

UAS

Mata Kuliah : Pembelajaran Mesin
Kelas/Tahun Ajar : Indralaya/Genap 2021-2022
Dosen Pengampu : Alvi Syahrini, M.Kom/Annisa Darmawahyuni, M.Kom
Hari/Tanggal : Kamis/28 April 2022

NIM : 09021181924006

Nama Mahasiswa : Zananda Aditya

Nama perangkat lunak							
Sistem Pendukung Keputusan Resiko Kesehatan Ibu Hamil							
Deskripsi perangkat lunak (latar belakang memilih perangkat lunak dan tujuan yang dicapai)							
Sistem ini dapat mengelola data kesehatan ibu hamil yang berkaitan dengan faktor resiko kehamilan dengan penentuan metode menggunakan data yang sudah ada sebelumnya sehingga hal ini sangat akurat dalam menggunakan Metode Naïve Bayes serta hasil akhir terdapat 3 kemungkinan <i>low risk</i> , <i>mid risk</i> , dan <i>high risk</i> . Harapan yang ingin dicapai agar ibu hamil mengetahui apakah semasa kehamilan rentan akan terkena penyakit atau tidak.							
Metode <i>machine/deep learning</i> yang digunakan							
Metode Naïve Bayes							
Dataset yang digunakan (serta link unduh dataset jika tersedia)							
https://archive.ics.uci.edu/ml/machine-learning-databases/00639/Maternal%20Health%20Risk%20Data%20Set.csv							
Pra-pengolahan data yang diimplementasikan							
	Age	SystolicBP	DiastolicBP	BS	BodyTemp	HeartRate	RiskLevel
0	25	130	80	15.0	98.0	86	high risk
1	35	140	90	13.0	98.0	70	high risk
2	29	90	70	8.0	100.0	80	high risk
3	30	140	85	7.0	98.0	70	high risk
...
1010	55	120	90	18.0	98.0	60	high risk
1011	35	85	60	19.0	98.0	86	high risk
1012	43	120	90	18.0	98.0	70	high risk
1013	32	120	65	6.0	101.0	76	mid risk

Metodologi pengerjaan perangkat lunak (dari raw data – prapengolahan – metode *machine/deep learning*)

1. *Mengimport library yang dibutuhkan*

```
import pandas as pd → Berguna untuk mengload data
```

```
from sklearn.model_selection import train_test_split → Mengspliting data
```

```
from sklearn.pipeline import Pipeline
```

```
from sklearn.compose import ColumnTransformer
```

```
from jcopml.pipeline import num_pipe, cat_pipe
```

2. Membaca data dan menampilkan dataset

```
df = pd.read_csv("https://archive.ics.uci.edu/ml/machine-learning-databases/00639/Maternal%20Health%20Risk%20Data%20Set.csv")
```

```
df.head()
```

```
df
```

3. *Mendrop data target*

```
x = df.drop(columns="RiskLevel")
```

```
y = df.RiskLevel
```

4. *Splitting data yang membagi data testing dan data training*

```
X_train, X_test, y_train, y_test = train_test_split(x, y,  
test_size=50, random_state=100) X_train.shape,  
X_test.shape, y_train.shape, y_test.shape
```

5. Membagi data numerik dengan metode onehot encoding

```
preprocessor = ColumnTransformer([  
    ('numeric', num_pipe(),  
    ['Age', 'SystolicBP', 'DiastolicBP', 'BS', 'BodyTemp', 'HeartRate'])  
)
```

6. Inisialisasi model menggunakan metode naïve bayes classifier dan menampung data di Pipeline

```
from sklearn.naive_bayes import GaussianNB  
pipeline = Pipeline([  
    ('prep', preprocessor),  
    ('algo', GaussianNB())  
)
```

7. *Mentraining data*

```
pipeline.fit(X_train, y_train)
```

8. Melihat akurasi atau *performance score data training*

```
pipeline.score(X_train, y_train)
```

9. Melihat akurasi atau *performance score data testing*

```
pipeline.score(X_test, y_test)
```

10. Melihat pesebaran data

```
from jcopml.plot import plot_confusion_matrix  
plot_confusion_matrix(X_train, y_train, X_test, y_test, pipeline)
```

11. Memasukkan data *testing* untuk di uji

```
X_pred = pd.read_csv("data/datasettugass.csv")  
X_pred
```

12. Memanggil pipeline

```
pipeline.predict(X_pred)
```

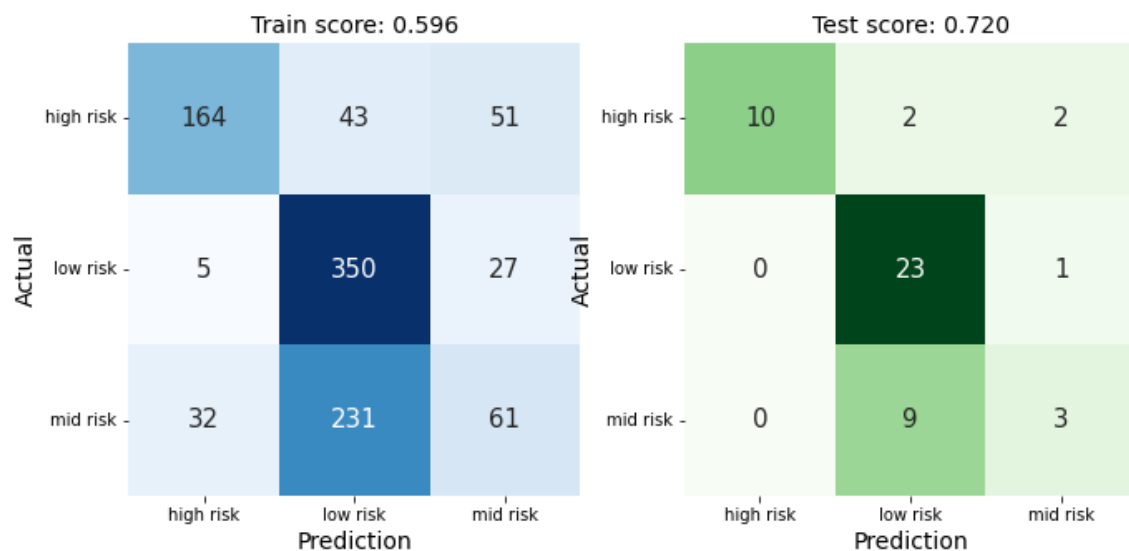
13. Membuat data hasil ke dalam tabel data *testing*

```
X_pred["RiskLevel"] =  
pipeline.predict(X_pred)  
X_pred
```

Hasil uji dataset dengan metode yang diusulkan

`pipeline.score(X_train, y_train)` → Prediksi akurasi pada data *training* sebesar **0.596**

`pipeline.score(X_test, y_test)` → Prediksi akurasi pada data *testing* sebesar **0.720**



Bahasa pemrograman yang dipakai

Python

Screenshot *Graphical User Interface* (jika ada)

-

Link Github

https://github.com/zanandaaditya/MachineLearning/tree/main/UAS_Machine%20Learning

LAMPIRAN

A. datasettugass.csv

Age	SystolicBP	DiastolicBP	BS	BodyTemp	HeartRate
19	120	70	13	102	76
82	140	80	4	98	78
24	90	100	17	95	45
30	120	68	7	100	77
29	118	40	4	98	80
30	120	60	6.5	98	77
24	129	100	10	70	80
35	129	80	6	99	80
35	120	60	6	98	80
25	120	60	6.2	100	67
19	130	60	6.1	88	70
40	140	80	7	98	60
52	120	80	5	99	70
33	120	70	6	99	80
32	140	70	6.5	98	80
39	120	60	7.3	98	70
45	130	80	6	98	60
32	120	60	2	87	30
21	120	80	6	87	50
29	110	80	6	79	60
22	110	80	6	99	60

B. Hasil datasettugass.csv

	Age	SystolicBP	DiastolicBP	BS	BodyTemp	HeartRate	RiskLevel
0	19	120	70	13	102	76	mid risk
1	82	140	80	4	98	78	high risk
2	24	90	100	17	95	45	high risk
3	30	120	68	7	100	77	low risk
4	29	118	40	4	98	80	mid risk
5	30	120	60	6.5	98	77	low risk
6	24	129	100	10	70	80	high risk
7	35	129	80	6	99	80	mid risk
8	35	120	60	6	98	80	mid risk
9	25	120	60	6.2	100	67	mid risk
10	19	130	60	6.1	88	70	high risk
11	40	140	80	7	98	60	low risk
12	52	120	80	5	99	70	mid risk
13	33	120	70	6	99	80	mid risk
14	32	140	70	6.5	98	80	mid risk
15	39	120	60	7.3	98	70	low risk
16	45	130	80	6	98	60	mid risk
17	32	120	60	2	87	30	high risk
18	21	120	80	6	87	50	high risk
19	29	110	80	6	79	60	high risk
20	22	110	80	6	99	60	mid risk