

Nama: Dimas Anggoro Sakti

NIM: 2110511115

Kelas: C

- Bernoulli Naïve Bayes
input:

```
import pandas as pd
import numpy as np
from sklearn import preprocessing
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import BernoulliNB
from sklearn.impute import SimpleImputer

direktori = "C:/Users/DIMAS ANGGORO SAKTI/Dropbox/My PC (LAPTOP-FMDVD92V)/Downloads/C_2110511115_Dimas Anggoro sakti naive bayes/fi

header = ['Class', 'AGE', 'SEX', 'STEROID', 'ANTIVIRALS', 'FATIGUE', 'MALAISE', 'ANOREXIA',
          'LIVER BIG', 'LIVER FIRM', 'SPLEEN PALPABLE', 'SPIDERS', 'ASCITES',
          'VARICES', 'BILIRUBIN', 'ALK PHOSPHATE', 'SGOT', 'ALBUMIN', 'PROTIME', 'HISTOLOGY']

# membaca data dengan library panda
data = pd.read_csv(direktori, names=header)

# mengganti '?' dengan nan
data.replace('?', np.nan, inplace = True)

# panggil nilai dataset
arrayData = data.values

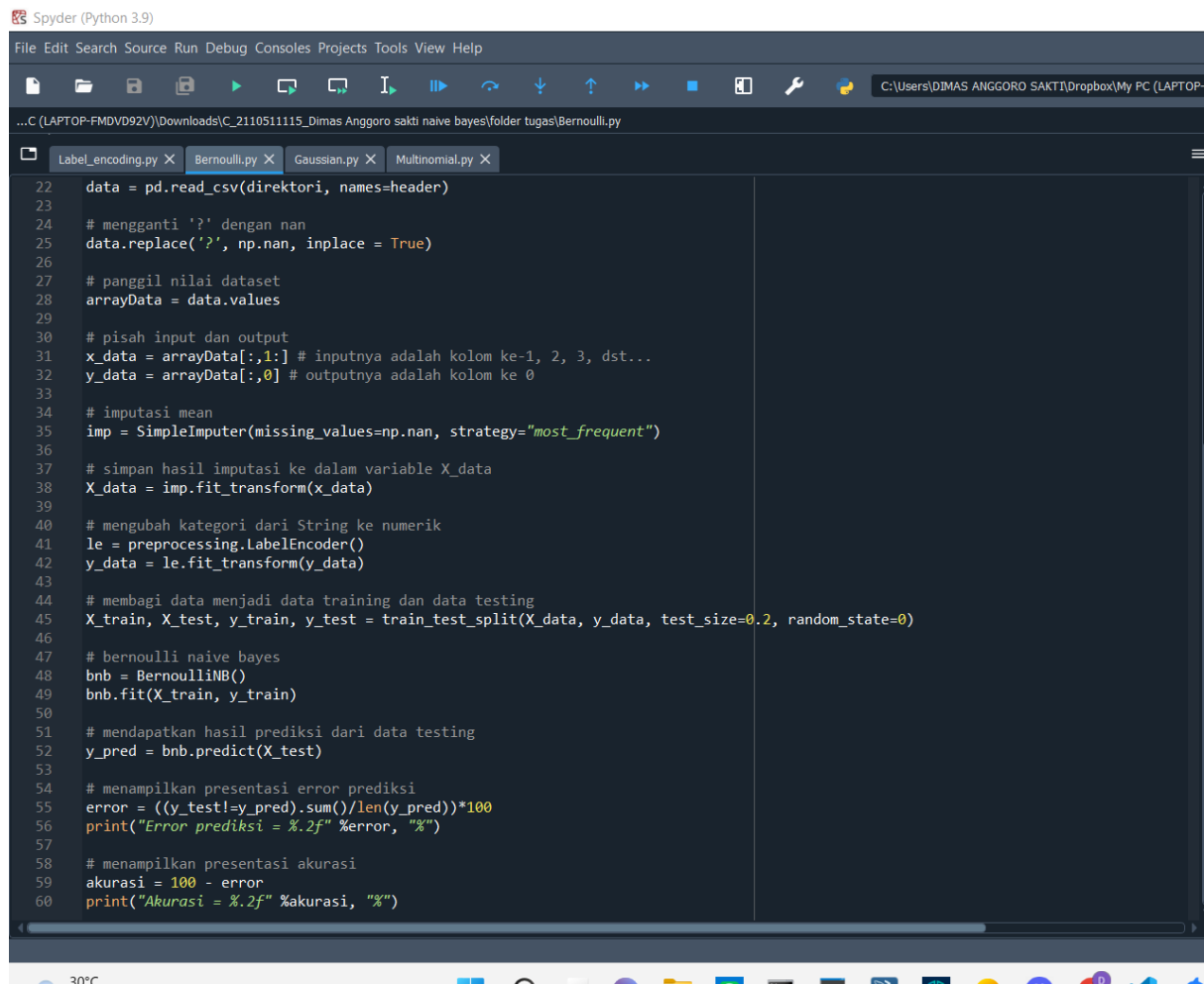
# pisah input dan output
x_data = arrayData[:,1:] # inputnya adalah kolom ke-1, 2, 3, dst...
y_data = arrayData[:,0] # outputnya adalah kolom ke 0

# imputasi mean
imp = SimpleImputer(missing_values=np.nan, strategy="most_frequent")

# simpan hasil imputasi ke dalam variable X_data
X_data = imp.fit_transform(x_data)

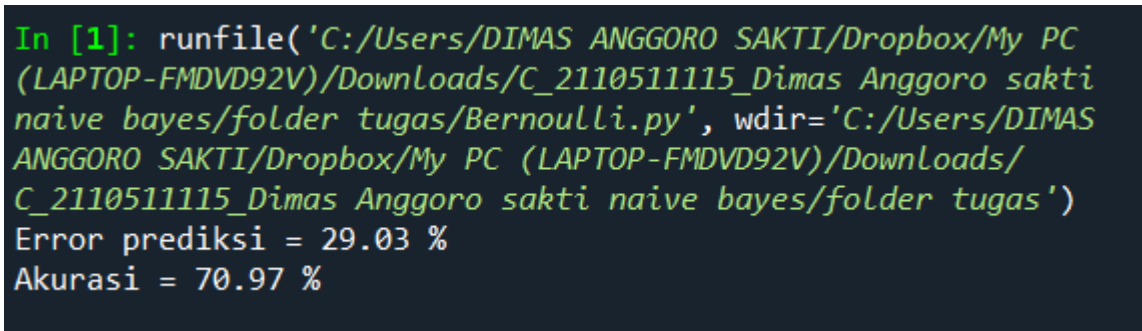
# mengubah kategori dari String ke numerik
le = preprocessing.LabelEncoder()
y_data = le.fit_transform(y_data)

# membagi data menjadi data training dan data testing
X_train, X_test, y_train, y_test = train_test_split(X_data, y_data, test_size=0.2, random_state=0)
```



```
22 data = pd.read_csv(direktori, names=header)
23
24 # mengganti '?' dengan nan
25 data.replace('?', np.nan, inplace = True)
26
27 # panggil nilai dataset
28 arrayData = data.values
29
30 # pisah input dan output
31 x_data = arrayData[:,1:] # inputnya adalah kolom ke-1, 2, 3, dst...
32 y_data = arrayData[:,0] # outputnya adalah kolom ke 0
33
34 # imputasi mean
35 imp = SimpleImputer(missing_values=np.nan, strategy="most_frequent")
36
37 # simpan hasil imputasi ke dalam variable X_data
38 X_data = imp.fit_transform(x_data)
39
40 # mengubah kategori dari String ke numerik
41 le = preprocessing.LabelEncoder()
42 y_data = le.fit_transform(y_data)
43
44 # membagi data menjadi data training dan data testing
45 X_train, X_test, y_train, y_test = train_test_split(X_data, y_data, test_size=0.2, random_state=0)
46
47 # bernoulli naive bayes
48 bnb = BernoulliNB()
49 bnb.fit(X_train, y_train)
50
51 # mendapatkan hasil prediksi dari data testing
52 y_pred = bnb.predict(X_test)
53
54 # menampilkan presentasi error prediksi
55 error = ((y_test!=y_pred).sum())/len(y_pred))*100
56 print("Error prediksi = %.2f" %error, "%")
57
58 # menampilkan presentasi akurasi
59 akurasi = 100 - error
60 print("Akurasi = %.2f" %akurasi, "%")
```

- output:



```
In [1]: runfile('C:/Users/DIMAS ANGGORO SAKTI/Dropbox/My PC
(LAPTOP-FMDVD92V)/Downloads/C_2110511115_Dimas Anggoro sakti
naive bayes/folder tugas/Bernoulli.py', wdir='C:/Users/DIMAS
ANGGORO SAKTI/Dropbox/My PC (LAPTOP-FMDVD92V)/Downloads/
C_2110511115_Dimas Anggoro sakti naive bayes/folder tugas')
Error prediksi = 29.03 %
Akurasi = 70.97 %
```

- Analisa:

Disini program mengilustrasikan probabilitas dataset hepatitis dengan metode Bernoulli Naïve Bayes

- Gaussian Naïve Bayes
- input:

```
import pandas as pd
import numpy as np
from sklearn import preprocessing
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.impute import SimpleImputer

direktori = "C:/Users/DIMAS ANGGORO SAKTI/Dropbox/My PC (LAPTOP-FMDVD92V)/Downloads/C_2110511115_Dimas Anggoro sakti naive bayes/fi

header = ['Class', 'AGE', 'SEX', 'STEROID', 'ANTIVIRALS', 'FATIGUE', 'MALAISE', 'ANOREXIA',
          'LIVER BIG', 'LIVER FIRM', 'SPLEEN PALPABLE', 'SPIDERS', 'ASCITES',
          'VARICES', 'BILIRUBIN', 'ALK PHOSPHATE', 'SGOT', 'ALBUMIN', 'PROTIME', 'HISTOLOGY']

# membaca data dengan library panda
data = pd.read_csv(direktori, names=header)

# mengganti '?' dengan nan
data.replace('?', np.nan, inplace = True)

# panggil nilai dataset
arrayData = data.values

# pisah input dan output
x_data = arrayData[:,1:] # inputnya adalah kolom ke-1, 2, 3, dst...
y_data = arrayData[:,0] # outputnya adalah kolom ke 0

# imputasi mean
imp = SimpleImputer(missing_values=np.nan, strategy="most_frequent")

# simpan hasil imputasi ke dalam variable X_data
X_data = imp.fit_transform(x_data)

# mengubah kategori dari String ke numerik
le = preprocessing.LabelEncoder()
y_data = le.fit_transform(y_data)

# membagi data menjadi data training dan data testing
X_train, X_test, y_train, y_test = train_test_split(X_data, y_data, test_size=0.2, random_state=0)
```

```
# membagi data menjadi data training dan data testing
X_train, X_test, y_train, y_test = train_test_split(X_data, y_data, test_size=0.2, random_state=0)

# gaussian naive bayes
gnb = GaussianNB()
gnb.fit(X_train, y_train)

# mendapatkan hasil prediksi dari data testing
y_pred = gnb.predict(X_test)

# Menampilkan presentasi error prediksi
error = ((y_test!=y_pred).sum()/len(y_pred))*100
print("Error prediksi = %.2f" %error, "%")

# menampilkan presentasi akurasi
akurasi = 100 - error
print("Akurasi = %.2f" %akurasi, "%")
```

- output:

```
In [2]: runfile('C:/Users/DIMAS ANGGORO SAKTI/Dropbox/My PC
(LAPTOP-FMDVD92V)/Downloads/C_2110511115_Dimas Anggoro sakti
naive bayes/folder tugas/Gaussian.py', wdir='C:/Users/DIMAS
ANGGORO SAKTI/Dropbox/My PC (LAPTOP-FMDVD92V)/Downloads/
C_2110511115_Dimas Anggoro sakti naive bayes/folder tugas')
Error prediksi = 22.58 %
Akurasi = 77.42 %
```

- Analisa:
 Disini program mengilustrasikan probabilitas dataset hepatitis dengan metode Gaussian Naïve Bayes
- Multinomial Naïve Bayes
- input:

Spyder (Python 3.9)

```
File Edit Search Source Run Debug Consoles Projects Tools View Help

...(\LAPTOP-FMDVD92V)\Downloads\C_2110511115_Dimas Anggoro sakti naive bayes\folder tugas\Multinomial.py

Label_encoding.py X Bernoulli.py X Gaussian.py X Multinomial.py X

8 import pandas as pd
9 import numpy as np
10 from sklearn import preprocessing
11 from sklearn.model_selection import train_test_split
12 from sklearn.naive_bayes import MultinomialNB
13 from sklearn.impute import SimpleImputer
14
15 direktori = "C:/Users/DIMAS ANGGORO SAKTI/Dropbox/My PC (LAPTOP-FMDVD92V)/Downloads/C_2110511115_Dimas Anggoro sakti naive bayes/fi
16
17 header = ['Class', 'AGE', 'SEX', 'STEROID', 'ANTIVIRALS', 'FATIGUE', 'MALAISE', 'ANOREXIA',
18 'LIVER BIG', 'LIVER FIRM', 'SPLEEN PALPLABLE', 'SPIDERS', 'ASCITES',
19 'VARICES', 'BILIRUBIN', 'ALK PHOSPHATE', 'SGOT', 'ALBUMIN', 'PROTIME', 'HISTOLOGY']
20
21 # membaca data dengan library panda
22 data = pd.read_csv(direktori, names=header)
23
24 # mengganti '?' dengan nan
25 data.replace('?', np.nan, inplace = True)
26
27 # panggil nilai dataset
28 arrayData = data.values
29
30 # pisah input dan output
31 x_data = arrayData[:,1:] # inputnya adalah kolom ke-1, 2, 3, dst...
32 y_data = arrayData[:,0] # outputnya adalah kolom ke 0
33
34 # imputasi mean
35 imp = SimpleImputer(missing_values=np.nan, strategy="most_frequent")
36
37 # simpan hasil imputasi ke dalam variable X_data
38 X_data = imp.fit_transform(x_data)
39
40 # mengubah kategori dari String ke numerik
41 le = preprocessing.LabelEncoder()
42 y_data = le.fit_transform(y_data)
43
44 # membagi data menjadi data training dan data testing
45 X_train, X_test, y_train, y_test = train_test_split(X_data, y_data, test_size=0.2, random_state=0)
46
47 # multinomial naive bayes
```

```
Spyder (Python 3.9)
File Edit Search Source Run Debug Consoles Projects Tools View Help

...(LAPTOP-FMDVD92V)\Downloads\C_2110511115_Dimas Anggoro sakti naive bayes\folder tugas\Multinomial.py

Label_encoding.py X Bernoulli.py X Gaussian.py X Multinomial.py X

22 data = pd.read_csv(direktori, names=header)
23
24 # mengganti '?' dengan nan
25 data.replace('?', np.nan, inplace = True)
26
27 # panggil nilai dataset
28 arrayData = data.values
29
30 # pisah input dan output
31 x_data = arrayData[:,1:] # inputnya adalah kolom ke-1, 2, 3, dst...
32 y_data = arrayData[:,0] # outputnya adalah kolom ke 0
33
34 # imputasi mean
35 imp = SimpleImputer(missing_values=np.nan, strategy="most_frequent")
36
37 # simpan hasil imputasi ke dalam variable X_data
38 X_data = imp.fit_transform(x_data)
39
40 # mengubah kategori dari String ke numerik
41 le = preprocessing.LabelEncoder()
42 y_data = le.fit_transform(y_data)
43
44 # membagi data menjadi data training dan data testing
45 X_train, X_test, y_train, y_test = train_test_split(X_data, y_data, test_size=0.2, random_state=0)
46
47 # multinomial naive bayes
48 mnb = MultinomialNB()
49 mnb.fit(X_train, y_train)
50
51 # mendapatkan hasil prediksi dari data testing
52 y_pred = mnb.predict(X_test)
53
54 # Menampilkan presentasi error prediksi
55 error = ((y_test!=y_pred).sum())/len(y_pred))*100
56 print("Error prediksi = %.2f" %error, "%")
57
58 # menampilkan presentasi akurasi
59 akurasi = 100 - error
60 print("Akurasi = %.2f" %akurasi, "%")
```

- output:

```
In [3]: runfile('C:/Users/DIMAS ANGGORO SAKTI/Dropbox/My PC
(LAPTOP-FMDVD92V)/Downloads/C_2110511115_Dimas Anggoro sakti
naive bayes/folder tugas/Multinomial.py', wdir='C:/Users/DIMAS
ANGGORO SAKTI/Dropbox/My PC (LAPTOP-FMDVD92V)/Downloads/
C_2110511115_Dimas Anggoro sakti naive bayes/folder tugas')
Error prediksi = 32.26 %
Akurasi = 67.74 %

In [4]:
```

- Analisa: Disini program mengilustrasikan probabilitas dataset hepatitis dengan metode Multinomial Naïve Bayes