

LOGISTIC REGRESSION, K- NEAREST NEIGHBOR, NÄÏVE BAYES, DECISION TREE

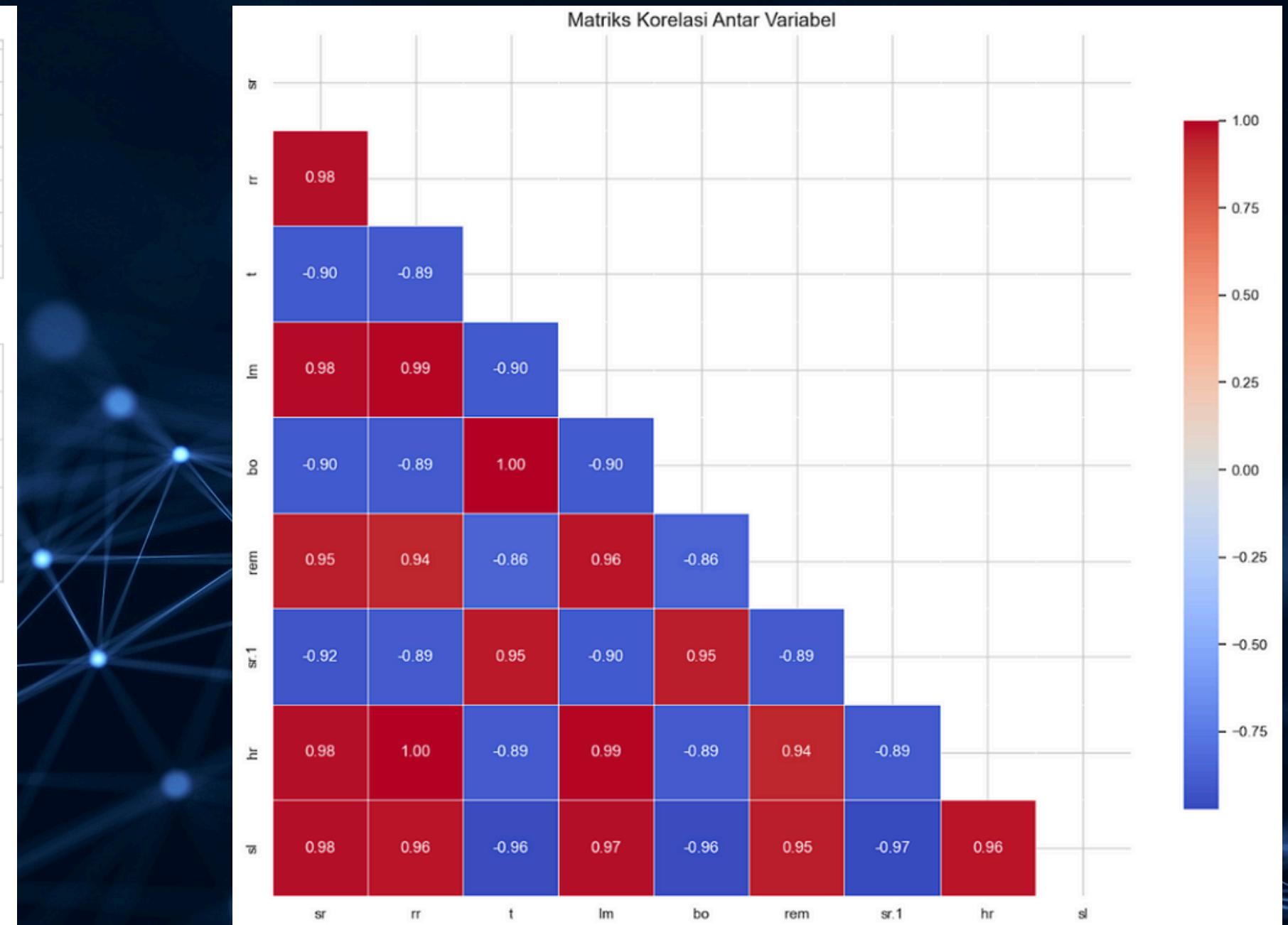
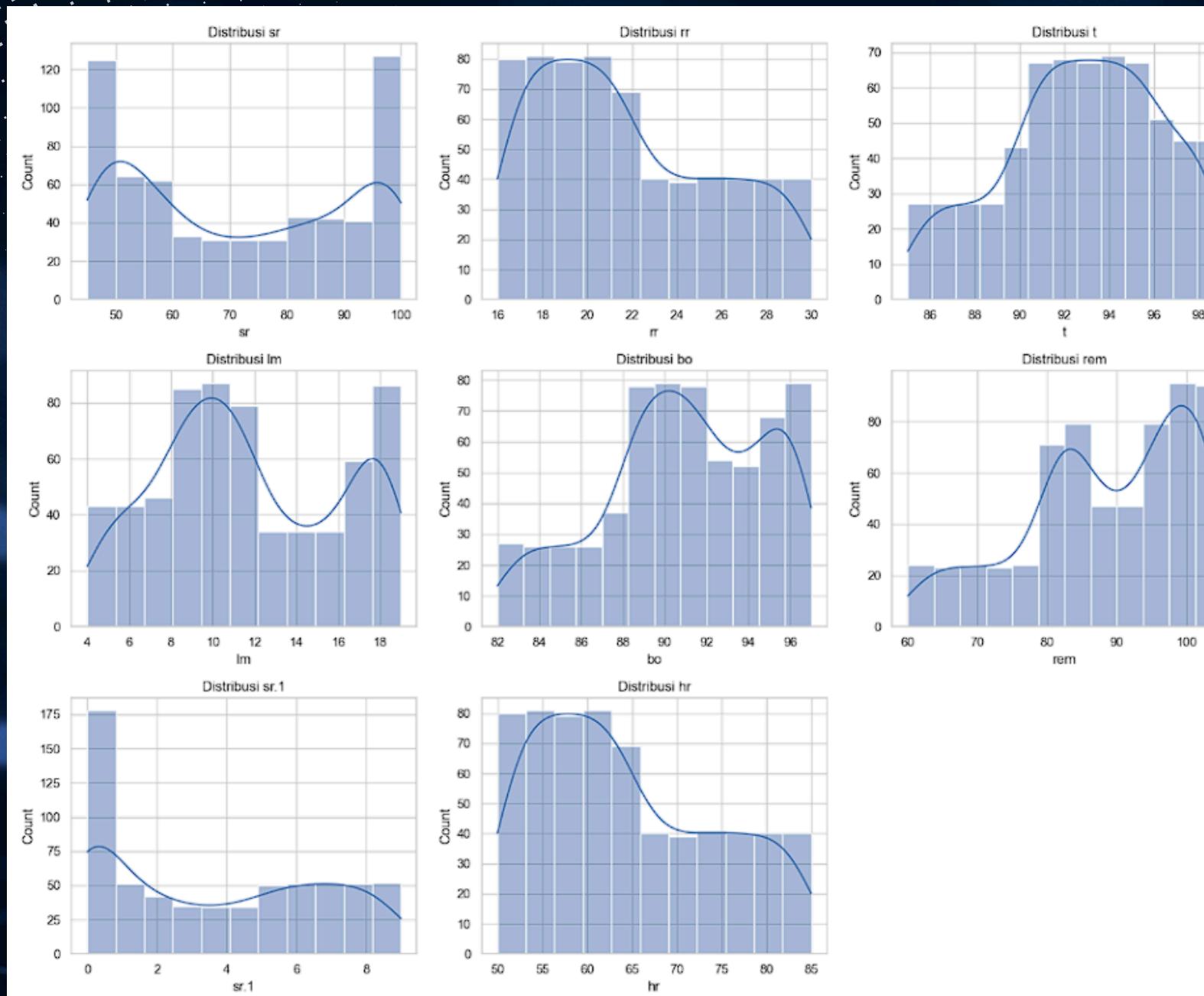
Tugas 3 - Kelompok 10

DESKRIPSI DATASET

Dataset bernama "**SayoPillow**" dan berisi data mengenai berbagai parameter fisiologis dan kebiasaan tidur yang berkaitan dengan tingkat stres seseorang. Memiliki **630 baris** dan **9 kolom**.

DATA EXPLORATION

DISTRIBUSI DATA KORELASI FITUR



DATA PREPROCESSING

MISSING VALUE

Jumlah Missing Values per Kolom:

```
sr      0  
rr      0  
t       0  
lm      0  
bo      0  
rem     0  
sr.1    0  
hr      0  
sl      0  
dtype: int64
```

NORMALISASI

```
# Membagi data menjadi fitur (X) dan target (y)  
X = df.drop('sl', axis=1)  
y = df['sl']  
  
# Standarisasi fitur  
scaler = StandardScaler()  
X_scaled = scaler.fit_transform(X)  
X_scaled_df = pd.DataFrame(X_scaled, columns=X.columns)  
  
# Melihat hasil standarisasi  
print("Data setelah standarisasi:")  
X_scaled_df.head()
```

```
# Membagi data menjadi set training dan testing (80% training, 20% testing)  
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2, random_state=42, stratify=y)  
  
# Memeriksa distribusi kelas pada data training dan testing  
print("Distribusi Target pada Training Set:")  
print(pd.Series(y_train).value_counts(normalize=True) * 100)
```

Data setelah standarisasi:

	sr	rr	t	lm	b
0	1.146845	0.979066	-0.272195	1.140539	-0.27183
1	1.035260	0.833720	-0.353853	0.972949	-0.34569
2	-0.599252	-0.454206	0.907316	-0.395697	1.05144
3	0.731501	0.438056	-0.576145	0.516734	-0.54675
4	-1.212970	-1.148636	1.438095	-1.211299	1.37149

Distribusi Target pada Training Set:

```
sl  
0  20.039683  
4  20.039683  
1  20.039683  
2  20.039683  
3  19.841270
```

Name: proportion, dtype: float64

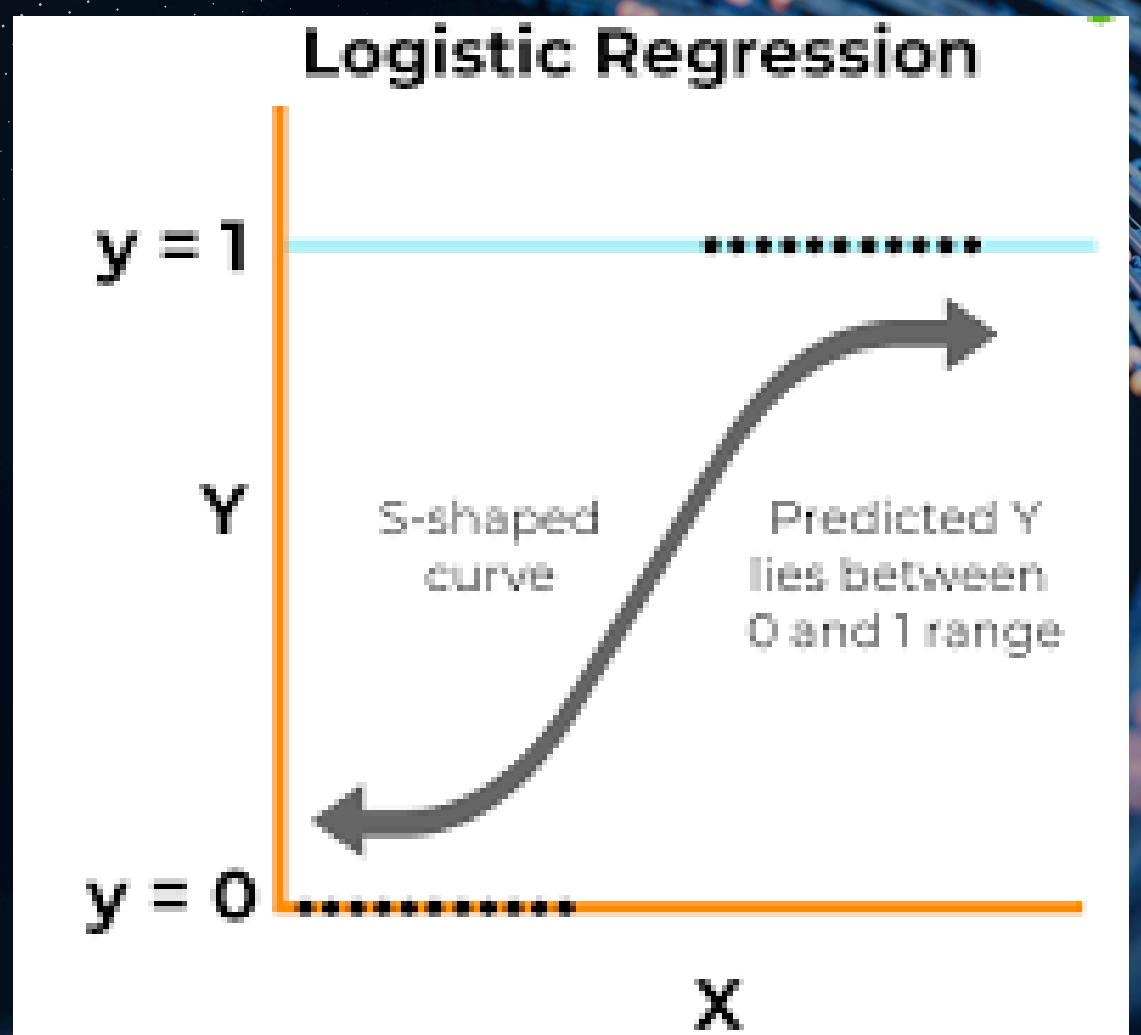
Distribusi Target pada Testing Set:

```
sl  
3  20.634921  
0  19.841270  
1  19.841270  
4  19.841270  
2  19.841270
```

Name: proportion, dtype: float64

SPLIT 80:20

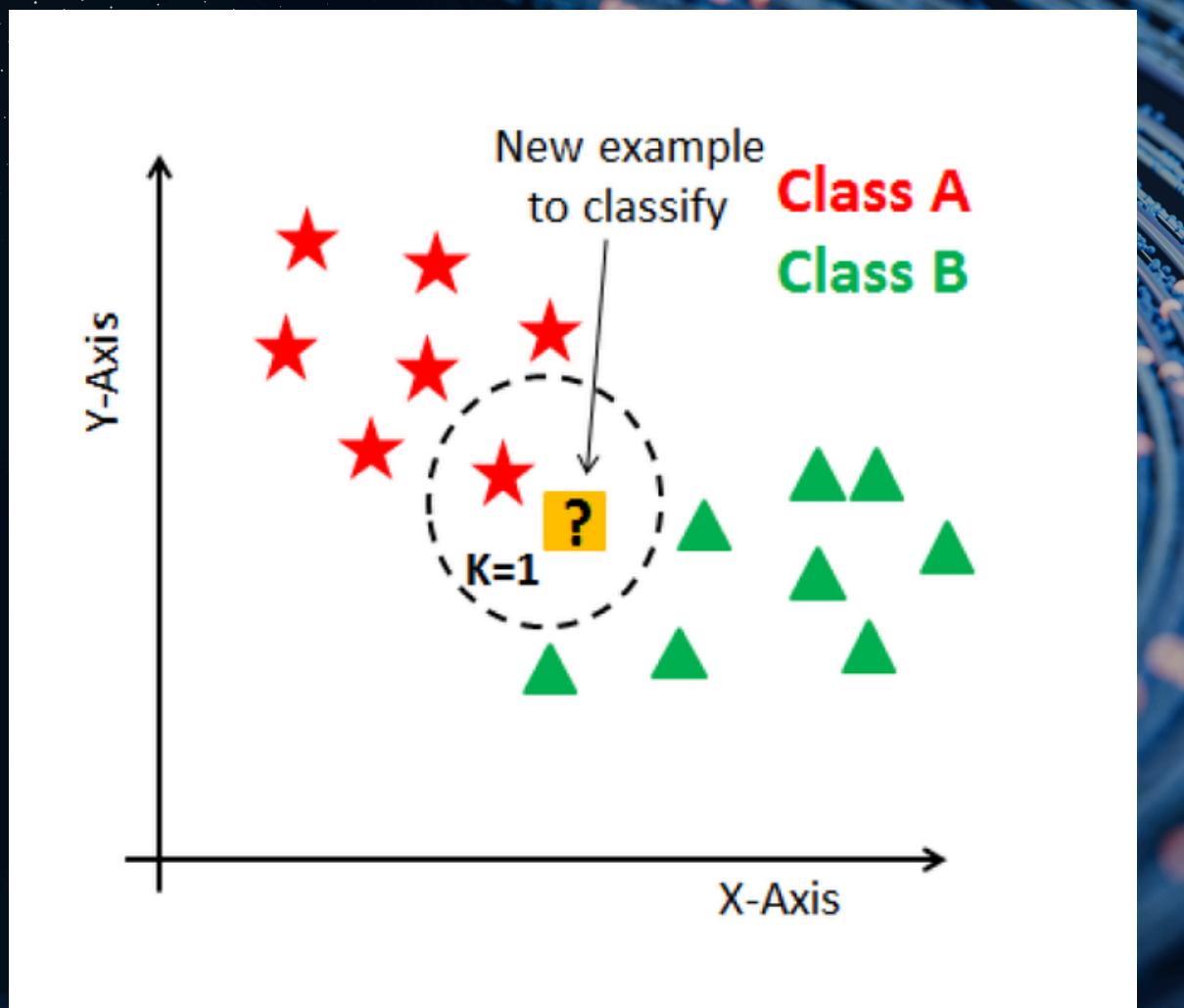
IMPLEMENTASI MODEL LOGISTIC REGRESSION



Classification Report - Logistic Regression:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	25
1	1.00	1.00	1.00	25
2	1.00	1.00	1.00	25
3	1.00	1.00	1.00	26
4	1.00	1.00	1.00	25
accuracy			1.00	126
macro avg	1.00	1.00	1.00	126
weighted avg	1.00	1.00	1.00	126

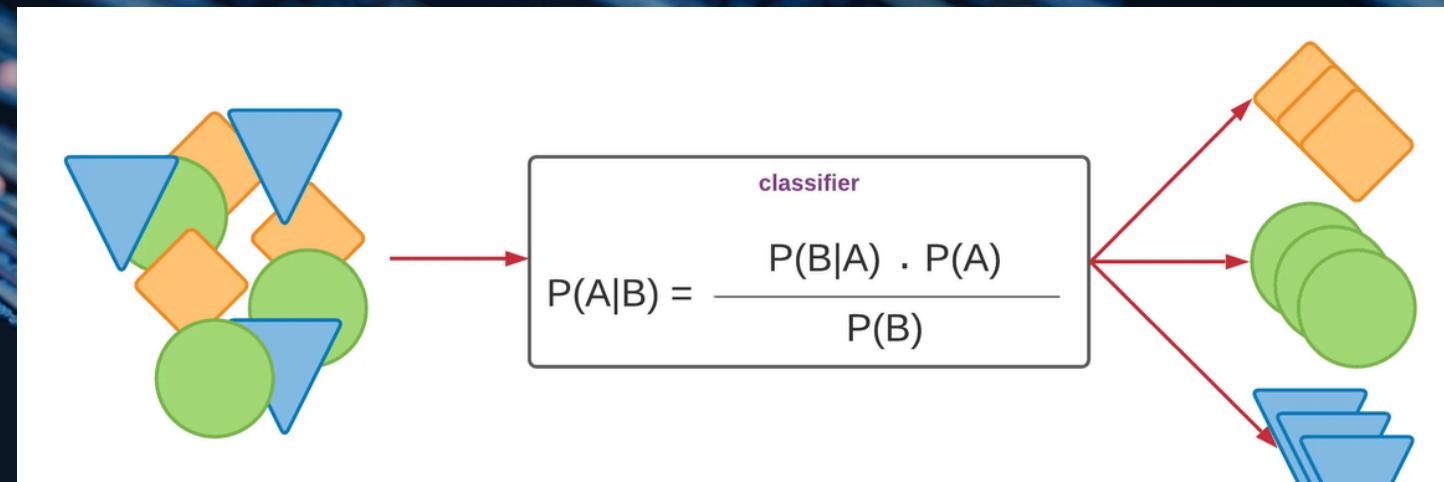
IMPLEMENTASI MODEL K-NEAREST NEIGHBORS (KNN)



Classification Report - KNN:

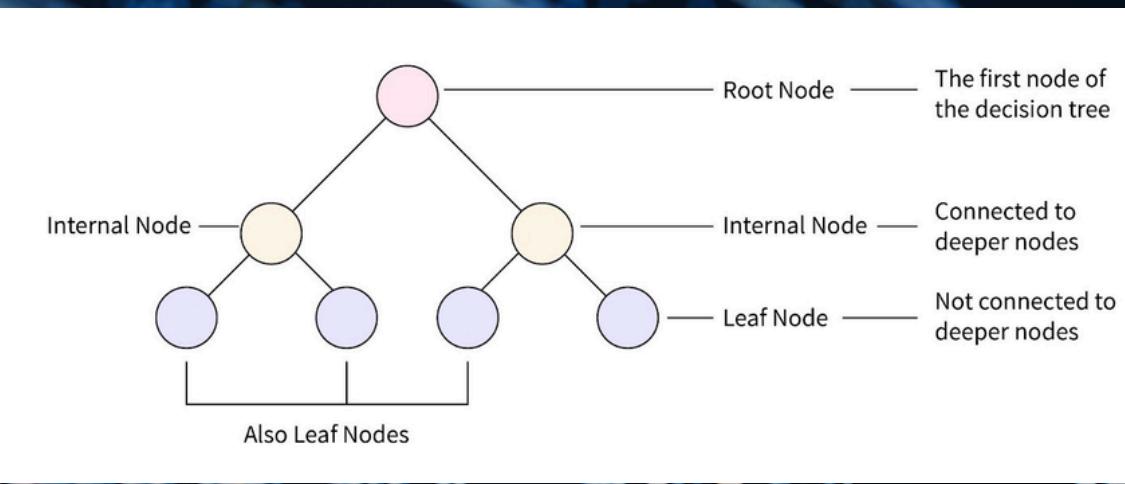
	precision	recall	f1-score	support
0	1.00	1.00	1.00	25
1	1.00	1.00	1.00	25
2	1.00	1.00	1.00	25
3	1.00	1.00	1.00	26
4	1.00	1.00	1.00	25
accuracy			1.00	126
macro avg	1.00	1.00	1.00	126
weighted avg	1.00	1.00	1.00	126

IMPLEMENTASI MODEL NAIVE BAYES



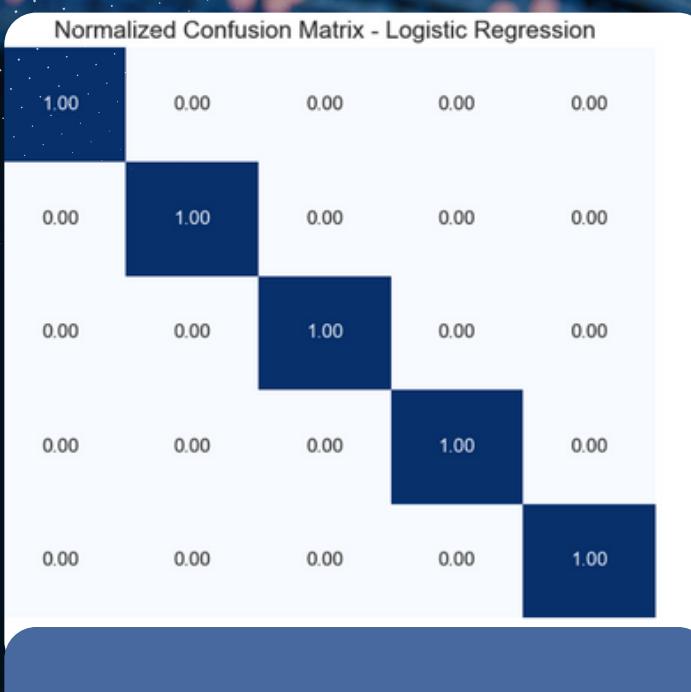
Classification Report - Naive Bayes:				
	precision	recall	f1-score	support
0	1.00	1.00	1.00	25
1	1.00	1.00	1.00	25
2	1.00	1.00	1.00	25
3	1.00	1.00	1.00	26
4	1.00	1.00	1.00	25
accuracy			1.00	126
macro avg	1.00	1.00	1.00	126
weighted avg	1.00	1.00	1.00	126

IMPLEMENTASI MODEL DECISION TREE

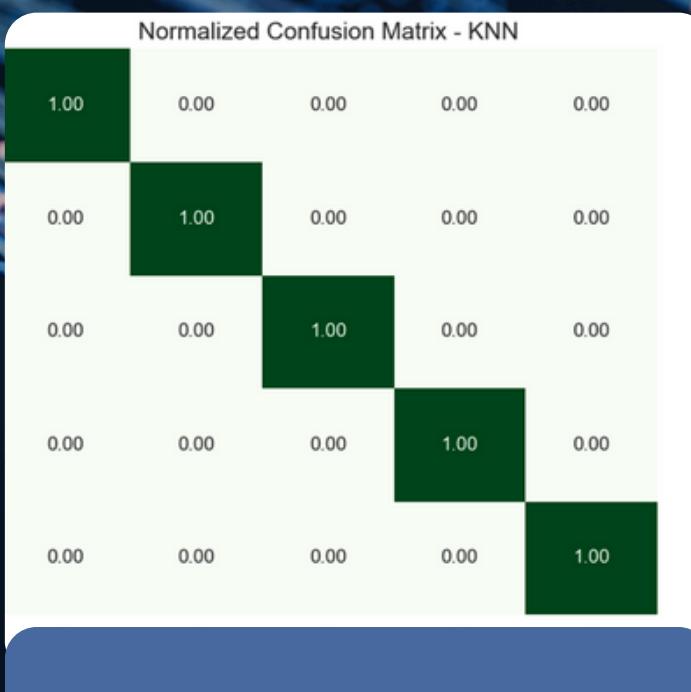


Classification Report - Decision Tree:				
	precision	recall	f1-score	support
0	1.00	0.96	0.98	25
1	0.96	0.96	0.96	25
2	0.93	1.00	0.96	25
3	0.96	0.96	0.96	26
4	1.00	0.96	0.98	25
accuracy			0.97	126
macro avg	0.97	0.97	0.97	126
weighted avg	0.97	0.97	0.97	126

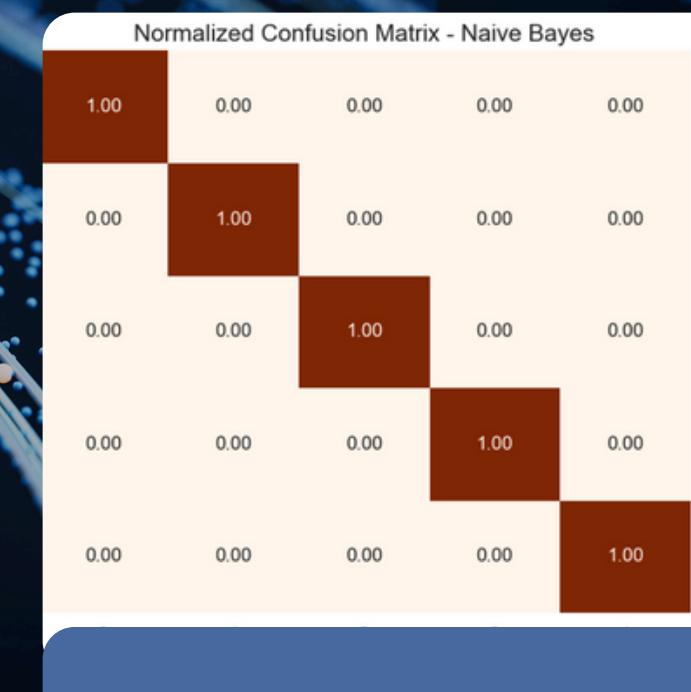
PERBANDING PERFORMA MODEL



Logistic
Regression



KNN



Naive Bayes

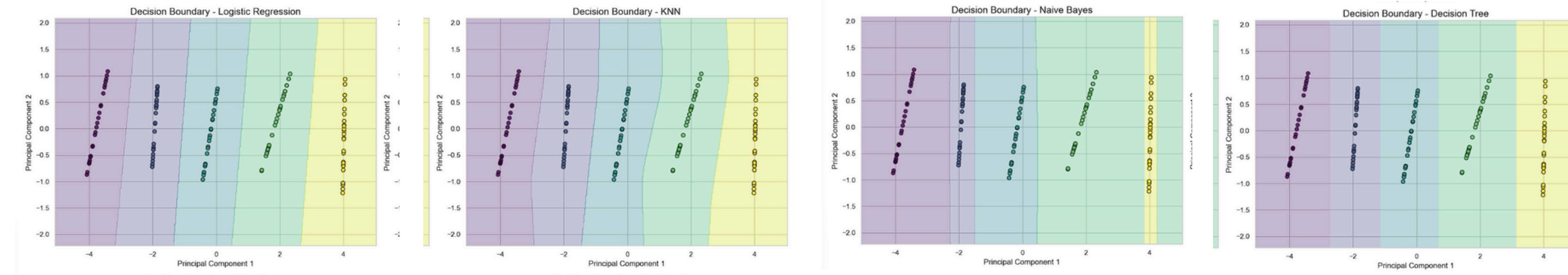


Product
Adaptation

Model	Accuracy	Log Loss
Logistic Regression	1.000000	6.280729e-02
KNN	1.000000	2.220446e-16
Naive Bayes	1.000000	2.396672e-16
Decision Tree	0.968254	1.144243e+00

ANALISIS HASIL

Desisioun Boundary



Grafik decision boundary memperlihatkan perbandingan empat algoritma klasifikasi (Logistic Regression, KNN, Naive Bayes, dan Decision Tree) setelah dimensi data direduksi menggunakan **PCA**, di mana dua komponen utama menjelaskan 97,79% varians total (93,71% oleh PC1 dan 4,07% oleh PC2). **Logistic Regression** dan **KNN** menampilkan batas keputusan yang sangat bersih dan konsisten, sementara **Naive Bayes** menunjukkan sedikit perbedaan di area kanan atas. **Decision Tree**, meskipun memiliki akurasi sedikit lebih rendah (96,83%), masih menghasilkan separasi yang baik.

TERIMA KASIH