

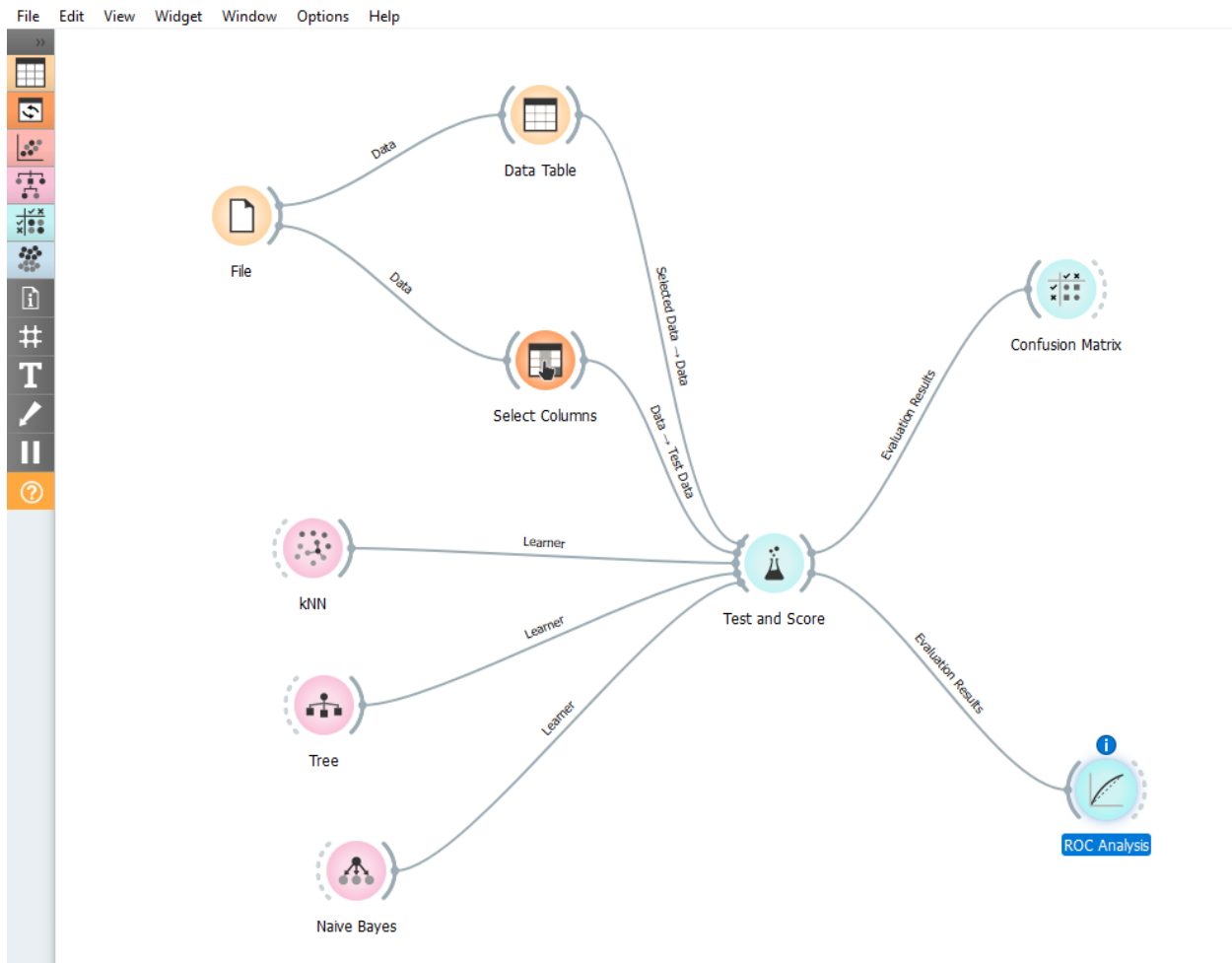
Nama: M.Rizky Fadillah

NIM: 1103210259

## TUGAS 3 PEMBELAJARAN MESIN 2425-1

Orange Data Mining:

- Decision Tree dan k-NN menggunakan orange data mining



## 1. Mengimpor dataset ke dalam Orange

File Type

Automatically detect type

Info

768 instances  
9 features (no missing values)  
Data has no target variable.  
0 meta attributes

Columns (Double click to edit)

	Name	Type	Role	Values
3	BloodPressure	N numeric	feature	
4	SkinThickness	N numeric	feature	
5	Insulin	N numeric	feature	
6	BMI	N numeric	feature	
7	DiabetesPedigr...	N numeric	feature	
8	Age	N numeric	feature	
9	Outcome	C categorical	target	0, 1

Reset Apply

Browse documentation datasets

768

## 2. Mengevaluasi distribusi data berdasarkan fitur-fitur yang ada

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

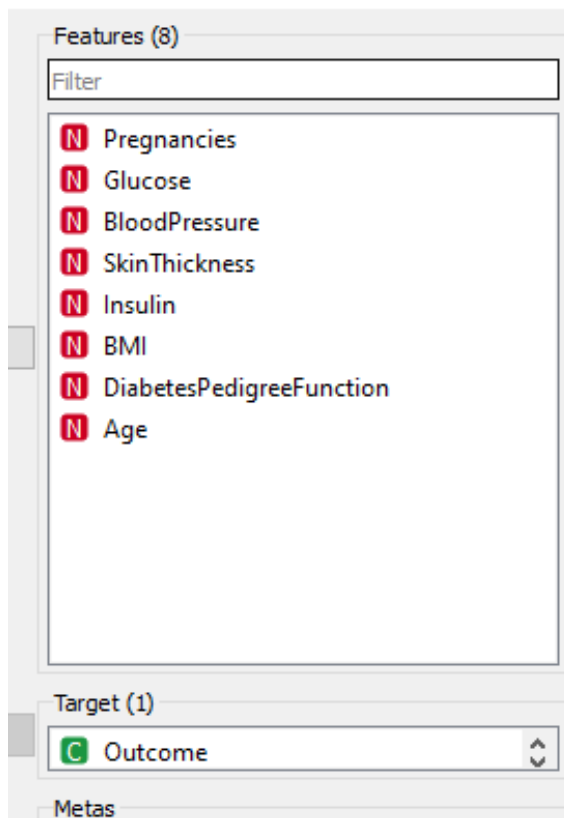
☒ Select full rows

Restore Original Order

☒ Send Automatically

	Outcome	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age
630	0	4	94	65	22	0	24.7	0.148	21
631	1	7	114	64	0	0	27.4	0.732	34
632	1	0	102	78	40	90	34.5	0.238	24
633	0	2	111	60	0	0	26.2	0.343	23
634	0	1	128	82	17	183	27.5	0.115	22
635	0	10	92	62	0	0	25.9	0.167	31
636	1	13	104	72	0	0	31.2	0.465	38
637	0	5	104	74	0	0	28.8	0.153	48
638	0	2	94	76	18	66	31.6	0.640	23
639	1	7	97	76	32	91	40.9	0.871	32
640	0	1	100	74	12	46	19.5	0.149	28
641	0	0	102	86	17	105	29.3	0.695	27
642	0	4	128	70	0	0	34.3	0.303	24
643	1	6	147	80	0	0	29.5	0.178	50
644	0	4	90	0	0	0	28.0	0.610	31
645	0	3	103	72	30	152	27.6	0.730	27
646	0	2	157	74	35	440	39.4	0.134	30
647	1	1	167	74	17	144	23.4	0.447	33
648	1	0	179	50	36	159	37.8	0.455	22
649	1	11	136	84	35	130	28.3	0.260	42
650	0	0	107	60	25	0	26.4	0.133	23
651	0	1	91	54	25	100	25.2	0.234	23
652	0	1	117	60	23	106	33.8	0.466	27
653	0	5	123	74	40	77	34.1	0.269	28
654	0	2	120	54	0	0	26.8	0.455	27
655	0	1	106	70	28	135	34.2	0.142	22
656	1	2	155	52	27	540	38.7	0.240	25
657	0	2	101	58	35	90	21.8	0.155	22
658	0	1	120	80	48	200	38.9	1.162	41
659	0	11	127	106	0	0	39.0	0.190	51
660	1	3	80	82	31	70	34.2	1.292	27
661	0	10	162	84	0	0	27.7	0.182	54
662	1	1	199	76	43	0	42.9	1.394	22
663	1	8	167	106	46	231	37.6	0.165	43
664	1	9	145	80	46	130	37.5	0.637	40
665	1	6	115	60	39	0	33.7	0.245	40
666	0	1	112	80	45	132	34.8	0.217	24
667	1	4	145	82	18	0	32.5	0.235	70
668	1	10	111	70	27	0	27.5	0.141	40
669	0	6	98	58	33	190	34.0	0.430	43
670	0	9	154	78	30	100	30.9	0.164	45
671	0	6	165	68	26	168	33.6	0.631	49
672	0	1	99	58	10	0	25.4	0.551	21
673	0	10	68	106	23	49	35.5	0.285	47
674	0	3	123	100	35	240	57.3	0.880	22

3. Memilih mana saja atribut yang digunakan sebagai fitur (input) dan mana yang dijadikan target (output) untuk proses analisis atau pembuatan model

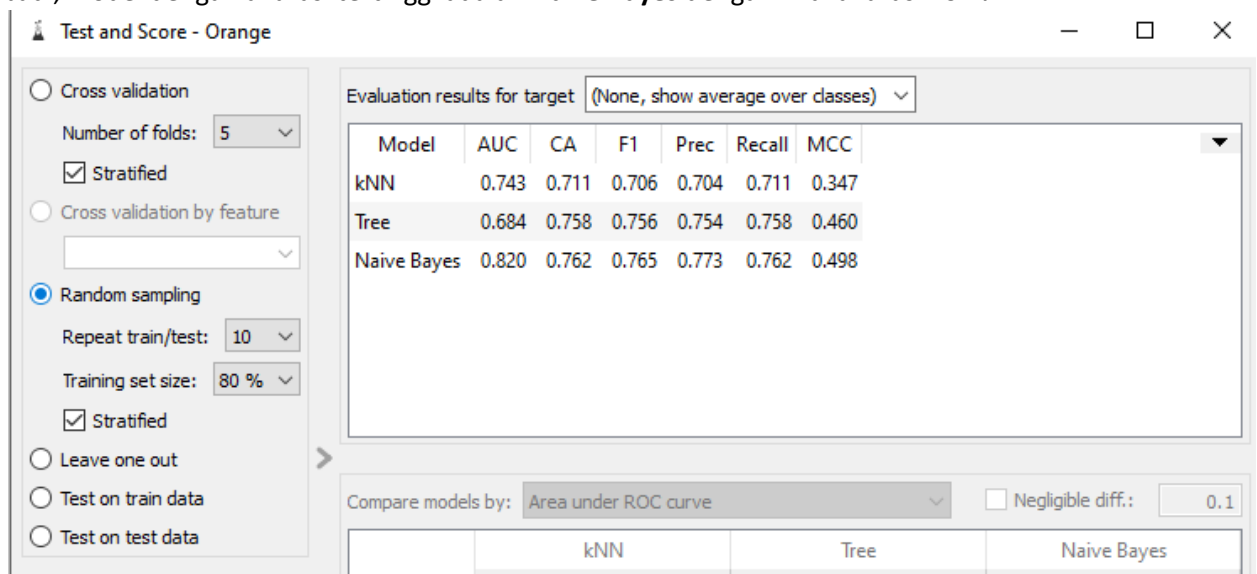


4. **k-NN: 71.1% (0.711)**

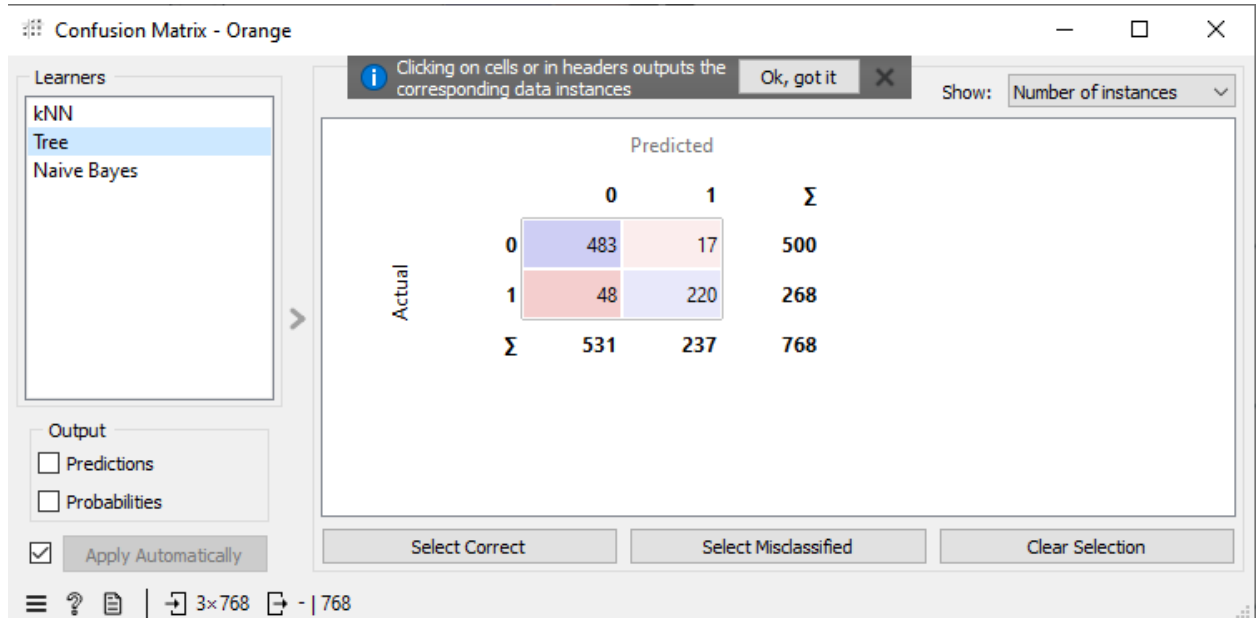
**Decision Tree: 75.8% (0.758)**

**Naive Bayes: 76.2% (0.762)**

Jadi, model dengan akurasi tertinggi adalah **Naive Bayes** dengan nilai akurasi **76.2%**.



5. Confusion Matrix untuk model yang dipilih, yaitu Tree



6. Analisa ROC dengan target "1" atau positif dari suatu kondisi

