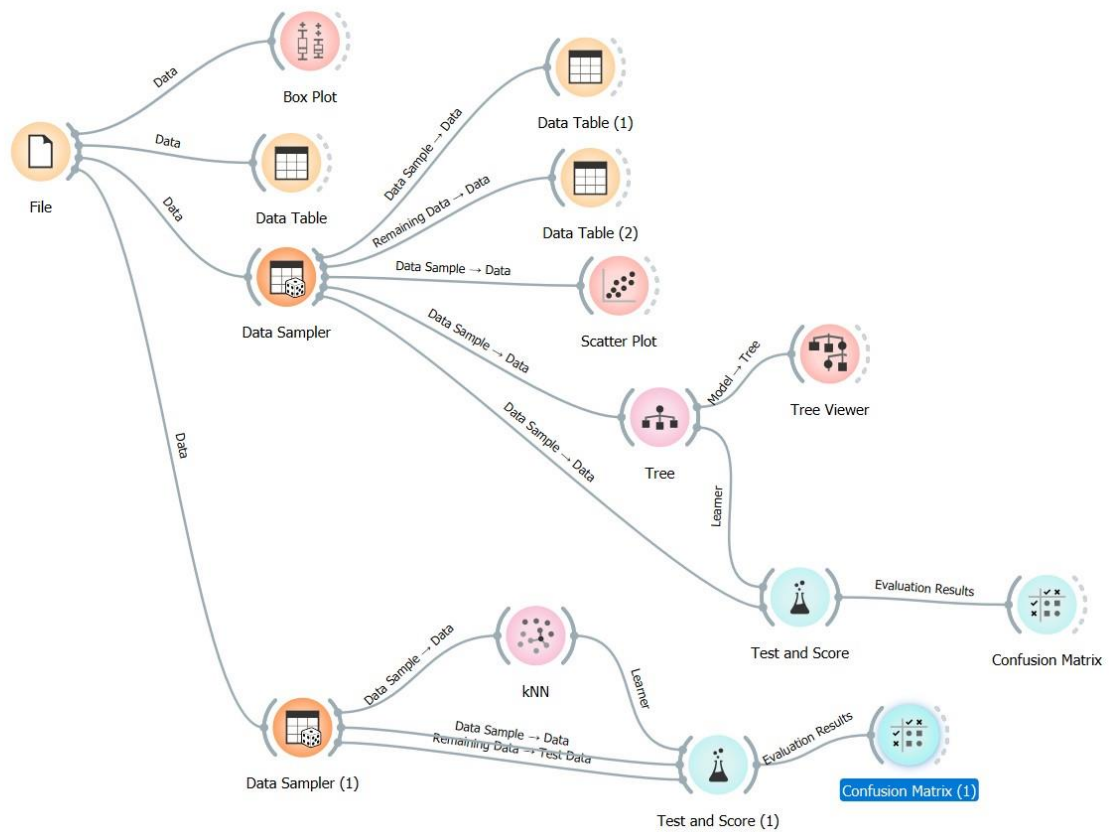


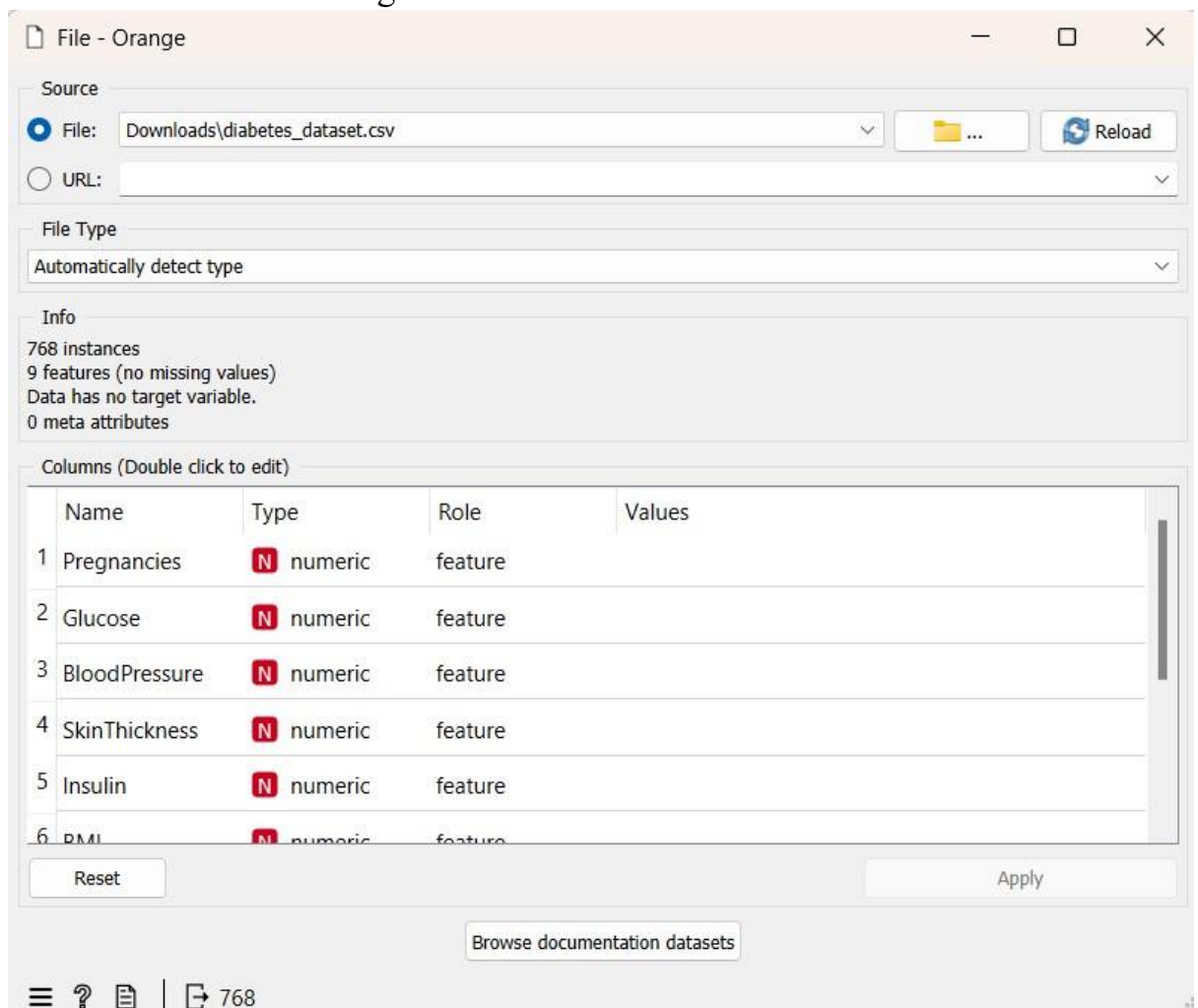
Orange Minggu 3

Nama : Ferdinant Hutajulu

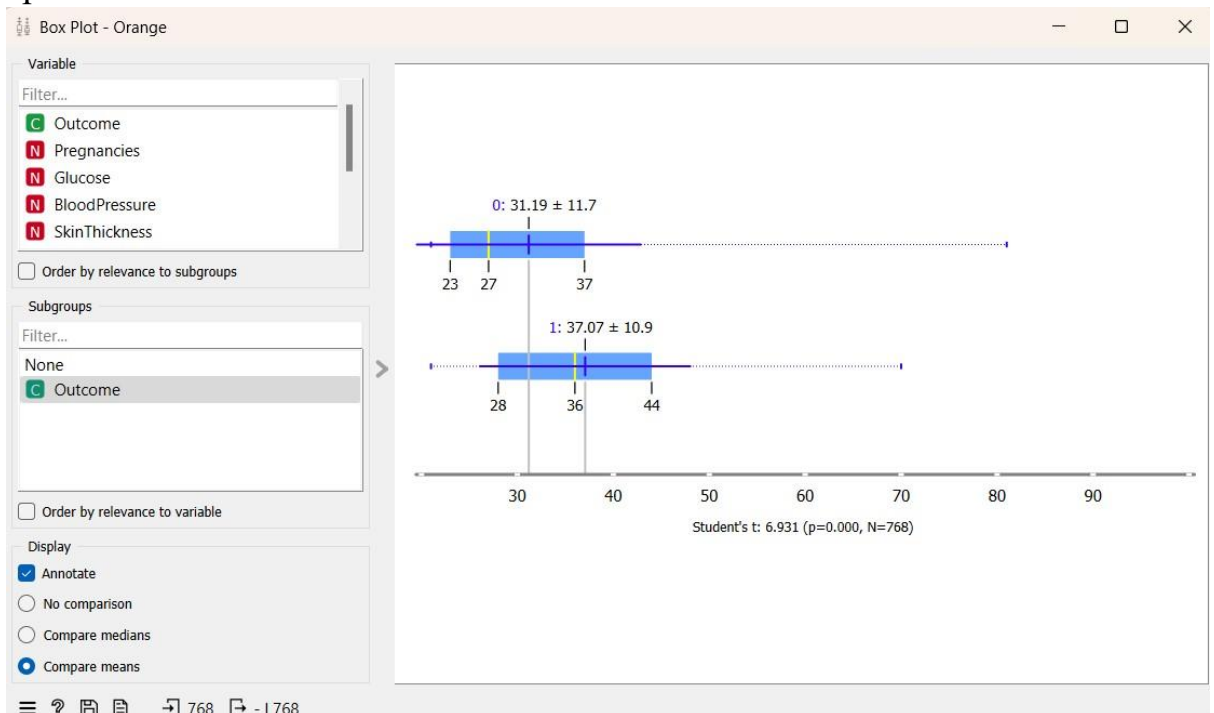
NIM : 1103213120



1. Memuat 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

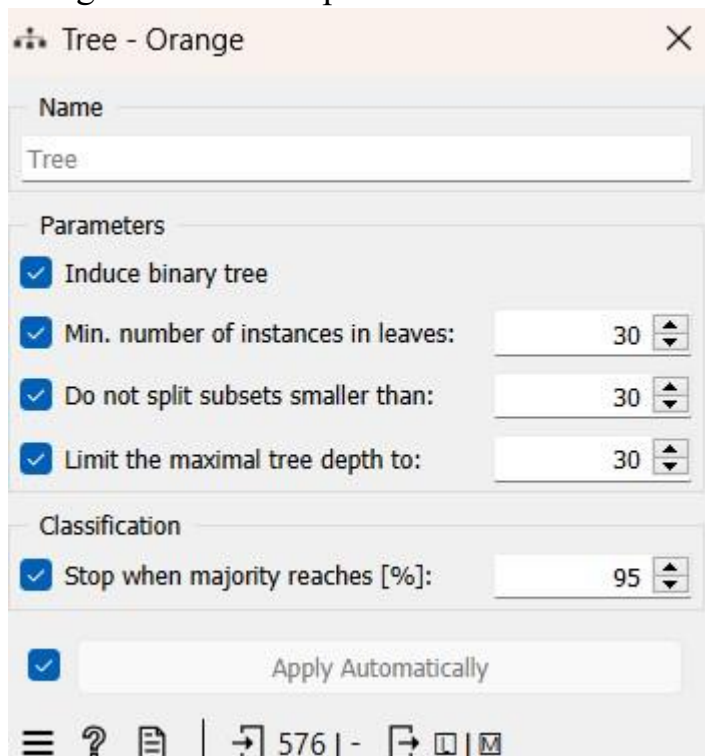
☒ Replicable (deterministic) sampling

☐ Stratify sample (when possible)

Sample Data

768 576

5. Mengatur Parameter pada model Tree



6. Melihat evaluasi hasil dari model tree

The 'Test and Score - Orange' window displays the following settings and results:

- Left Panel (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
- Right Panel (Evaluation results):**
 - Target: (None, show average over classes)
 - Table:
- Bottom Panel (Compare models):**
 - Compare models by: Area under the curve
 - Negligible diff.: 0.1
 - Table:

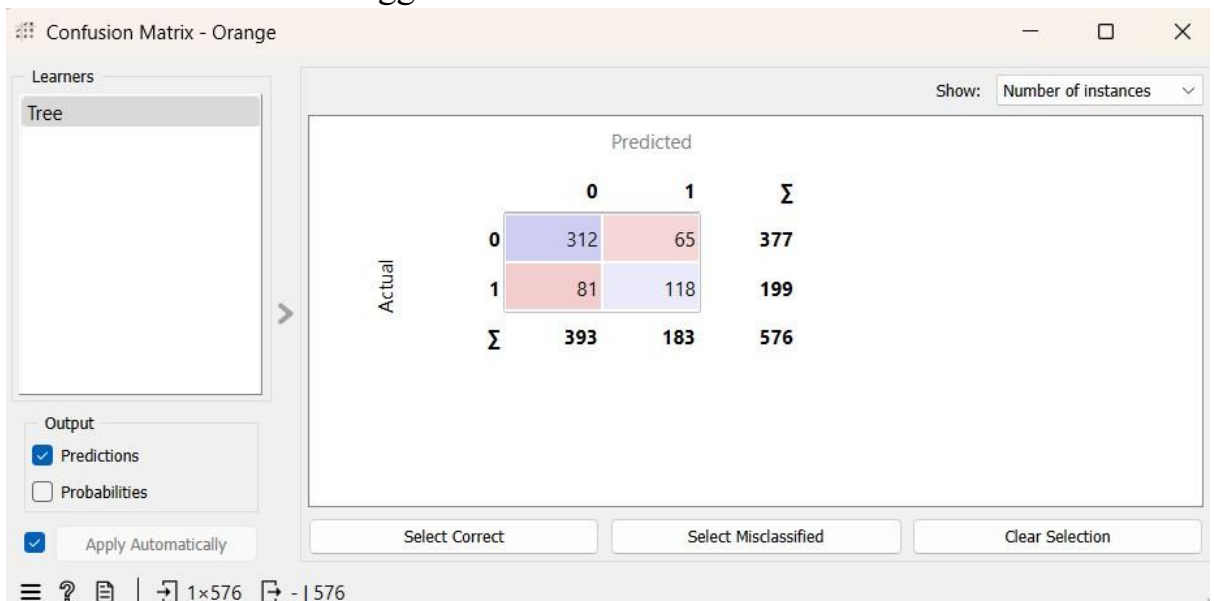
The bottom status bar shows 576 instances and a 1x576 grid.

Model	AUC	CA	F1	Prec	Recall	MCC
Tree	0.815	0.747	0.744	0.742	0.747	0.430

	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.

7. Melihat hasil model menggunakan confusion matrix



8. Sampling data untuk model KNN

Data Sampler (1) - 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

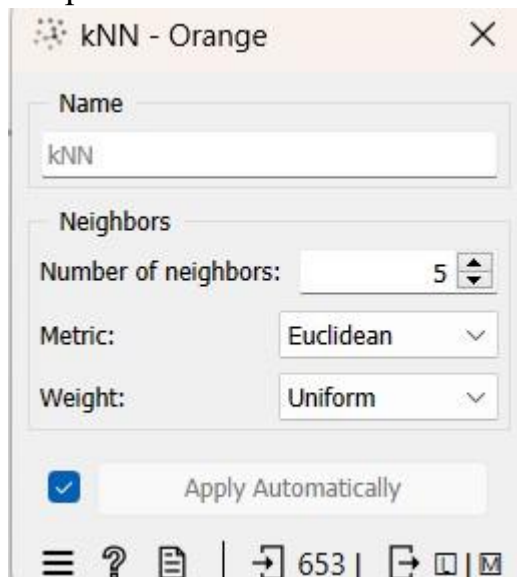
☒ Replicable (deterministic) sampling

☐ Stratify sample (when possible)

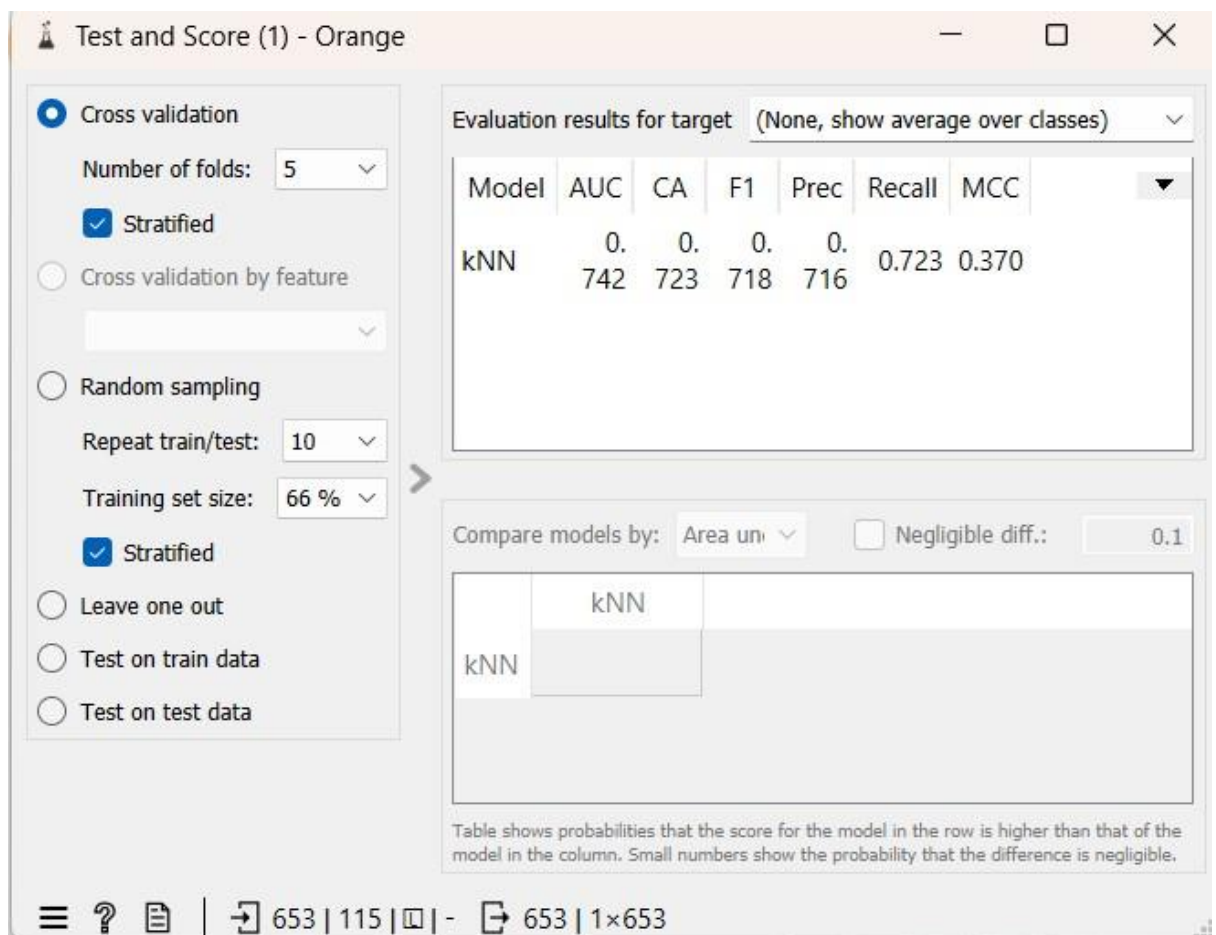
Sample Data

768 653

9. Set parameter untuk model KNN



10. Melihat hasil test dan hasil evaluasi model KNN



11. Melihat hasil train test menggunakan confusion matrix

