



**Capstone
Project**

DIABETES DIAGNOSIS

Setya Kristendy Cianjur - 71210678

Dhea Angelina - 71210693

Angelene Nadine - 71210725

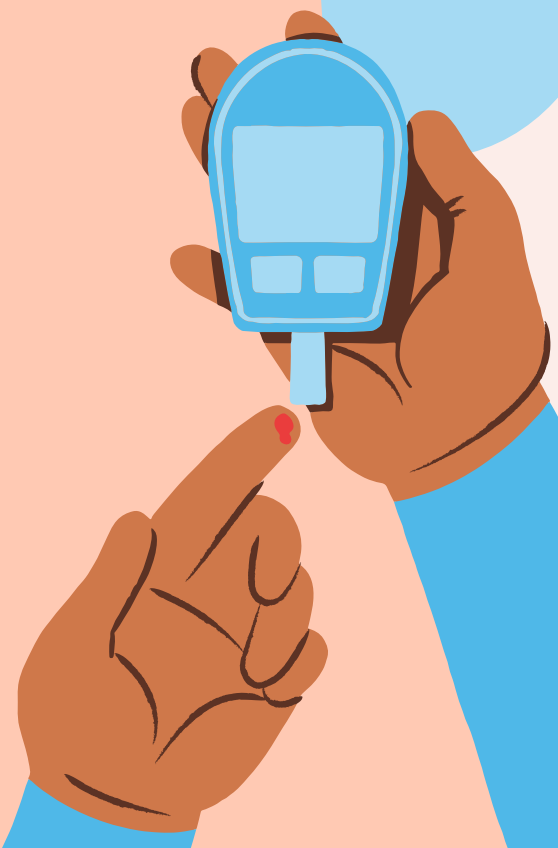
PEMBATASAN MASALAH

Masalah apa yang ingin dicari penyelesaiannya?

- Kurangnya kesadaran masyarakat akan penyakit diabetes.
- Banyak penderita diabetes yang tidak terdeteksi dan terlambat mendapatkan penanganan sehingga kesehatannya menurun perlahan-lahan dan mempengaruhi produktivitas.

Solusi yang ingin dihasilkan?

Membuat diagnosis yang akurat tentang apakah seseorang mengidap diabetes atau tidak berdasarkan faktor-faktor yang relevan, seperti kehamilan, glukosa darah, tekanan darah, ketebalan kulit, insulin, indeks massa tubuh (BMI), riwayat diabetes, dan usia



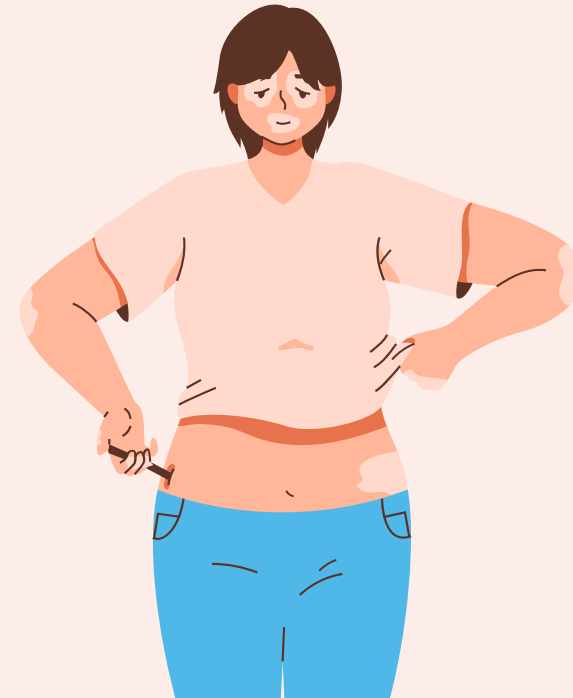
METODE 4W CANVAS

Para [stakeholders]	Masyarakat Indonesia	WHO
Masalah yang dihadapi [isu, masalah, kebutuhan]	Kurangnya kesadaran akan penyakit diabetes	WHAT
Ketika [konteks dan situasi]	Saat pengguna menderita ataupun tidak menderita gejala diabetes	WHERE
Solusi yang diharapkan	Mendeteksi diabetes sedini mungkin	WHY



AKUISISI DATA

- Apakah ada data untuk dianalisis? **Ya**
- Apakah memilih data primer atau sekunder atau keduanya? **Data Sekunder**
- Apakah ada literatur yang mendukung solusi? **Ya**
- Apakah ada sumber data lain untuk melengkapi data sekunder? **Tidak**



EKSPLORASI

No	Nama Parameter	Status (Input/Output)	Tipe Data (Nominal/Numerik)
1	Pregnancies	Input	Numerik
2	Glucose	Input	Numerik
3	Blood Pressure	Input	Numerik
4	Skin Thickness	Input	Numerik
5	Insulin	Input	Numerik
6	Body Mass Index (BMI)	Input	Numerik
7	Diabetes Pedigree Function	Input	Numerik
8	Age	Input	Numerik
9	Outcome	Output	Numerik

EKSPLORASI

File - Orange

Source

☒ File: diabetes.csv

☐ URL:

File Type

Automatically detect type

Info

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

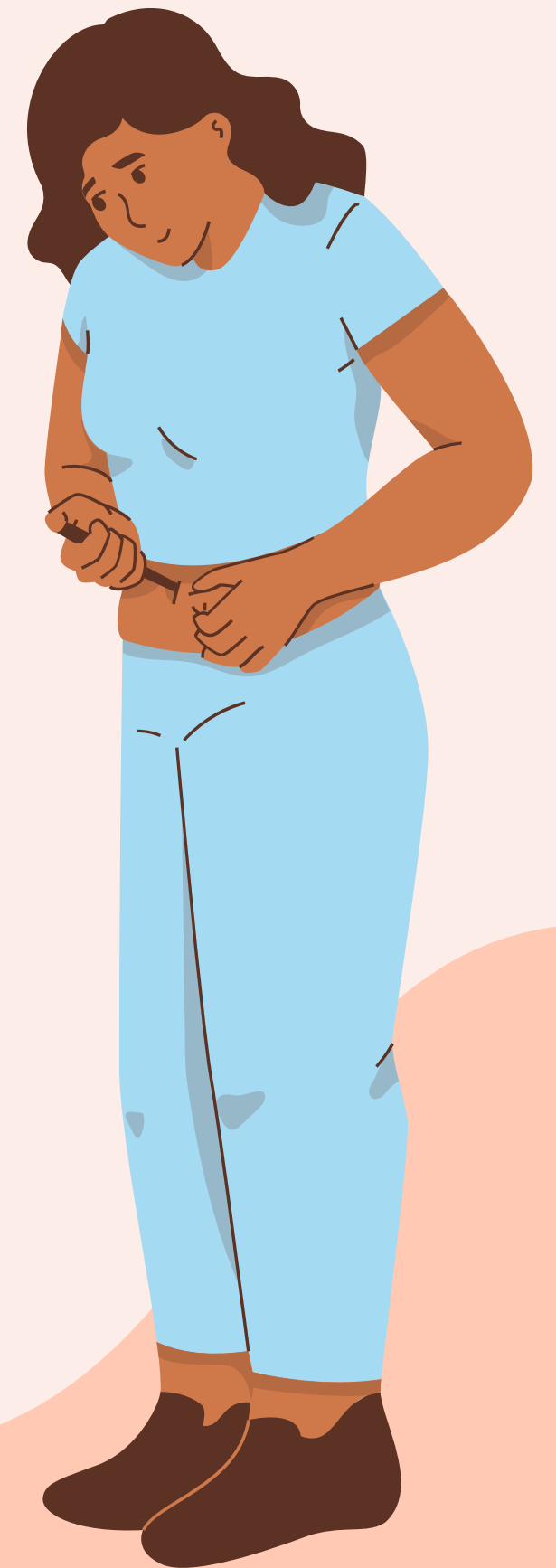
Columns (Double click to edit)

	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

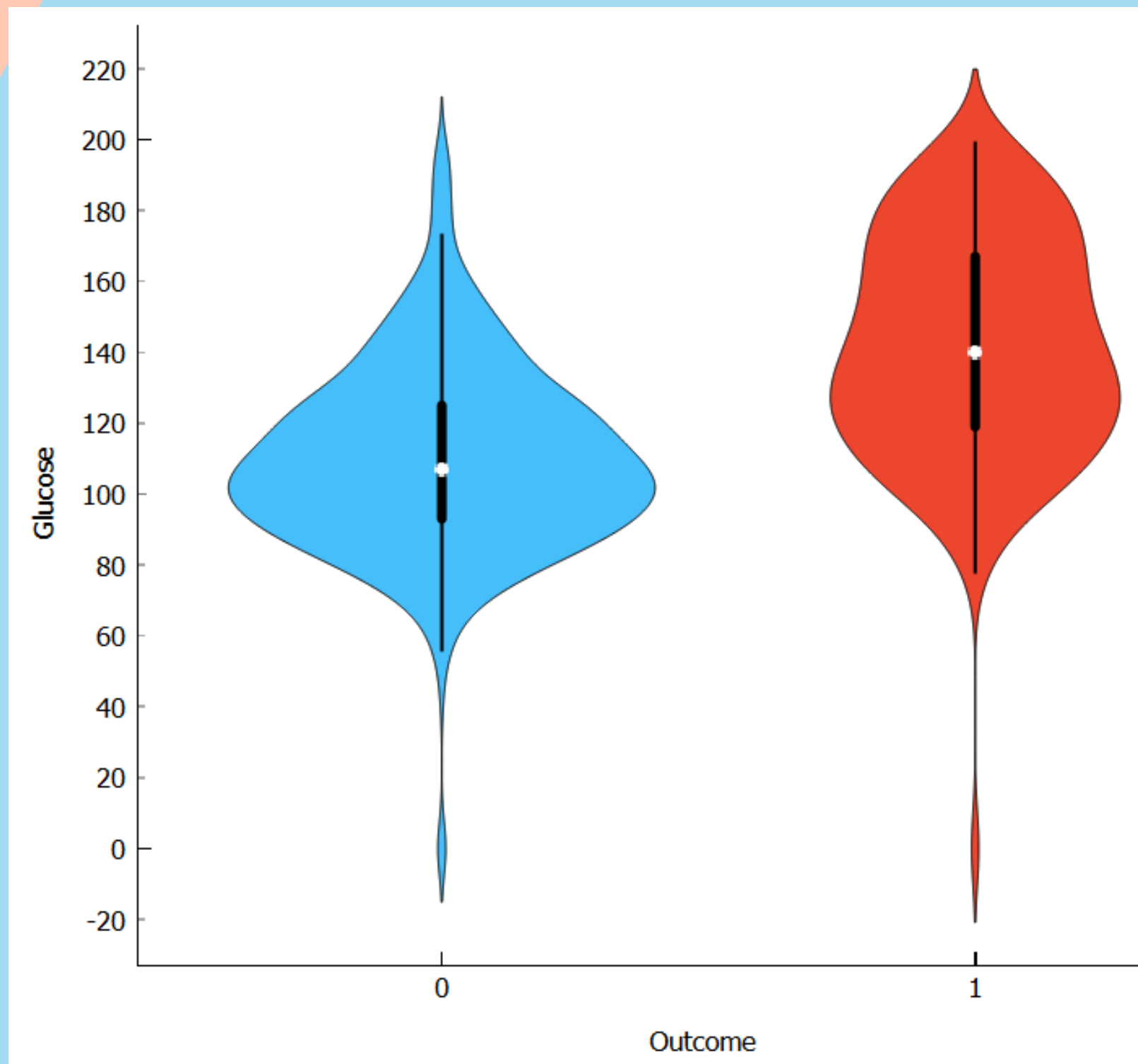
	Outcome	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age
1	1	6	148	72	35	0	33.6	0.627	50
2	0	1	85	66	29	0	26.6	0.351	31
3	1	8	183	64	0	0	23.3	0.672	32
4	0	1	89	66	23	94	28.1	0.167	21
5	1	0	137	40	35	168	43.1	2.288	33
6	0	5	116	74	0	0	25.6	0.201	30
7	1	3	78	50	32	88	31.0	0.248	26
8	0	10	115	0	0	0	35.3	0.134	29
9	1	2	197	70	45	543	30.5	0.158	53
10	1	8	125	96	0	0	0.0	0.232	54
11	0	4	110	92	0	0	37.6	0.191	30
12	1	10	168	74	0	0	38.0	0.537	34
13	0	10	139	80	0	0	27.1	1.441	57
14	1	1	189	60	23	846	30.1	0.398	59
15	1	5	166	72	19	175	25.8	0.587	51
16	1	7	100	0	0	0	30.0	0.484	32
17	1	0	118	84	47	230	45.8	0.551	31
18	1	7	107	74	0	0	29.6	0.254	31
19	0	1	103	30	38	83	43.3	0.183	33
20	1	1	115	70	30	96	34.6	0.529	32
21	0	3	126	88	41	235	39.3	0.704	27

EKSPLORASI DATA

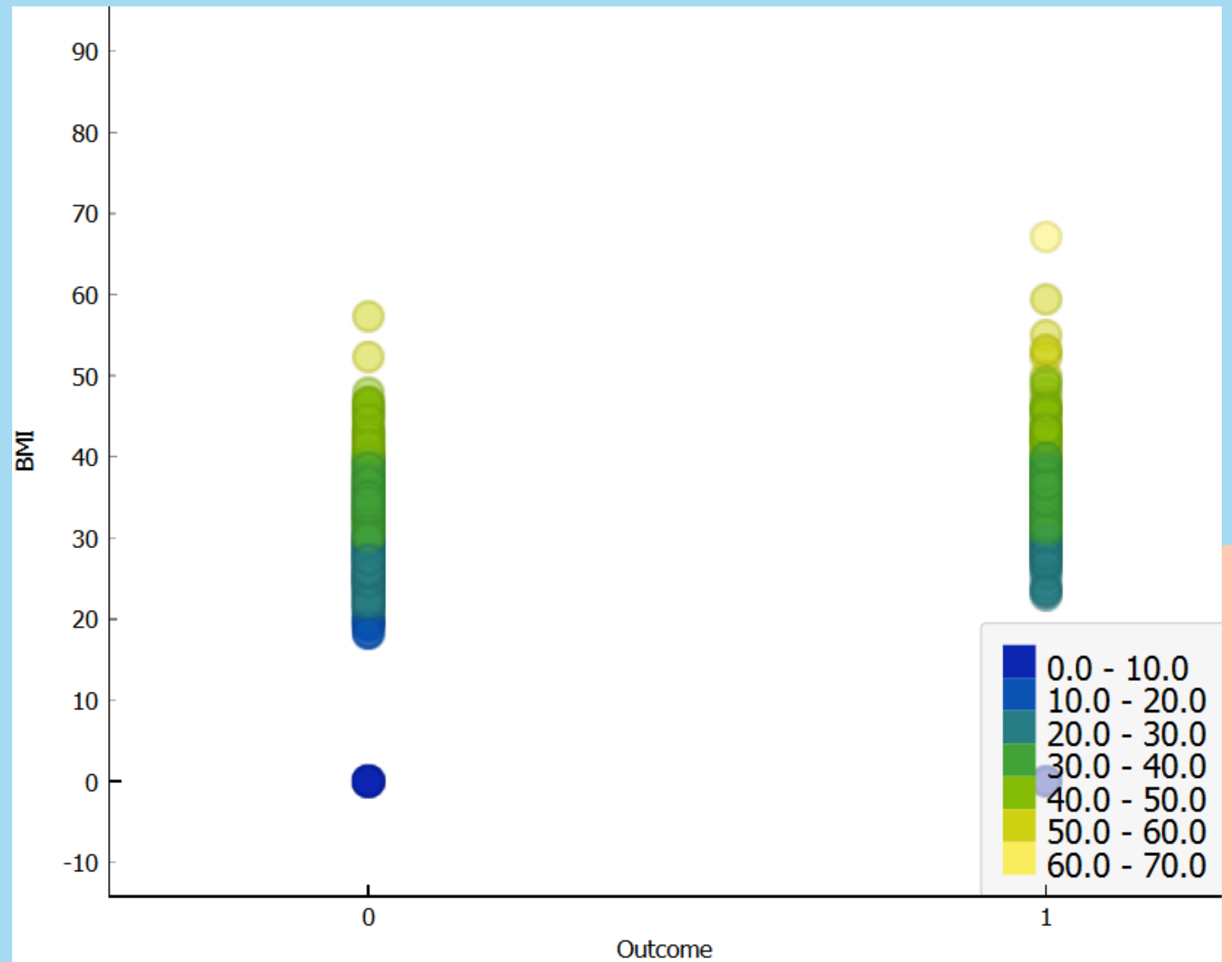
- Visualisasi yang kami gunakan
 - Violin Plot
 - Scatter Plot
- Pilihan Metode Klasifikasi



VIOLIN PLOT



SCATTER PLOT



MODELLING

- Algoritma untuk menghasilkan model (pengetahuan)
Logistic Regression (LogR)

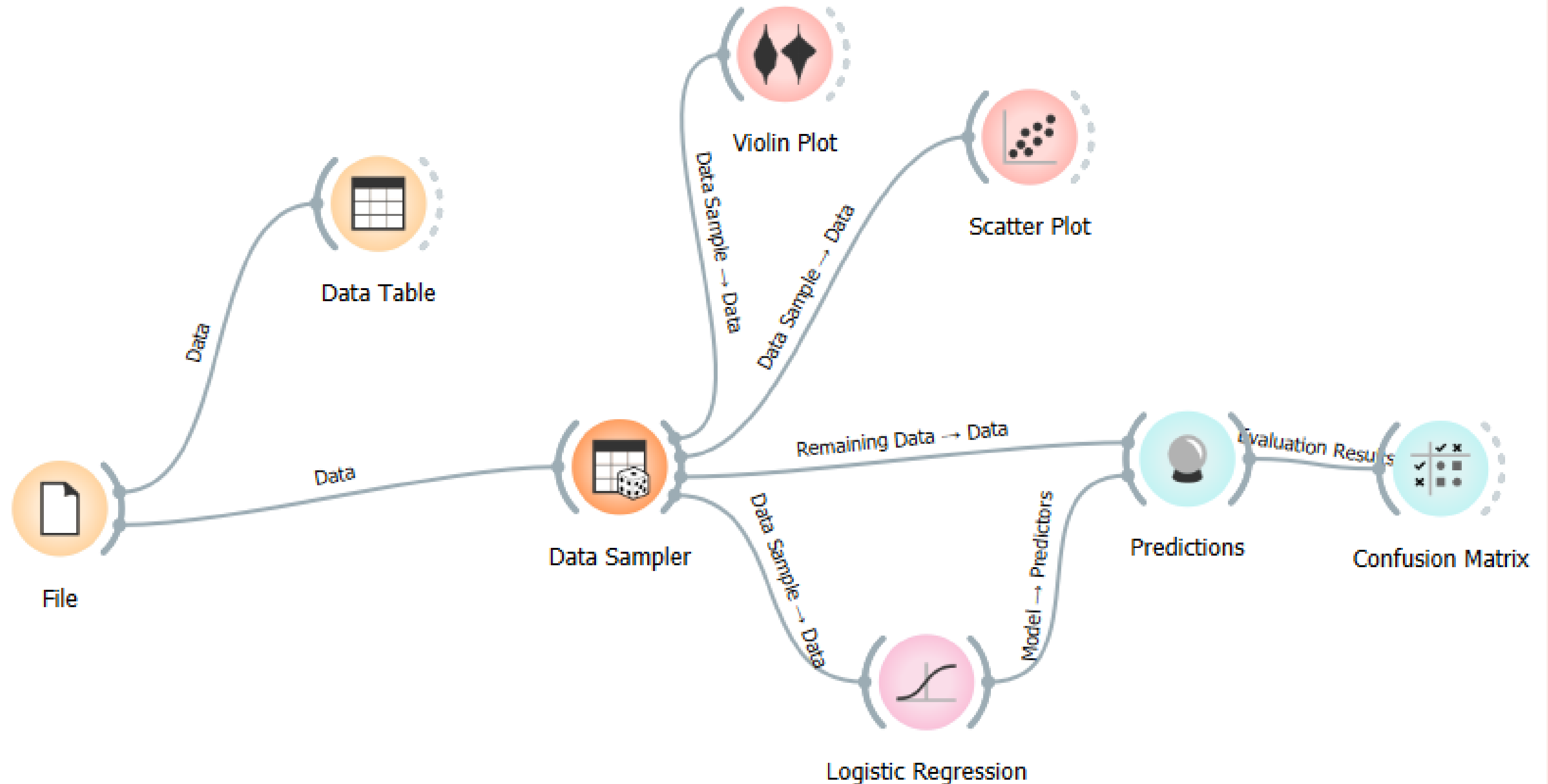
- Pilihan Model Evaluasi

Confusion Matrix

Gambaran jumlah diagnosis benar dan salah yang mencakup empat matriks evaluasi utama (True Positive, True Negative, False Positive, dan False Negative)



MODELLING USING ORANGE



LOGISTIC REGRESSION

Predictions - Orange

Show probabilities for 1 ☐ Show classification errors

Restore Original Order

	Logistic Regression	Outcome	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	etesPedigreeFunc	Age
1	<u>0.48</u> → 0	0	10	122	68	0	0	31.2	0.258	41
2	<u>0.54</u> → 1	1	3	173	84	33	474	35.7	0.258	22
3	<u>0.31</u> → 0	1	4	125	80	0	0	32.3	0.536	27
4	<u>0.12</u> → 0	1	0	105	84	0	0	27.9	0.741	62
5	<u>0.46</u> → 0	1	5	136	84	41	88	35.0	0.286	35
6	<u>0.07</u> → 0	0	1	111	62	13	182	24.0	0.138	23
7	<u>0.21</u> → 0	0	6	96	0	0	0	23.7	0.190	28
8	<u>0.12</u> → 0	0	0	99	0	0	0	25.0	0.253	22
9	<u>0.34</u> → 0	1	5	112	66	0	0	37.8	0.261	41
10	<u>0.24</u> → 0	1	3	130	78	23	79	28.4	0.323	34
11	<u>0.82</u> → 1	1	3	176	86	27	156	33.3	1.154	52
12	<u>0.74</u> → 1	1	1	168	88	29	0	35.0	0.905	52
13	<u>0.07</u> → 0	0	0	67	76	0	0	45.3	0.194	46
14	<u>0.36</u> → 0	1	5	116	74	29	0	32.3	0.660	35
15	<u>0.93</u> → 1	0	13	153	88	37	140	40.6	1.174	39

☒ Show performance scores

Target class: 0

Model	AUC	CA	F1	Prec	Recall	MCC
Logistic Regression	0.818	0.729	0.803	0.752	0.862	0.385

CONFUSION MATRIX

Confusion Matrix - Orange

Show: Number of instances

		Predicted		
		0	1	Σ
Actual	0	106	17	123
	1	35	34	69
Σ		141	51	192

HASIL PROGRAM

Diabetes Diagnosis

Kehamilan (Pregnancies)

Glukosa (Glucose)

Tekanan Darah (BloodPressure)

Ketebalan Kulit (SkinThickness)

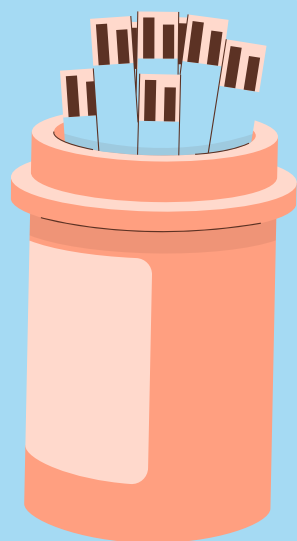
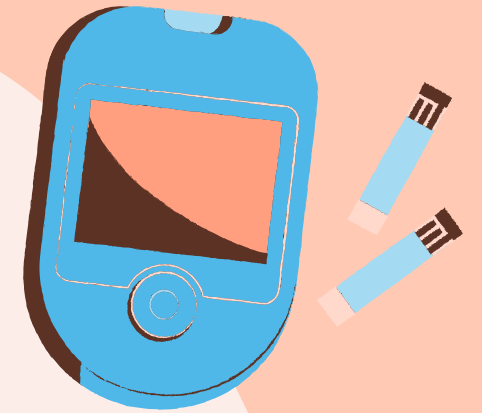
Insulin (Insulin)

Indeks Massa Tubuh (BMI)

Riwayat Diabetes (DiabetesPedigreeFunction)

Umur (Age)

Test



Diabetes Diagnosis

Kehamilan (Pregnancies)

6

Glukosa (Glucose)

148

Tekanan Darah (BloodPressure)

72

Ketebalan Kulit (SkinThickness)

35

Insulin (Insulin)

0

Indeks Massa Tubuh (BMI)

33.6

Riwayat Diabetes (DiabetesPedigreeFunction)

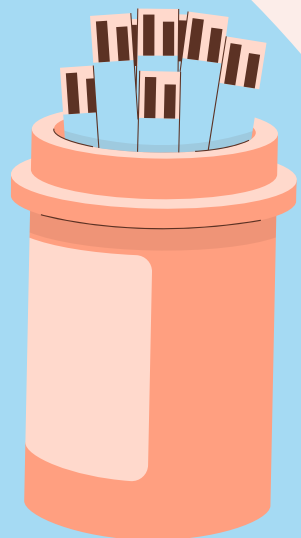
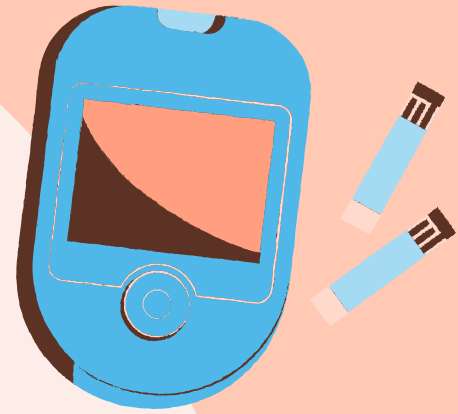
0.627

Umur (Age)

50

Test

Anda TERDETEKSI diabetes (You HAVE Diabetes)



Diabetes Diagnosis

Kehamilan (Pregnancies)

5

Glukosa (Glucose)

116

Tekanan Darah (BloodPressure)

74

Ketebalan Kulit (SkinThickness)

0

Insulin (Insulin)

0

Indeks Massa Tubuh (BMI)

25.6

Riwayat Diabetes (DiabetesPedigreeFunction)

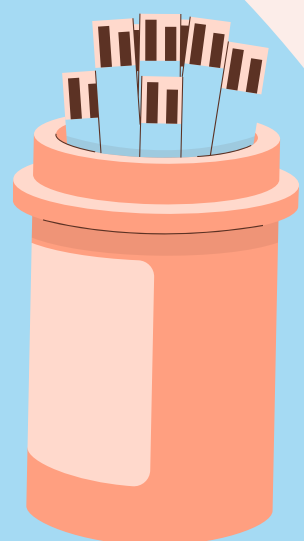
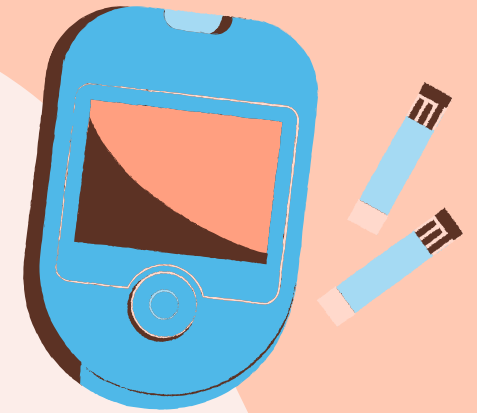
0.201

Umur (Age)

30

Test

Anda TIDAK TERDETEKSI diabetes (You DON'T HAVE Diabetes)



KESIMPULAN

Pendekatan AI untuk mendiagnosis diabetes (**Diabetes Diagnosis**) diimplementasikan menggunakan **Model Klasifikasi** dengan algoritma **Logistic Regression** dan dievaluasi dengan menggunakan **Confusion Matrix** sehingga memiliki kinerja yang baik dalam mendeteksi diabetes. Eksplorasi data yang digunakan divisualisasikan dalam bentuk **Violin Plot dan Scatter Plot** yang dapat membantu dalam pemahaman distribusi variabel yang relevan, sehingga dapat memperkaya interpretasi hasil diagnosis.





THANK YOU!