

A decorative graphic on the right side of the slide, composed of numerous thin, red, wavy lines that create a sense of movement and depth, resembling a stylized wave or a series of overlapping curves.

TRAVEL & TOURISM **PREDICTION (STAGE 2)**

ABOUT US

Unison (Unity Scientist One) adalah komunitas data scientist yang dibentuk dalam rangka menyelesaikan tugas bootcamp di Rakamin Academy. Kami terdiri dari individu-individu yang bersemangat dalam analisis data, machine learning, dan artificial intelligence, dengan tujuan mengembangkan keterampilan dan menciptakan solusi berbasis data yang berdampak nyata.

UNISON

MEET THE TEAM



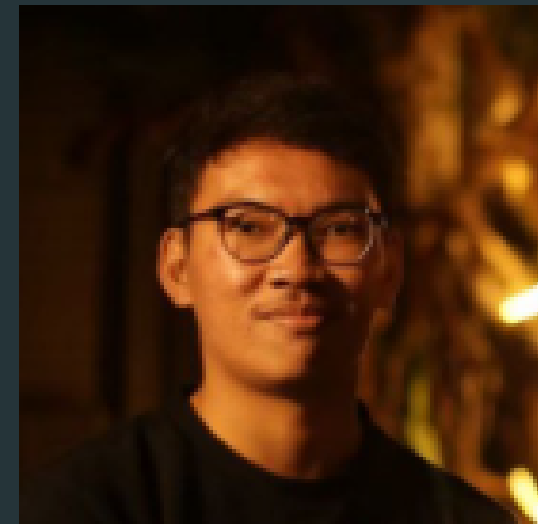
Farah Nelia Satriani

Project manager



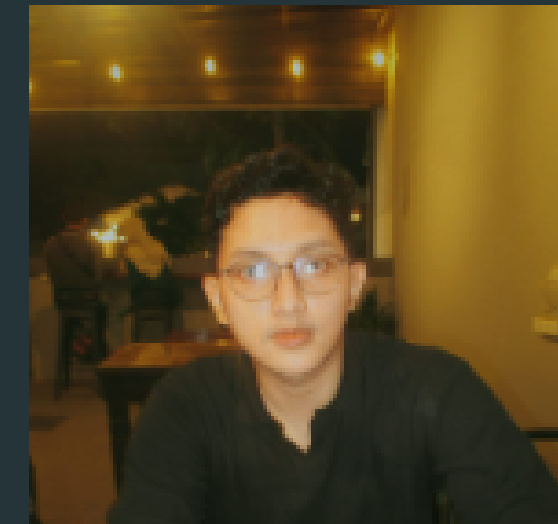
Sutan Faisal

Data Scientist



Hafizh Fakhri Alfarisy

Data Engineer



M. Ath Thaariq A. C

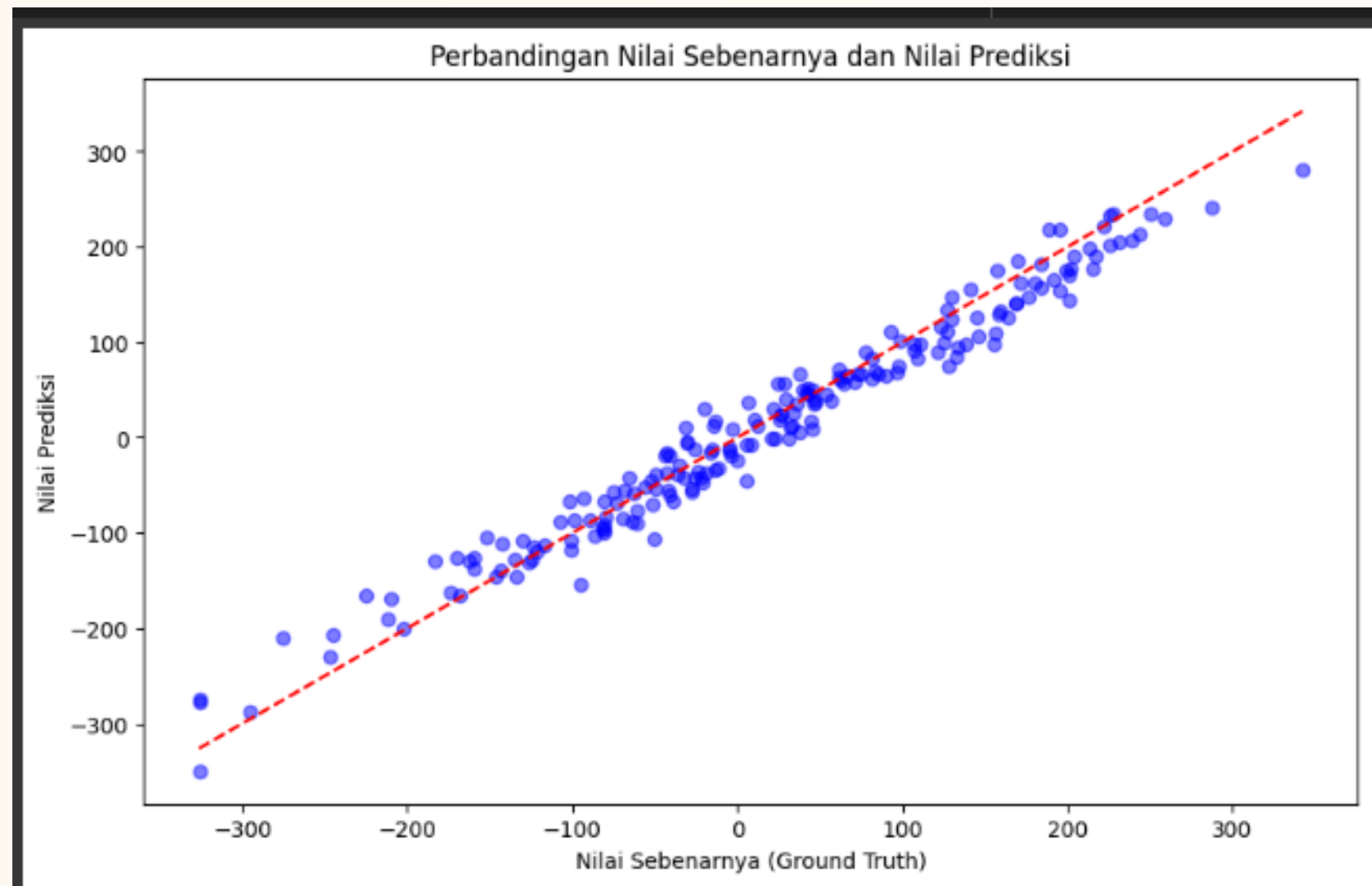
Data Analyst

Final Model Testing

1. Digunakan XGBoost Regresor untuk memprediksi nilai dari data.
2. Data dievaluasi menggunakan MAE, RMSR, dan MAPE. (Default Hypertuning)
3. Validasi kinerja menggunakan K-Fold Cross Validation untuk menilai kestabilan performa dari model
4. Meninjau hasil dari evaluasi
5. Melakukan perbandingan Nilai Sebenarnya dan Nilai Prediksi

```
Mean Absolute Error (MAE) for XGBoost: 33.272145604976856  
Root Mean Squared Error (RMSE) for XGBoost: 41.34799507599749  
Mean Absolute Percentage Error (MAPE) for XGBoost: 99.6015922003529  
K-Fold Cross-Validation MAE for XGBoost: 39.409734619147876  
K-Fold Cross-Validation RMSE for XGBoost: 51.74705559115449  
K-Fold Cross-Validation MAPE for XGBoost: 107.98479892662618
```

Final Model Testing



1. Grafik ini berguna untuk mengevaluasi seberapa baik model memprediksi nilai yang sebenarnya. Jika titik-titik berada sangat dekat dengan garis merah, itu menandakan bahwa model memprediksi dengan akurat. Jika titik-titik tersebar jauh dari garis merah, itu menunjukkan bahwa model kurang akurat dalam memprediksi.

Discuss Model Evaluation

Dilakukan evaluasi terhadap nilai testnya didapat hasil sebagai berikut :

```
RMSE pada dataset uji: 25.00907384585698  
MAE pada dataset uji: 20.30564023030516  
MAPE pada dataset uji: 74.0909036674338%  
RMSE pada dataset pelatihan: 8.957257004307746  
MAE pada dataset pelatihan: 6.94261643314077  
MAPE pada dataset pelatihan: 27.791858241369894%  
Model kemungkinan mengalami overfitting
```

Error Analysis

1. Linear Regression:

MAE: 37.71

RMSE: 59.83

MAPE: 33.65

2. Random Forest:

MAE: 34.35

RMSE: 55.99

MAPE: 28.21

3. Gradient Boosting:

MAE: 34.04

RMSE: 56.56

MAPE: 28.35

4. XGBoost:

MAE: 52.46

RMSE: 66.48

MAPE: 156.92

Model Terbaik Berdasarkan Hasil:

Random Forest adalah model terbaik berdasarkan MAE, RMSE, dan MAPE. Ini memiliki MAE terkecil (34.35), RMSE terkecil (55.99), dan MAPE terkecil (28.21), yang menunjukkan bahwa Random Forest menghasilkan prediksi harga yang lebih akurat daripada model lainnya.

XGBoost memiliki nilai MAE dan RMSE yang lebih tinggi, serta MAPE yang sangat tinggi, yang menandakan bahwa model ini kurang akurat dibandingkan yang lainnya.

Kesimpulan:

Random Forest setelah GridSearchCV atau RandomizedSearchCV tampaknya memberikan hasil terbaik berdasarkan kinerja pada semua metrik utama.

TERIMAKASIH |

LINK GOOGLE COLAB :