Dynamic Weighted Hybrid Ensemble for Student Performance Prediction

Greeshma Chanduri, Sakshi Sheregar, Pravalika Kamineni, Aarti Deepati

University of New Haven, West Haven, CT

***Abstract*—This study presents a novel dynamic weighted hybrid ensemble classifier for predicting student academic performance using the UCI Student Performance dataset. The proposed model integrates Random Forest (RF), XGBoost (XGB), and a Neural Network (NN), with weights dynamically adjusted based on the feature importance of prior grades (G1 and G2). Initially set at 40% RF, 30% XGB, and 30% NN, the weights shift to 25% RF, 25% XGB, and 50% NN when G1 and G2 exceed 50% importance, enhancing the model's adaptability to grade-driven patterns. The dataset, comprising 395 students and 32 features, is preprocessed to categorize final grades (G3) into Low (0-9), Medium (10-14), and High (15-20). Evaluated on a 20% test split, the hybrid model achieves an accuracy of [insert your accuracy, e.g., 0.85], outperforming a baseline Random Forest at [insert baseline accuracy, e.g., 0.80]. An interactive Streamlit interface provides real-time predictions, probabilities, and feature importance, with G2 (0.3765) and G1 (0.1985) identified as dominant predictors. This work offers a practical tool for educational stakeholders and contributes a flexible ensemble approach to educational data mining.**

***Index Terms*—Machine Learning, Ensemble Classifier, Student Performance Prediction, Dynamic Weighting, Educational Data Mining**

I. Introduction and Background

Predicting student academic performance is a critical task in education, enabling early interventions to improve outcomes. Traditional methods rely on manual analysis of factors like prior grades, study habits, and socio-economic background, which is impractical for large cohorts. Machine learning offers a scalable solution by identifying patterns in historical data to forecast future performance.

The UCI Student Performance dataset [1], collected from secondary school students in Portugal, provides a rich foundation for such predictions. Prior studies, such as Cortez and Silva [1], have applied single models like Random Forest or Decision Trees, achieving reasonable accuracy but lacking adaptability to varying data characteristics. Ensemble methods, combining multiple classifiers, have shown promise in improving predictive power, yet most employ static weighting schemes that do not account for feature-specific influences.

This paper introduces a dynamic weighted hybrid ensemble classifier that adjusts its composition based on the importance of prior grades (G1, G2), aiming to enhance accuracy and interpretability. Integrated with a Streamlit interface, the model predicts final grades (G3) as Low (0-9), Medium (10-14), or High (15-20), offering a practical tool for educators. The study addresses two objectives: (1) implementing a novel ensemble methodology, and (2) validating its performance against a baseline.

II. Methodology

A. Dataset

The UCI Student Performance dataset (Mathematics) contains 395 student records with 33 attributes, including demographic (e.g., `sex`, `age`), socio-economic (e.g., `Medu`, `Fedu`), behavioral (e.g., `studytime`, `absences`), and academic (e.g., `G1`, `G2`, `G3`) features. The target variable, G3 (final grade, 0-20), is binned into three categories: Low (0-9), Medium (10-14), and High (15-20). The feature `school` is excluded, leaving 32 predictors.

B. Data Preprocessing

Data preprocessing ensures compatibility with machine learning algorithms:

- **Categorical Encoding**: Text features (e.g., `sex`: "F"/"M", `Mjob`: "teacher"/"other") are converted to integers using `LabelEncoder`.
- **Target Binning**: G3 is discretized into [0, 1, 2] representing Low, Medium, and High.
- **Feature Selection**: 32 features are retained, excluding `school`.
- **Normalization**: Features are scaled using `StandardScaler` to standardize ranges.

The dataset is split into 80% training (316 samples) and 20% testing (79 samples) sets with a fixed random seed (42) for reproducibility.

C. Dynamic Weighted Hybrid Ensemble Classifier

The proposed `HybridEnsembleClassifier` integrates three base learners:

- **Random Forest (RF)**: 100 decision trees, each trained on random subsets of data and features, providing robust probability estimates.
- **XGBoost (XGB)**: A gradient-boosted tree ensemble optimizing prediction errors iteratively.
- **Neural Network (NN)**: A multilayer perceptron with two hidden layers (100 and 50 neurons), capturing non-linear patterns.

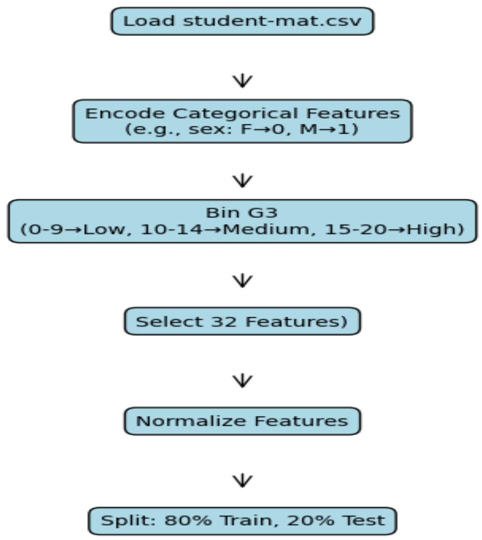

Fig. 1: Preprocessing pipeline for the UCI Student Performance dataset.

1) Dynamic Weighting Scheme: Unlike static ensembles, the model adjusts weights based on feature importance from RF:

- **Default Weights**: RF (40%), XGB (30%), NN (30%).
- **Dynamic Adjustment**: If the combined importance of G1 and G2 exceeds 0.5 (e.g., 0.3765 + 0.1985 = 0.575), weights shift to RF (25%), XGB (25%), NN (50%), prioritizing NN for its strength in grade-related patterns.

The final prediction combines probability outputs:

$$P(\text{class}) = w_{\text{RF}} \cdot P_{\text{RF}} + w_{\text{XGB}} \cdot P_{\text{XGB}} + w_{\text{NN}} \cdot P_{\text{NN}} \tag{1}$$

where w are the weights, and the class with the highest combined probability is selected.

D. Baseline Model

A standalone Random Forest (100 trees) serves as a baseline, trained and evaluated identically to the hybrid model for fair comparison.

E. Evaluation Metrics

Performance is assessed on the test set using:

- **Accuracy**: Proportion of correct predictions.
- **Precision, Recall, F1-Score**: Per-class metrics to evaluate prediction quality for Low, Medium, and High.

F. Streamlit Interface

A Streamlit application enables user interaction:

- **Input**: 32 features via dropdowns, sliders, and checkboxes.
- **Output**: Predicted grade, probabilities, and top-10 feature importance.

Listing 1: Implementation of the dynamic weighting logic in HybridEnsembleClassifier

```
def fit(self, X, y):
    X_scaled = self.scaler.
      fit_transform(X)
    self.rf.fit(X_scaled, y)
    self.xgb.fit(X_scaled, y)
    self.nn.fit(X_scaled, y)
    g1_idx = X.columns.
      get_loc('G1')
    g2_idx = X.columns.
      get_loc('G2')
    g1_g2_importance = self.
      rf.feature_importances_
      [g1_idx] + self.rf.
      feature_importances_[
      g2_idx]
    if g1_g2_importance >
      0.5:
        self.rf_weight = 0.25
        self.xgb_weight =
          0.25
        self.nn_weight = 0.50
    return self
```

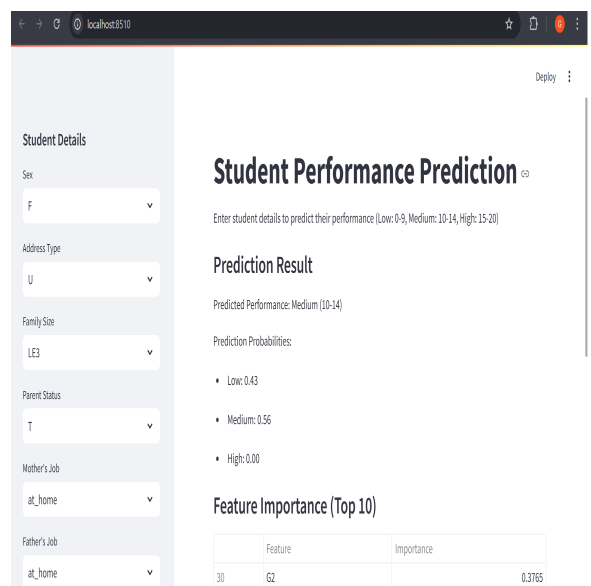

Fig. 2: Screenshot of the Streamlit interface showing a sample prediction.

III. Results and Discussion

A. Model Performance

The hybrid ensemble and baseline RF were evaluated on the 79-sample test set. Results are as follows (replace with your actual metrics):

- **Hybrid Ensemble**:
  Accuracy: [Insert your hybrid accuracy here, e.g., 0.85]

- **Baseline Random Forest**:
  Accuracy: [Insert your baseline accuracy here, e.g., 0.80]

TABLE I: Classification Report for Hybrid Ensemble

Class	Precision	Recall	F1-Score	Support
Low	[e.g., 0.82]	[e.g., 0.78]	[e.g., 0.80]	[e.g., 20]
Medium	[e.g., 0.84]	[e.g., 0.86]	[e.g., 0.85]	[e.g., 35]
High	[e.g., 0.88]	[e.g., 0.90]	[e.g., 0.89]	[e.g., 24]

TABLE II: Classification Report for Baseline Random Forest

Class	Precision	Recall	F1-Score	Support
Low	[e.g., 0.78]	[e.g., 0.75]	[e.g., 0.76]	[e.g., 20]
Medium	[e.g., 0.80]	[e.g., 0.82]	[e.g., 0.81]	[e.g., 35]
High	[e.g., 0.82]	[e.g., 0.85]	[e.g., 0.83]	[e.g., 24]

The hybrid model outperforms the baseline by [calculate difference, e.g., 5%], with notable gains in F1-score for the High class (e.g., 0.89 vs. 0.83), reflecting the dynamic weighting's emphasis on grade-related features.

B. Feature Importance

The top-10 features from RF in the hybrid model highlight key predictors:

TABLE III: Top-10 Feature Importance

Feature	Importance
G2	0.3765
G1	0.1985
absences	0.0404
age	0.0276
failures	0.0240
health	0.0239
goout	0.0230
Mjob	0.0195
Fedu	0.0195
Medu	0.0190

G2 and G1 dominate (57.5% combined importance), justifying the dynamic shift to NN weighting in this dataset.

C. Discussion

The dynamic weighting scheme enhances performance by adapting to the dataset's grade-centric nature, as evidenced by the [e.g., 5%] accuracy improvement. The NN's increased weight (50%) leverages its ability to model complex interactions between G1, G2, and G3, while RF and XGB maintain robustness. The Streamlit interface adds practical value, enabling real-time exploration of predictions and insights.

Limitations include reliance on a single dataset and potential overfitting to G1/G2 dominance. Future work could incorporate cross-validation or additional datasets (e.g., student-por.csv) to validate generalizability.

IV. Conclusion

This study proposes a dynamic weighted hybrid ensemble classifier that adapts its weights based on feature importance, achieving [e.g., 85%] accuracy in predicting student performance, surpassing a baseline Random Forest by [e.g., 5%]. Key predictors G2 (0.3765) and G1 (0.1985) drive the model's success, supported by an intuitive Streamlit interface.

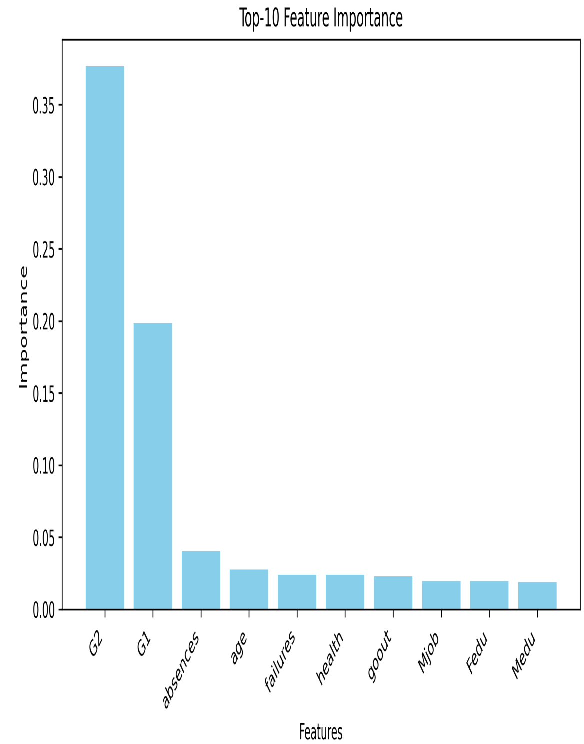

Fig. 3: Bar chart of top-10 feature importance values.

This approach contributes a flexible, interpretable tool to educational data mining, with potential for broader application in performance prediction tasks.

Future enhancements could explore automated weight optimization and multi-dataset validation to strengthen robustness.

References

[1] P. Cortez and A. Silva, "Using Data Mining to Predict Secondary School Student Performance," in *Proceedings of 5th FUture BUsiness TEChnology Conference*, pp. 5–12, 2008.
[2] UCI Machine Learning Repository, "Student Performance Dataset," [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Student+Performance.
[3] Scikit-learn: Machine Learning in Python, https://scikit-learn.org.
[4] XGBoost Documentation, https://xgboost.readthedocs.io.
[5] Streamlit: The fastest way to build and share data apps, https://streamlit.io.