

ORANGE MACHINE LEARNING MINGGU 1

Nama : Ferdinant Hutajulu

NIM : 1103213120

1. Upload File lalu setting Role parameternya

File - Orange

Source

☒ File: archive (1)\Salary_dataset.csv

☐ URL:

File Type

Automatically detect type

Info

30 instances
3 features (no missing values)
Data has no target variable.
0 meta attributes

Columns (Double click to edit)

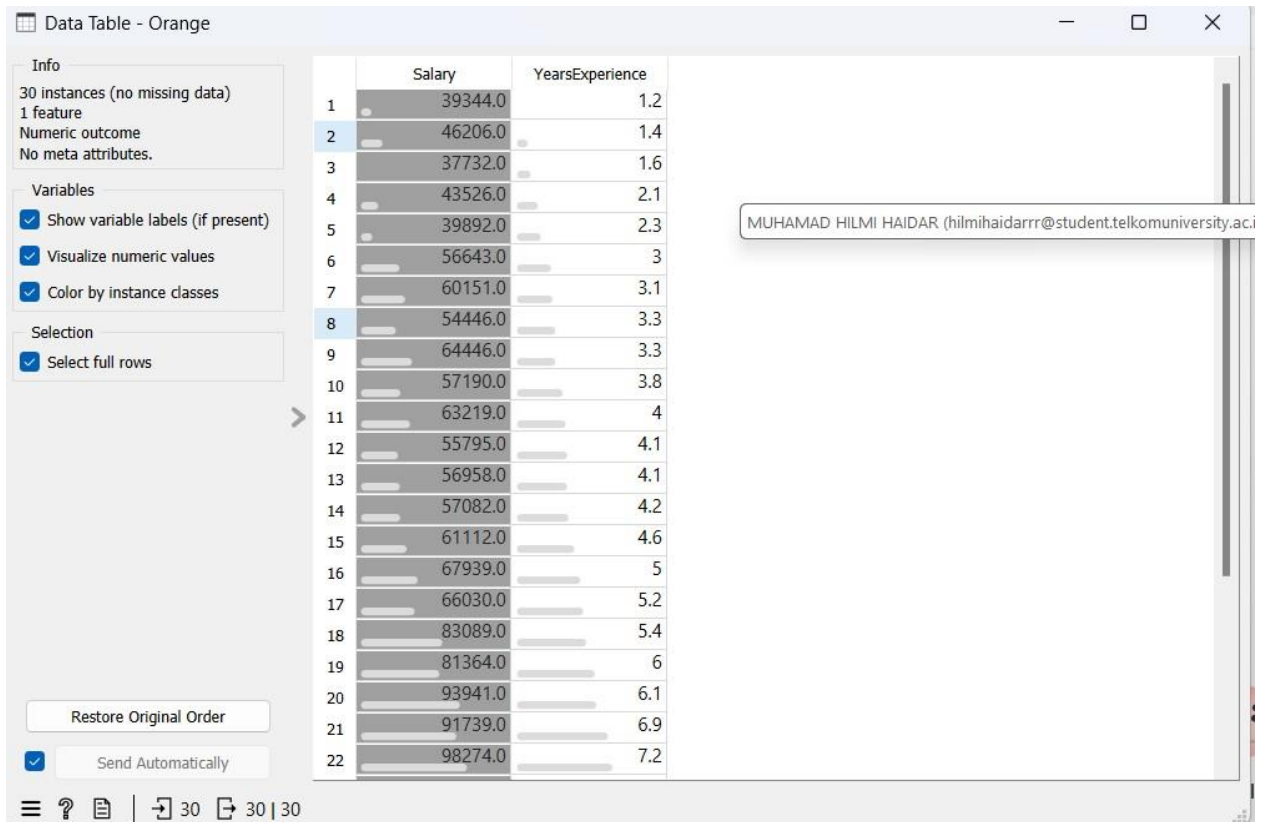
	Name	Type	Role	Values
1	Feature 1	N numeric	skip	
2	YearsExperience	N numeric	feature	
3	Salary	N numeric	target	

Reset

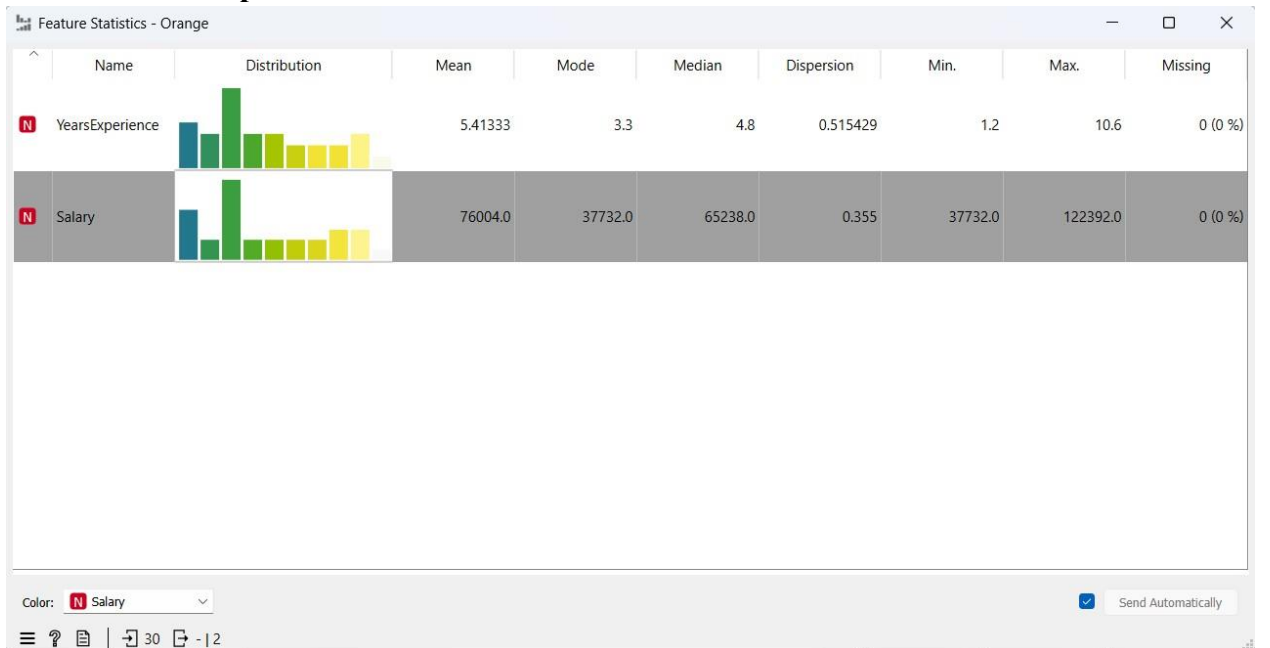
[Browse documentation datasets](#)

≡ ?

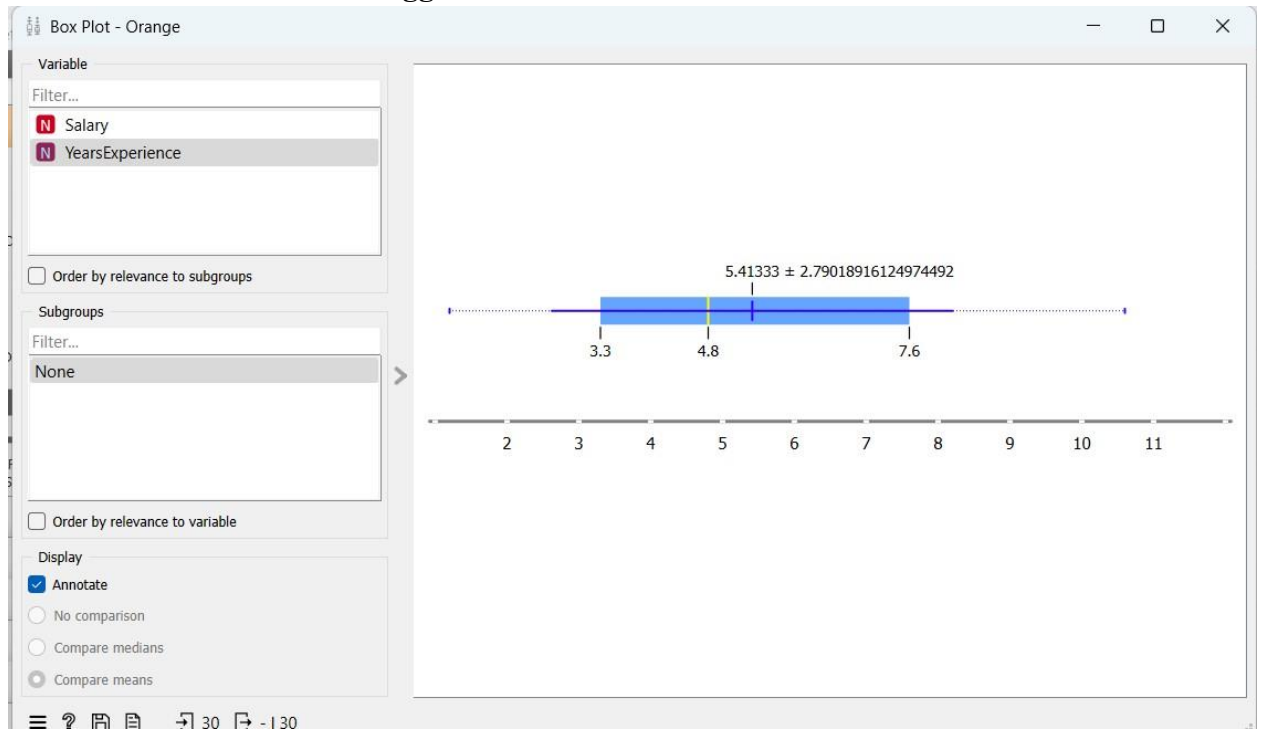
2. Load Data dalam bentuk Table



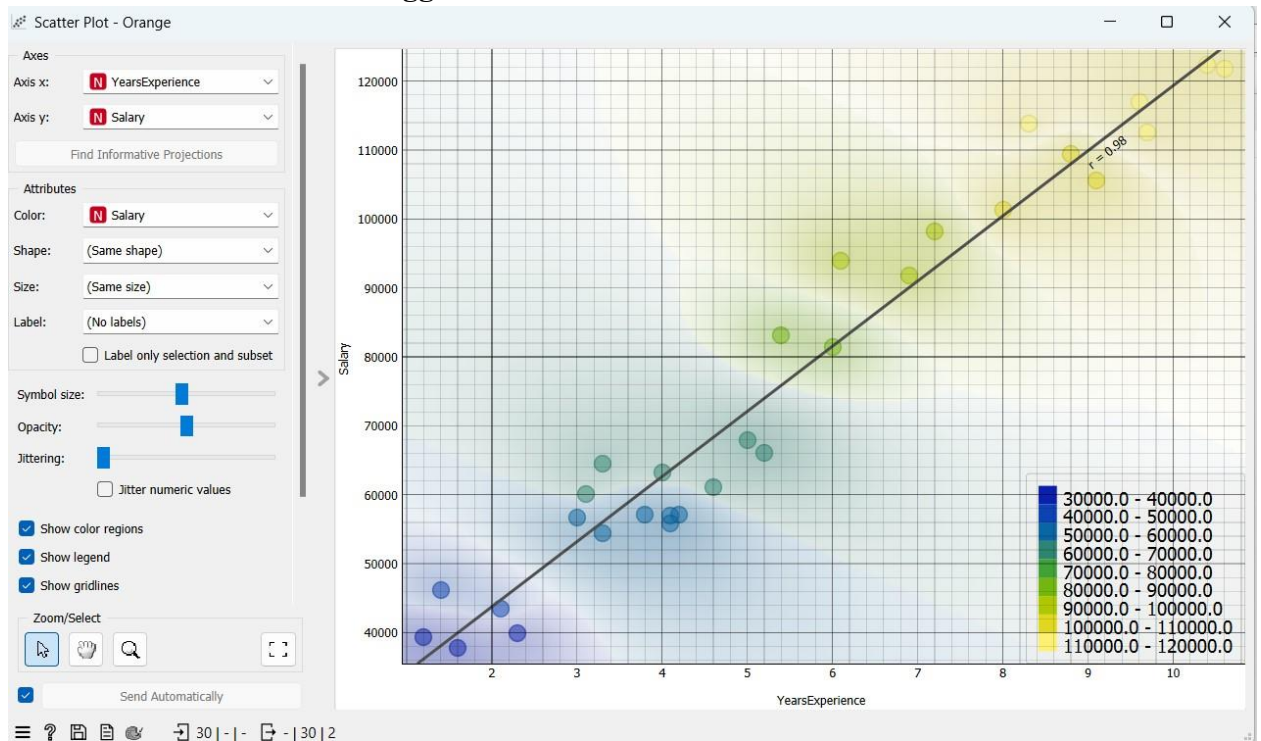
3. Feature Statistik pada Data



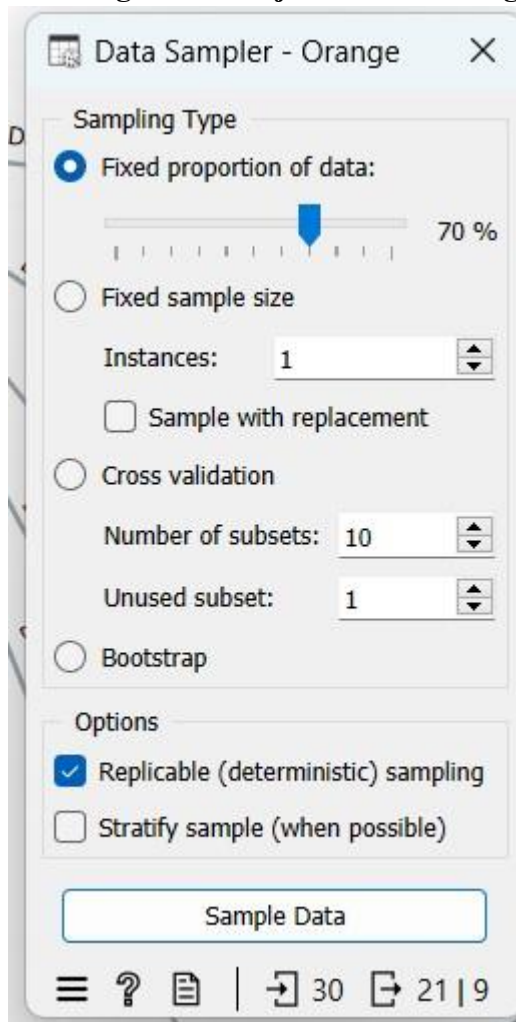
4. Memvisualisasikan data menggunakan BOX PLOT



5. Memvisualisasikan data menggunakan Scatter Plot



6. Membagi data menjadi 70% training dan 30% testing



7. Memuat Model Linear Regression

Linear Regression - Orange

Name
Linear Regression

Parameters
☒ Fit intercept (unchecking it fixes it to zero)

Regularization

☒ No regularization Regularization strength:
Alpha: 0.0001

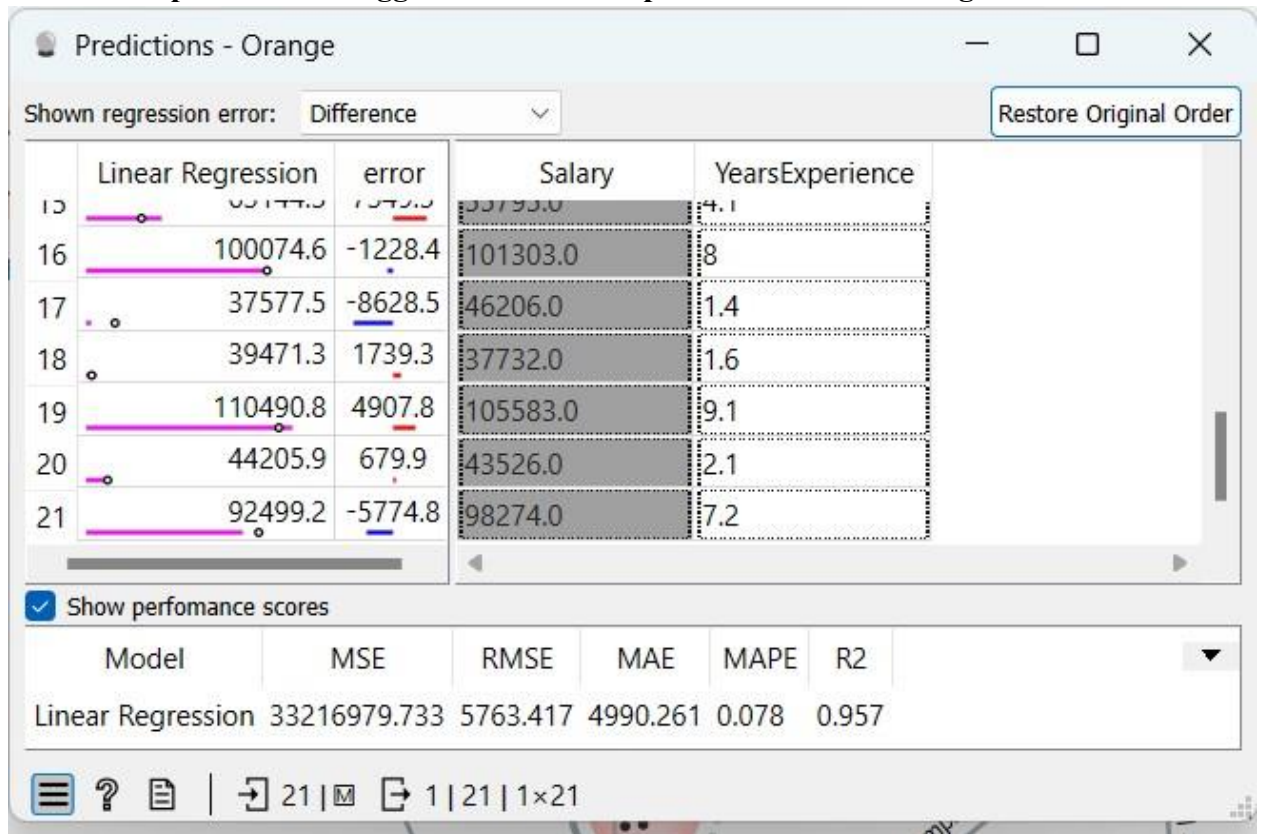
☐ Ridge regression (L2)

☐ Lasso regression (L1) Elastic net mixing:
L1 L2

☐ Elastic net regression 0.50 : 0.50

☒ Apply Automatically

8. Melakukan prediction menggunakan data sample dan model linear regression



9. Melakukan evaluasi model dengan melihat nilai RMSE , MSE, MAE, R2

Test and Score - Orange

☒ Cross validation

Number of folds: 5

☒ Stratified

☐ Cross validation by feature

☐ Random sampling

Repeat train/test: 10

Training set size: 66 %

☒ Stratified

☐ Leave one out

☐ Test on train data

☐ Test on test data

Model	MSE	RMSE	MAE	MAPE	R2
Linear Regression	41...	6445...	568...	0.091	0.946

Compare models by: Mean square error

☐ Negligible diff.: 0.1

Model	Linear Re...
Linear Regression	

Table shows probabilities that the score for the model in the row is higher than that of the model in the column. Small numbers show the probability that the difference is negligible.

Stratification is ignored for regression

10. Memvisualisasikan hasil prediksi model di test data menggunakan scatter plot

