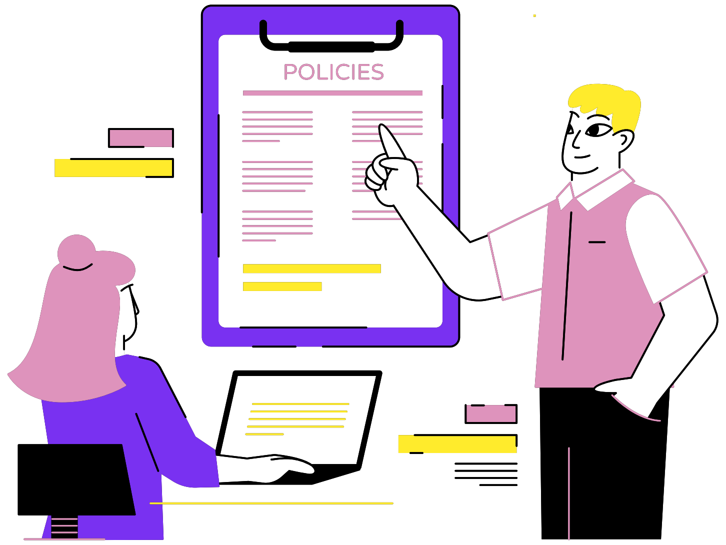


Decoding Instructor Performance with Data Mining and Machine Learning

Done By-
Sai Gopal Lanka - AP21110010042 - 0%
Namratha Addagada - AP21110010048
Haasitha Ambati - AP21110010049

Agenda

- Abstract
- Introduction
- Objectives
- Dataset Overview
    - Datasets Used
    - Data Preprocessing
    - Machine Learning Models applied on datasets
    - Cross-dataset Methodology Evaluation
- Website:KnowMyProf
- Conclusion

Abstract

Educational Data Mining is a growing subject that tries to enhance education quality and address administrative issues, such as predicting instructor effectiveness. This paper evaluates instructor performance using three datasets: Turkish Student Evaluation Dataset, Poland University Dataset, and Synthetic Dataset. Feature selection techniques such as recursive feature removal and Random Forest Classifier are utilized. The databases include Student Evaluation of Teaching survey results. Various classification techniques, including Decision Tree, Naive Bayes, and Multi Class Logistic Regression, have been employed to predict instructor performance.

Introduction

Evaluating instructor performance is crucial for maintaining educational quality. Our project leverages data mining and machine learning techniques to analyze student feedback and predict instructor performance. By utilizing real-world and synthetic datasets, we explore key factors like teaching effectiveness, student engagement, and course delivery.

We applied advanced machine learning models, including boosting, feature selection techniques, and classification algorithms, to improve accuracy and reliability. The insights from this study help institutions assess teaching quality, improve faculty development, and enhance student learning experiences.
Moving forward, we aim to develop a professor rating system based on these predictions, enabling students to make informed decisions while choosing courses and instructors.

Objectives

- **Apply Data Mining for Preprocessing:**
  - Clean and prepare data using techniques like feature selection, handling null values, and SMOTE.
- **Predict Instructor Performance:**
  - Use machine learning models to accurately evaluate instructor effectiveness based on student feedback.
- **Implement 3×3 Model-to-Dataset Comparison:**
  - Evaluate each machine learning method on all three datasets to identify the best-performing model-dataset combinations.
- **Develop a Professor Rating Webpage:**
  - Create a user-friendly webpage to display instructor ratings and feedback insights, helping students make better course choices.

This project aims to revolutionize instructor evaluation, making it more objective, reliable, and useful for both institutions and students.

Dataset Overview

Datasets **USED**

Turkiye Student Evaluation Dataset

- 5820 records, 33 columns.
- 28 questions assess instructor effectiveness.
- Data is numerical (1-5 scale) with no missing values.

Poland University Dataset

- 8015 records.
- 9 student feedback questions (1-5 scale).
- Additional features: instructor seniority, gender, pass rate, average SET score.

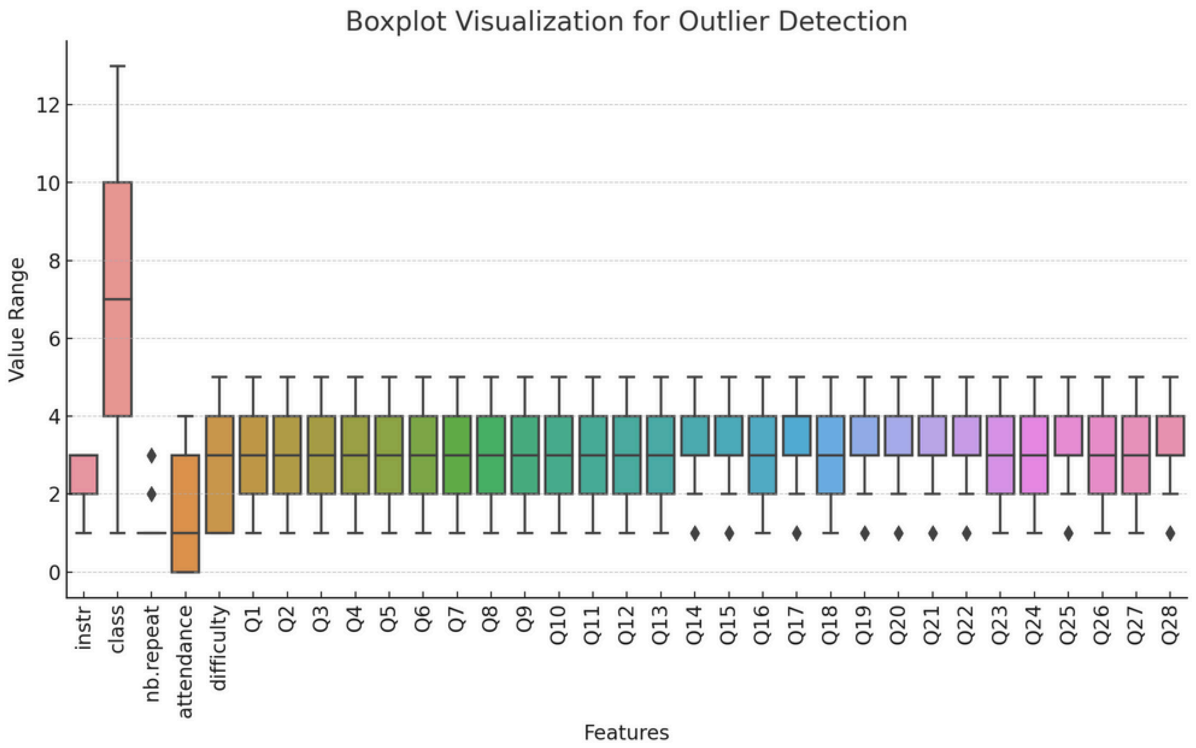
Synthetic Dataset

- Generated Data with 10,000 records.
- 11 evaluation questions based on academic surveys.
- Responses generated randomly for ML model analysis.

Data Preprocessing

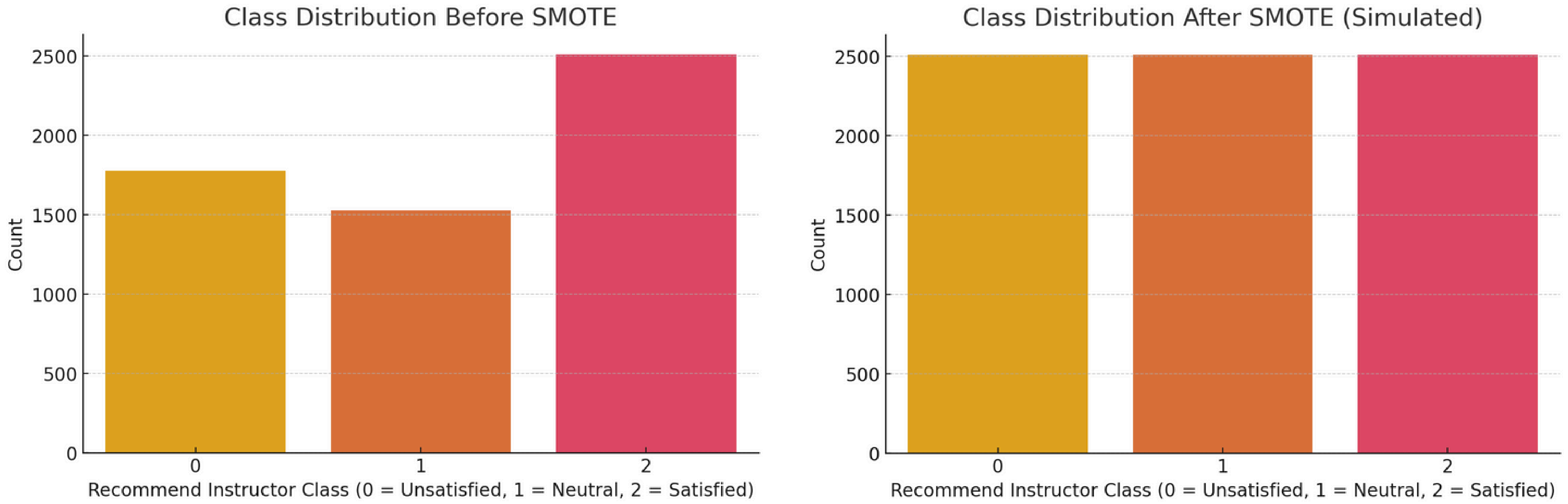
1.For Turkiye Studet Evaluation Dataset

Boxplot for Outliers Detection



- Checked for null values and outliers using boxplots.
- Dropped "nb.repeat" (constant values across data).
- Outlier removal caused data loss (1000+ points), so the original dataset was retained.
- Used before applying threshold-based outlier removal.

Distribution Plot of Satisfaction Index (Before & After SMOTE)



Before SMOTE:

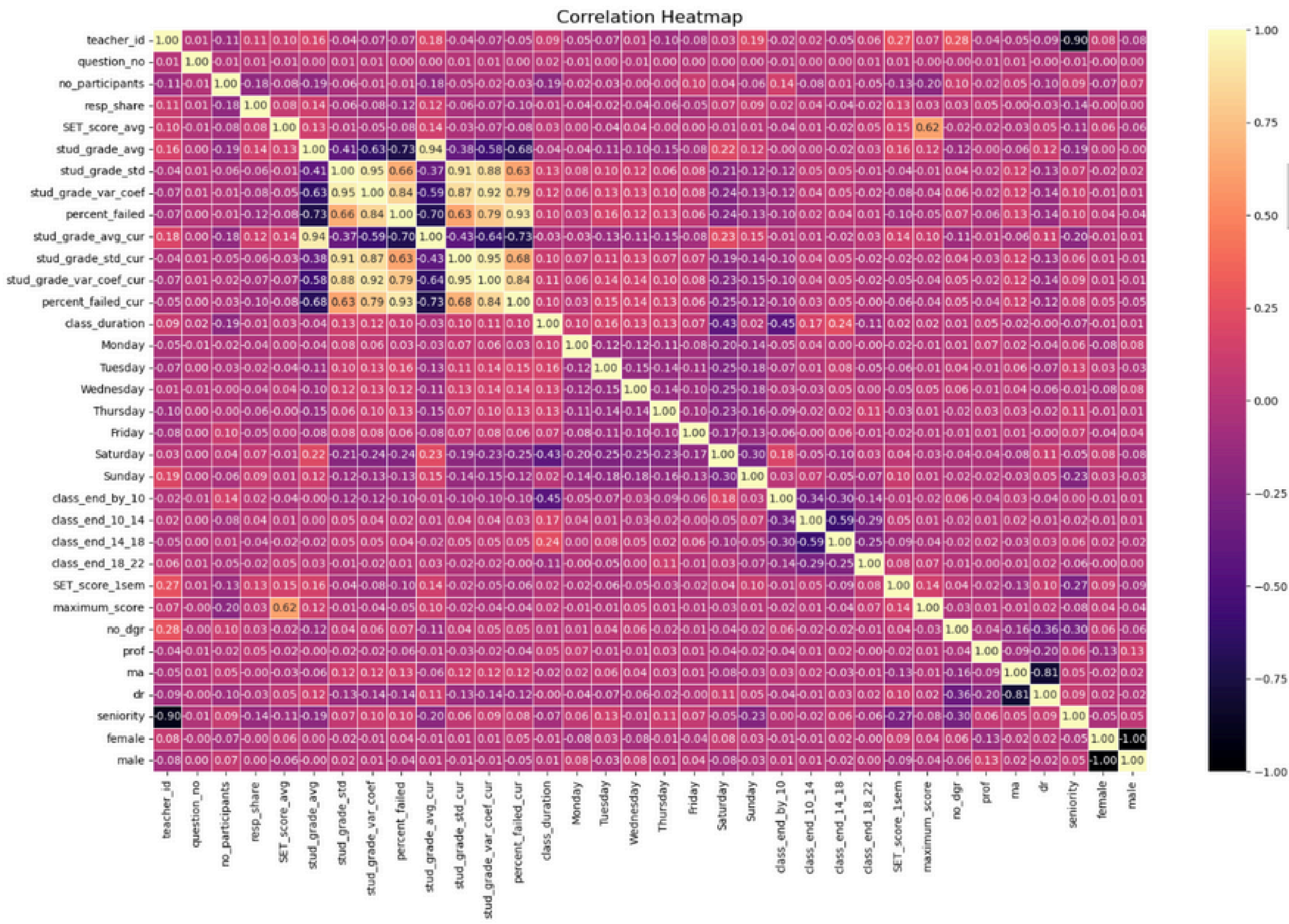
- The dataset shows class imbalance, with "Satisfied" instructors (Class 2) having the highest count.
- This imbalance can lead to biased model predictions.

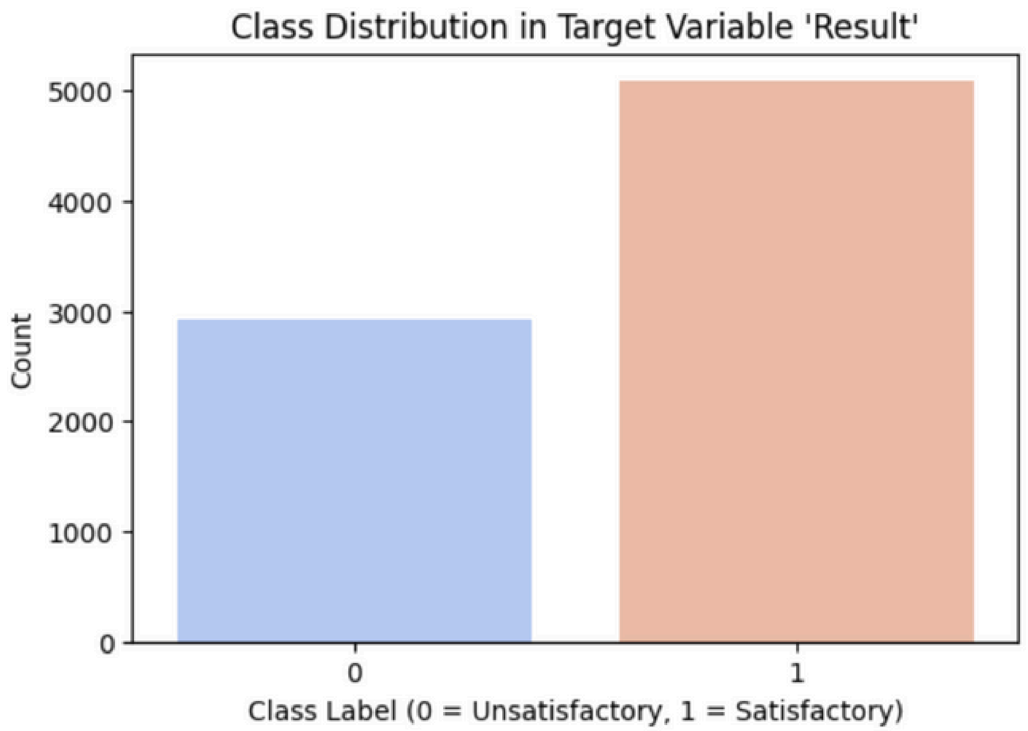
After SMOTE (Simulated):

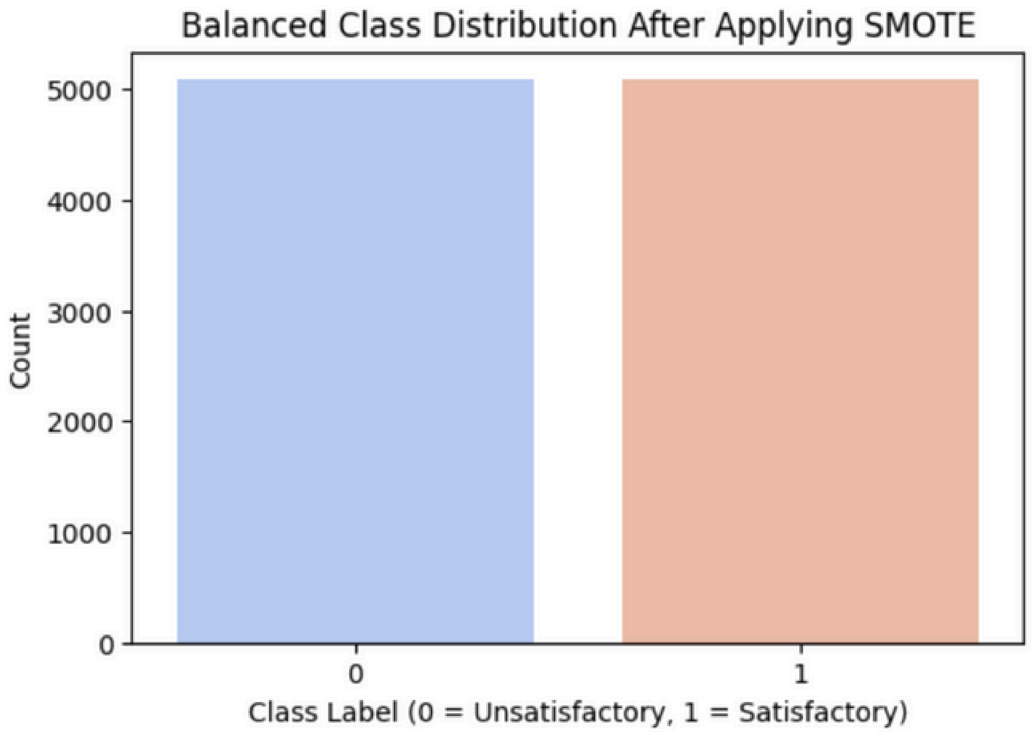
- The dataset is now balanced, ensuring that all classes have equal representation.
- This helps prevent model bias and improves classification performance.

2. For Poland University Dataset

- **Handling Missing Values**
  - Missing values filled using mean imputation for numeric attributes.
  - Gender attributes ("male", "female") combined into a single "Gender" column.
  - Correlation Heatmap Generated to remove highly correlated features (Threshold: 0.9).
- **Feature Engineering**
  - Created a new attribute "Result":
  - 1 (Satisfactory Instructor) → If SET_score_avg > mean SET_score_avg.
  - 0 (Unsatisfactory Instructor) → Otherwise.
  - Class Imbalance Identified, requiring SMOTE for balancing.







About Synthetic Dataset

DATASET GENERATION PROCESS

- Synthetic dataset created to simulate real-world student feedback responses.
- 9,000 entries generated where responses are balanced:
    - 33% High Ratings (Positive Feedback)
    - 33% Low Ratings (Negative Feedback)
    - 33% Mixed Ratings (Neutral/Mixed Feedback)
- The dataset was shuffled to prevent models from easily predicting patterns.

CHANGE IN TARGET VARIABLE

- New Target Variable: “Is a role model and life-long mentor for me”
- If the response for this question is **greater than the column mean**, target = 1 (Instructor is a good mentor).
- **Otherwise**, target = 0 (Instructor is not an ideal mentor).

INITIAL TARGET VARIABLE & OBSERVATIONS

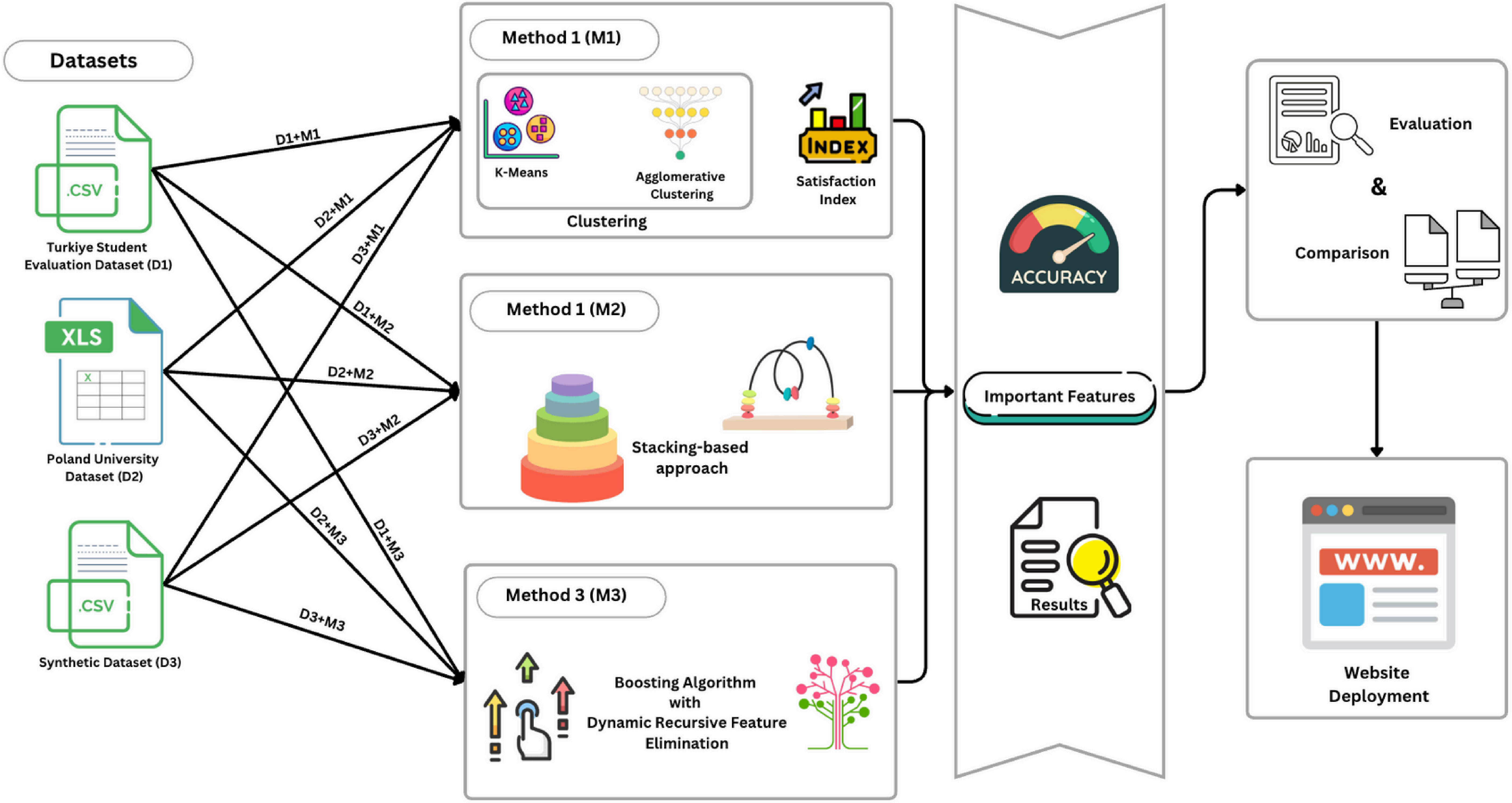
- Initial Target Variable:
    - Weighted Average of all responses.
- Issue Identified:
    - Produced unrealistically high accuracy (~99%), meaning models found patterns too easy to learn.
    - Indicated a need for better feature-target mapping.

WHY THIS CHANGE?

- More realistic instructor evaluation based on student feedback.
- Prevents overfitting and bias in predictions.

Machine Learning Models applied on Datasets

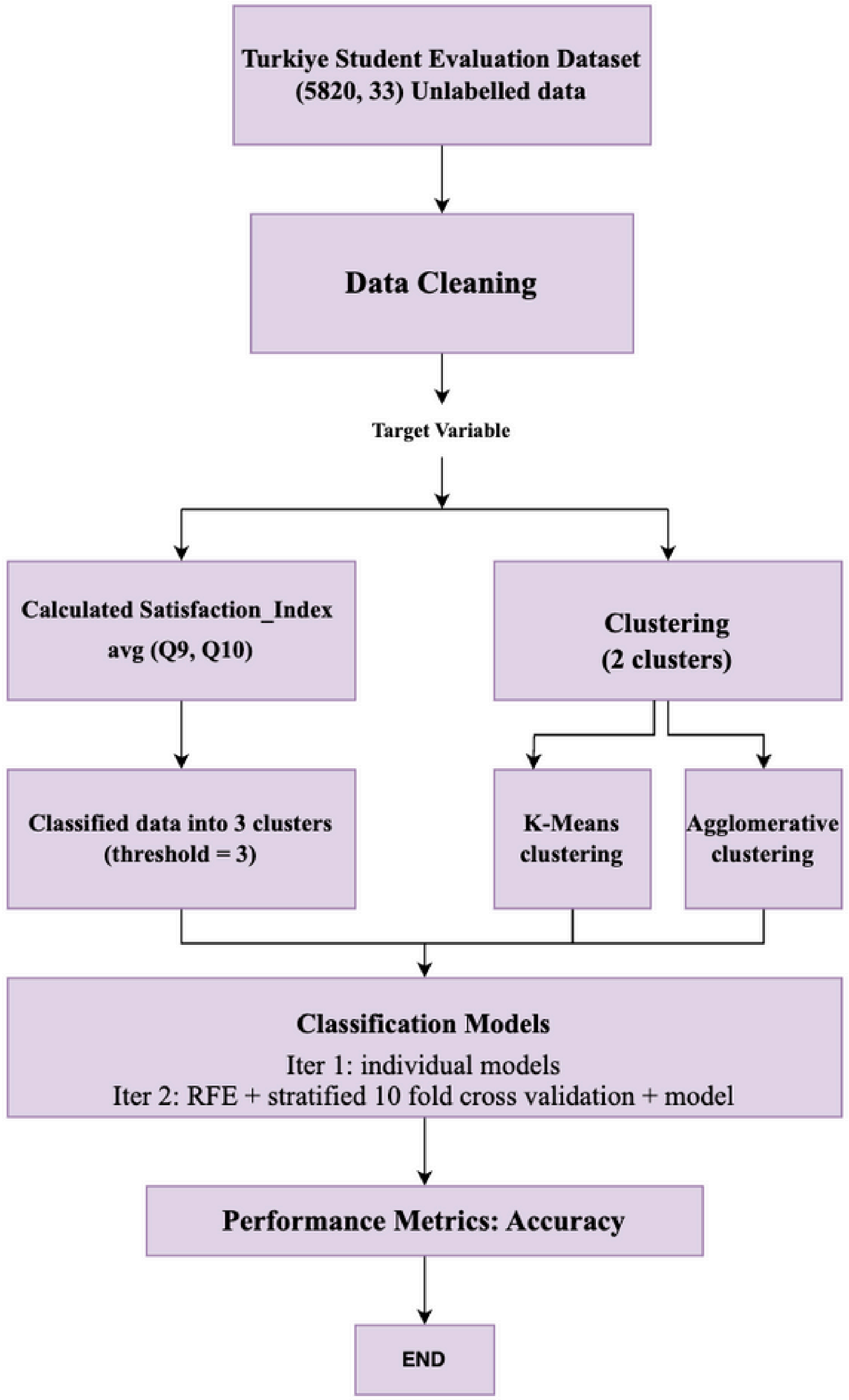
Proposed methodology workflow-



On Turkiye Student Evaluation Dataset

- **Approach 1: Satisfaction Index Approach:**
  - Dropped Satisfaction_Index, Q9, and Q10 from the dataset.
  - Target Variable: "Recommend_Instructor".
  - Models Applied:
    - Logistic Regression, Decision Tree, Bagging, Random Forest, AdaBoost, XGBoost, Voting (Hard & Soft), SVM.
  - Best Model: AdaBoost Classifier (84.4% accuracy).
  - Lowest Accuracy: Soft Voting Classifier (77.4% accuracy).
  - Best Model after Feature Selection & Cross-Validation: Random Forest Classifier (86.9% Accuracy).
  - Recursive Feature Elimination (RFE) + Stratified K-Fold improved results.

- **Approach 2: Clustering:**
  - Optimal cluster count: Determined using Elbow Method & Silhouette Scores.
  - K-Means Clustering:
    - Best Individual Model: Logistic Regression (99.4% Accuracy).
    - After RFE + Cross-Validation: Logistic Regression (98.5% Accuracy).
  - Agglomerative Clustering:
    - Best Individual Model: Voting Classifier (Hard Voting) (96.5% Accuracy).
    - After RFE + Cross-Validation: Logistic Regression (98.5% Accuracy).



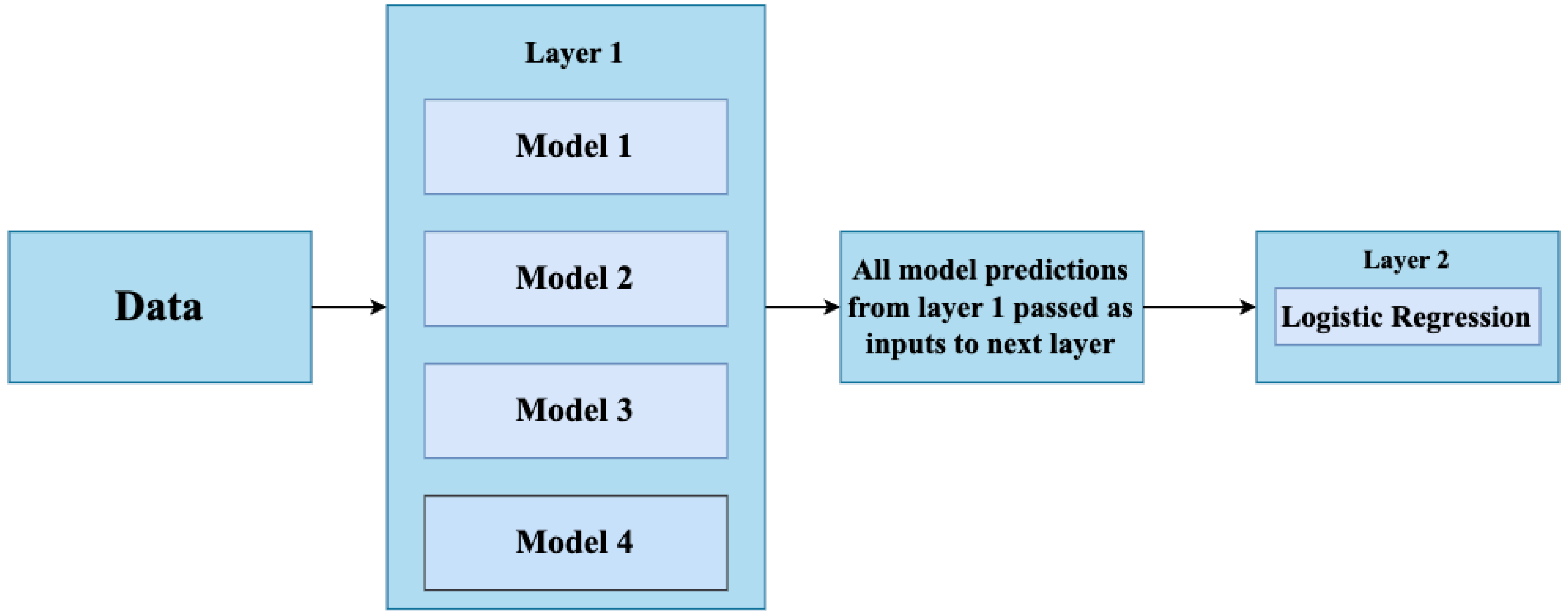
Cont..

- **Accuracies Obtained on Turkiye Student Evaluation Dataset:**

Classifier	**Satisfaction Index**		**K-Means Clustering**		**Agglomerative Clustering**	
	Ind. Model	**RFE + CV**	**Ind. Model**	**RFE + CV**	**Ind. Model**	**RFE + CV**
Logistic Regression	81.7	81.5	**99.4**	98.5	95.4	98.5
Decision Tree Classifier	77.4	80.6	94.8	95.1	92.5	95.1
Bagging Classifier	80.7	84.7	96.8	96.5	95.1	96.5
Random Forest Classifier	77.4	**86.9**	94.8	97.7	92.5	97.7
AdaBoost Classifier	**84.4**	83.0	97.4	97.8	94.8	97.8
XGBoost Classifier	83.6	86.1	98.1	97.6	96.4	97.6
Support Vector Machine	82.9	82.2	99.0	98.4	95.4	98.4
Voting Classifier (Soft)	83.7	–	98.8	–	96.1	–
Voting Classifier (Hard)	82.6	–	98.9	–	**96.5**	–

On Poland University Dataset

- **Stacking-Based Approach (Multi-Level Model)**
  - First Layer:
    - Decision Tree, Random Forest, KNN, XGBoost→ Generate predictions.
  - Second Layer:
    - Logistic Regression applied on predictions of the first layer.
    - Uses first-layer predictions to make final classification.
  - Highest Accuracy Achieved:
    - 78.5% (Stacking Approach), outperforming individual models.
- **ML Models Applied:**
  - Individual Classifiers Used:
    - Logistic Regression, Decision Tree, Naïve Bayes, Random Forest, Bagging Classifier, AdaBoost, XGBoost, KNN.
  - Best Individual Model:
    - Random Forest with Recursive Feature Elimination (RFE) → 66% Accuracy.



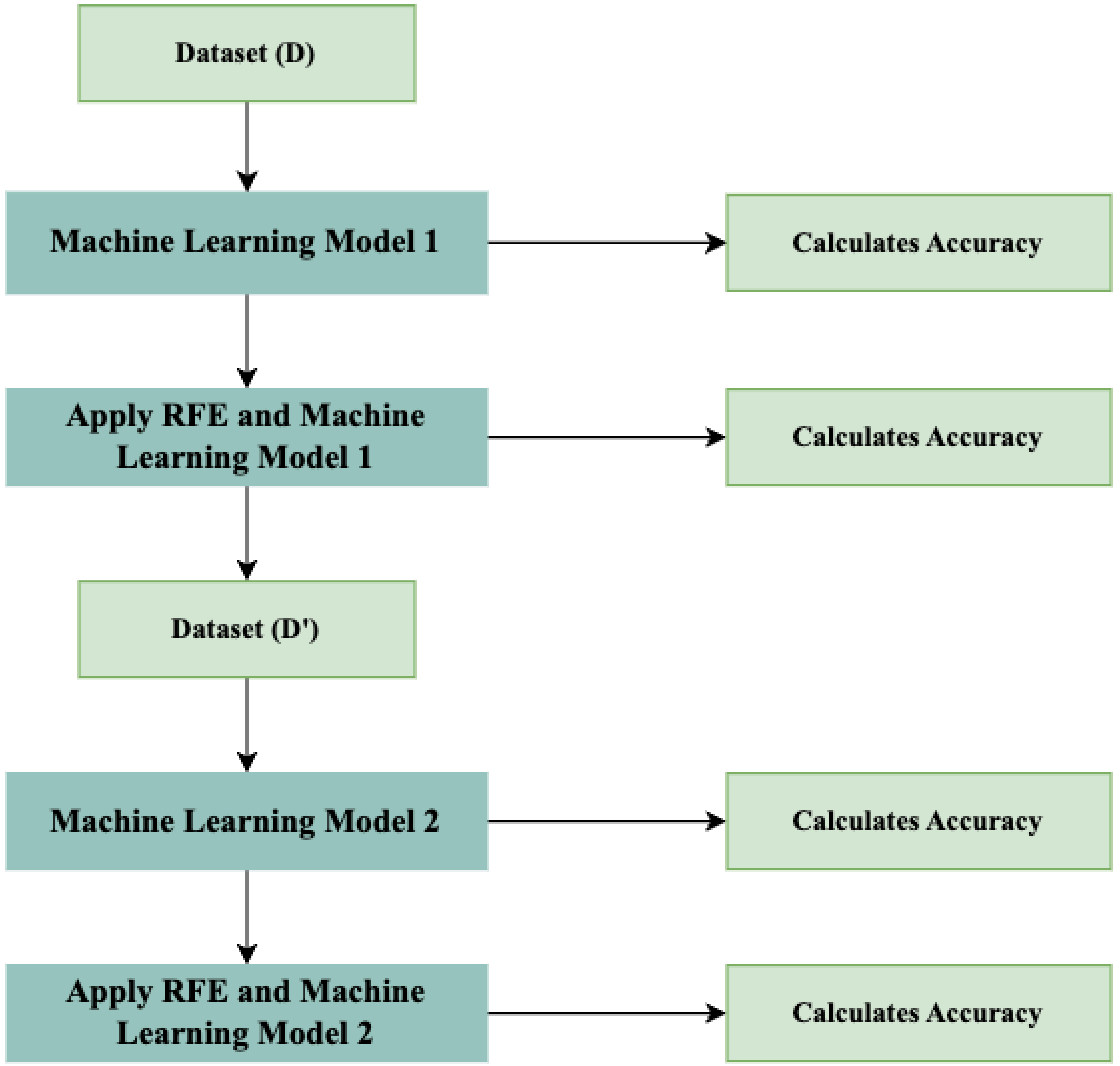
Cont..

- **Accuracies Obtained on Poland University Dataset:**

Classifier	**Performance (Accuracy)**	
		n_splits = 10, n_repeats = 3
Logistic Regression	56	61
Decision Tree Classifier	63	64
Bagging Classifier	**67**	66
Random Forest Classifier	66.5	**66.6**
AdaBoost Classifier	58	61
XGBoost Classifier	54	66
SVM	54	59
Voting Classifier	64	–
Stacking Classifier	**78.5**	–

On Synthetic Dataset

- **Boosting Approach with Recursive Feature Elimination:**
  - Hybrid Model = Boosting + RFE for better feature selection & performance.
  - Pipeline Approach:
    - First Model: Predicts values → Adds predictions as new feature ("First_Prediction").
    - Second Model: Uses updated dataset for final prediction.
- **Dynamic Recursive Feature Elimination (DRFE):**
  - Uses Linear Regression & R-Squared values to iteratively remove least significant features.
  - Stopping Criteria:
    - Accuracy drops compared to the previous iteration.
    - Accuracy reaches 100%.
    - No more features left for elimination.
- **Experimental Results:**
  - Initial Model Accuracy: ~67% - 72%.
  - After Dynamic RFE: Increased to 90.15%.
  - 49 Model Combinations Tested using 7 ML Models:
    - Decision Tree, Logistic Regression, SVM, AdaBoost, MLP Classifier, Gradient Boosting, XGBoost.



Cont..

- **Accuracies Obtained on Synthetic Dataset:**

S.No	First Model	Acc. 1	Acc. 1 + RFE	Second Model	Acc. 2	Acc. 2 + RFE
2	Decision Tree	67.25	67.73	SVM	**90.15**	**90.15**
3	Decision Tree	67.46	67.66	Adaboost Classifier	90.04	90.04
4	Decision Tree	67.23	67.37	MLP Classifier	89.90	89.90
1	Decision Tree	67.59	67.59	Logistic Regression	89.85	89.85
5	Decision Tree	67.37	67.73	Gradient Boosting	89.69	89.69
6	Decision Tree	67.84	67.84	XGBoost	88.38	88.38
23	Adaboost Classifier	72.05	72.12	SVM	73.89	73.89
30	MLP Classifier	71.46	72.51	SVM	74.62	74.62
37	Gradient Boosting	72.01	72.01	SVM	74.12	74.12
38	Gradient Boosting	72.01	72.01	Adaboost Classifier	74.03	74.03

Cross-Dataset Methodology Evaluation

The 3×3 model-to-dataset comparison framework is established by running each machine learning method on all three datasets.

Accuracies Obtained by Applying Turkiye Student Evaluation Dataset Methodologies on Poland Dataset

Classifier	Satisfaction Index		K-Means Clustering		Agglomerative Clustering	
	Ind. Model	RFE + CV	Ind. Model	RFE + CV	Ind. Model	RFE + CV
Logistic Regression	71.31	71.39	98.71	98.33	94.89	94.12
Decision Tree Classifier	97.30	97.98	**100.00**	**100.00**	**100.00**	**100.00**
Bagging Classifier	98.00	98.48	**100.00**	**100.00**	99.83	**100.00**
Random Forest Classifier	98.96	**99.14**	**100.00**	**100.00**	**100.00**	**100.00**
AdaBoost Classifier	76.55	74.61	98.46	98.85	94.18	94.93
XGBoost Classifier	98.79	98.43	**100.00**	**100.00**	99.83	**100.00**
Support Vector Machine	54.93	65.56	98.09	97.98	96.17	96.31
Voting Classifier (Soft)	98.25	98.59	**100.00**	**100.00**	99.83	99.98

Accuracies Obtained by Applying Turkiye Student Evaluation Dataset Methodologies on Synthetic Dataset

Classifier	Satisfaction Index		K-Means Clustering		Agglomerative Clustering	
	Ind. Model	RFE + CV	Ind. Model	RFE + CV	Ind. Model	RFE + CV
Logistic Regression	50.63	50.61	50.97	50.61	50.97	50.61
Decision Tree Classifier	50.07	49.51	50.97	49.61	50.97	49.56
Random Forest Classifier	48.73	49.48	50.97	49.45	50.97	49.10
Bagging Classifier	48.23	–	50.97	–	50.97	–
AdaBoost Classifier	49.40	50.94	50.97	50.94	50.97	50.94
XGBoost Classifier	48.67	49.65	50.97	49.65	50.97	49.65
Support Vector Machine	50.97	50.95	50.97	50.95	50.97	50.95
Voting Classifier (Soft)	48.80	–	50.97	–	50.97	–
Voting Classifier (Hard)	48.60	–	50.97	–	50.97	–

Accuracies Obtained by Applying Poland Dataset Methodologies on Turkiye Student Evaluation Dataset

Classifier	Model Accuracy (%)
Decision Tree Classifier	100.00
XGBoost Classifier	100.00
Bagging Classifier	100.00
Random Forest Classifier	100.00
Adaboost Classifier	100.00
Stacking Classifier	100.00
Voting Classifier	100.00
Logistic Regression	98.92
SVM	98.38

Accuracies Obtained by Applying Poland Dataset Methodologies on Synthetic Dataset

Classifier	**Model Accuracy (%)**
Logistic Regression	50.63
Decision Tree	50.03
Random Forest	48.93
Bagging Classifier	49.43
Adaboost Classifier	49.40
XGBoost	48.67
SVM	50.97
Voting Classifier (Soft Voting)	47.93
Voting Classifier (Hard Voting)	48.53
Stacking Classifier	**51.10**

Accuracies Obtained by Applying Synthetic Dataset Methodologies on Turkiye Student Evaluation Dataset

S.No	First Model	Acc. 1	Acc. 1 + RFE	Second Model	Acc. 2	Acc. 2 + RFE
0	Decision Tree	100.00	100.00	Decision Tree	100.00	100.00
1	Decision Tree	100.00	100.00	Logistic Regression	100.00	100.00
2	Decision Tree	100.00	100.00	SVM	100.00	100.00
3	Decision Tree	100.00	100.00	Adaboost Classifier	100.00	100.00
4	Decision Tree	100.00	100.00	MLP Classifier	100.00	100.00
5	Decision Tree	100.00	100.00	Gradient Boosting	100.00	100.00
6	Decision Tree	100.00	100.00	XGBoost	100.00	100.00
21	Adaboost Classifier	100.00	100.00	Decision Tree	100.00	100.00
24	Adaboost Classifier	100.00	100.00	Adaboost Classifier	100.00	100.00
40	Gradient Boosting	100.00	100.00	Gradient Boosting	100.00	100.00

Accuracies Obtained by Applying Synthetic Dataset Methodologies on Poland Dataset

S.No	First Model	Acc. 1	Acc. 1 + RFE	Second Model	Acc. 2	Acc. 2 + RFE
42	XGBoost	98.21	98.71	Decision Tree	99.29	99.29
48	XGBoost	98.21	98.71	XGBoost	99.04	98.67
13	Logistic Regression	62.87	62.74	XGBoost	98.34	98.71
20	SVM	63.58	63.28	XGBoost	97.71	97.75
27	Adaboost Classifier	67.73	67.73	XGBoost	97.96	97.96
34	MLP Classifier	83.37	–	XGBoost	98.67	98.05
41	Gradient Boosting	78.00	77.34	XGBoost	97.30	96.59
5	Decision Tree	99.29	99.21	Gradient Boosting	99.29	99.29
2	Decision Tree	99.29	99.21	SVM	99.29	99.29
3	Decision Tree	99.29	99.21	Adaboost Classifier	99.29	99.29

Website: KnowMyProf

Website

Know Your Professors

Prof. John Doe
Professor, Department of Computer Science

Dr. Alice Smith
Assistant Professor, Department of Mathematics

Dr. Michael Brown
Associate Professor, Department of Physics

See more >>

Professor Cards:

Professors

Prof. John Doe
Professor, Department of Computer Science

Dr. Alice Smith
Assistant Professor, Department of Mathematics

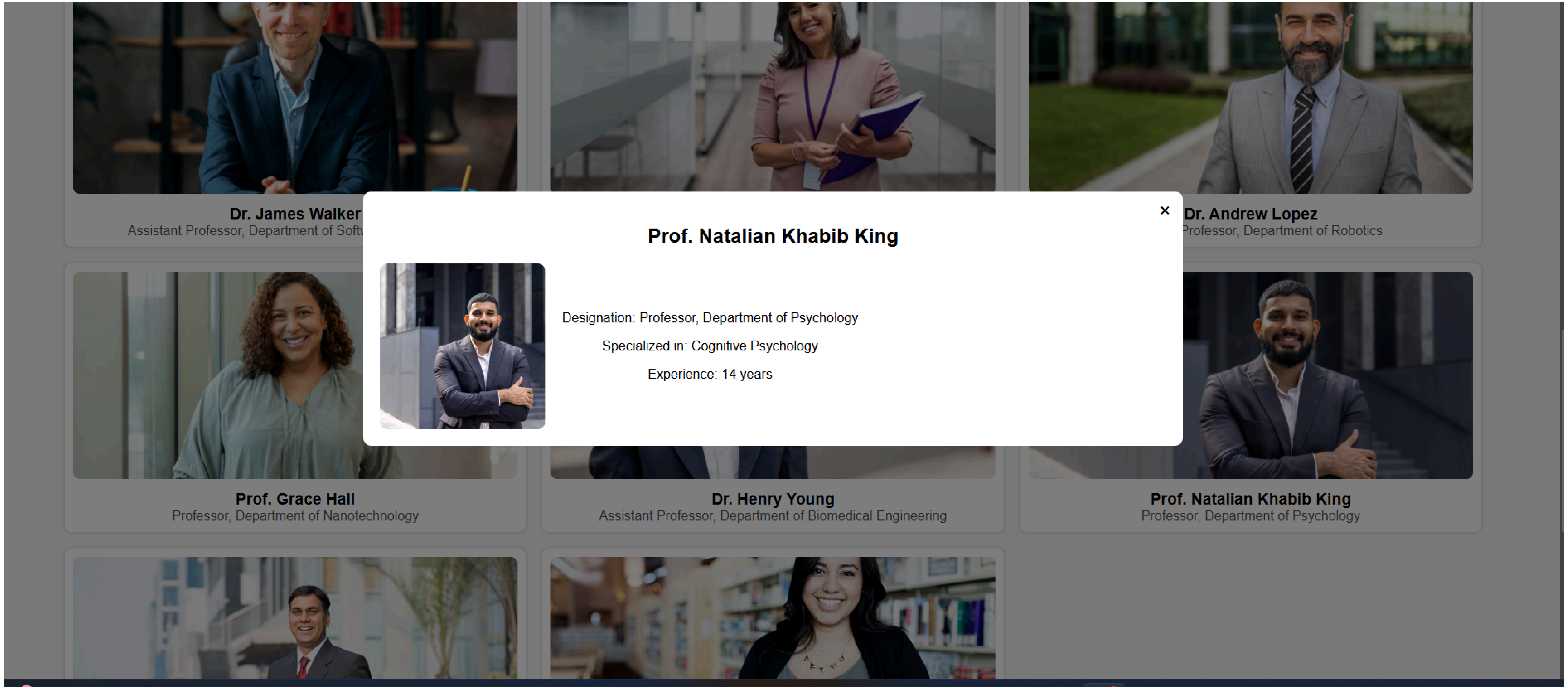
Dr. Michael Brown
Associate Professor, Department of Physics

Prof. Emily Davis
Professor, Department of Chemistry

Dr. Robert Johnson
Assistant Professor, Department of Biology

Dr. Sarah Wilson
Professor, Department of Civil Engineering

Professor Details:



Professor Rating:

Prof. John Doe

- Over 15 Years of Teaching Experience: With more than a decade and a half of experience in higher education, Professor John Doe has successfully taught over 2000 students across various undergraduate and graduate courses.
- Published 30+ Research Papers: He has authored and co-authored more than 30 peer-reviewed research papers in reputable academic journals, contributing significantly to his field of expertise.
- Supervised 10+ PhD Dissertations: Professor Doe has served as a primary advisor for over 10 doctoral students, guiding them through their research projects and contributing to the advancement of their academic careers.

Conclusion:

This project uses machine learning to predict instructor performance based on student feedback, making evaluations more objective and reliable. It helps improve teaching quality and supports students in choosing the right instructors. Moving forward, we aim to develop a professor rating system for better transparency and decision-making in education.

Thank you!