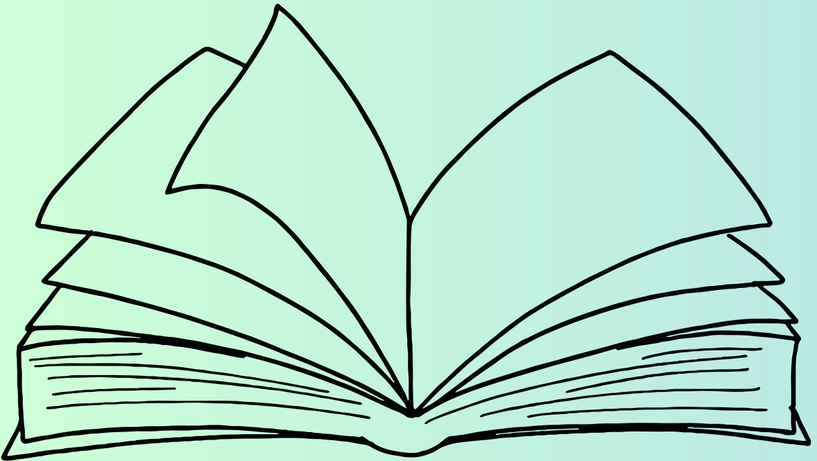
Students Score Prediction

Based on Study Habits....

Data-Driven Score Predictions Using Machine Learning



By Shreya Anand

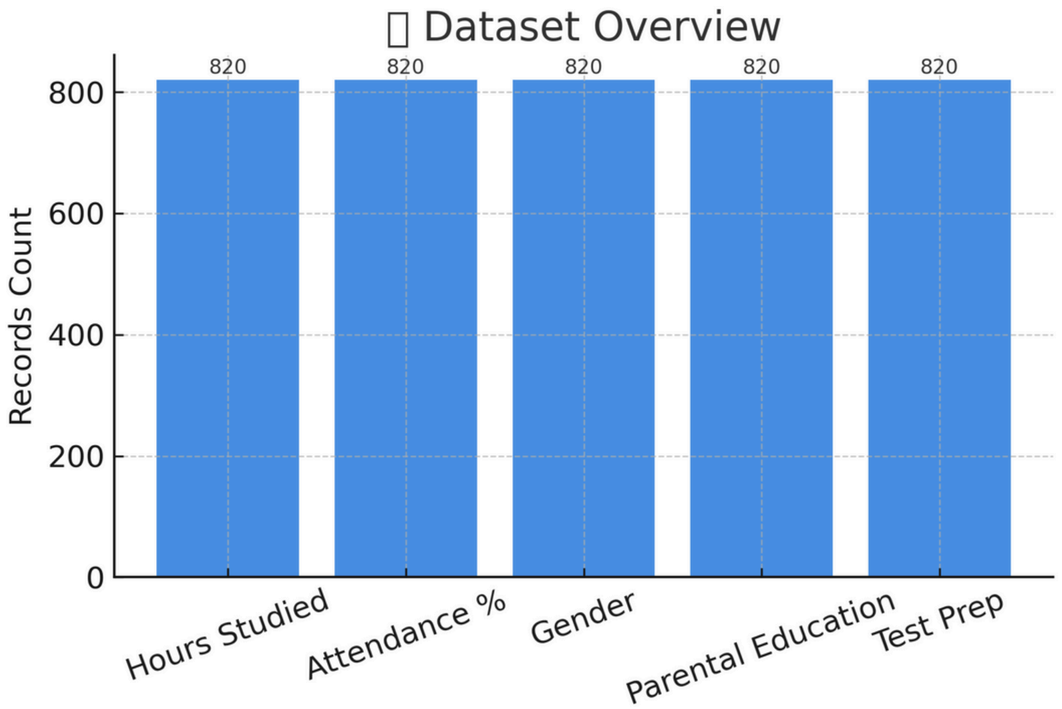
Contents:

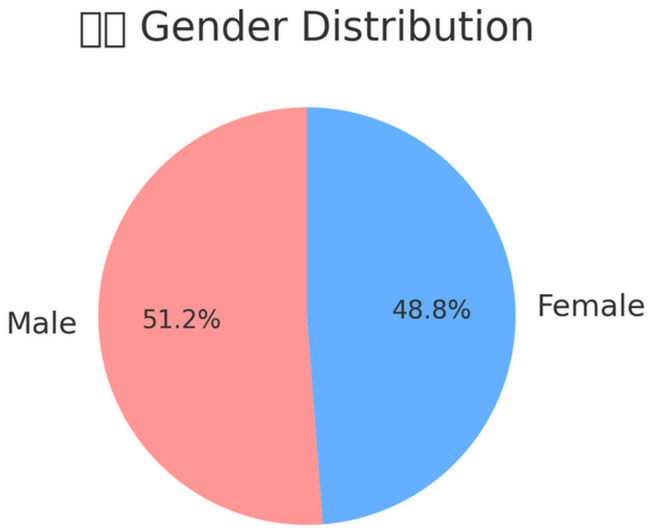
Project Objective:

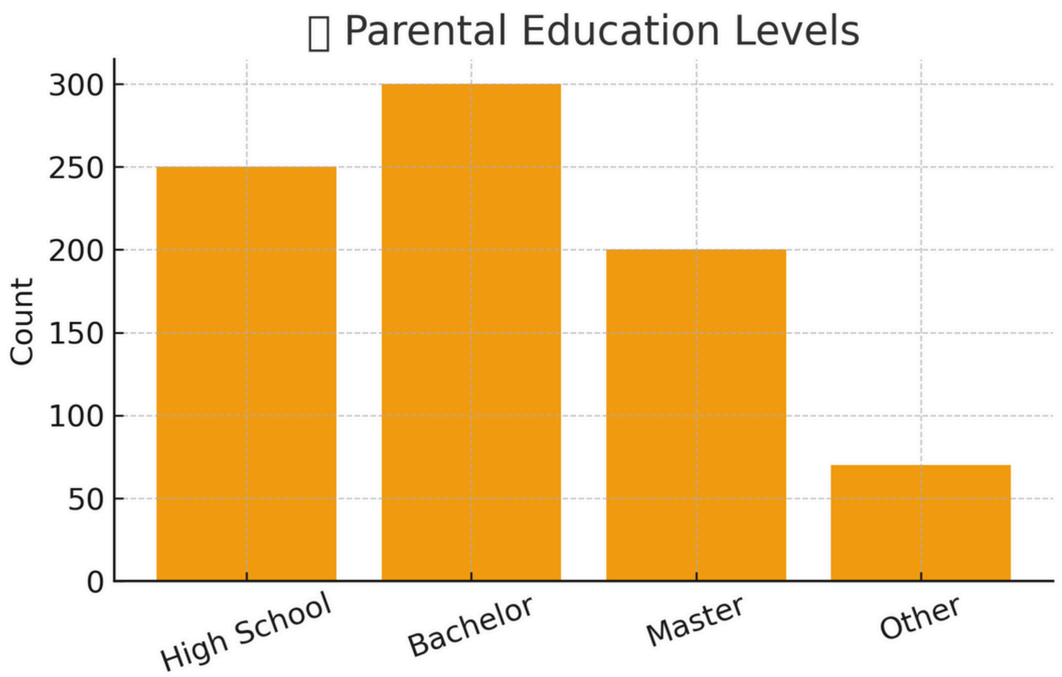
To predict student performance (exam scores) using machine learning (Linear Regression), based on study habits and demographics.

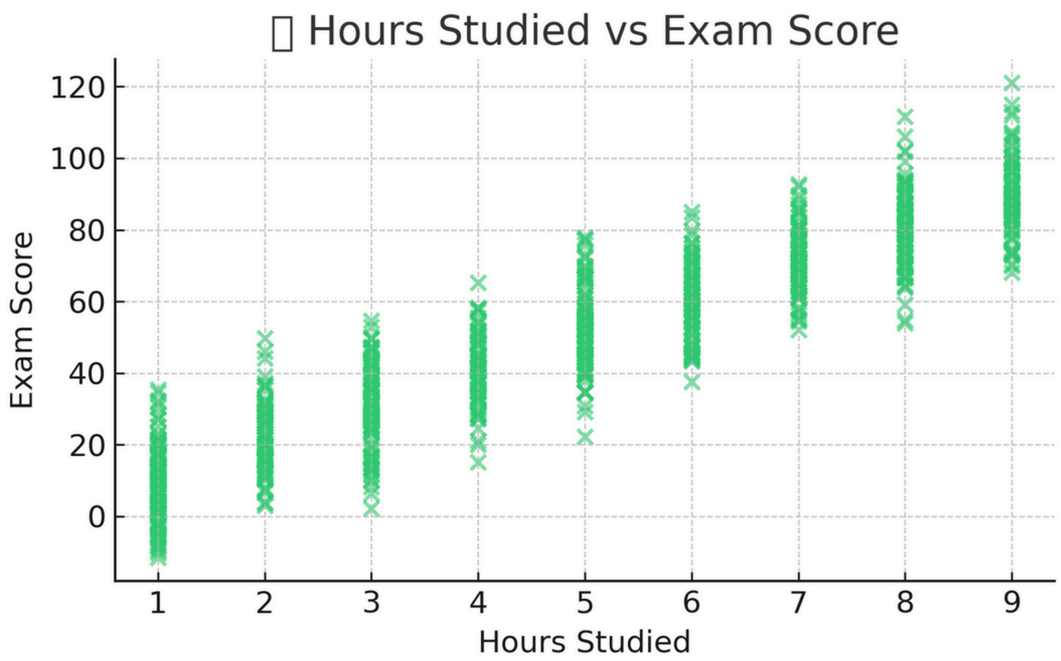
Dataset Overview:

📦 ~800+ student records (example dataset)
📚 Attributes: Hours Studied, Attendance%, Gender, Parental Education, Test Preparation Course
🎯 Target: Exam Score (out of 100)
🧹 Cleaned & encoded categorical data, normalized numeric features
📂 Source: Open dataset (modified for project)


Dataset Overview
Records Count
0
200
400
600
800
820
820
820
820
820
Hours Studied
Attendance %
Gender
Parental Education
Test Prep


Gender Distribution
Male
51.2%
48.8%
Female


Parental Education Levels
Count
0
50
100
150
200
250
300
High School
Bachelor
Master
Other


Hours Studied vs Exam Score
Exam Score
0
20
40
60
80
100
120
1
2
3
4
5
6
7
8
9
Hours Studied

Model Workflow:

📝 Input Student Data

⚙️ Data Preprocessing (Encoding, Scaling)

📈 Model Training (Linear Regression)

🎯 Prediction & Evaluation

📐Metrics:MAE = 10.97, MSE = 189.33

R^2 = 0.07

🎮 App Demo (Streamlit):

✨ Interactive Features:

📝 Input Panel → Enter study hours, attendance, and background info

🎛️ One-Click Predict → Instant score estimation on screen

📊 Smart Visuals → See Regression Plot

Student Final Score Predictor 🎓

Attendance %:

80 − +

Hours Studied:

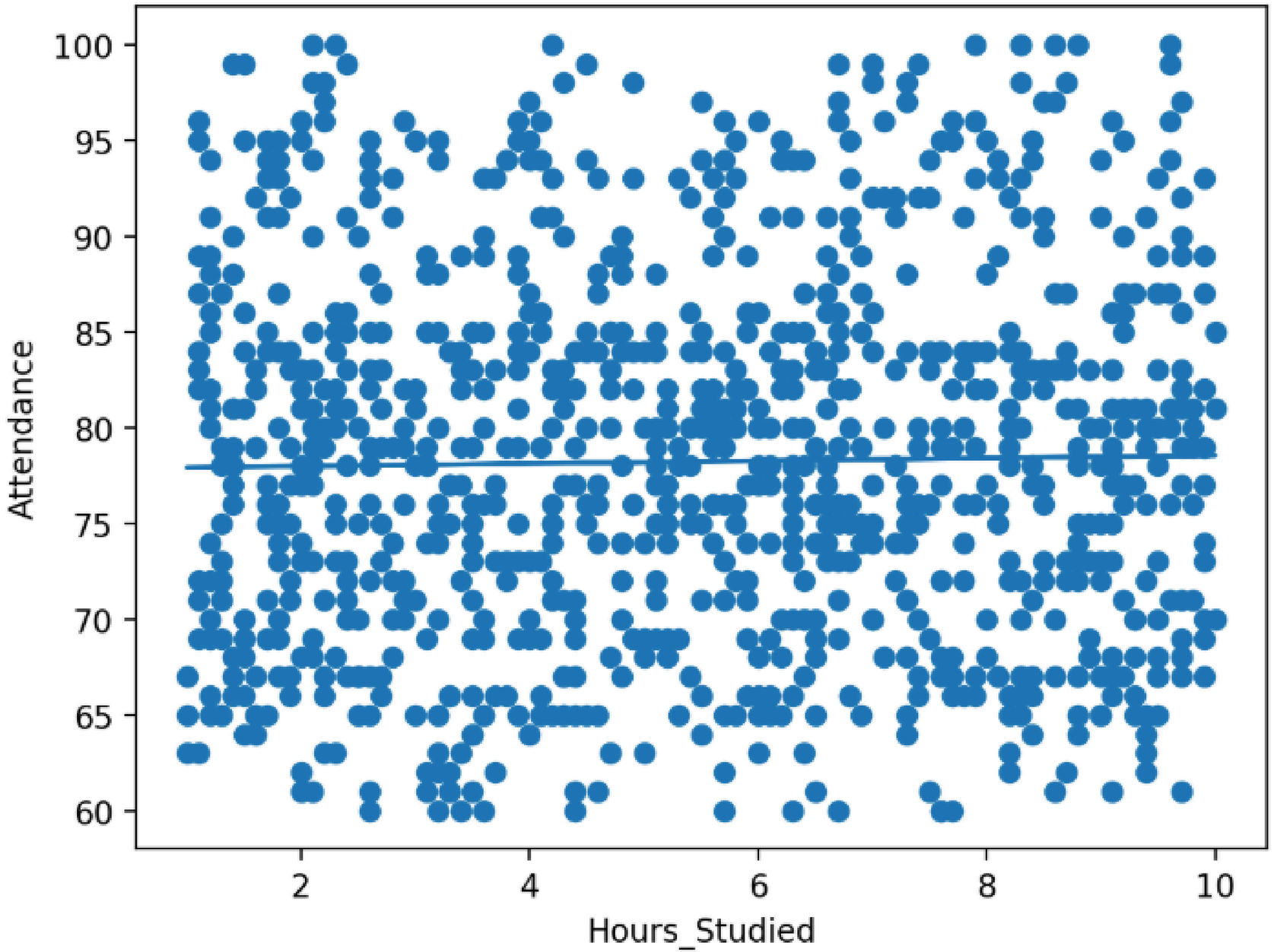
4.00 − +

Predict

Predicted Final Score: 69.40

Model Performance on Dataset

MSE: 189.33, MAE: 10.97, R^2: 0.07



🔑 Key Insights:

📘 Study Hours Matter → More hours = higher predicted score

🕒 Attendance is Critical → Strong correlation with performance

📚 Test Prep Works → Boosts scores consistently

👨‍👩‍👧 Background Influence → Gender & parental education = moderate effect

🎯 Conclusion :

- 🔹 Built a Linear Regression model to predict student scores
- 🔹 Designed an interactive Streamlit app for real-time use
- 🔹 Identified key drivers: study hours, attendance, and test prep
- 🔹 Project shows how ML can support smarter education decisions.

🔮 Future Work:

🚀 Expand Features → Include past scores, attendance logs, study resources

🌲 Advanced Models → Try Random Forest, XGBoost, and Neural Networks

☁️ Cloud Deployment → Host app with authentication & scalability

📊 Teacher Dashboards → Personalized insights for students & classrooms

🤝 Integration → Link with Learning Management Systems (LMS)

"📌 Scan to view GitHub Repository"

Thank You