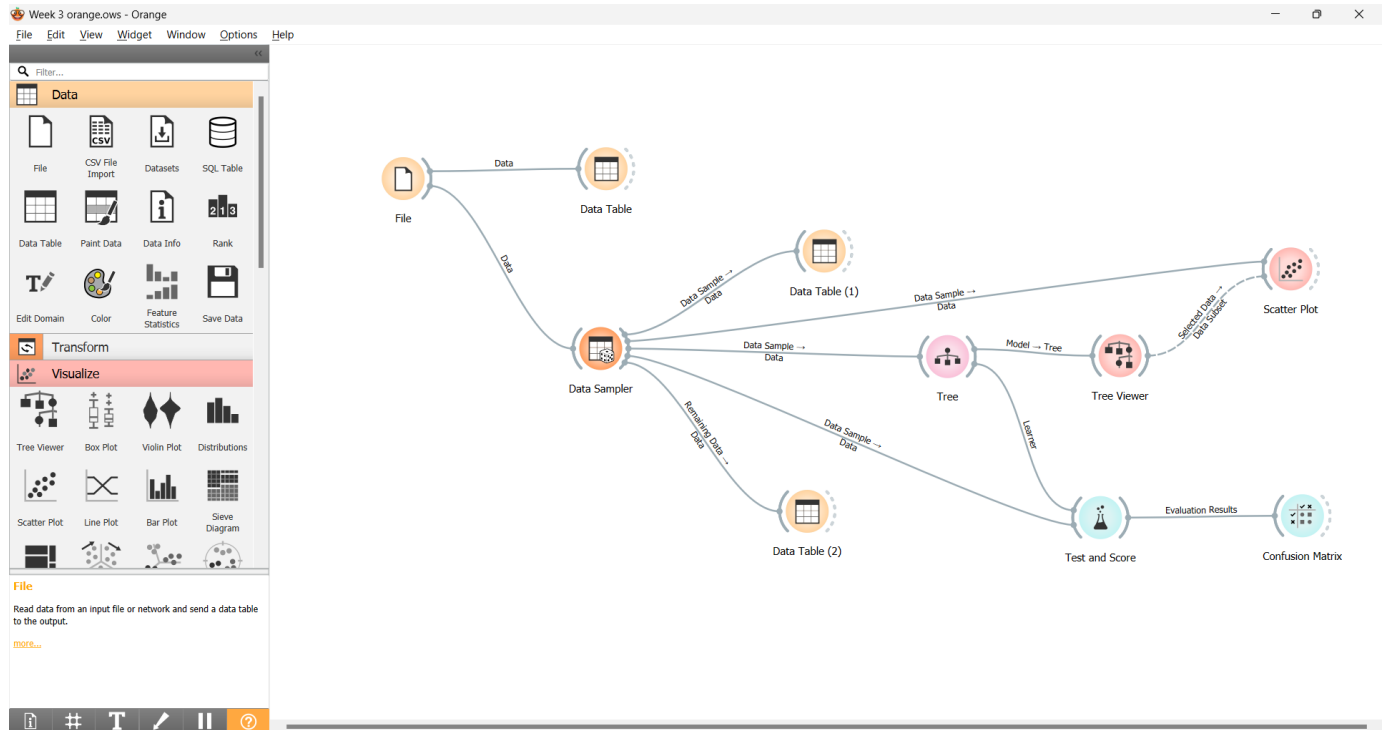
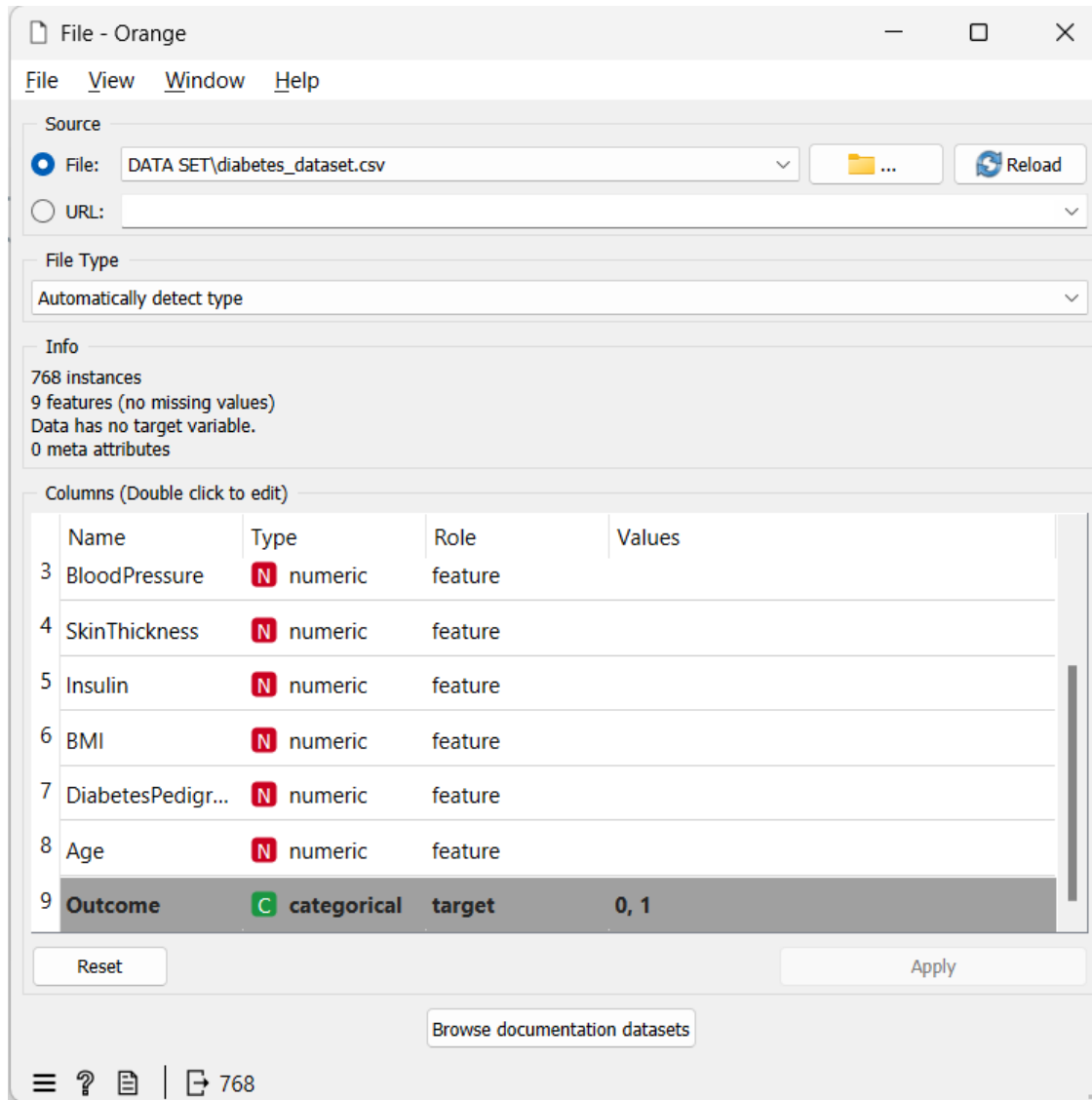


Tugas Machine Learning 3

Decision Tree dan K-NN menggunakan orange data mining

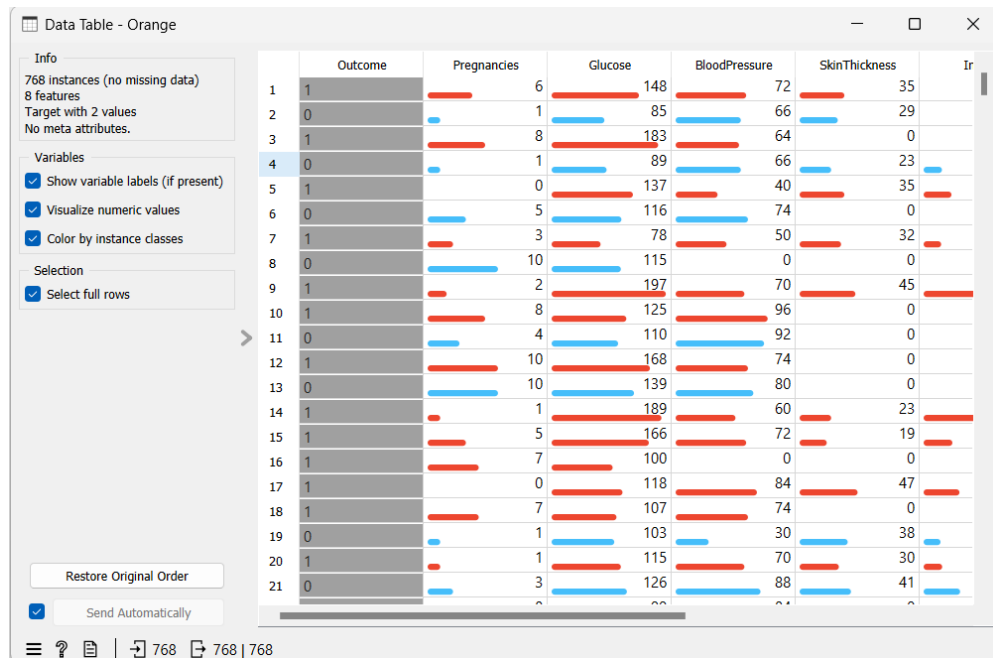


1. File



Kita menggunakan dataset **diabetes_dataset.csv**, lalu kita jadikan *Outcome* sebagai target.

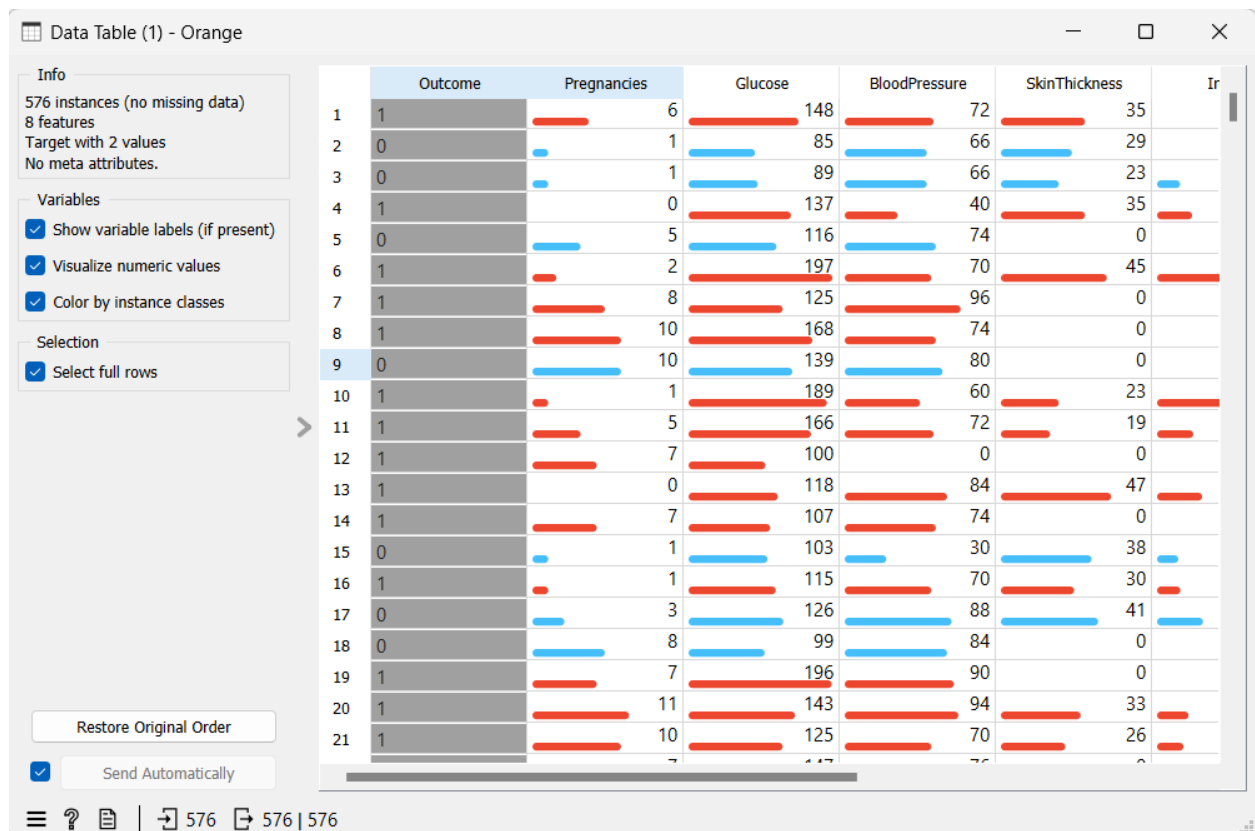
2. Data Table



3. Data Sampler

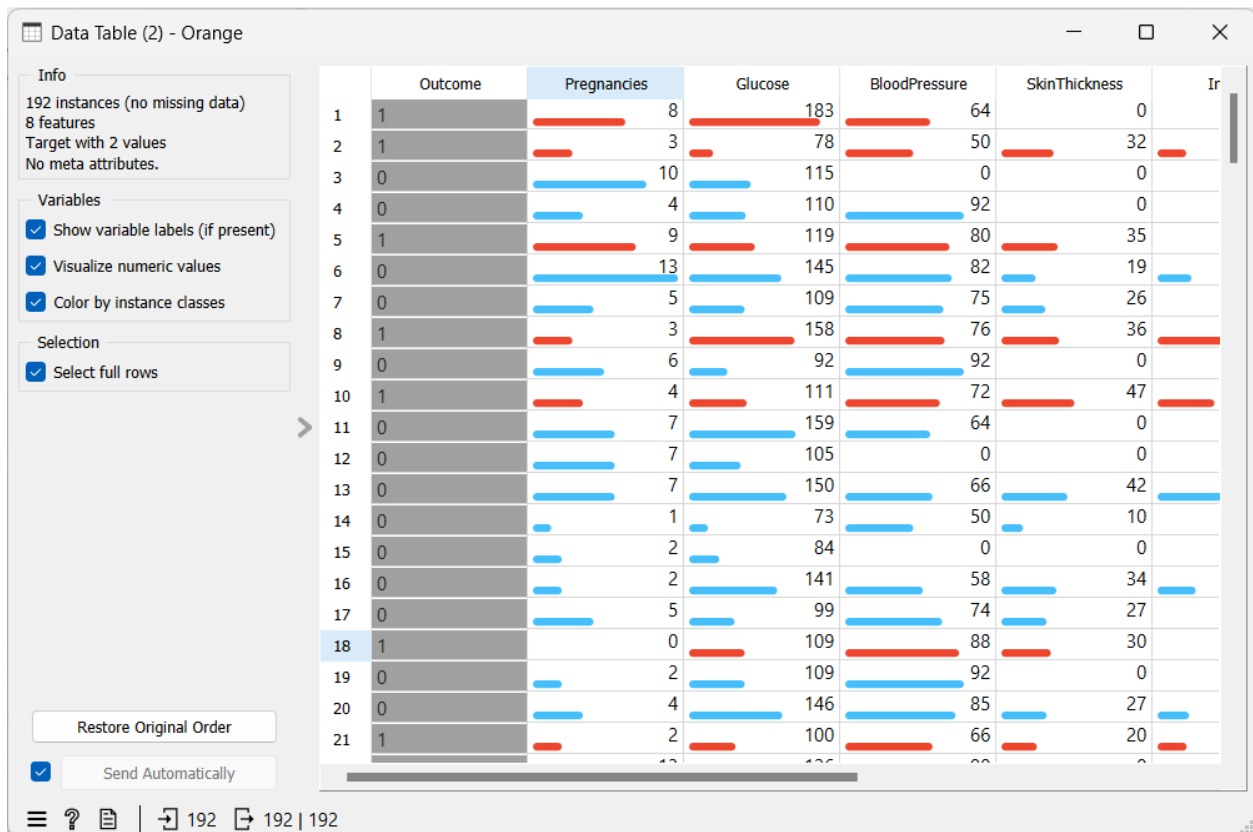
The screenshot shows the 'Data Sampler - Orange' window. It has a 'Sampling Type' section with four options: 'Fixed proportion of data' (80%), 'Fixed sample size' (700 instances), 'Cross validation' (4 subsets, 1 unused subset), and 'Bootstrap'. The 'Options' section has two checkboxes: 'Replicable (deterministic) sampling' (checked) and 'Stratify sample (when possible)' (unchecked). A 'Sample Data' button is at the bottom. The status bar at the bottom shows 768 instances and 576 features.

4. Data Table (1)



Menampilkan Data Sample yang sudah di sampling

5. Data Table (2)



Menampilkan Remaining Data yang sudah di sampling

6. Tree

Tree - Orange

Name
Tree

Parameters
☒ Induce binary tree
☒ Min. number of instances in leaves: 2
☒ Do not split subsets smaller than: 5
☒ Limit the maximal tree depth to: 3


Classification
☒ Stop when majority reaches [%]: 95

☒ Apply Automatically

576 | -

Kita menggunakan Min. number of intances in leaves 2 dengan limit depth 3

7. Test and Score



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

Evaluation results for target

(None, show average over classes)




Model	AUC	CA	F1	Prec	Recall	MCC
Tree	0.761	0.738	0.708	0.740	0.738	0.380



Compare models by:

Area under ROC curve

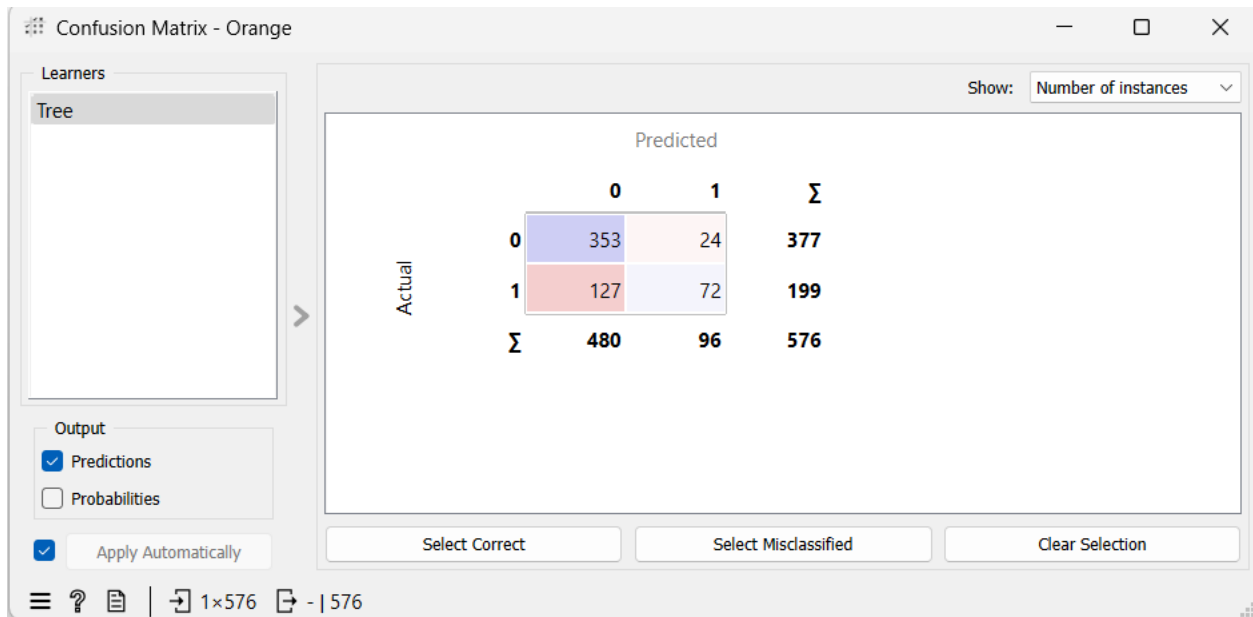
	Tree
Tree	

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 th

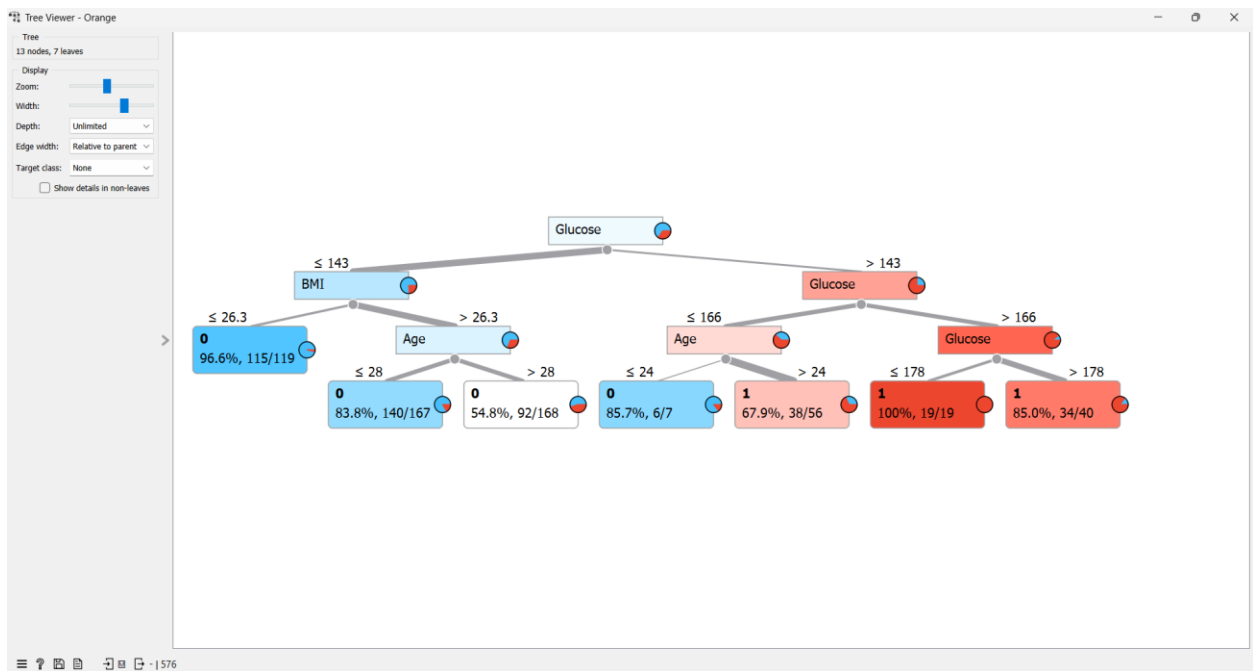


 576 | - |  576 | 1×576

8. Confusion Matrix



9. Tree Viewer



10. Scatter Plot

