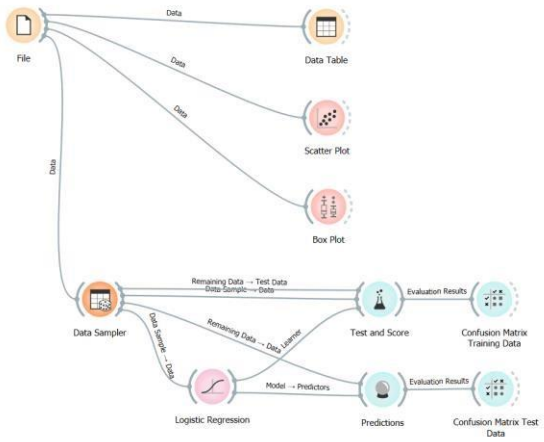
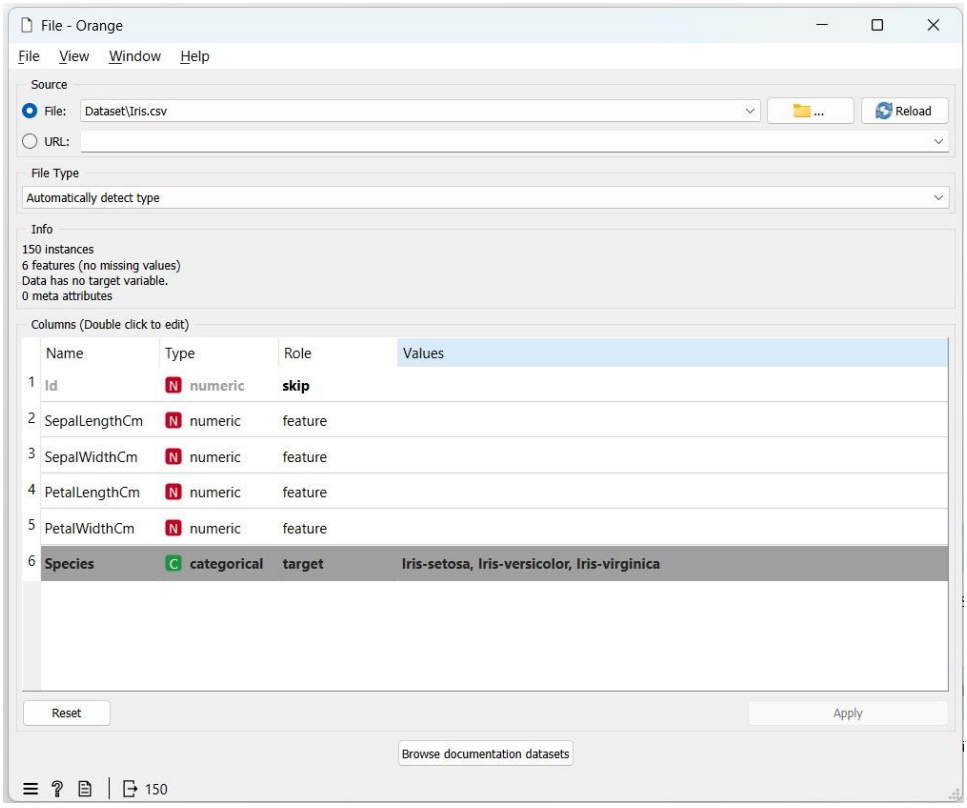
Orange Minggu 2

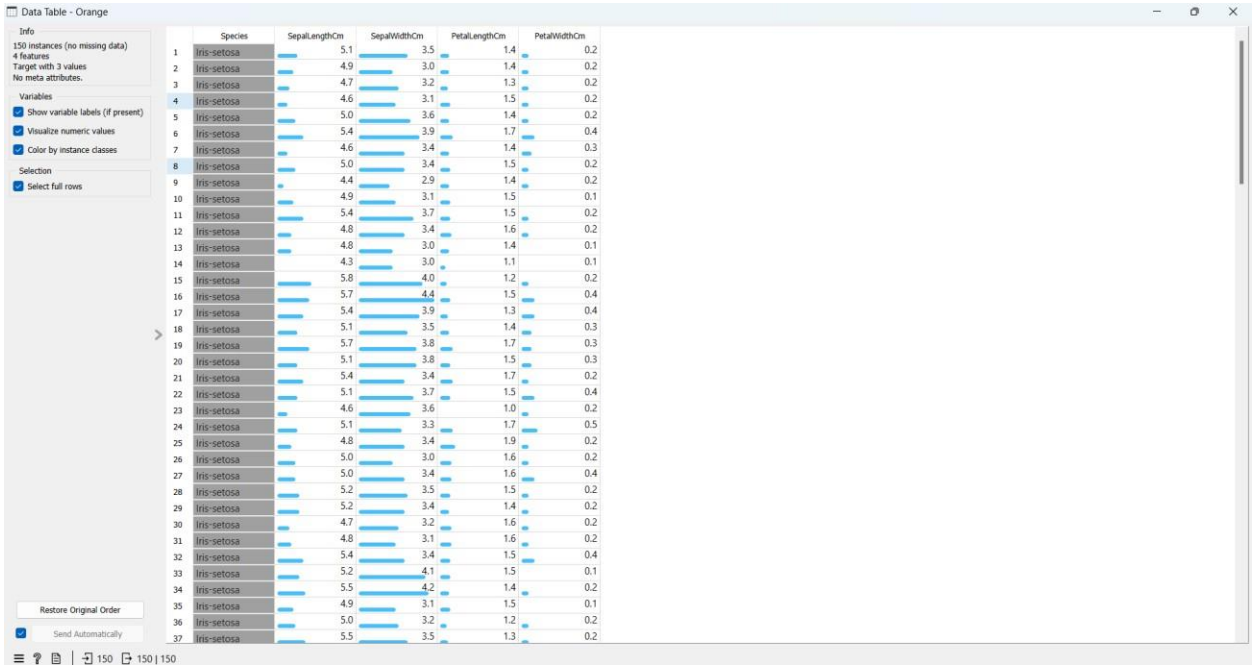
Nama : Nicholas Alvito Diandra

NIM : 1103210207

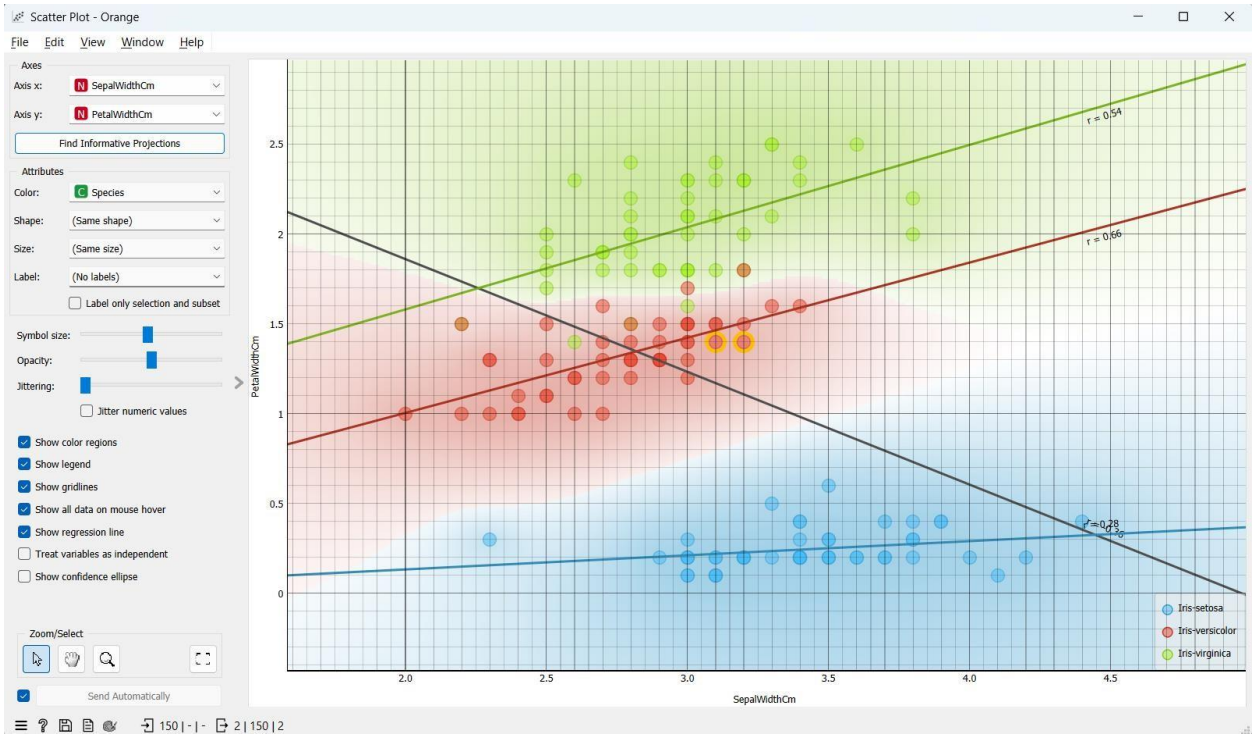


1. Memuat dataset dan set target dan skip yang tidak digunakan



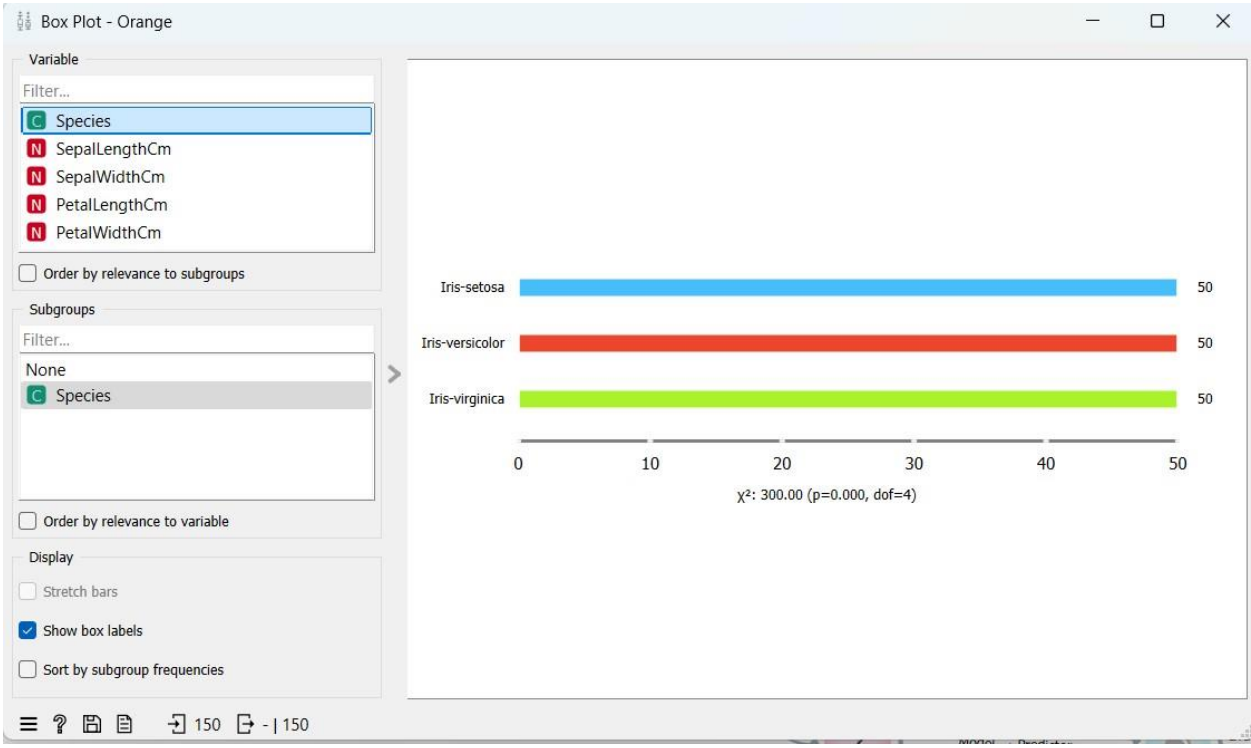


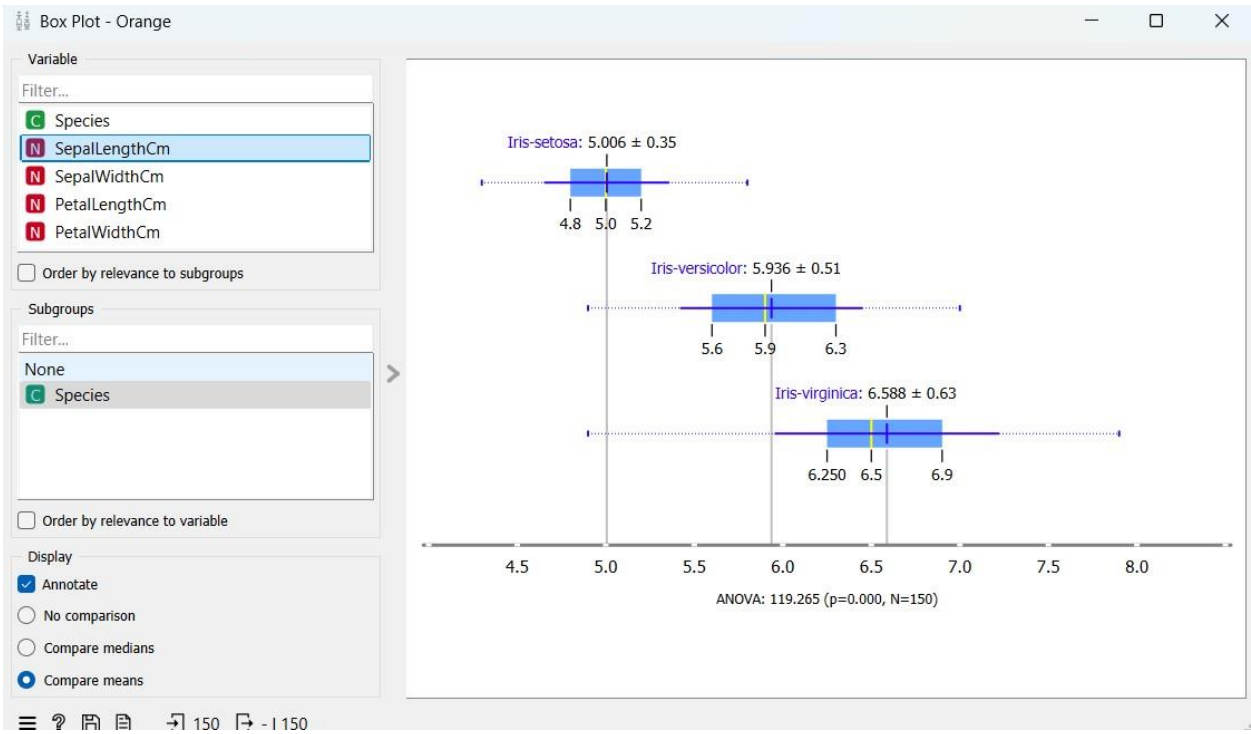
2. Memeriksa isi dataset menggunakan fungsi data table



3. Memeriksa sebaran data menggunakan scatter plot Memeriksa outliers dalam dataset menggunakan box plot

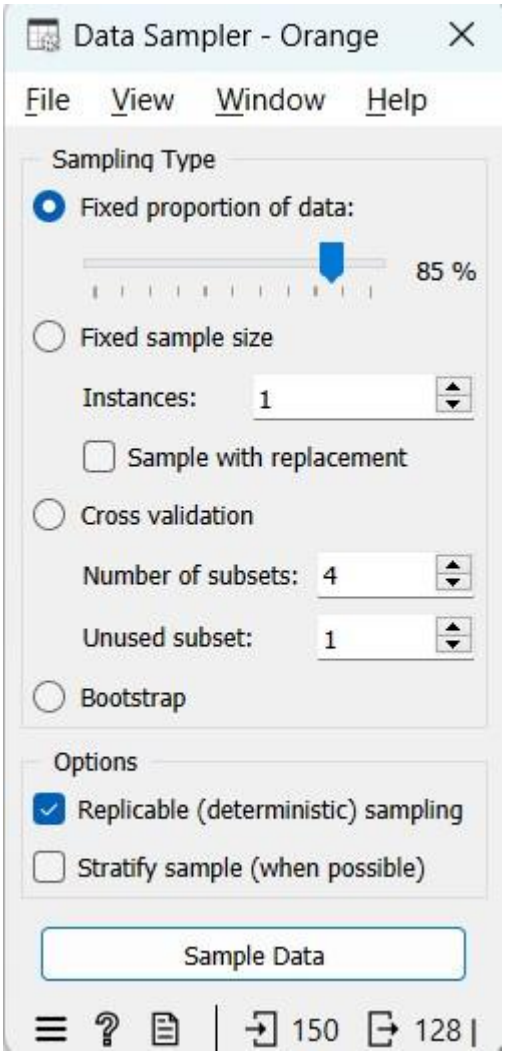
4.


Box Plot - Orange
Variable
Filter...
Species
SepalLengthCm
SepalWidthCm
PetalLengthCm
PetalWidthCm
Order by relevance to subgroups
Subgroups
Filter...
None
Species
Order by relevance to variable
Display
Stretch bars
Show box labels
Sort by subgroup frequencies
150
- | 150
Iris-setosa
Iris-versicolor
Iris-virginica
50
50
50
0
10
20
30
40
50
χ²: 300.00 (p=0.000, dof=4)


Box Plot - Orange
Variable
Filter...
Species
SepalLengthCm
SepalWidthCm
PetalLengthCm
PetalWidthCm
Order by relevance to subgroups
Subgroups
Filter...
None
Species
Order by relevance to variable
Display
Annotate
No comparison
Compare medians
Compare means
150
- | 150
Iris-setosa: 5.006 ± 0.35
4.8
5.0
5.2
Iris-versicolor: 5.936 ± 0.51
5.6
5.9
6.3
Iris-virginica: 6.588 ± 0.63
6.250
6.5
6.9
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
ANOVA: 119.265 (p=0.000, N=150)

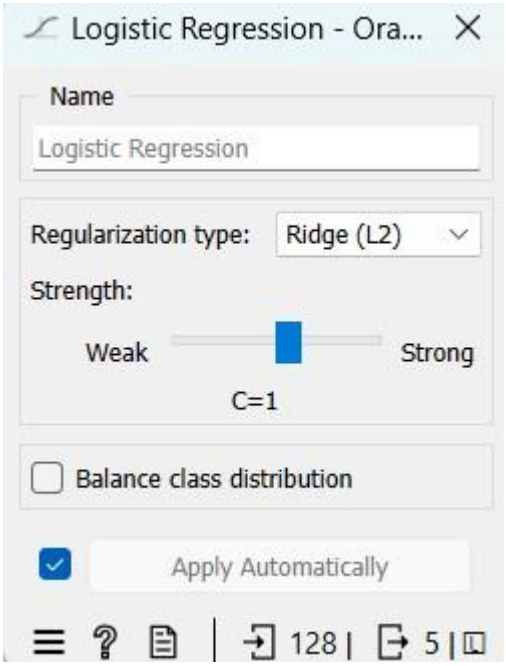
5.

Membagi dataset menjadi training data dan test data (85% training data dan 15% test data)

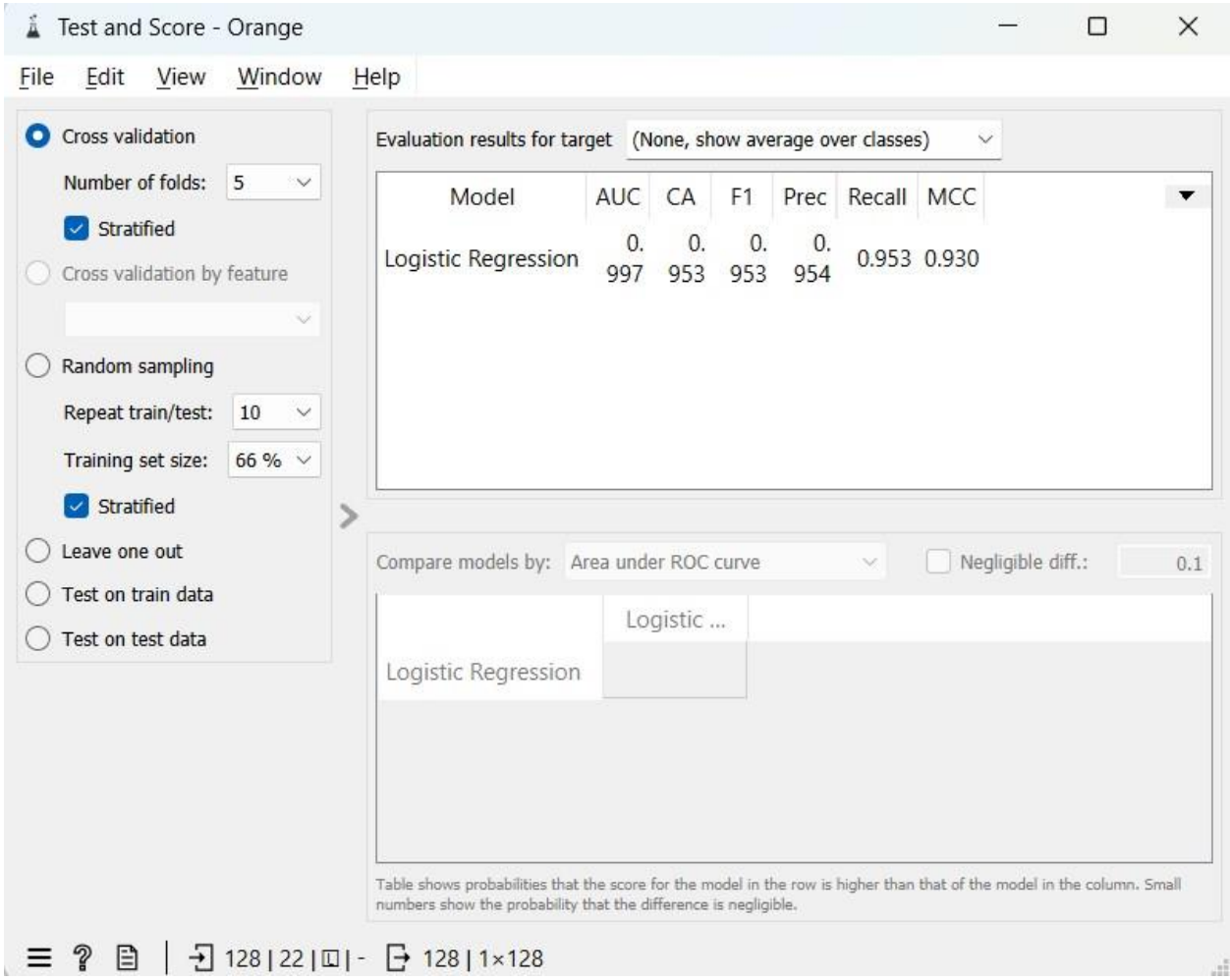


Memuat Model Logistic Regression

6.

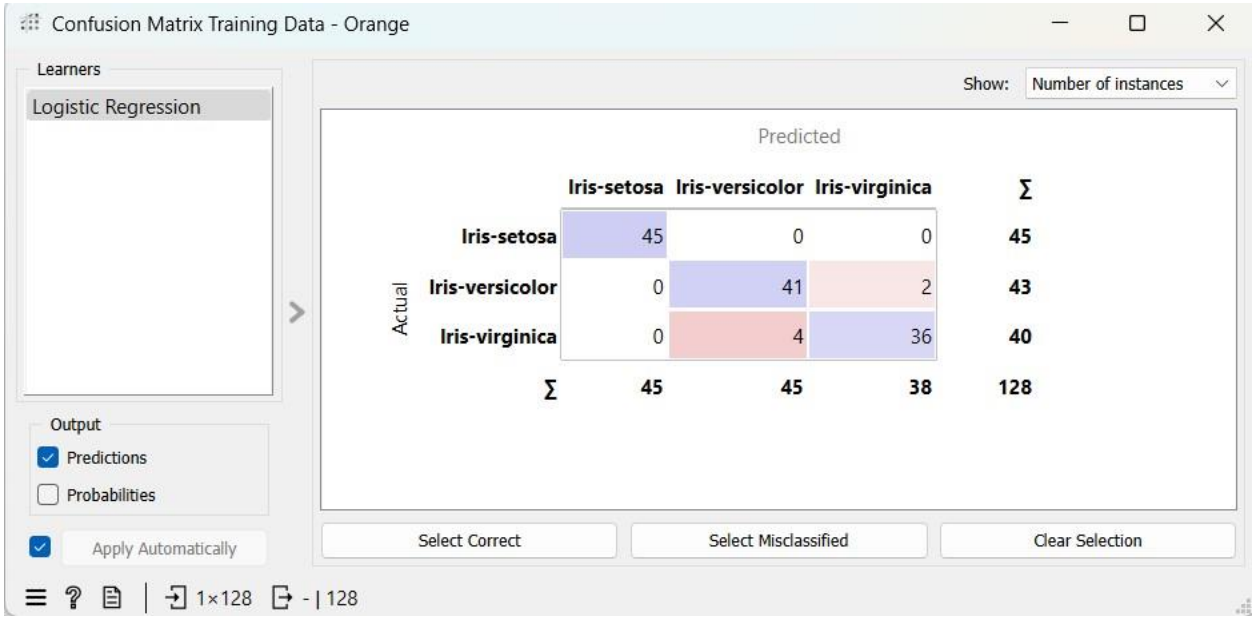

Logistic Regression - Ora...
Name
Logistic Regression
Regularization type:
Ridge (L2)
Strength:
Weak
Strong
C=1
Balance class distribution
Apply Automatically
128 |
5 |

7.

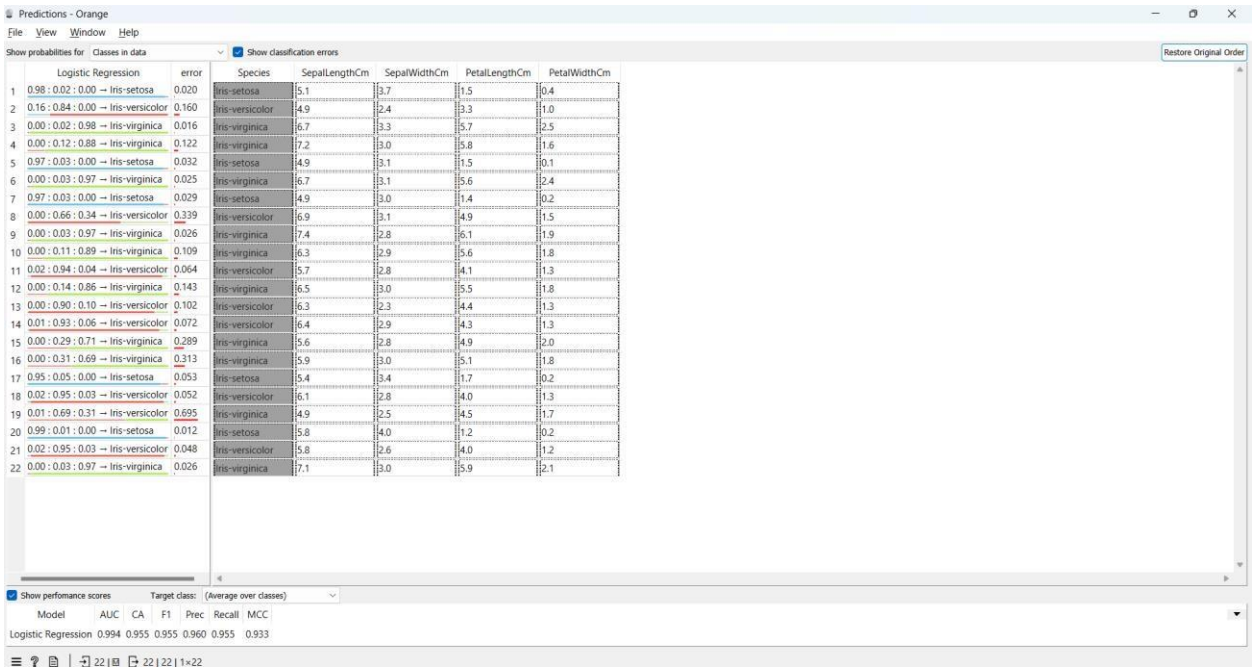


Melakukan pelatihan model dan memeriksa hasilnya Melakukan visualisasi pada hasil training menggunakan confusion matrix

8.



9. Melakukan prediksi dan membandingkan hasilnya dengan test data



10. Melihat hasil prediksi menggunakan test data dengan visualisasi menggunakan confusion matrix