

Decoding Instructor Performance with Data Mining and Machine Learning

Done By-

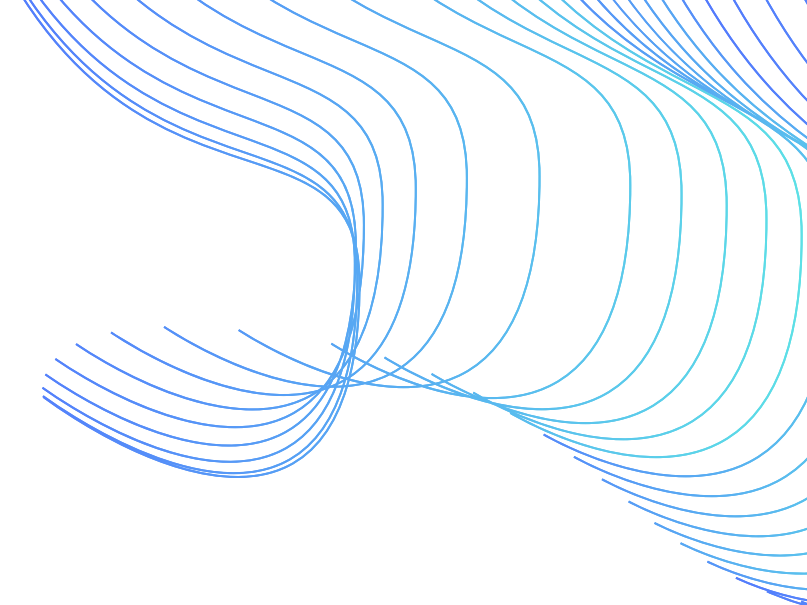
L. Sai Gopal - AP21110010042 - Model evaluation and web app integration

A. Namratha - AP21110010048 - Performance tuning and report compilation

A .Haasitha - AP21110010049 - Literature review and feature engineering



Problem Description



- **The Importance of Instructor Evaluations:** Universities often rely on student evaluations to assess teaching effectiveness. However, these evaluations can be subjective and influenced by various external factors.
- **The Need for Objectivity:** There's a growing need to analyze these evaluations systematically to derive objective insights into instructor performance.
- **Our Approach:** We aim to leverage data mining and machine learning techniques to predict instructor performance more accurately, providing actionable feedback for educational improvement.



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.

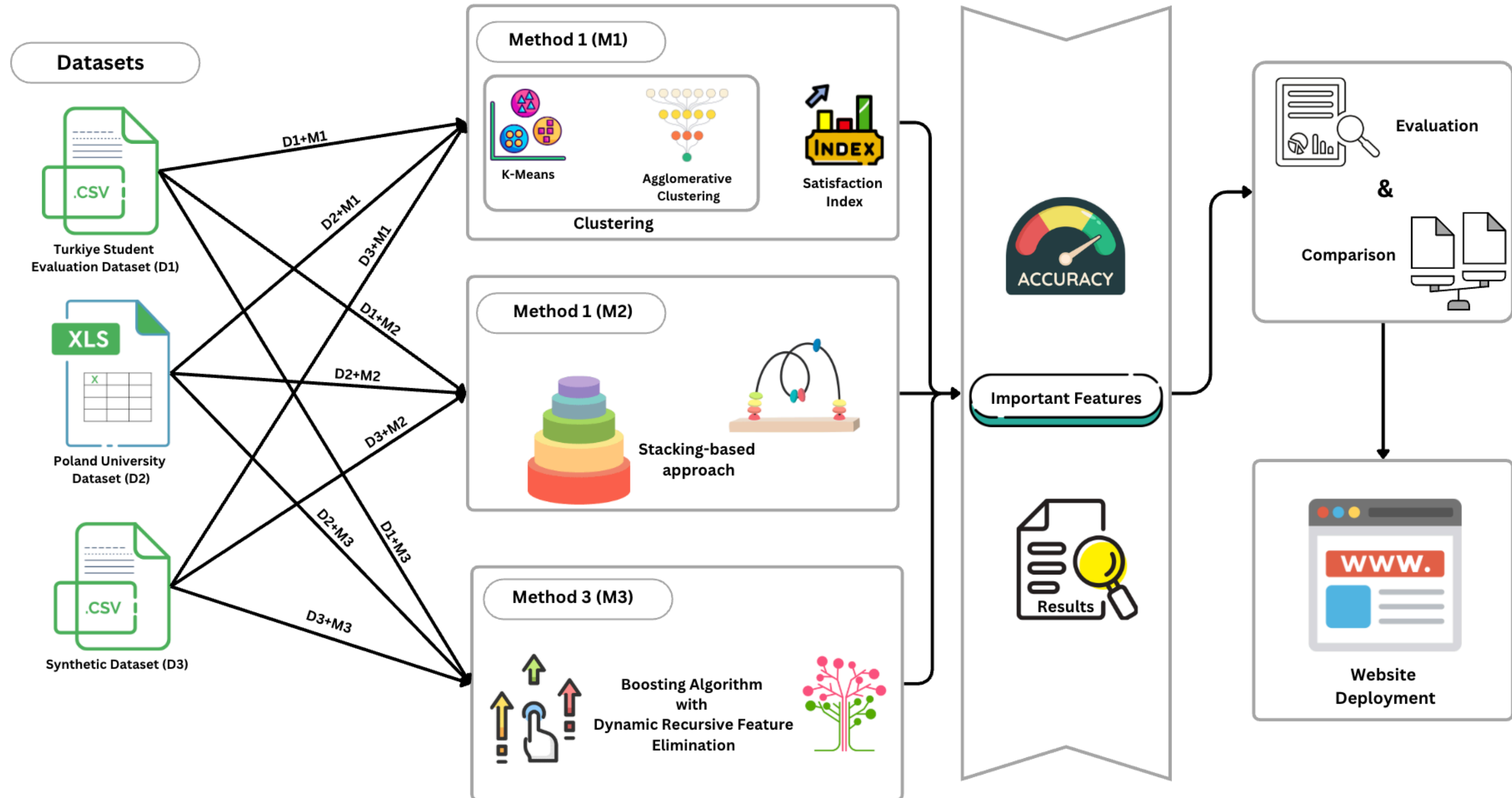
Poland University Dataset

- 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.

Method developed:



Performance Metrics



Turkiye Dataset:

- Best Accuracy: 86.9% (Random Forest + RFE)
- Clustering approaches (KMeans + Logistic Regression): 99.4%

Poland Dataset:

- Best Accuracy: 78.5% (Stacking Classifier)

Synthetic Dataset:

- Best Accuracy: 90.15% (Decision Tree → SVM with RFE)

When Synthetic Dataset Methodology when applied on other two datasets we have obtained accuracies as follows:

- Türkiye - 82.67%
- Poland - 72.14%

Website Deployment:

Website: <https://knowmy.prof.netlify.app/>

Concluding Remark

- We developed a machine learning-based system to evaluate college instructors using student feedback. By analyzing factors like teaching experience, preparedness, and fairness, our models achieved high accuracy: 78% with the Poland dataset, 99.4% with the Turkiye dataset, and 90.15% with a synthetic dataset we created.
- We chose a combination of Decision Tree and Support Vector Machine (SVM) models for deployment due to their strong performance. To make the results accessible, we built a web application where users can search for professors by name or ID to view their ratings and performance details.
- We have also prepared a paper in the IEEE conference format, and our professor has mentioned that if there are any suitable conferences aligned with our work, we will proceed with submitting it for consideration.

A decorative graphic on the left side of the slide, consisting of a dense field of small purple dots that form a triangular shape pointing towards the center.

Thank you!