



TUGAS UTS MACHINE LEARNING

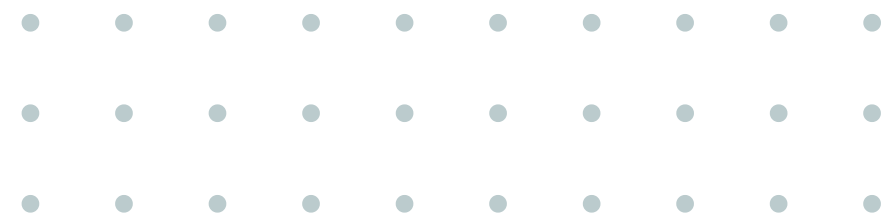
Muhamad Hilmi Haidar
1103213005





01. REGRESSION MODEL

02. CLASSIFICATION MODEL



**TABLE OF
CONTENT**



01.

REGRESSION MODEL



TAHAPAN YANG DILAKUKAN

PREPROCESSING

Menghilangkan nilai duplikat, memberinama pada setiap kolom, dan menentukan target dan fiturnya berdasarkan korelasi

1

EDA

Memvisualisasikan distribusi, korelasi, dan boxplot.

2

MODELLING

Melakukan modelling pipeline serta hyperparameter tuning

3

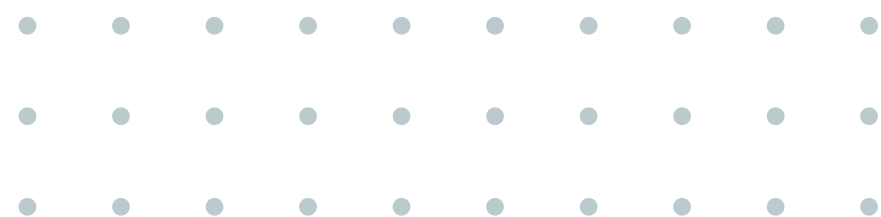
HASIL EVALUASI YANG DIDAPAT

Best Polynomial Regression Model Metrics:
MSE: 0.0608, RMSE: 0.2467, MAE: 0.1885, R^2 : 0.7463

Best Decision Tree Model Metrics:
MSE: 0.0790, RMSE: 0.2811, MAE: 0.2156, R^2 : 0.6705

Best k-Nearest Neighbors Model Metrics:
MSE: 0.0772, RMSE: 0.2779, MAE: 0.2140, R^2 : 0.6779

Best XGBoost Model Metrics:
MSE: 0.0528, RMSE: 0.2297, MAE: 0.1783, R^2 : 0.7799



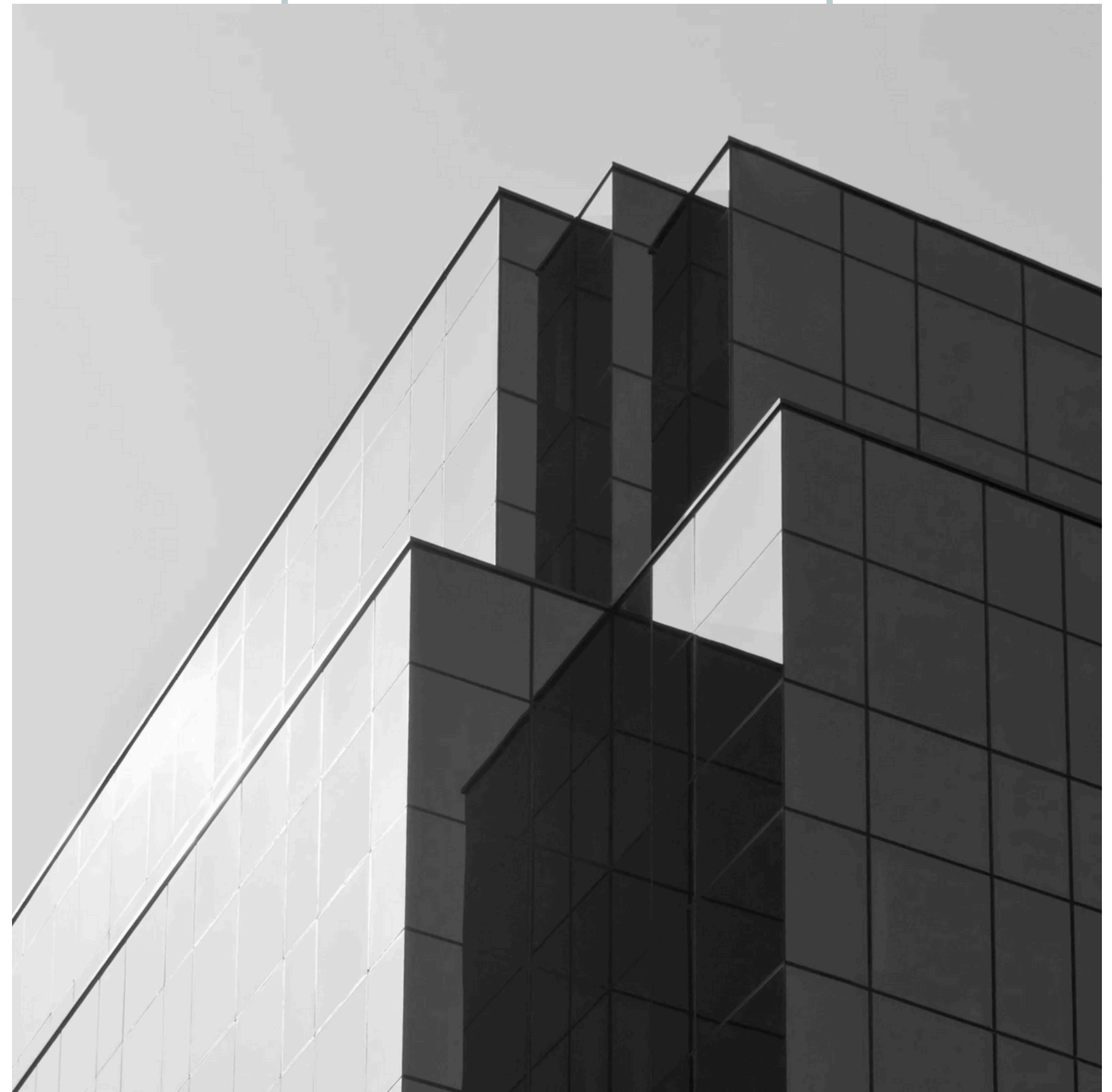
KESIMPULAN

Kesimpulan dari output ini menunjukkan bahwa XGBoost Regression memberikan performa terbaik dibandingkan model lainnya. Dengan MSE 0.0528, RMSE 0.2297, MAE 0.1783, dan R^2 0.7799, XGBoost memiliki error paling rendah dan kemampuan prediktif terbaik dalam menjelaskan variabilitas data.

Polynomial Regression berada di urutan kedua dengan MSE 0.0608, RMSE 0.2467, MAE 0.1885, dan R^2 0.7463, menunjukkan performa yang baik tetapi tidak seakurat XGBoost.

k-Nearest Neighbors (KNN) mencatat MSE 0.0772, RMSE 0.2779, MAE 0.2140, dan R^2 0.6779, sedikit lebih baik dibandingkan Decision Tree, yang memiliki MSE 0.0790, RMSE 0.2811, MAE 0.2156, dan R^2 0.6705.

Secara keseluruhan, XGBoost adalah pilihan terbaik untuk meminimalkan error dan menghasilkan prediksi yang paling akurat.



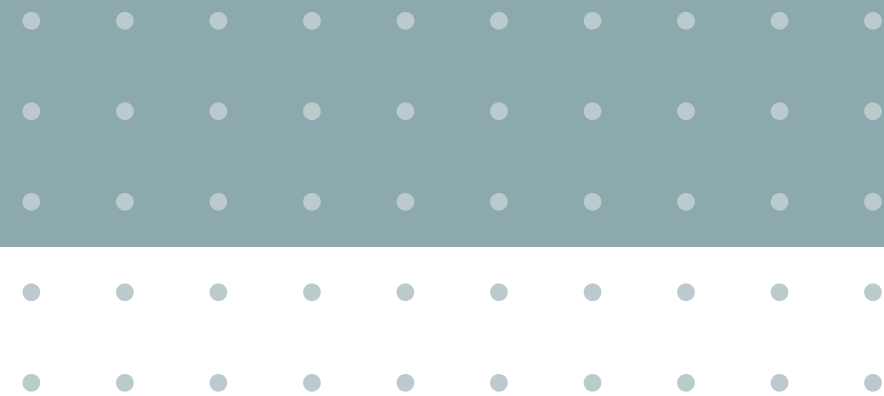


HOW?



02.

CLASSIFICATION MODEL



TAHAPAN YANG DILAKUKAN

EDA

Memvisualisasikan distribusi, korelasi, dan boxplot.

1

MODELLING

Melakukan modelling pipeline untuk model klasifikasi

2

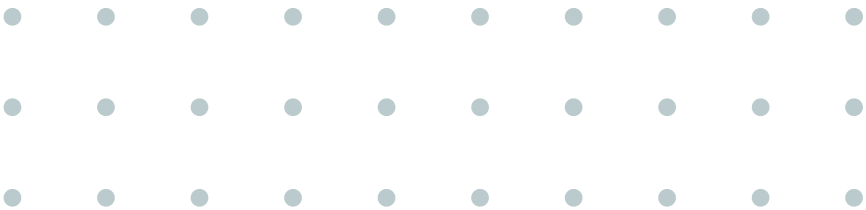
TUNING

Melakukan tuning hyperparameter untuk model-model

3

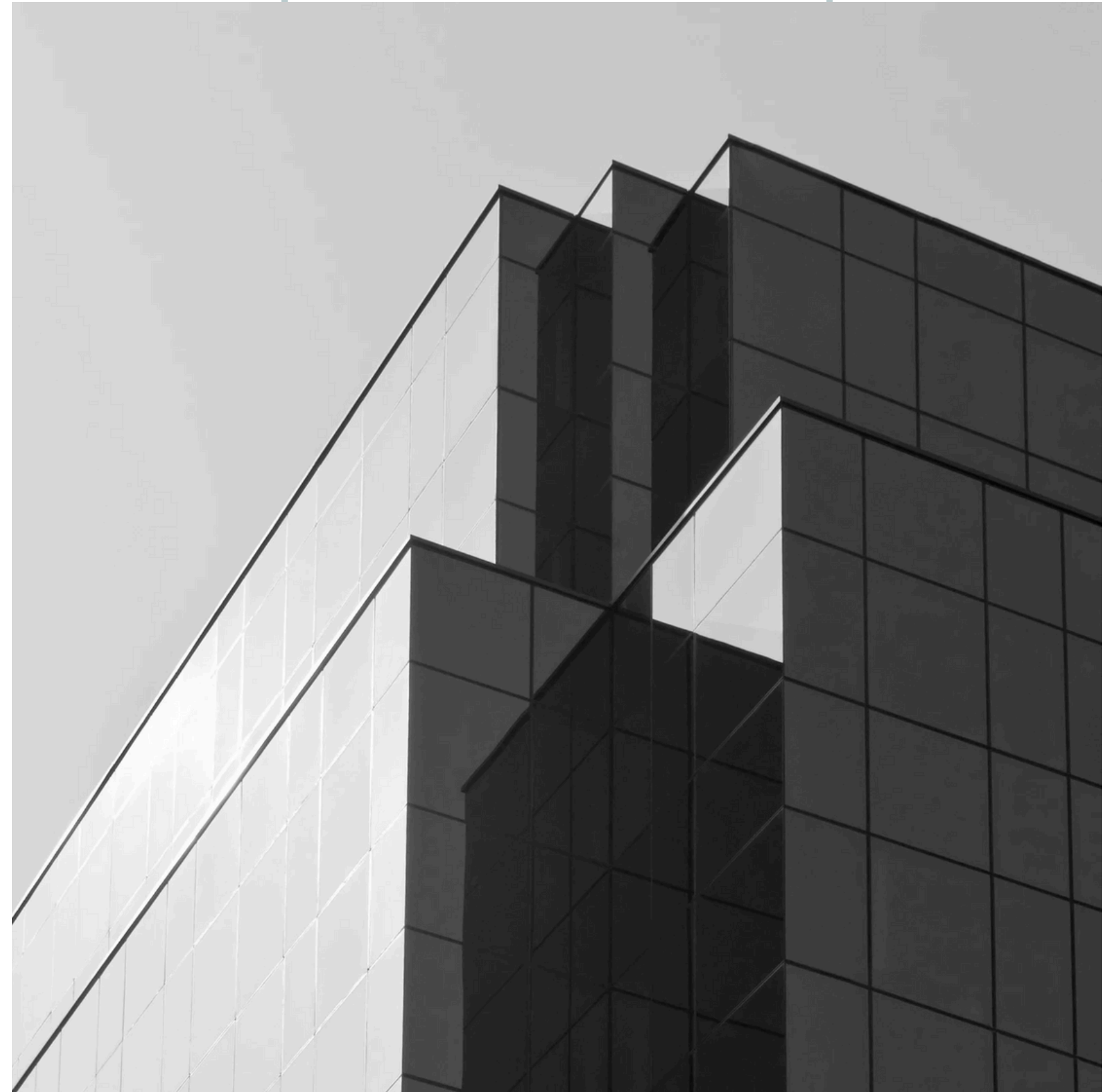
HASIL EVALUASI YANG DIDAPAT

Hasil Evaluasi Setelah Hyperparameter Tuning:						
	Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
0	Logistic Regression	0.916592	0.953184	0.940963	0.947034	0.948345
1	Decision Tree	0.999265	0.999768	0.999305	0.999536	0.999323
2	k-NN	0.999490	0.999923	0.999434	0.999678	0.999670
3	XGBoost	0.999510	0.999897	0.999485	0.999691	0.999976



KESIMPULAN

Hyperparameter tuning menghasilkan sedikit peningkatan pada beberapa model, terutama k-NN dan XGBoost, yang kini menunjukkan performa terbaik dengan akurasi, F1-Score, dan AUC-ROC hampir sempurna. Sementara itu, Logistic Regression tidak mengalami peningkatan signifikan, dan Decision Tree tetap sangat akurat tetapi memiliki risiko overfitting. Secara keseluruhan, XGBoost menjadi pilihan terbaik berdasarkan hasil evaluasi setelah tuning.





HOW?





THANK YOU

