

STUDENT SCORE PREDICTION

A Machine Learning Project to Predict Student Scores Based on Study Hours

By : Ira Revaliawanti

TUJUAN PROYEK

Memprediksi skor siswa berdasarkan jumlah jam yang dihabiskan untuk belajar dengan menggunakan model regresi.

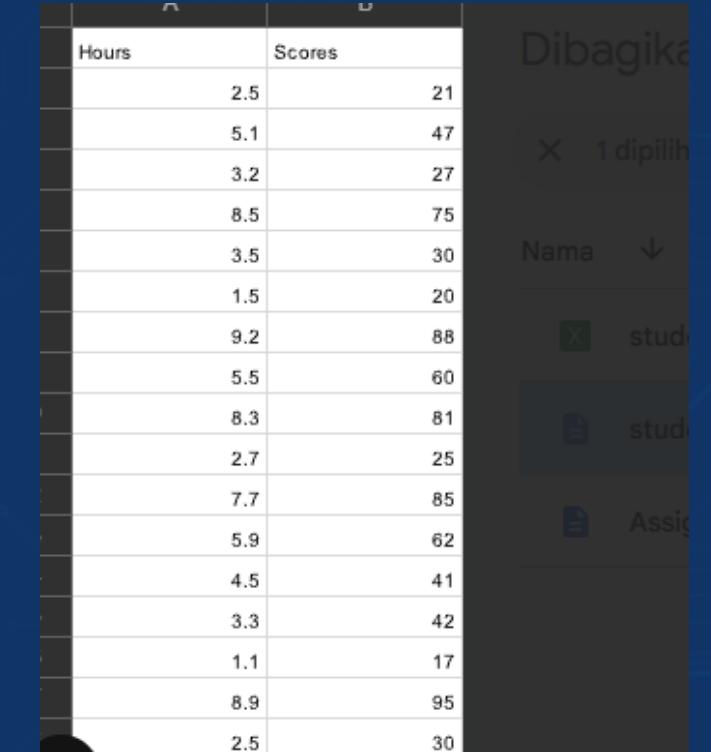
Mengidentifikasi hubungan antara waktu belajar dan hasil yang dicapai oleh siswa.

MENGAPA INI PENTING?

- Memahami faktor-faktor yang memengaruhi skor ujian dapat membantu siswa dan pengajar dalam merencanakan waktu belajar yang lebih efektif.
- Proyek ini dapat membantu mendemonstrasikan bagaimana data dapat digunakan untuk mendukung keputusan dalam pendidikan.

DATASET

- Dataset yang Digunakan:
- Dataset yang digunakan adalah `student_scores.csv`, yang berisi informasi tentang jam belajar dan skor ujian siswa.
- Fitur dalam Dataset:
- Study Hours: Menunjukkan jumlah jam yang dihabiskan oleh siswa untuk belajar.
- Scores: Menunjukkan skor ujian yang diperoleh oleh siswa.



Hours	Scores
2.5	21
5.1	47
3.2	27
8.5	75
3.5	30
1.5	20
9.2	88
5.5	60
8.3	81
2.7	25
7.7	85
5.9	62
4.5	41
3.3	42
1.1	17
8.9	95
2.5	30

EXPLORATORY DATA ANALYSIS (EDA)

- Tujuan EDA:
- Memahami pola data dan mencari hubungan antara fitur (Study Hours) dan target (Scores).
- Mencari apakah ada outliers atau kesalahan dalam data.
- Visualisasi Data:
- Scatter Plot: Tampilkan scatter plot yang menunjukkan hubungan antara Study Hours dan Scores.
- Deskripsi Statistik: Rata-rata, median, standar deviasi dari Study Hours dan Scores.
- Insight dari EDA:
- Terdapat hubungan positif antara jam belajar dan skor.
- Outliers: Beberapa data yang jauh di luar pola umum, seperti siswa yang belajar sangat sedikit tapi mendapatkan skor tinggi, atau sebaliknya.

MACHINE LEARNING MODELS



Model yang Digunakan:

- Linear Regression: Model yang mencari hubungan linear antara Study Hours dan Scores.
- Decision Tree Regressor: Model berbasis pohon keputusan yang membagi data berdasarkan fitur untuk memprediksi skor.
- Random Forest Regressor: Model ensemble yang menggunakan banyak pohon keputusan untuk meningkatkan akurasi.



EVALUASI MODEL

- Evaluasi Model:
- Linear Regression:
 - R² Score: [nilai R²]
 - Mean Squared Error: [nilai MSE]
- Decision Tree Regressor:
 - R² Score: [nilai R²]
 - Mean Squared Error: [nilai MSE]
- Random Forest Regressor:
 - R² Score: [nilai R²]
 - Mean Squared Error: [nilai MSE]
- Hasil Terbaik:
- Model terbaik adalah [sebutkan model terbaik] dengan R² score = [nilai terbaik].

**SEKIAN
TERIMAKASIH**