

Project Report: Intelligent Student Performance Dashboard

1. Project Overview

The **Intelligent Student Performance Dashboard** is an interactive data analytics and AI-powered platform built using **Streamlit**, designed to monitor, analyze, and predict student academic performance.

It integrates **machine learning models**, **interactive visualization**, and **What-if analysis** to support teachers and decision-makers in understanding students' learning behavior and predicting outcomes.

2. Technologies Used

Category	Tools & Libraries
Programming Language	Python 3.11
Framework	Streamlit
Visualization	Plotly, Plotly Express
Machine Learning	Scikit-learn, NumPy, Pandas
Model Storage	Pickle (student_multi_model.pkl)
Mappings Storage	JSON (label_mappings.json)

3. Dashboard Architecture

The dashboard consists of **five main tabs**, each representing a different analytical perspective:

Tab	Name	Function
1	Overview	Displays overall statistics about students' performance and attendance.
2	Subjects	Shows detailed academic results per subject (Math, Reading, Writing).
3	Attendance	Visualizes attendance distribution and its effect on performance.
4	Insights	Provides advanced analytical summaries and patterns across the dataset.
5	AI Prediction (What-If Analysis)	Allows users to simulate input changes and predict outcomes using the trained ML model.

4. Data Overview

The dataset (DATA.csv) includes several academic and behavioral variables:

- **Academic scores:** Math, Reading, Writing, OverallPerformance, AcademicIndex
- **Behavioral attributes:** AttendanceRate, StudyHours, SleepHours, SocialIndex, BehaviorIndex
- **Socio-demographic:** Parent Education, Marital Status, Transport Means, etc.

Data preprocessing includes:

- Dropping unnecessary identifiers (e.g., student_id, student_name).
- Encoding categorical variables using LabelEncoder.
- Splitting data for model training and testing.



5. Machine Learning Model

Model Type:

MultiOutputRegressor(RandomForestRegressor)

Purpose:

Predicts both:

1. Academic Average (AcademicIndex)
2. Overall Performance

Model Configuration:

Parameter	Value
n_estimators	150
max_depth	12
random_state	42

Model Performance:

Metric	Overall Performance	Academic Average
R ² Score	0.971	1.000
MAE	1.551	0.166

Excellent accuracy was achieved due to the robustness of Random Forests and good dataset quality.

The trained model is saved as:

- student_multi_model.pkl
 - Label encodings in label_mappings.json
-

6. Dashboard Functionalities

Tab 1: Overview

- Shows total number of students, averages of performance and attendance.
- Displays best and lowest scores.
- Summarizes students above and below the performance average.
- Provides interactive KPIs (using st.metric) and visuals (plotly.express).

Tab 2: Subjects

- Compares each student's scores (Math, Reading, Writing) to the dataset average.
- Interactive bar chart visualizes subject-wise performance gaps.

Tab 3: Attendance

- Displays histogram of attendance rates.
- Interprets attendance as a direct factor influencing performance.

Tab 4: Insights

- Summarizes overall averages (study hours, attendance, performance).
- Identifies top and lowest performers.
- Highlights score gaps between writing and math.
- Calculates percentages of:
 - High-attendance students
 - Active learners (≥ 10 h study/week)
 - High performers ($\geq 70\%$)

Includes a **Smart Summary** that narratively explains key findings.

Tab 5: AI Prediction (What-If Analysis)

- Loads mappings from label_mappings.json.
- Users can input or modify variables such as:
 - Math, Reading, Writing scores
 - Attendance Rate
 - Study and Sleep Hours
 - Behavioral indexes

- The app encodes inputs automatically and predicts:
 - **Predicted Academic Average**
 - **Predicted Overall Performance**

Outputs include metrics, progress bars, and interactive feedback (balloons and highlights).

7. Prediction Workflow

1. User selects or enters custom student attributes.
 2. System encodes categorical variables based on the same label mapping used during training.
 3. Model predicts both **Academic Average** and **Overall Performance** simultaneously.
 4. Results are displayed instantly on-screen with visual indicators.
-

8. Files and Structure

Project Folder

```
|  
|--- DATA.csv  
|--- student_multi_model.pkl  
|--- label_mappings.json  
|--- app_pro.py ← (Streamlit Dashboard)  
|--- model_training.py ← (ML Model Training Script)
```

9. Deployment and Access

The dashboard runs locally using:

```
streamlit run app_pro.py
```

To share it online:

1. Install **ngrok** (`pip install pyngrok`)
2. Authenticate with your token (`ngrok config add-authtoken <your_token>`)
3. Run your app:

4. streamlit run app_pro.py
 5. Start ngrok tunnel:
 6. ngrok http 8501
 7. Share the public URL generated.
-

10. Key Insights and Findings

- High attendance and consistent study habits strongly correlate with performance.
 - Students with better writing and reading abilities generally perform better overall.
 - AI predictions can effectively forecast academic performance with minimal error.
 - The dashboard helps educators **identify weak students early** and **take action**.
-

11. Future Improvements

- Enhance model explainability (e.g., SHAP or LIME visualization).
 - Integrate live database for real-time student data updates.
 - Add student grouping (e.g., clustering by performance type).
 - Deploy as a web app using Streamlit Cloud or Render for public access.
-

12. Conclusion

The **Intelligent Student Performance Dashboard** combines **AI-driven prediction** with **data visualization** to transform student analytics into actionable insights.

It demonstrates how machine learning can be embedded within interactive dashboards to make academic monitoring **smarter, faster, and more reliable**.