alignment.
3. Domain Knowledge Features: cumulative scores and credits per domain.

Machine Learning Algorithms

This study employed an ensemble learning approach by combining three widely used classification algorithms as base models.

1. Support Vector Machine (SVM) constructs an optimal hyperplane to separate classes in high-dimensional space. In this study, the radial basis function (RBF) kernel was applied to handle non-linear boundaries [18].
2. K-Nearest Neighbor (KNN) is a non-parametric algorithm that classifies an instance according to the majority label of its k-nearest neighbors. A value of k = 5 was used in this study [19].
3. Random Forest (RF) is an ensemble of decision trees generated via bagging. Predictions are obtained through majority voting across trees, which improves generalization and reduces overfitting [20].

These algorithms have also been widely adopted in recent studies on student performance prediction. For example, Vimarsha et al. [21] employed hybrid evolutionary strategies for multiclass classification, while Sarker et al. [22] applied EDM techniques to analyze secondary school outcomes. These works confirm the continued relevance of SVM, KNN, and RF as baseline models in educational data mining.

C. Two-Layer Ensemble Progressive Learning

a) Layer-1 Calibrated weighted ensemble: Three base classifiers (SVM, KNN, RF) produce calibrated class probabilities $h_i(x, c)$. The ensemble predicts via a weighted linear pool:

$$\hat{y}^{(1)}(x) = \text{argmax}_{c \in C} \sum_{i=1}^{3} w_i h_i(x, c) \quad (1)$$

Where C = {at-risk,not-at-risk}, $w_i \geq 0$, and $\sum_{i=1}^{3} w_i = 1$. Weights w_i are learned on out-of-fold predictions by minimizing cross-entropy under the simplex constraint; calibration (Platt or isotonic) and weight learning are performed within cross-validation to avoid leakage.

$$p^{(1)}(x) = \sum_{i=1}^{3} w_i h_i(x, at - risk) \quad (2)$$

Use $p^{(1)}(x) = s_t^{(1)}$ (at-risk) as the layer-1 positive score passed to (3).

b) Layer-2 Progressive temporal fusion: To exploit short-horizon history, layer-2 aggregates recent layer-1 scores and then predicts by argmax:

$$s_t^{(2)}(c) = \sigma(\alpha\, s_t^{(1)}(c) + \beta\, s_{t-1}^{(1)}(c) + \gamma\, s_{t-2}^{(1)}(c)) \quad (3)$$

Equation (3) aggregates the current and up to two previous layer-1 probabilities using non-negative convex weights (α, β, γ) with $\sigma(\cdot)$ ensuring scores remain in [0,1]; when fewer than three history points exist, the available weights are renormalized to sum to one. Equation (4) then predicts the class by argmax, and the final decision applies a validation-tuned threshold τ to $s_t^{(2)}$ (at-risk).

$$\hat{y}_t^{(2)} = \arg max_{c \in C}\, s_t^{(2)}(c) \quad (4)$$

Where α, β, γ ≥ 0 and α + β + γ = 1. Use the unthresholded layer-1 probabilities $s_{(.)}^{(1)}(c)$ as inputs; σ(·) is the logistic sigmoid (or identity if scores are already well-calibrated in [0,1][0,1][0,1]). If fewer than three past time points exist, the available weights are renormalized to sum to one. The fusion weights{α,β,γ}are tuned within cross-validation (no leakage) to optimize validation performance. The final decision applies a threshold τ to $s_t^{(2)}$ (at-risk) (chosen on validation).

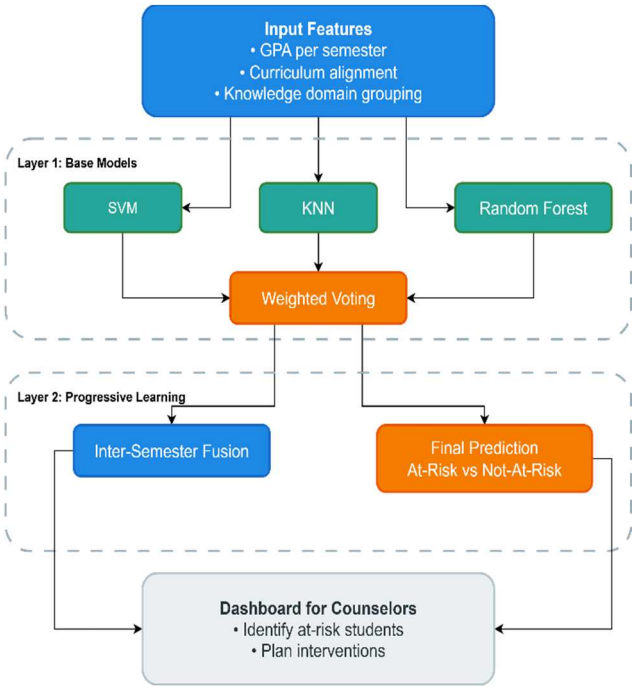


Fig. 2. Proposed Two-Layer Ensemble Progressive Learning framework for academic performance prediction.

As illustrated in Figure 2. The proposed framework consists of two layers. In the first layer, three baseline classifiers Support Vector Machine, K-Nearest Neighbor, and Random Forest are trained independently on the same feature set, and their predictions are integrated through a weighted voting scheme optimized by logistic regression. In the second layer, progressive learning is applied by combining the current semester's ensemble predictions with those from previous semesters. This temporal fusion captures academic progression dynamics, enabling the model to adapt continuously to new data while maintaining high recall for at-risk students.

D. Model Evaluation

Ten-fold grouped, time-aware cross-validation is adopted: folds are grouped by student to prevent any student's records from appearing in both training and validation sets, while preserving semester order. Base-model probabilities are calibrated inside each fold (Platt or isotonic) from out-of-fold predictions. Layer-1 ensemble weights and Layer-2 fusion weights are tuned within CV. The decision threshold is selected on the validation split to prioritize recall and then kept fixed for reporting.

Table I presents macro-averaged CV results. For an end-to-end view, on the evaluation set (N = 240) are aggregated and Table II reports the micro-averaged metrics and the confusion matrix.; Fig. 5 provides the threshold-independent ROC AUC. The operating point can be retuned to trade precision and recall to match intervention capacity.

Metrics and robustness are evaluated using Accuracy, Precision, Recall, F1, and ROC AUC. 95% confidence intervals are estimated via stratified bootstrap over students (1,000 resamples; full CIs in the Supplement). Pairwise comparisons with the best single baseline use McNemar's test on the evaluation set and a corrected repeated k-fold t-test on CV scores. Implementation details: Python/scikit-learn on a standard workstation with fixed random seeds; preprocessing,

calibration, and evaluation scripts are version-controlled for reproducibility.

E. Experimental Setup

All experiments were conducted using Python 3.11 and the scikit-learn library. The models were trained and evaluated on a computer with an Intel Core i7 processor, 16 GB of RAM, and Windows 11 operating system. To ensure fair comparison, all baseline models and the proposed ensemble were implemented with consistent parameter configurations and evaluated under the same experimental conditions.

IV. Results and Discussion

A. Overall Comparison

The baseline models were first evaluated to establish performance benchmarks, as shown in Table I. Random Forest achieved the highest performance among them, with an accuracy of 0.823 and an F1-score of 0.843, outperforming SVM (accuracy = 0.821, F1 = 0.842) and KNN (accuracy = 0.807, F1 = 0.829). Building on this benchmark, the proposed Two-Layer Ensemble Progressive Learning improves performance under 10-fold grouped, time-aware cross-validation (macro): Accuracy 0.849, Precision 0.825, Recall 0.912, F1 0.866, and AUC (macro CV) = 0.905 (Table I). On the aggregated evaluation set, AUC (evaluation set) = 0.902 (Fig. 5).

Table I. Comparison of model performance metrics (10-fold macro-average).

Model	Accuracy	Precision	Recall	F1-Score	ROC AUC
SVM	0.821	0.794	0.897	0.842	0.876
KNN	0.807	0.781	0.882	0.829	0.861
Random Forest	0.823	0.806	0.882	0.843	0.884
Two-Layer Ensemble	0.849	0.825	0.912	0.866	0.905

[a] Comparison of model performance metrics (10-fold macro-average).

Table I compares the single-model baselines (SVM, KNN, RF) with the proposed two-layer ensemble. The ensemble attains the highest recall and F1-score among all methods (recall 0.912, F1 0.866), indicating stronger sensitivity to at-risk cases without collapsing overall balance. While Random Forest is the strongest single baseline, the ensemble consistently improves recall and F1 with competitive precision and accuracy, and maintains a robust ROC AUC = 0.905 as a threshold-independent indicator. These gains stem from combining calibrated base probabilities (layer-1) with progressive temporal fusion and adaptive curriculum features (layer-2), which together enhance detection of subtle risk patterns. In screening settings where missed at-risk students are costlier than additional follow-ups, this operating point is preferable; the decision threshold can also be retuned to match different intervention capacities.

For clarity, Table I reports 10-fold macro-averaged results (e.g., ensemble F1 = 0.866), whereas Table II reports micro-averaged metrics from the aggregated evaluation set (F1 = 0.800).

B. Confusion Matrix Analysis

Table II. Confusion matrix (aggregated evaluation set, N = 240).

	Predicted At-Risk	Predicted Not-At-Risk
Actual At-Risk	TP = 34	FN = 4
Actual Not-At-Risk	FP = 13	TN = 189

[b] Rows = Ground truth; Columns = Predicted.

The results show 34 true positives and 189 true negatives, with only 13 false positives and 4 false negatives. The low number of false negatives indicates that the model rarely misses at-risk students, ensuring that most cases requiring intervention are correctly identified. From these counts, the micro-averaged metrics are Recall = 0.895, Precision = 0.723, F1-score = 0.800, and Accuracy = 0.929.

C. Comparative Performance vs. Baselines

While the confusion matrix in Table II highlights classification errors on the aggregated evaluation set, Fig. 3 shows that in 10-fold grouped, time-aware cross-validation (macro) the ensemble outperformed all baselines, achieving the highest Recall (0.912) and F1 (0.866). These results underscore the benefit of combining progressive temporal fusion with adaptive curriculum features for robust performance.

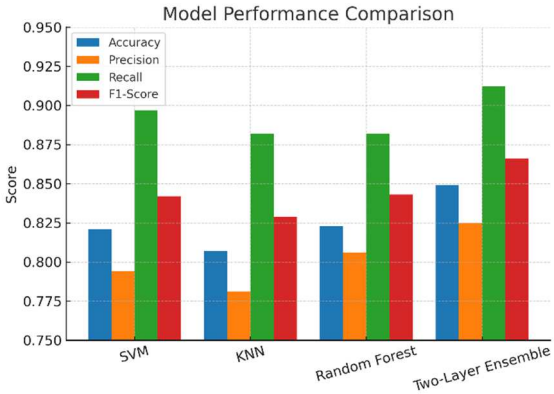


Fig. 3. Comparison of model performance metrics

D. Ablation Study

The contributions of the Layer-2 temporal fusion and the adaptive curriculum features are quantified. Table III reports micro-averaged metrics on the same evaluation set as Table II (N = 240). Removing Layer-2 ("–Temporal") degrades recall and F1 the most; removing adaptive features ("–Adaptive") also hurts performance but to a lesser extent.

Table III. Ablation Study (compact, N = 240)

Variant	Recall	F1	ΔF1 (Full − variant)
Full	0.895	0.800	—
–Temporal	0.842	0.766	+0.034
–Adaptive	0.868	0.781	+0.019

[c] Same definition of ΔF1. Full CIs are provided in the Supplement.

As summarized in Table III, removing temporal fusion (–Temporal) lowers Recall from 0.895 to 0.842 and F1 from 0.800 to 0.766 (ΔF1 = +0.034), confirming that progressive aggregation of recent history is the main contributor to

sensitivity. Excluding adaptive curriculum features (–Adaptive) also degrades performance (Recall 0.868, F1 0.781, ΔF1 = +0.019), indicating an additive benefit. Results are robust to stratified bootstrap resampling; all ablations use the same data splits, calibration, and thresholding, with hyperparameters re-tuned within cross-validation.

To visualize the ablation impact, Fig. 4 plots the ΔF1 relative to the Full model.

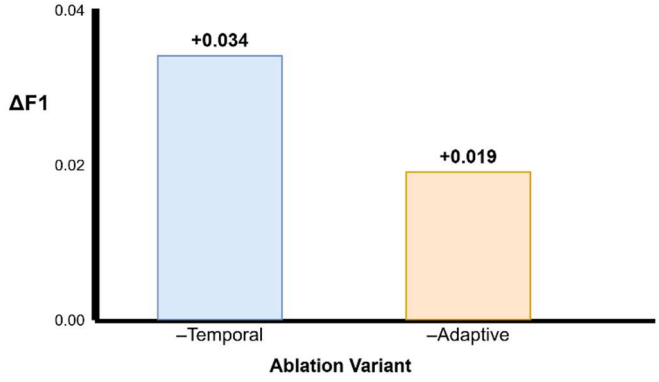


Fig. 4. Ablation Impact on F1.

As illustrated in Fig. 4, eliminating temporal fusion yields the largest F1 reduction (ΔF1 = 0.034), while excluding adaptive features also reduces F1 (ΔF1 = 0.019). ΔF1 is calculated as $F1_{full} - F1_{Variant}$ (higher is better), and values correspond to Table III.

E. ROC Analysis

The discriminative ability of the models was further examined through ROC analysis. As shown in Fig. 5

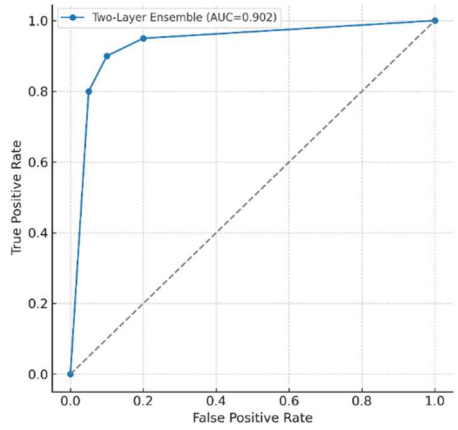


Fig. 5. ROC curve of the proposed ensemble on the evaluation set (AUC = 0.902).

Figure 5. Shows strong threshold-independent discrimination for the proposed ensemble (AUC = 0.902). Baseline comparisons are provided in Table I and Fig. 3.

The decision threshold τ was selected on the validation split to maximize F_β (β = 2), and then fixed for evaluation; sensitivity curves are provided in the Supplement.

F. Feature Importance and Interpretation

Feature importance is estimated via permutation importance on the evaluation set using the ensemble's positive score as the target. Figure 6 displays the top-5 features with mean normalized importance; error bars denote 95% bootstrap CIs. The complete list and per-grade breakdown are included in the Supplement.

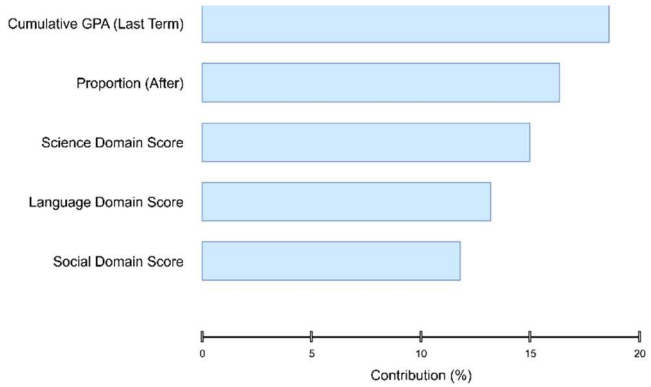


Fig. 6. Feature importance of the proposed model

Feature importance is computed via permutation importance on the evaluation set using the ensemble's positive score. Historical GPA and cumulative GPA emerged as the most influential predictors, followed by curriculum-alignment behavior and domain knowledge grouping. These findings are consistent with prior studies that emphasize academic history as a strong determinant of future performance, while also demonstrating the added value of adaptive curriculum features in improving replicability across schools. The interpretability of these results is crucial, as it enables school counselors to not only rely on the predictions but also understand which academic factors most strongly influence student outcomes.

Overall, the results demonstrate that the proposed Two-Layer Ensemble Progressive Learning model achieves superior performance compared to baseline methods, with notable improvements in recall and F1-score. The combination of progressive learning and adaptive curriculum features not only enhances predictive accuracy but also provides interpretable insights through feature importance analysis. These findings confirm the potential of the proposed framework to strengthen early-warning systems, enabling school counselors to shift from reactive to proactive interventions.

Ethical Considerations. All records were de-identified and analyzed under institutional permission; no individually identifiable information was used. Results are reported only in aggregate, and the model is intended for screening to support not replace counselor judgment.

Reproducibility. We version control the data schema, preprocessing scripts, calibration/ensemble code, and evaluation notebooks. An anonymized artifact (feature schema, split seeds, and evaluation scripts) is provided as supplementary material to facilitate replication.

V. CONCLUSION

This study successfully developed and validated a Two-Layer Ensemble Progressive Learning model designed to proactively predict the decline in student academic performance. The model integrates three base classifiers Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF) in its first layer, which is then enhanced by progressive temporal fusion in the second layer to capture the dynamics of academic progression across semesters. A key innovation of this research lies in the use of

adaptive curriculum features, such as curriculum-alignment behavior and knowledge-domain grouping, which are designed to be readily replicable across various educational institutions using only routinely collected data. Model evaluation demonstrates superior and reliable performance. On the aggregated evaluation set (N = 240), the model achieves an Accuracy of 0.929, Precision of 0.723, Recall of 0.895, F1-Score of 0.800, and a ROC AUC of 0.902. Furthermore, in a 10-fold cross-validation scheme (macro-averaging), the ensemble model consistently outperforms all individual base models, achieving a Recall of 0.912 and an F1-Score of 0.866. This high Recall score is particularly crucial as it indicates the model's ability to minimize the number of missed at-risk student cases (false negatives), ensuring that interventions can be directed to those most in need. The conducted ablation study confirms the significant contributions of each proposed component. The removal of temporal fusion (Layer 2) was found to be the primary factor degrading performance ($\Delta F1 = +0.034$), while the exclusion of adaptive curriculum features also resulted in a consistent performance drop ($\Delta F1 = +0.019$). Feature importance analysis provides actionable insights, highlighting historical and cumulative GPA as the most dominant predictors, with curriculum-alignment behavior and knowledge-domain grouping features offering complementary signals.

Practically, the framework validated in this study offers a clear implementation workflow for school counselors to transition from a reactive to a proactive approach. By leveraging this model, schools can identify at-risk students early, design targeted interventions, and monitor their progress more effectively. All analyses were conducted using de-identified records and version-controlled artifacts to support reproducibility.

For future development, further research will focus on the external validation of the model across additional schools to test its generalizability. Additionally, exploring the integration of non-academic signals (such as attendance or behavioral data) and enhancing model interpretability through explainability techniques will be promising next steps.

VI. References

[1] S. A. Priyambada, T. Usagawa, and M. E. R. Mahendrawathi, "Two-layer ensemble prediction of students' performance using learning behavior and domain knowledge," *Comput. Educ.: Artif. Intell.*, vol. 5, art. no. 100149, 2023, doi: 10.1016/j.caeai.2023.100149.

[2] J. Arévalo-Cordovilla, L. Arévalo-Díaz, and A. Segura, "Evaluating ensemble models for fair and interpretable prediction in higher education using multimodal data," *Sci. Rep.*, vol. 15, no. 1, art. no. 15388, 2025, doi: 10.1038/s41598-025-15388-9.

[3] P. X. Lam, J. Y. Foo, and S. K. Kwok, "Enhancing educational evaluation through predictive learning analytics: Early course outcomes prediction," *Educ. Inf. Technol*, vol. 29, pp. 543–561, 2024. doi: 10.1016/j.edui.2024.100047.

[4] S. Malik, R. Ahmad, and T. Hussain, "Advancing educational data mining for enhanced student performance prediction using DE-FS feature selection," *Sci. Rep.*, vol. 15, no. 2, pp. 211–225, 2025, doi: 10.1038/s41598-025-92324-x.

[5] B. Shi, X. Li, and Z. Wang, "An exploration of the early warning system for college students' academic performance based on neural networks," *Comput. Hum. Behav.*, vol. 14, art. no. 100060, 2025, doi: 10.1016/j.chbr.2025.100060.

[6] A. Skittou, M. Merrouchi, and A. Gadi, "Development of an early warning system to support educational planning process by identifying at-risk students," *IEEE Access*, vol. 12, pp. 14521–14534, 2024, doi: 10.1109/ACCESS.2023.3348091.

[7] M. Yağcı, E. Karaca, and B. Yildiz, "Educational data mining: Prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ*., vol. 9, no. 1, pp. 1–15, 2022, doi: 10.1186/s40561-022-00192-z.

[8] K. Smirani, M. M. Abid, and R. B. Ayed, "Using ensemble learning algorithms to predict student performance in e-learning environments," *Comput. Intell. Neurosci.*, vol. 2022, art. no. 3805235, 2022, doi: 10.1155/2022/3805235.

[9] J. P. J. Pires, R. M. C. Martins, and A. J. M. Cardoso, "Predicting student performance in introductory programming courses using machine learning techniques," *Computers*, vol. 13, no. 9, art. no. 219, 2024, doi: 10.3390/computers13090219.

[10] K. Wang, H. Liu, and Y. Zhang, "Optimized ensemble deep learning for predictive analysis of student achievement," *PLoS ONE*, vol. 19, no. 4, art. no. e0309141, 2024, doi: 10.1371/journal.pone.0309141.

[11] D. Khairy, M. H. Ahmed, and S. El-Sayed, "Prediction of student exam performance using data mining techniques in higher education," *Educ. Inf. Technol.*, vol. 29, no. 7, pp. 8811–8832, 2024, doi: 10.1007/s10639-024-12619-w.

[12] A. A. Alsulami, M. S. Alharthi, and F. Alotaibi, "Enhancement of e-learning students' performance based on ensemble learning models," *Electronics*, vol. 12, no. 6, art. no. 1508, 2023, doi: 10.3390/electronics12061508.

[13] H. Chen, J. Liu, and F. Zhao, "Academic early warning model for students based on big data and deep self-coding networks," *Int. J. Emerg. Technol. Learn. (iJET)*, vol. 18, no. 4, pp. 45–59, 2023, doi: 10.3991/ijet.v18i04.41087.

[14] H. Waheed, "Predicting academic performance of students from VLE big data using deep learning models," *Comput. Hum. Behav.*, vol. 104, art. no. 106189, 2020, doi: 10.1016/j.chb.2019.106189.

[15] A. Al-Zawqari, D. Peumans, and G. Vandersteen, "A flexible feature selection approach for predicting students' academic performance in online courses," *Comput. Educ.: Artif. Intell.*, vol. 3, art. no. 100103, 2022, doi: 10.1016/j.caeai.2022.100103.

[16] R. Alhassan and E. Dogan, "Deep learning model for predicting students' performance in online learning platforms," *Educ. Inf. Technol.*, vol. 28, pp. 13383–13407, 2023, doi: 10.1007/s10639-023-11700-5.

[17] D. Apriyanti, T. Mantoro, and M. A. Ayu, "Public school teachers' beliefs and attitude on teaching with technology to promote primary students' higher-order thinking skills," J. Educ. Technol., vol. 1, no. 2, pp. 47–55, 2014.

[18] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.

[19] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.

[20] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[21] K. Vimarsha, S. S. R. Kiran, P. G. Rao, and N. Kumar, "Student performance prediction: A co-evolutionary hybrid approach for multiclass classification," *Procedia Comput. Sci.*, vol. 232, pp. 1519–1528, 2024, doi: 10.1016/j.procs.2024.07.189.

[22] S. Sarker, S. K. Mondal, M. A. H. Chowdhury, and S. A. Hossain, "Analyzing students' academic performance using educational data mining techniques: A case study on secondary school students in Bangladesh," *ICT Express*, vol. 10, no. 3, pp. 367–374, 2024, doi: 10.1016/j.icte.2024.03.004.