



# AI | Machine Learning

with final project: salary prediction

# What is AI/ML

Artificial Intelligence (AI) adalah simulasi proses kecerdasan manusia dalam mesin yang diprogram untuk berpikir seperti manusia dan meniru tindakannya. AI memungkinkan sistem untuk belajar dari data, menyesuaikan diri dengan situasi baru, dan menyelesaikan masalah.

Machine Learning (ML) adalah subbidang dari AI yang memungkinkan sistem komputer belajar dari data tanpa diprogram secara eksplisit. Algoritma ML dapat mengidentifikasi pola dalam data, membuat prediksi, dan mengambil keputusan.

# ML Chatagory

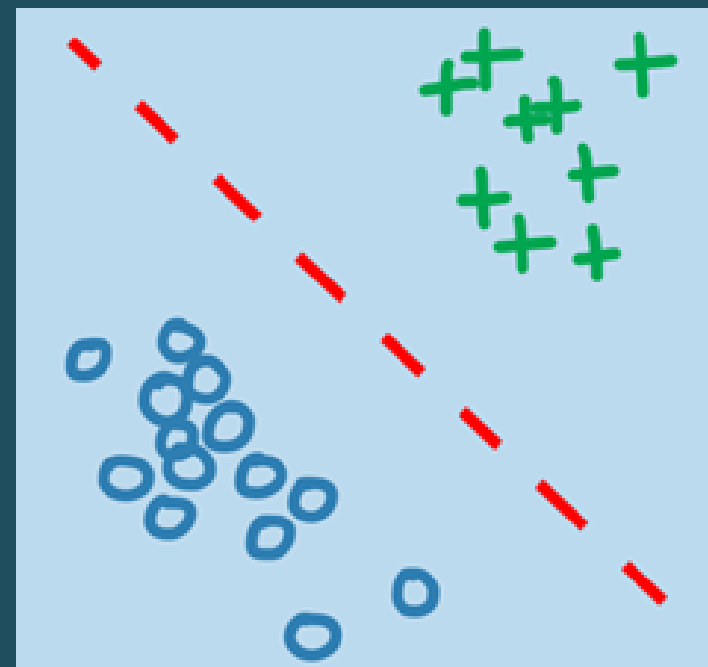
## 1. Unsupervised Learning

Model mencari pola dalam data tanpa adanya label. C



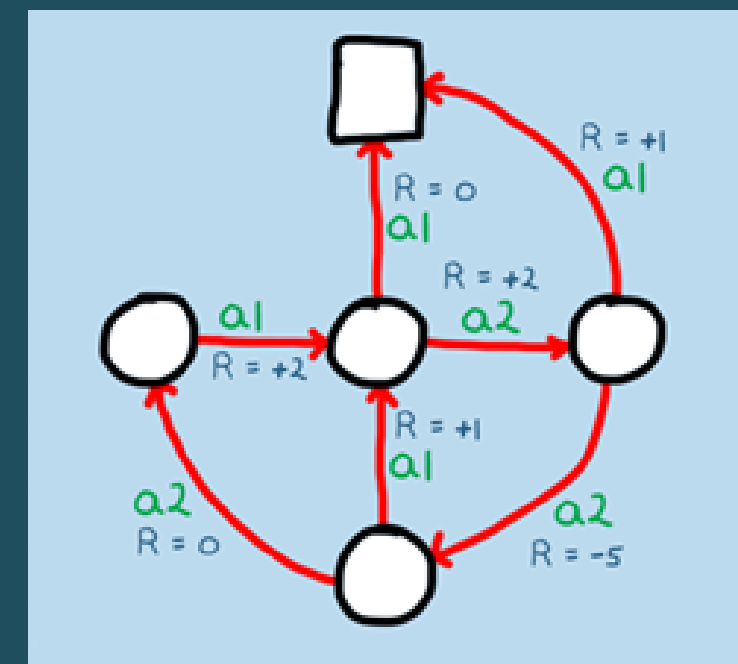
## 2. Supervised Learning

Model mencari pola dalam data tanpa adanya label.



## 3. Reinforcement Learning

Model belajar dengan dapat reward atau punishment berdasarkan tindakannya.



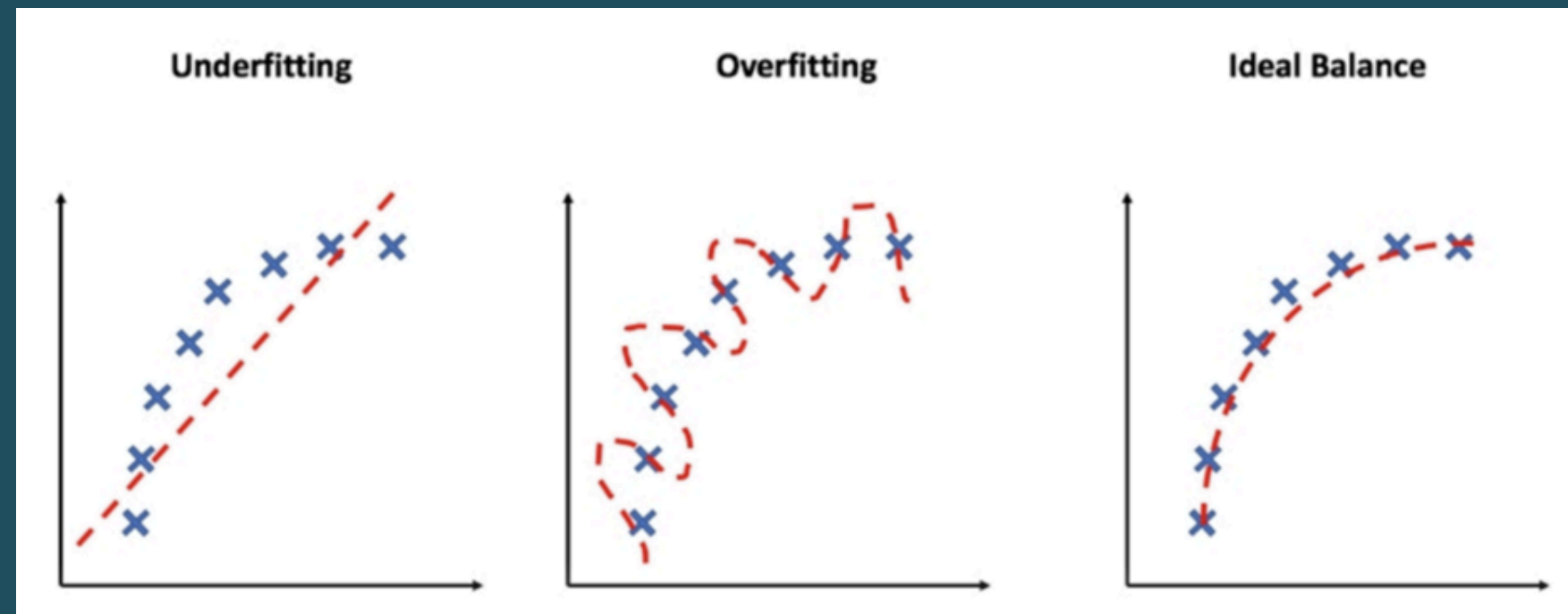
# Machine Learning procces



# What is underfitting and Overfitting

**Overfitting** adalah kondisi di mana model machine learning terlalu "hafal" data pelatihan. Akibatnya, model menjadi terlalu kompleks dan tidak dapat memprediksi data baru dengan baik.

**Underfitting** adalah kondisi di mana model machine learning terlalu sederhana. Model ini tidak mampu menangkap pola yang ada dalam data, sehingga performanya buruk baik pada data pelatihan maupun data baru

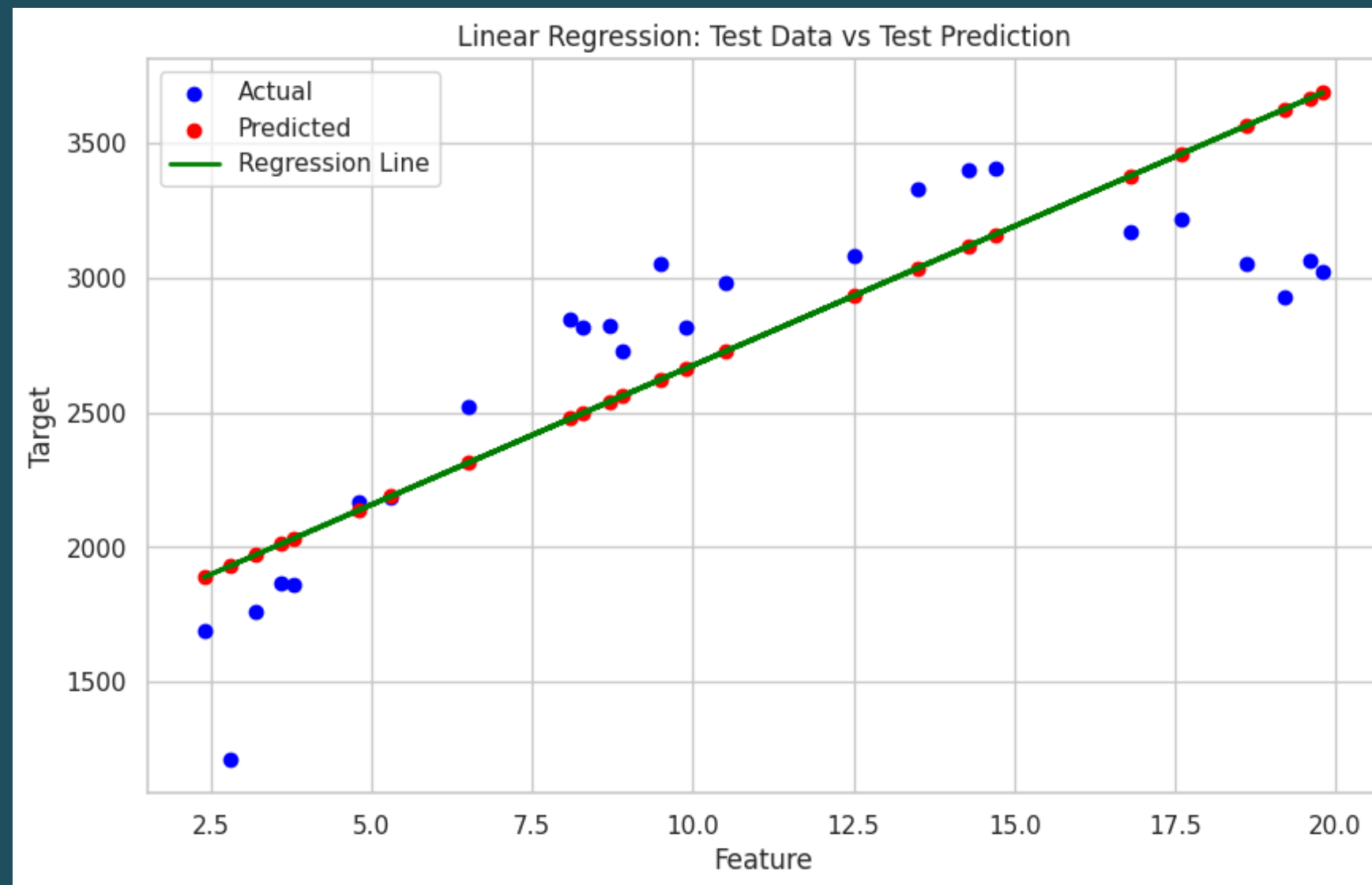




# Final Project: Supervised- Salary prediction

 [https://github.com/amdihsann/Supervised\\_Regression\\_SalaryPrediction](https://github.com/amdihsann/Supervised_Regression_SalaryPrediction)

# Salary prediction-Linear Regression



## Results

Mean Squared Error:

Train: 107699.85

Test : 128111.12

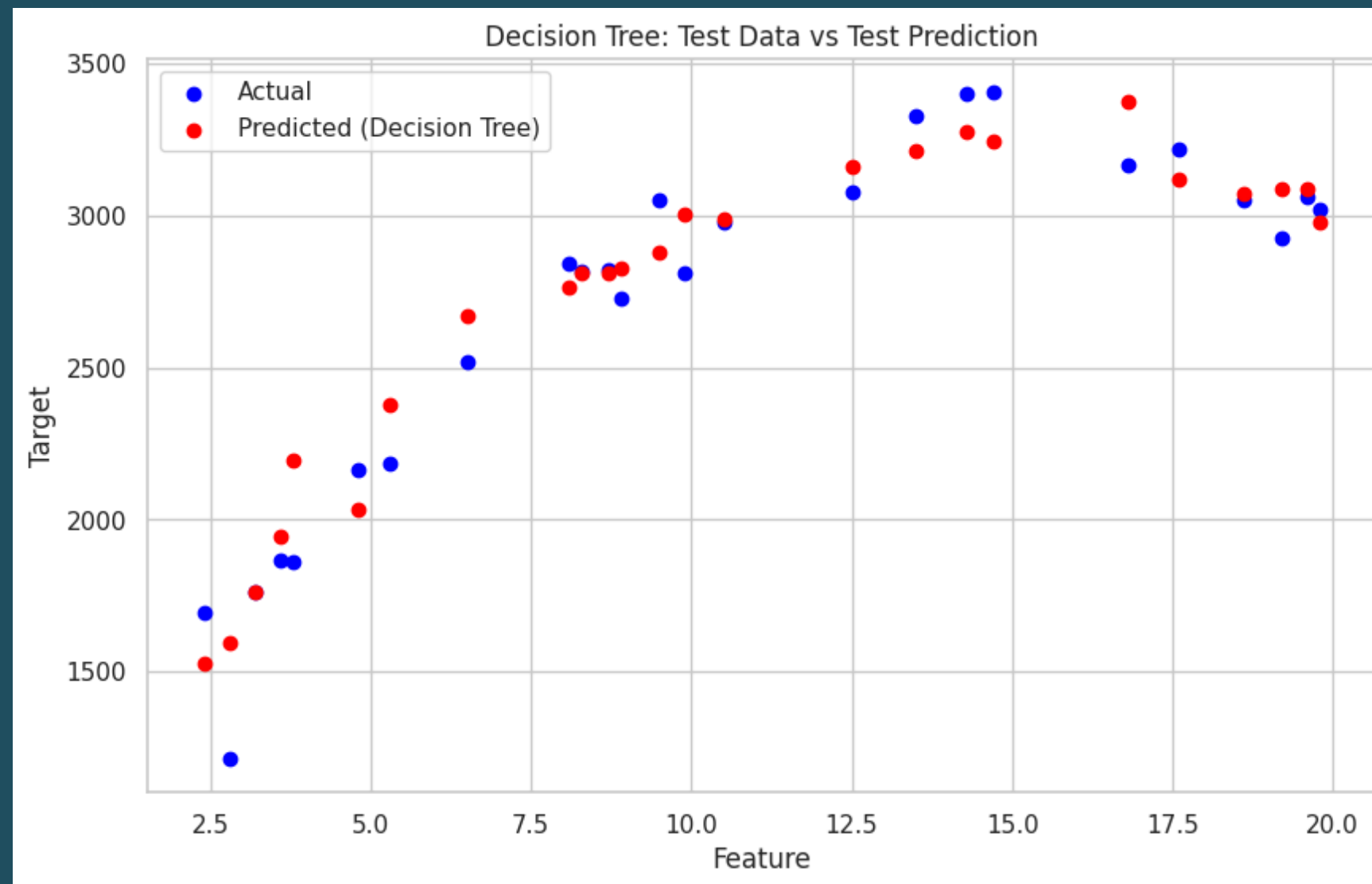
Gap : 20411.27

R<sup>2</sup> Score:

Train: 0.77

Test : 0.63

# Salary prediction-Decision Tree



## Results

Mean Squared Error:

Train: 88.12

Test : 23627.99

Gap : 23539.87

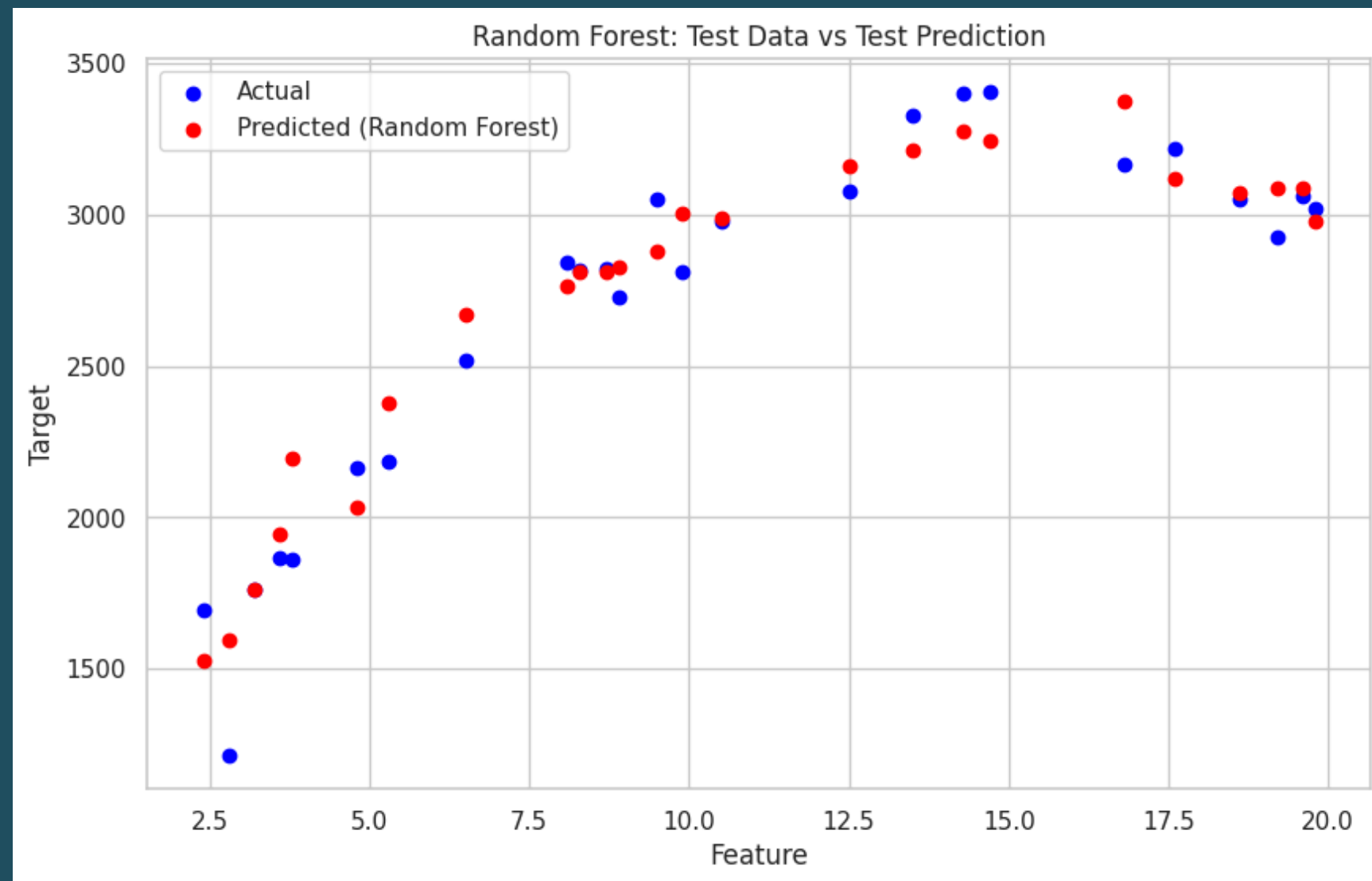
R<sup>2</sup> Score:

Train: 1.00

Test : 0.93



# Salary prediction-Random Forest



## Results

Mean Squared Error:

Train: 3737.44

Test : 21744.73

Gap : 18007.29

R<sup>2</sup> Score:

Train: 0.99

Test : 0.94

# Comperation

## Linear-regression

### Results

Mean Squared Error:

Train: 107699.85

Test : 128111.12

Gap : 20411.27

R<sup>2</sup> Score:

Train: 0.77

Test : 0.63

## Decision Tree

### Results

Mean Squared Error:

Train: 88.12

Test : 23627.99

Gap : 23539.87

R<sup>2</sup> Score:

Train: 1.00

Test : 0.93

## Random Forest

### Results

Mean Squared Error:

Train: 3737.44

Test : 21744.73

Gap : 18007.29

R<sup>2</sup> Score:

Train: 0.99

Test : 0.94

# Conclusion

Di antara ketiga model yang diuji, Random Forest menunjukkan kinerja terbaik karena beberapa alasan berikut:

1. **Performa yang Konsisten:** Random Forest memiliki nilai  $R^2$  yang tinggi pada data latih (0.99) dan data uji (0.94), yang menunjukkan model mampu memprediksi dengan akurasi tinggi.
2. **Error yang Lebih Rendah:** Mean Squared Error (MSE) pada data uji lebih rendah (21,744.73) dibandingkan Decision Tree, sehingga memberikan hasil prediksi yang lebih akurat.
3. **Keseimbangan Optimal:** Random Forest memiliki gap error yang kecil antara data latih dan uji (18,007.29), menandakan model tidak mengalami overfitting seperti Decision Tree atau underfitting seperti Linear Regression.



# Thank You

Feel free to reach out if you'd like to collaborate,  
discuss ideas, or explore exciting opportunities.



ahmadihsan506@gmail.com



<https://www.linkedin.com/in/ahmadihsan-/>



<https://github.com/amdihsann>