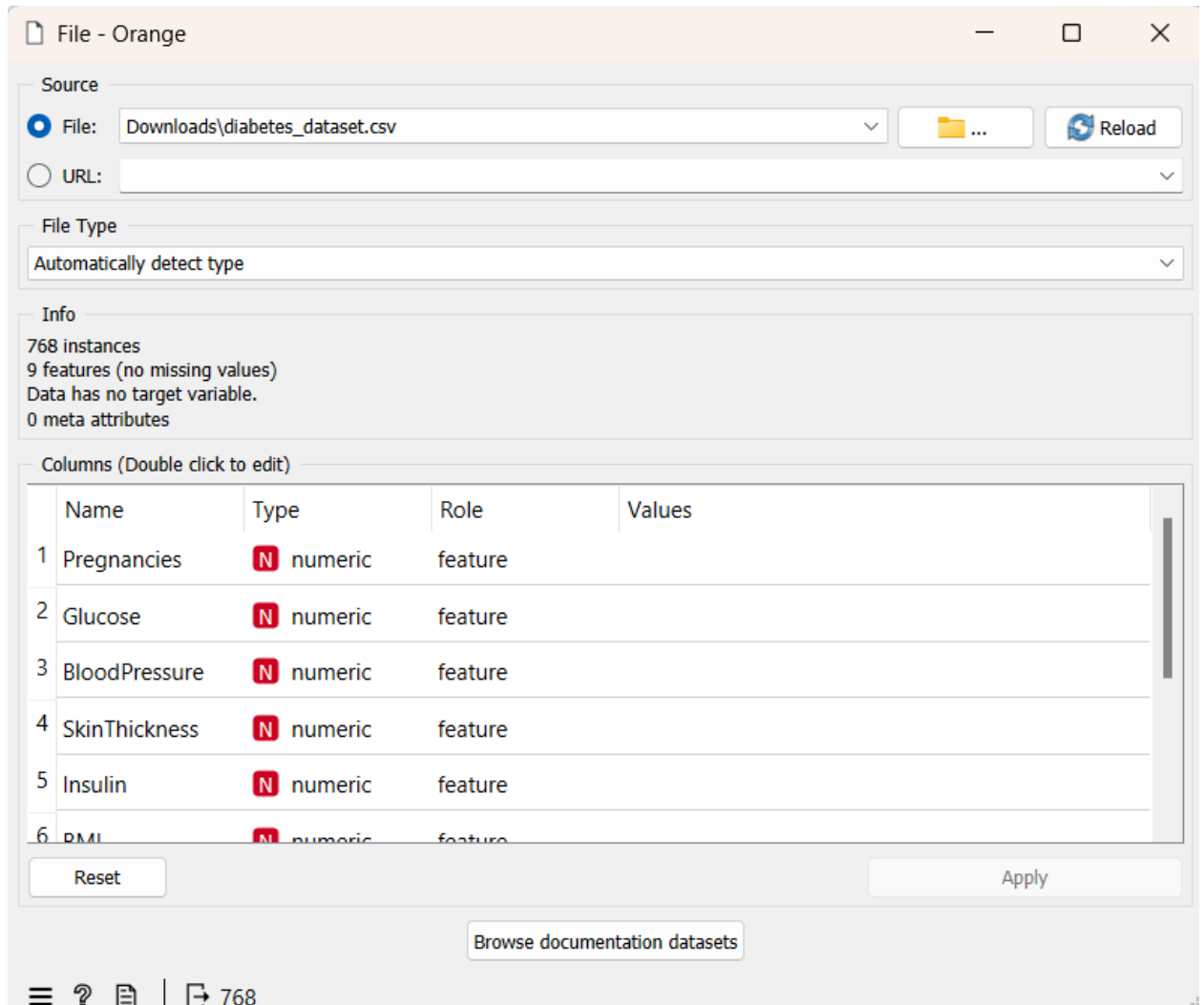


Nama : Muhamad Hilmi Haidar

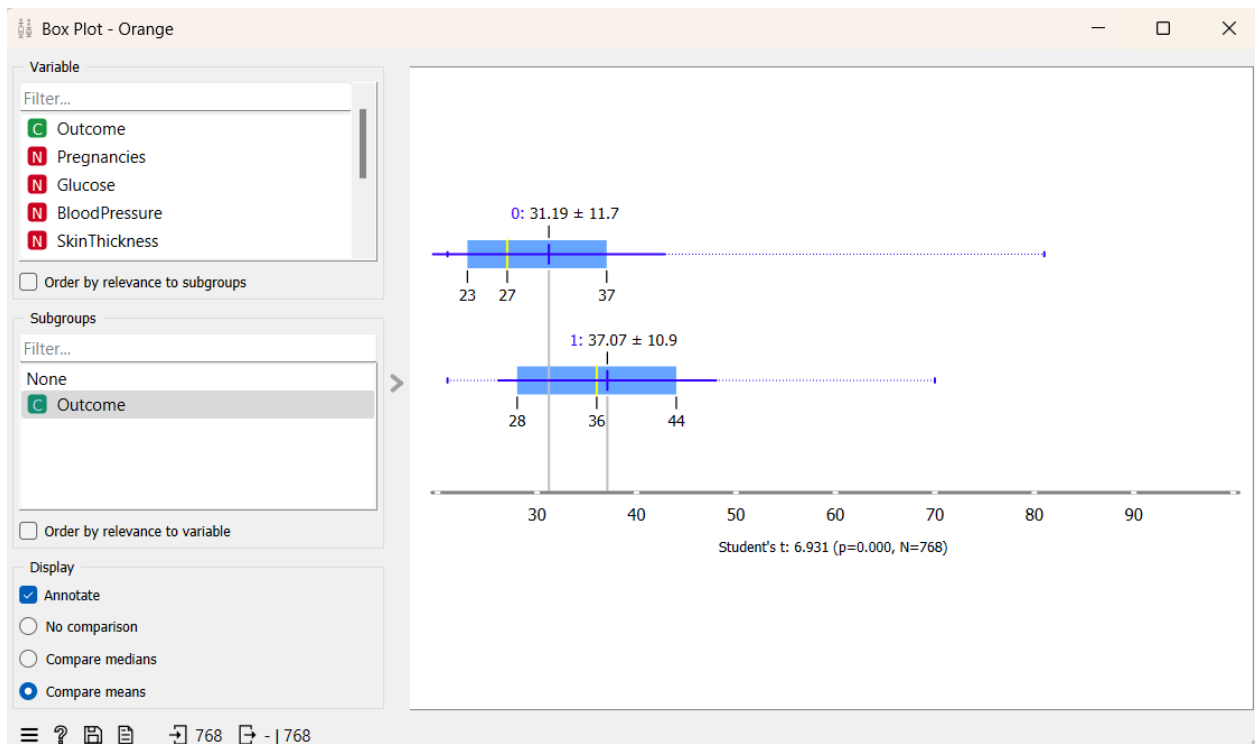
NIM : 1103213005

Praktikal OWS Minggu 3

1. MeLoad Dataset ke Orange



2. Memvisualisasikan dataset menggunakan BoxPlot untuk memeriksa apakah ada outlier di dataset tersebut



3. Memeriksa isi Data Table

Data Table - Orange

Info
768 instances (no missing data)
8 features
Target with 2 values
No meta attributes.

Variables
☒ Show variable labels (if present)
☒ Visualize numeric values
☒ Color by instance classes

Selection
☒ Select full rows

Restore Original Order

☒ Send Automatically

	Outcome	Pregnancies	Glucose	BloodPressure	SkinThickness	Ir
1	1	6	148	72	35	
2	0	1	85	66	29	
3	1	8	183	64	0	
4	0	1	89	66	23	
5	1	0	137	40	35	
6	0	5	116	74	0	
7	1	3	78	50	32	
8	0	10	115	0	0	
9	1	2	197	70	45	
10	1	8	125	96	0	
11	0	4	110	92	0	
12	1	10	168	74	0	
13	0	10	139	80	0	
14	1	1	189	60	23	
15	1	5	166	72	19	
16	1	7	100	0	0	
17	1	0	118	84	47	
18	1	7	107	74	0	
19	0	1	103	30	38	
20	1	1	115	70	30	
21	0	3	126	88	41	

768 | 768

4. Sampling data untuk digunakan model tree

Data Sampler - Orange

Sampling Type

☐ Fixed proportion of data:

85 %

☐ Fixed sample size

Instances: 1

☐ Sample with replacement

☒ Cross validation

Number of subsets: 4

Unused subset: 1

☐ Bootstrap

Options

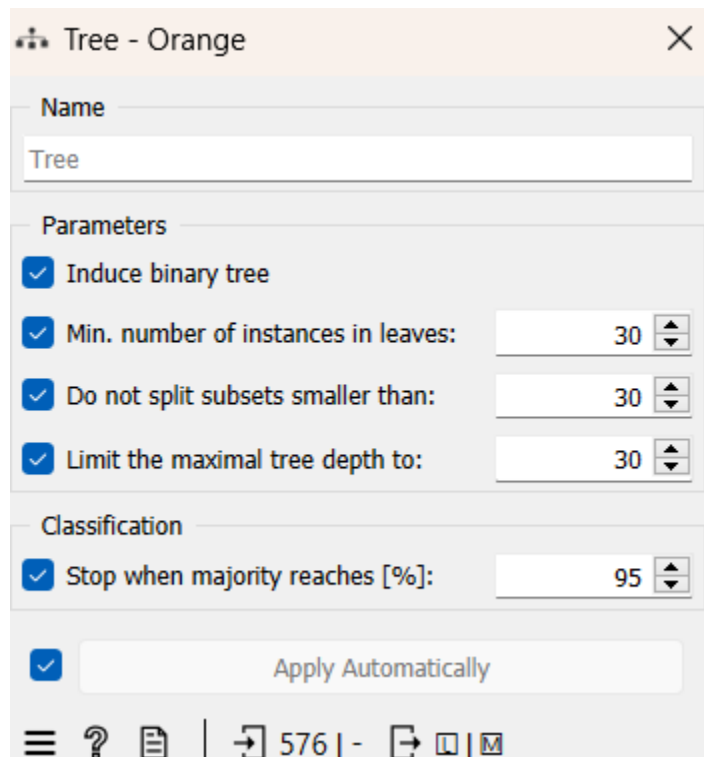
☒ Replicable (deterministic) sampling

☐ Stratify sample (when possible)

Sample Data

768 576

5. Mengatur Parameter pada model Tree



The screenshot shows the 'Tree - Orange' dialog box with the following settings:

- Name:** Tree
- Parameters:**
 - ☒ Induce binary tree
 - ☒ Min. number of instances in leaves: 30
 - ☒ Do not split subsets smaller than: 30
 - ☒ Limit the maximal tree depth to: 30
- Classification:**
 - ☒ Stop when majority reaches [%]: 95
- ☒ Apply Automatically

The bottom status bar shows icons for menu, help, and document, followed by a separator, a refresh icon, and the text '576 | -' followed by icons for zoom in, zoom out, and a magnifying glass.

6. Melihat evaluasi hasil dari model tree

The screenshot shows the 'Test and Score - Orange' window. On the left, the 'Cross validation' section is active, with 'Number of folds' set to 5, 'Stratified' checked, and 'Training set size' at 66%. The 'Evaluation results for target' section on the right displays a table of metrics for the 'Tree' model. Below this, the 'Compare models by' section shows a comparison table for the 'Tree' model against itself. The status bar at the bottom indicates a file named '576' with dimensions '1x576'.

Test and Score - Orange

Cross validation

- ☒ 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	0.815	0.747	0.744	0.742	0.747	0.430

Compare models by: Area under the curve

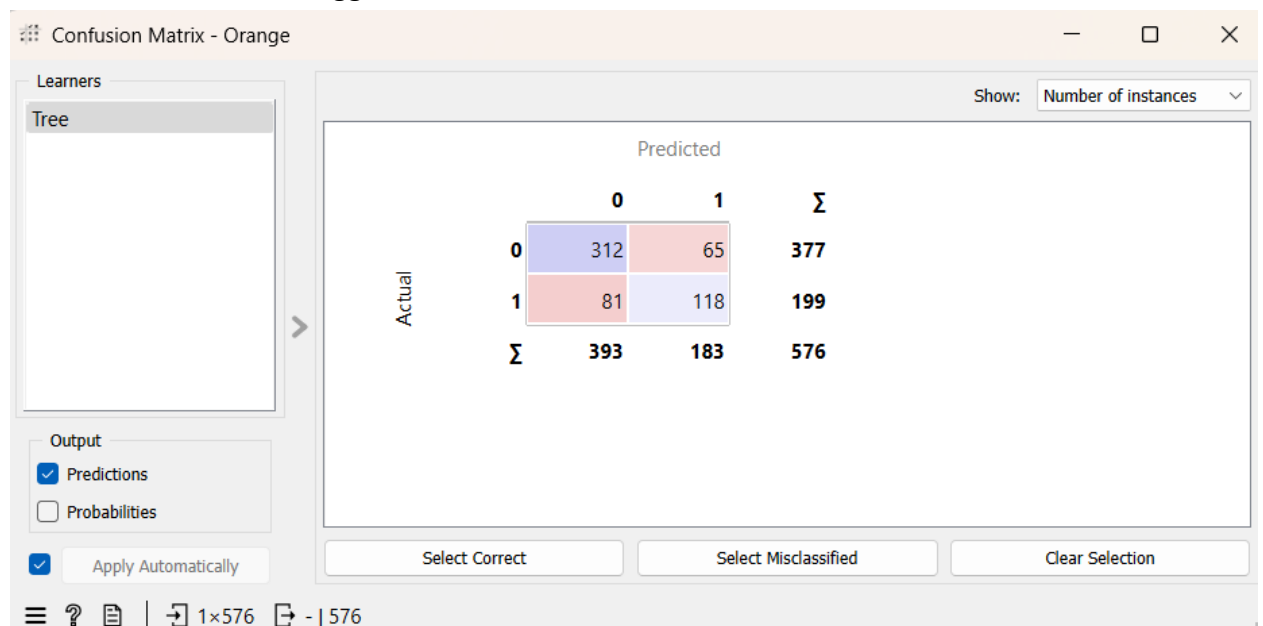
☐ Negligible diff.: 0.1

	Tree
Tree	

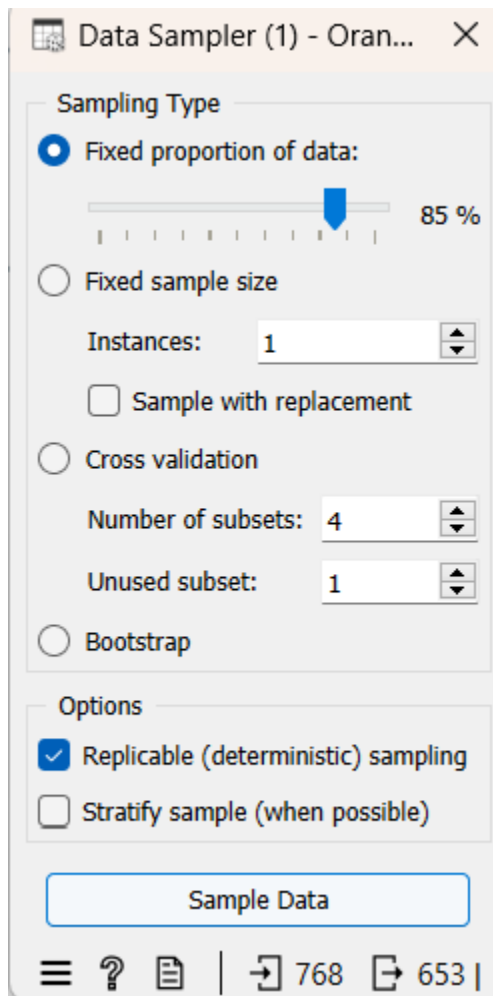
Table shows probabilities that the score for the model in the row is higher than that of the model in the column. Small numbers show the probability that the difference is negligible.

576 | 1x576

7. Melihat hasil model menggunakan confusion matrix



8. Sampling data untuk model KNN



The 'Data Sampler (1) - Orange3' dialog box is shown. It has a title bar with a close button. The 'Sampling Type' section contains four radio buttons: 'Fixed proportion of data:' (selected), 'Fixed sample size', 'Cross validation', and 'Bootstrap'. The 'Fixed proportion of data:' option has a slider set to 85%. The 'Fixed sample size' option has an 'Instances:' spinner set to 1 and a 'Sample with replacement' checkbox. The 'Cross validation' option has a 'Number of subsets:' spinner set to 4 and an 'Unused subset:' spinner set to 1. 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 shows icons for menu, help, and document, followed by a refresh icon and the number 768, and a save icon and the number 653.

Data Sampler (1) - Oran...

Sampling Type

- ☒ Fixed proportion of data:
85 %
- ☐ Fixed sample size
Instances: 1
☐ Sample with replacement
- ☐ Cross validation
Number of subsets: 4
Unused subset: 1
- ☐ Bootstrap

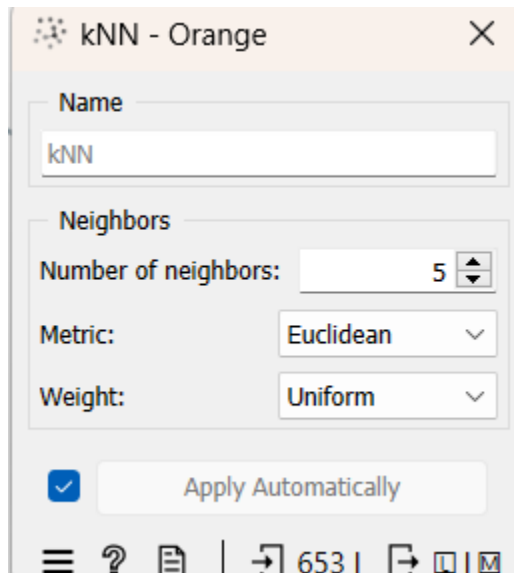
Options

- ☒ Replicable (deterministic) sampling
- ☐ Stratify sample (when possible)

Sample Data

768 | 653

9. Set parameter untuk model KNN



The 'kNN - Orange3' dialog box is shown. It has a title bar with a close button. The 'Name' field contains 'kNN'. The 'Neighbors' section has a 'Number of neighbors:' spinner set to 5, a 'Metric:' dropdown menu set to 'Euclidean', and a 'Weight:' dropdown menu set to 'Uniform'. There is a checked checkbox and an 'Apply Automatically' button. The status bar shows icons for menu, help, and document, followed by a refresh icon and the number 653, and a save icon and the number 653.

kNN - Orange

Name
kNN

Neighbors

Number of neighbors: 5

Metric: Euclidean

Weight: Uniform

☒ Apply Automatically

653 | 653

10. Melihat hasil test dan hasil evaluasi model KNN

Test and Score (1) - 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
kNN	0.742	0.723	0.718	0.716	0.723	0.370

Compare models by: Area un ☐ Negligible diff.: 0.1

	kNN
kNN	

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.

653 | 115 | 653 | 1×653

11. Melihat hasil train test menggunakan confusion matrix

