

Laporan UTS

Kyan Dillan Verado – 223400001

Tugas 1

```
Multiple Linear Regression:  
R-squared: 0.5925593134440169  
RMSE: 8511116020.235831  
MSE: 7.243909590991502e+19  
  
Polynomial Regression:  
R-squared: 0.41277065547439284  
RMSE: 9352097086.462294  
MSE: 8.74617199146165e+19
```

Gambar diatas adalah hasil perhitungan evaluasi model **Multiple Linear Regression** dan **Polynomial Regression**.

1. R-squared

Semakin nilai R-squared mendekati angka 1, maka menunjukkan semakin baik model dalam menjelaskan varians variabel dependen (HARGA) menggunakan variabel independen (LT, LB, JKT, JKM, GRS, KOTA). Dari analisa perhitungan gambar diatas, **Multiple Linear Regression** memiliki nilai yang lebih tinggi daripada **Polynomial Regression**.

2. RMSE

Semakin rendah nilai RMSE, maka semakin akurat model dalam melakukan prediksi. **Multiple Linear Regression** memiliki nilai yang lebih rendah daripada **Polynomial Regression**.

3. MSE

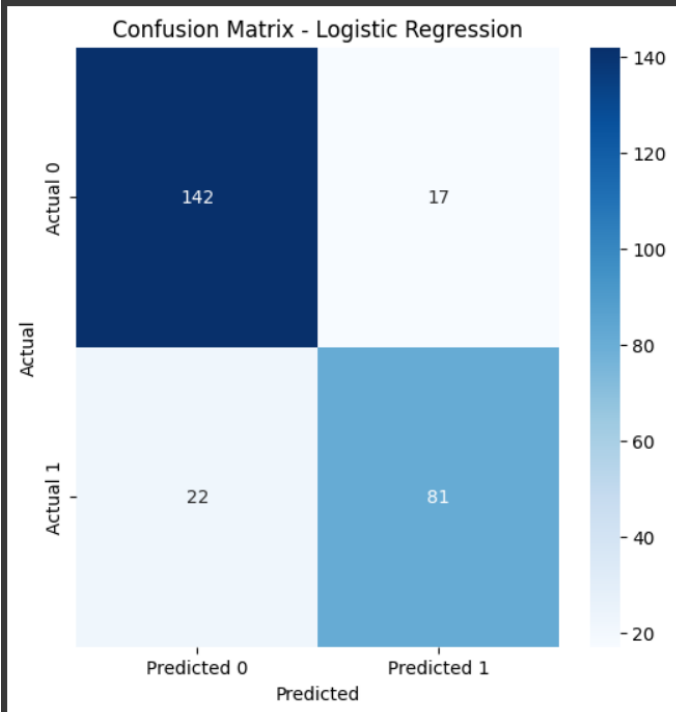
MSE mirip dengan RMSE, tetapi menggunakan kuadrat dari selisih. Nilai MSE yang lebih rendah menunjukkan model yang lebih akurat. **Multiple Linear Regression** memiliki nilai yang lebih rendah daripada **Polynomial Regression**.

4. Kesimpulan

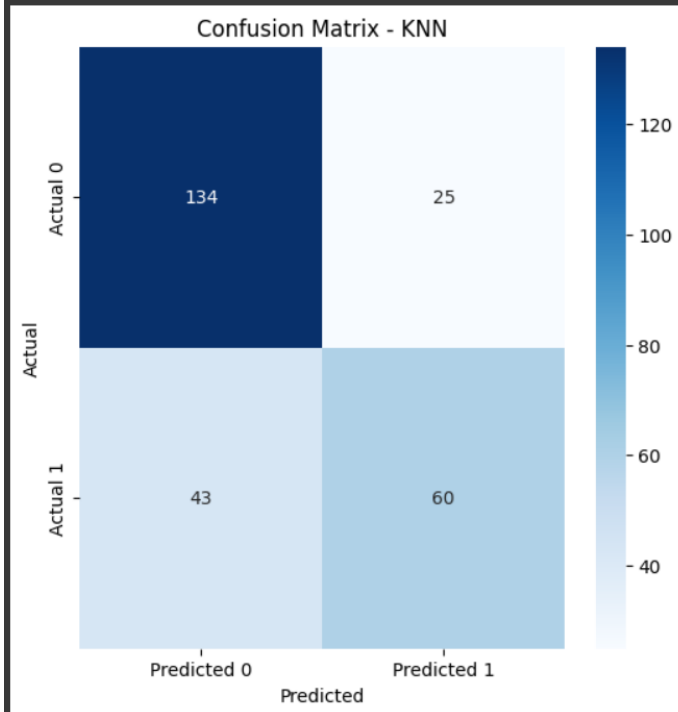
Dari 3 perhitungan evaluasi model yang diambil dari model **Multiple Linear Regression** dan **Polynomial Regression**, dapat dinyatakan perhitungan **Multiple Linear Regression** lebih cocok digunakan untuk dataset Harga Rumah Jaksel.

Tugas 2

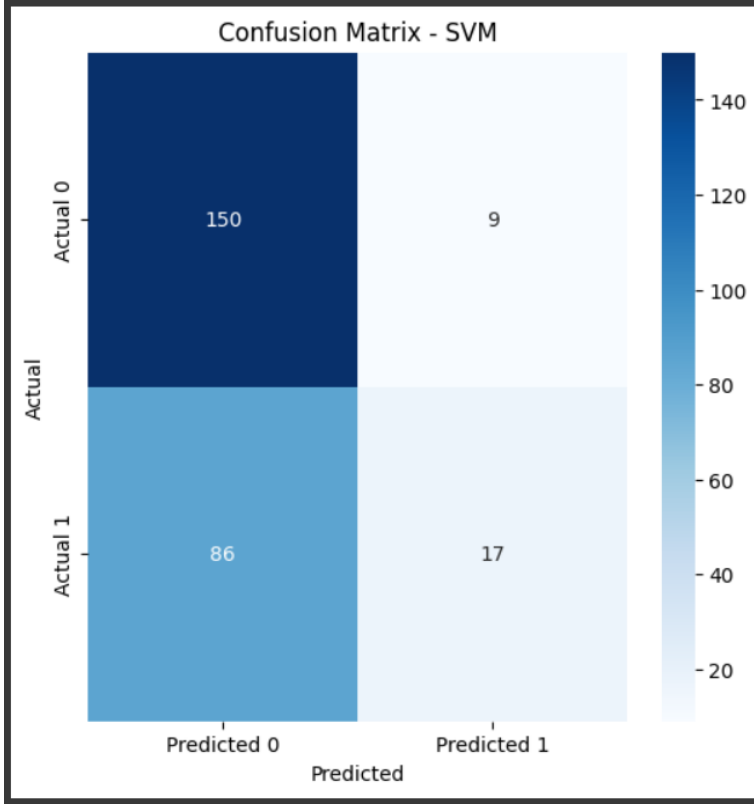
--- Logistic Regression ---
Accuracy: 0.851145038167939
Precision: 0.826530612244898
Recall: 0.7864077669902912
F1-score: 0.8059701492537313



--- KNN ---
Accuracy: 0.7404580152671756
Precision: 0.7058823529411765
Recall: 0.5825242718446602
F1-score: 0.6382978723404256



--- SVM ---
Accuracy: 0.6374045801526718
Precision: 0.6538461538461539
Recall: 0.1650485436893204
F1-score: 0.26356589147286824



Gambar di atas adalah hasil Evaluasi model perhitungan dari **logistic regression, kNN, dan SVM**. Fitur yang tidak digunakan adalah **PassengerId, Name, Ticket, dan Cabin**. Dan melakukan Encoding pada fitur **Sex dan Embarked**.

1. Accuracy

logistic regression = 0.851145038167939

kNN = 0.7404580152671756

SVM = 0.6374045801526718

Akurasi yaitu proporsi prediksi yang benar (baik positif maupun negatif) terhadap total jumlah data. Semakin mendekati angka 1, maka nilai semakin baik. Dari sisi Akurasi **logistic regression** memiliki nilai yang lebih baik dari yang lain.

2. Precision

logistic regression = 0.826530612244898

kNN = 0.7058823529411765

SVM = 0.6538461538461539

Presisi yaitu proporsi prediksi positif yang benar terhadap semua prediksi positif. Precision menunjukkan seberapa banyak prediksi positif yang benar. Semakin mendekati angka 1, maka nilai semakin baik. Dari sisi Presisi **logistic regression** memiliki nilai yang lebih baik dari yang lain.

3. Recall

logistic regression = 0.7864077669902912

kNN = 0.5825242718446602

SVM = 0.1650485436893204

Recall yaitu proporsi data positif yang teridentifikasi dengan benar oleh model. Semakin mendekati angka 1, maka nilai semakin baik. Dari sisi Recall **logistic regression** memiliki nilai yang lebih baik dari yang lain.

4. F1-score

logistic regression = 0.8059701492537313

kNN = 0.6382978723404256

SVM = 0.26356589147286824

F1-score yaitu rata-rata antara Precision dan Recall. F1-Score digunakan untuk menemukan keseimbangan antara Precision dan Recall, terutama pada dataset yang tidak seimbang. Semakin mendekati angka 1, maka nilai semakin baik. Dari sisi F1-score **logistic regression** memiliki nilai yang lebih baik dari yang lain.