

# Analisis Prediksi dengan Machine Learning

Alif Fauzan

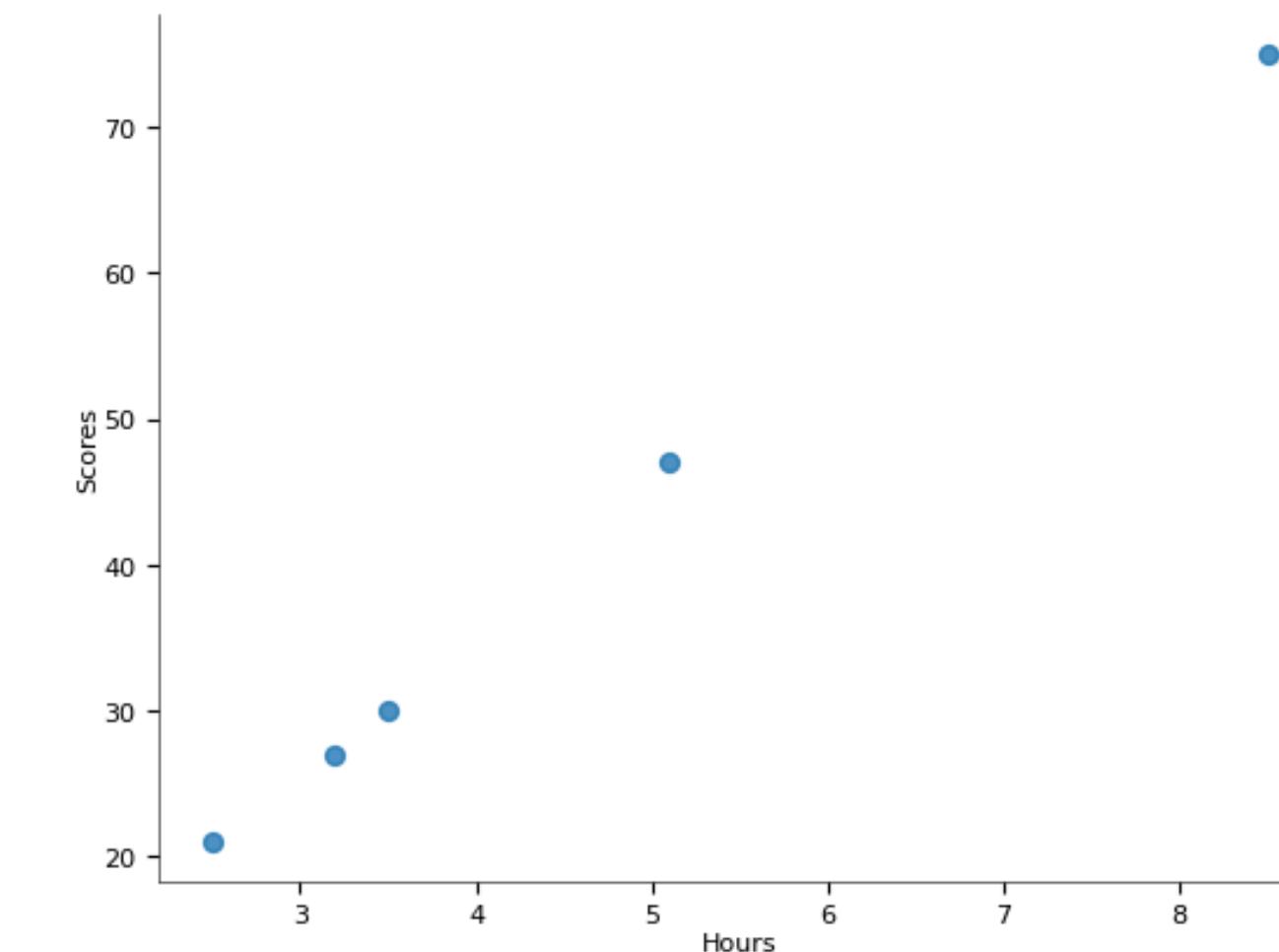
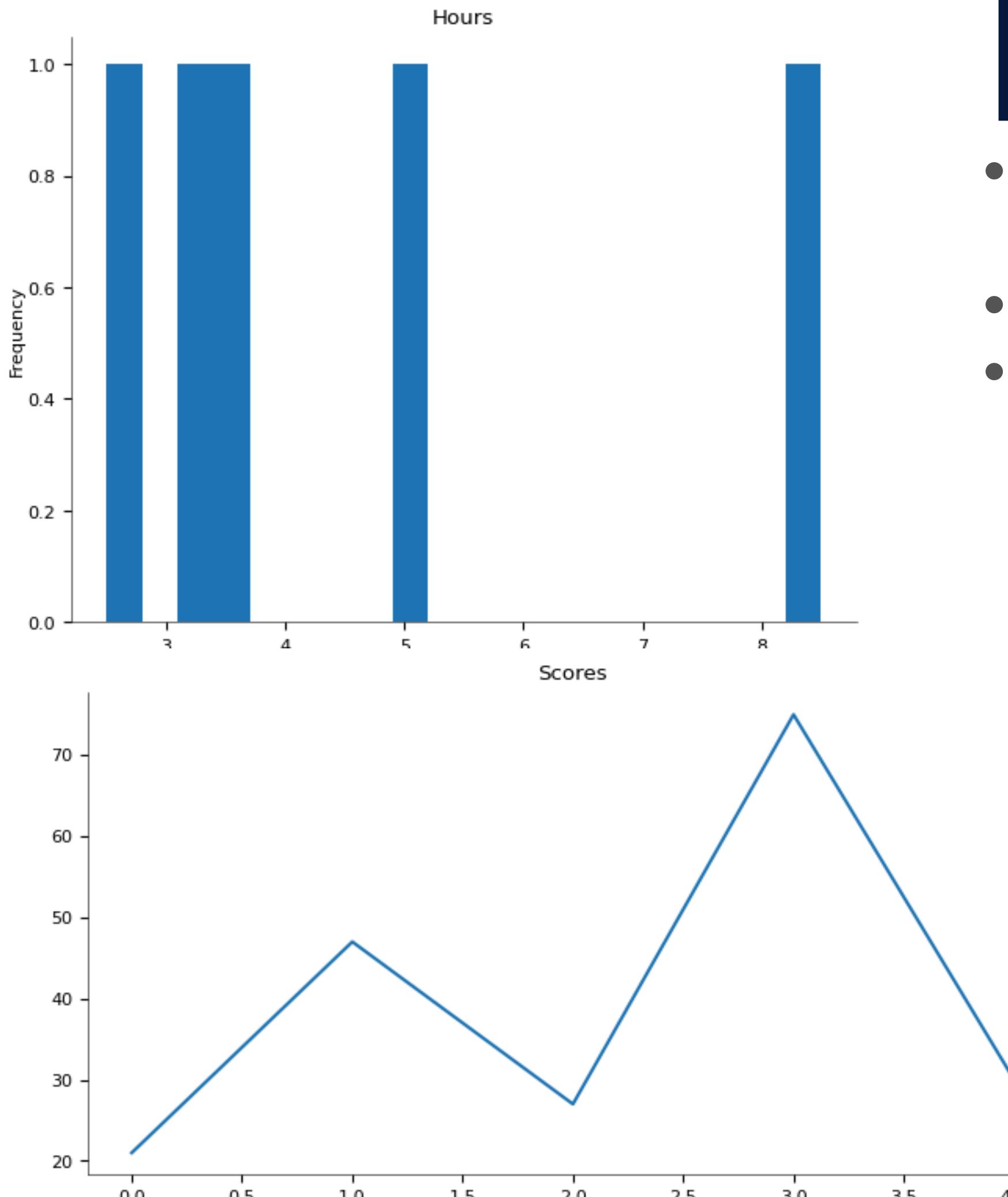
*Studi kasus analisis prediksi berdasarkan  
dataset yang diberikan.*

# Latar Belakang

- Masalah: Prediksi nilai berdasarkan jumlah jam belajar.
- Tujuan: Menganalisis hubungan antara jam belajar dan nilai dengan berbagai model machine learning.
- Manfaat: Membantu memahami efektivitas waktu belajar terhadap pencapaian akademik.

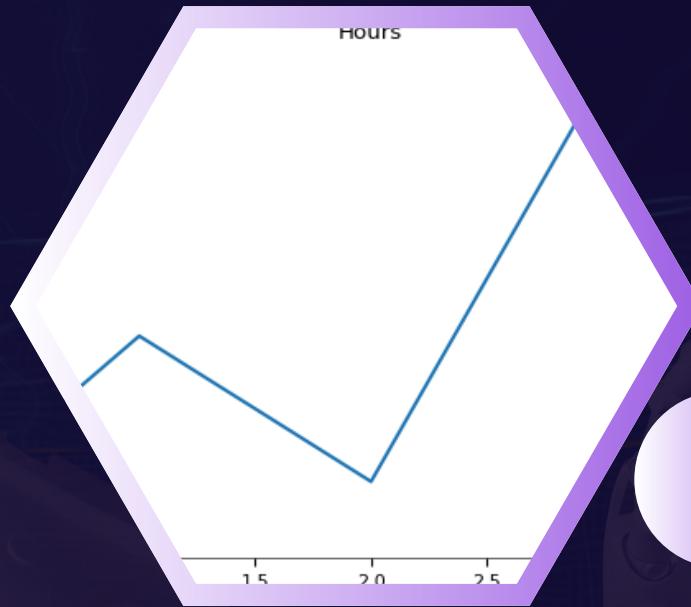
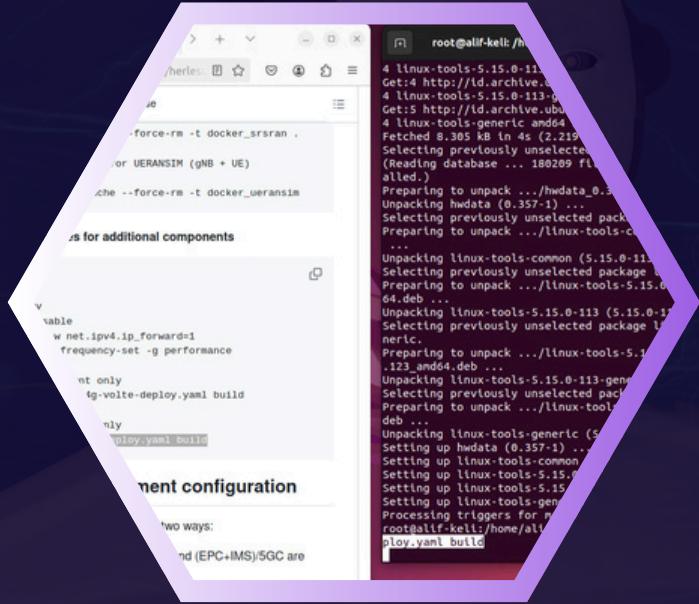
# Dataset

- Deskripsi dataset: Dataset berisi kolom Hours (jam belajar) dan Scores (nilai ujian).
- Jumlah Data: [2 kolom 25 baris]
- Pembersihan Data:
  - Tidak ada data duplikat
  - Tidak ada nilai yang hilang (missing values)

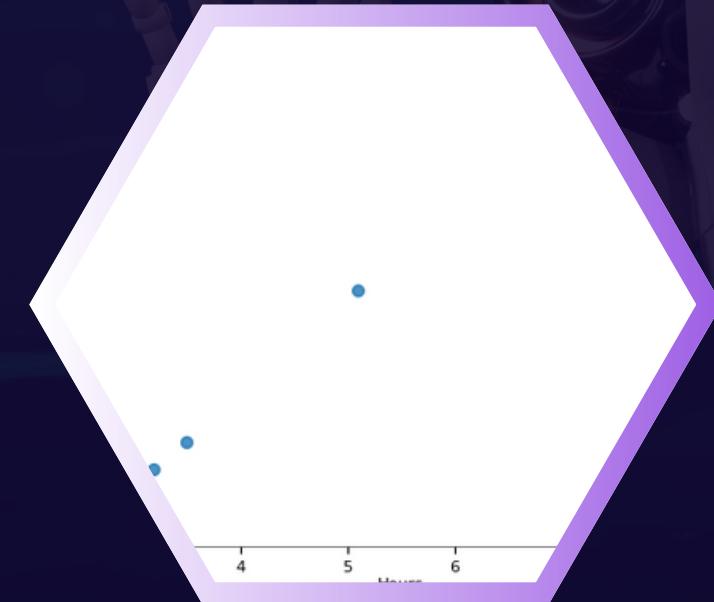


# Flowchart Proses Analisis

04



Pengumpulan data → Pembersihan data → Eksplorasi Data → Pemilihan Model → Pelatihan Model →



```
DecisionTreeRegressor
Decision Tree Regr Loading...
reeRegressor(random_state=42)

X_train, y_train)
data uji
model_dt.predict(X_test)

1
absolute_error(y_test, y_pred_dt)
ared_error(y_test, y_pred_dt)
test, y_pred_dt)

Regressor - MAE: {mae_dt}, MSE: {ms
- MAE: 5.4, MSE: 31.7, R2 Score:
```

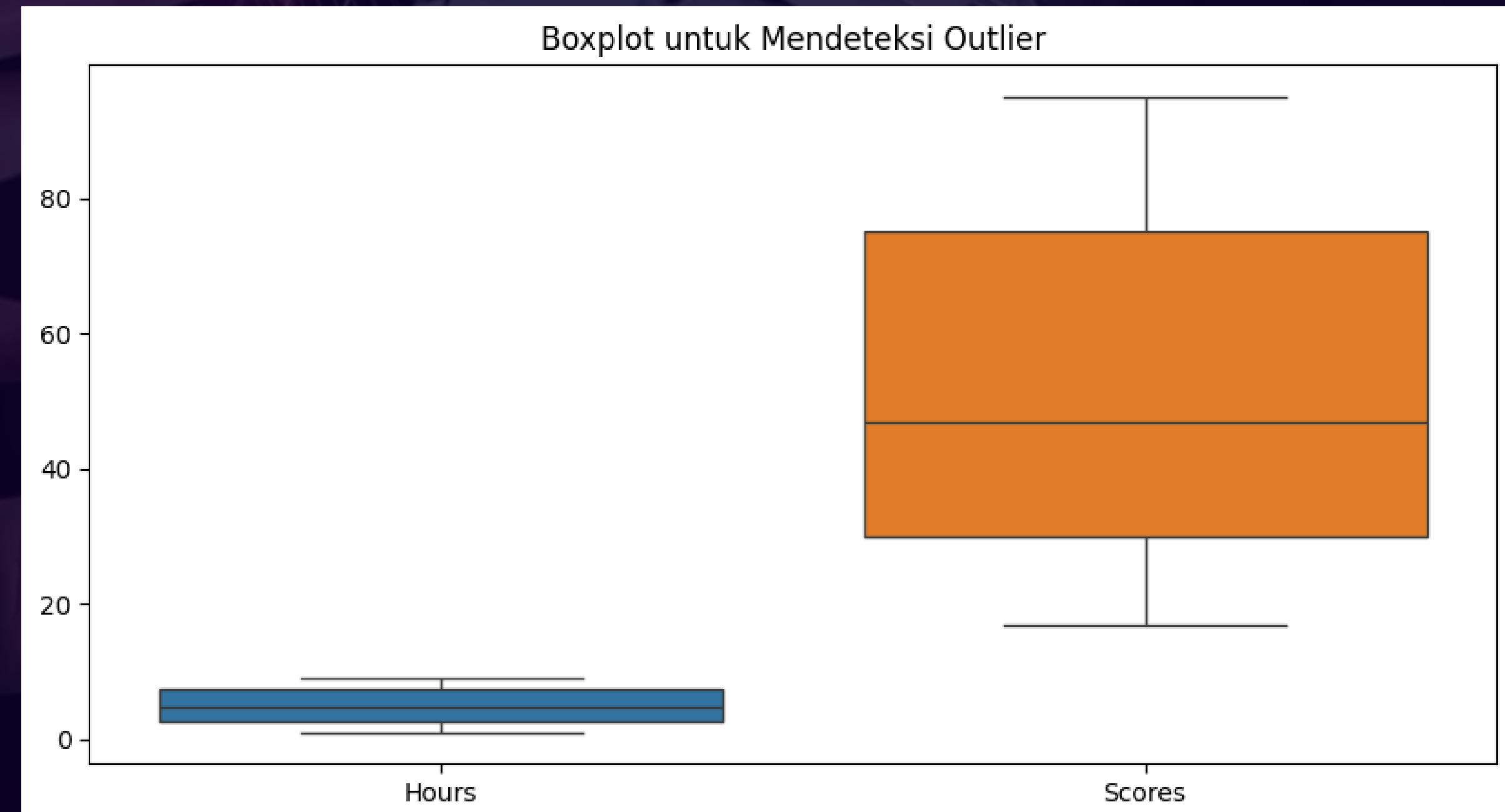
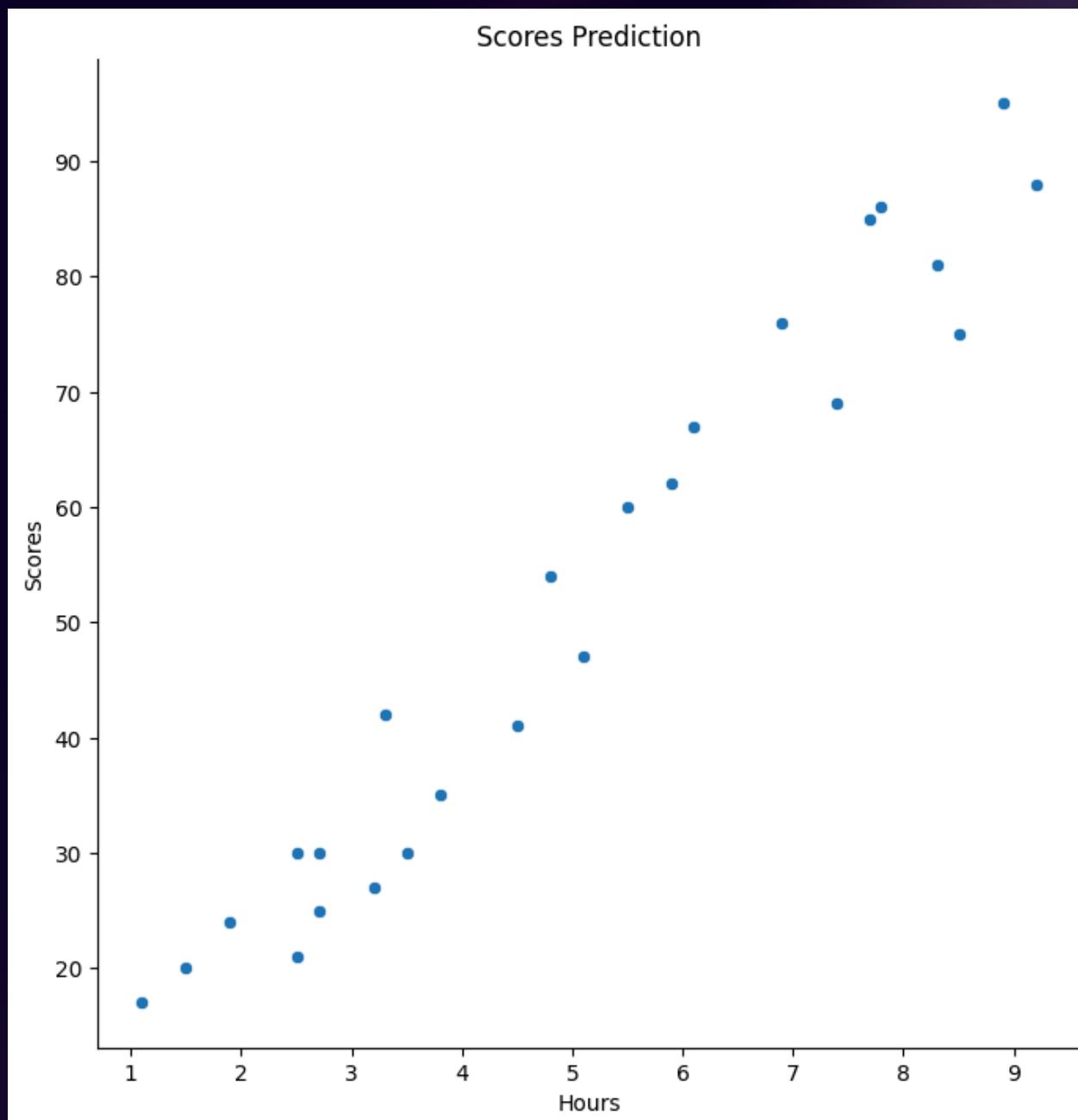
Evaluasi Model

Prediksi

# Visualisasi Data

05

- Boxplot: Deteksi outlier pada dataset.
- Scatter Plot: Hubungan antara jam belajar dan nilai ujian.



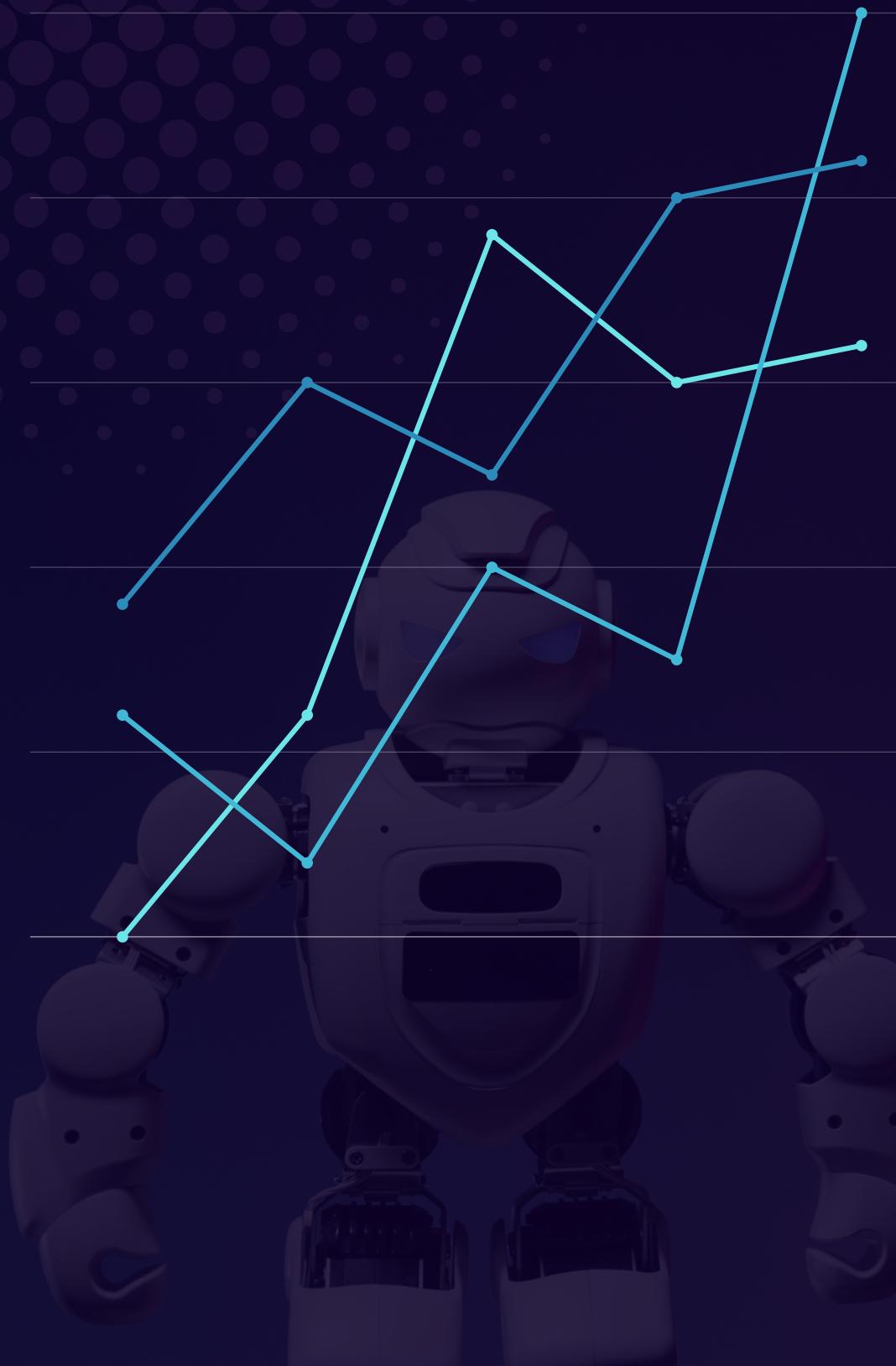
# Model yang Digunakan

06

- Linear Regression
- Decision Tree Regressor
- Random Forest Regressor



# Evaluasi Model



Model	MAE	MSE	R <sup>2</sup> Score
Linear Regression	3.92	18.94	0.96
Decision Tree Regressor	5.4	31.7	0.94
Random Forest Regressor	2.77	13.04	0.97

Random Forest Regressor - MAE: 2.778, MSE: 13.045153611111104, R2 Score: 0.9778294466160586

Decision Tree Regressor - MAE: 5.4, MSE: 31.7, R2 Score: 0.9461250849762066

Linear Regression - MAE: 3.9207511902099244, MSE: 18.943211722315272, R2 Score: 0.9678055545167994

# Hasil dan Analisis

- Model terbaik: Random Forest Regressor dengan R<sup>2</sup> Score = 0.97
- Insight:
  - Model ini lebih akurat dibandingkan Linear Regression dan Decision Tree.
  - Outlier berpotensi mempengaruhi hasil prediksi.

# Kesimpulan & Pengembangan Selanjutnya

## Kesimpulan:

- Model machine learning dapat memprediksi nilai berdasarkan jumlah jam belajar.
- Random Forest memberikan performa terbaik.
- Pengembangan:
  - Menambahkan fitur lain (misalnya metode belajar, kondisi psikologis, dsb.).
  - Menggunakan teknik preprocessing lebih lanjut untuk meningkatkan akurasi.

# Terima Kasih

Q&A



Alif Fauzan



fauzanalif738@gmail.com