

Student Score Prediction Based on Study Habits

Description:-

This project focuses on predicting students' final exam scores based on their study habits, specifically study hours and class attendance. Using machine learning techniques such as linear regression, the model analyzes how these independent variables influence academic performance. The analysis shows that study hours have a strong positive impact on scores, while attendance plays a moderate but supportive role. Correlation analysis and data visualization highlight the importance of consistent study habits in improving exam results. The model helps in understanding learning patterns and can guide students to optimize their study strategies for better academic success.

Problem Statement

Can we predict a student's final exam score using study hours and attendance?

Data Visualization

Scatter Plot:-

This plot shows the relationship between study hour, attendance, and final exam score. Each point represents a student, and we can observe a trend of higher scores with more study hours.

Correlation Heatmap:-

A color-coded grid that shows the correlation coefficient between all pairs of variables. Red indicates a strong positive correlation, while blue indicates a negative or weak correlation.

Linear Regression Plot (Implot):-

This plot combines a scatter plot with a regression line. It visually represents the linear relationship and shows how well the independent variables (study hour and attendance) predict the final exam score.

3D Scatter Plot:-

A three-dimensional plot that displays the data points for all three variables simultaneously. It helps to visualize the combined effect of study hour and attendance on the final exam score.

Methodology

Import → Explore → Visualize → Split → Train →
Evaluate → Predict

Model Training & Evaluation:-

Train/Test Split

R² Score, Mean Absolute Error

Technologies Used:-

Python-(Programming Language)

Pandas-(Data Handling)

Matplotlib & Seaborn-(Visualization)

Scikit-Learn (sklearn)-(Machine Learning Model)

Jupyter Notebook-(Development Environment)

Prediction Example Screenshot:-

```
new_data = pd.DataFrame({'study_hour': [5], 'attendance': [80]})  
print("Predicted Score:", model.predict(new_data)[0])
```

Predicted Score: 60.22215778184802

Calculate R² score and Mean Absolute Error:-

```
print("R2 Score:", r2_score(y_test, y_pred))  
print("Mean Absolute Error:", mean_absolute_error(y_test, y_pred))
```

R² Score: 0.9075808162800836

Mean Absolute Error: 3.7539148588148294

Conclusion:-

The analysis shows that study hours strongly influence student exam scores, while attendance has a moderate impact. Together, they significantly affect performance. Therefore, consistent study habits combined with regular class attendance can enhance academic success, making study time the most critical factor in predicting student achievement.

THE END.....