

# Perancangan Website Pemilihan Obat Menggunakan Algoritma Naive Bayes

Muhammad Risalah Naufal<sup>1</sup>, Reyhan Rainaldy<sup>2</sup>

<sup>1,2</sup>(Teknik Informatika/S1 Sistem Informasi, Universitas Negeri Surabaya)

<sup>1</sup>[muhammadrисalah.21008@mhs.unesa.ac.id](mailto:muhammadrисalah.21008@mhs.unesa.ac.id)

<sup>2</sup>[reyhan.21030@mhs.unesa.ac.id](mailto:reyhan.21030@mhs.unesa.ac.id)

**Abstrak**— Penelitian ini bertujuan untuk mengembangkan website klasifikasi obat menggunakan machine learning untuk memberikan rekomendasi obat berdasarkan data pasien. Data yang digunakan pada penelitian ini berasal dari dataset Kaggle, dan metode penelitian melibatkan pencarian data, preprocessing, analisis, dan implementasi berbagai algoritma klasifikasi. Naive Bayes terpilih sebagai model terbaik dengan akurasi tinggi. Website ini memungkinkan input data pasien, termasuk usia, jenis kelamin, tekanan darah, kolesterol, dan rasio sodium ke potasium, memberikan rekomendasi obat yang sesuai. Pengujian langsung pada website memastikan output sesuai dengan input. Diharapkan, website ini dapat memudahkan tenaga kesehatan dalam pengambilan keputusan pengobatan, meningkatkan efektivitas terapi, dan memberikan layanan kesehatan yang lebih personal. Implementasi website menggunakan Visual Studio Code juga disertakan, menambah kontribusi dalam penerapan solusi teknologi informasi untuk meningkatkan kualitas layanan kesehatan.

**Kata Kunci**— Klasifikasi Obat, Dataset, Naive Bayes, Website, Machine Learning.

## I. PENDAHULUAN

Pembuatan sebuah website klasifikasi obat yang mampu menentukan obat yang tepat berdasarkan beberapa faktor seperti usia, jenis kelamin, tekanan darah, tingkat kolesterol, dan rasio sodium ke potasium dalam darah merupakan sebuah aplikasi yang dapat memberikan manfaat penting di dunia kesehatan. Dengan menggunakan data-data ini, website ini dapat memberikan rekomendasi terhadap jenis obat yang paling sesuai untuk pasien, yang akan memudahkan dokter atau tenaga medis dalam membuat keputusan pengobatan yang lebih akurat.

Website ini akan memanfaatkan teknik-teknik machine learning dan pengolahan data untuk melakukan klasifikasi obat berdasarkan data pasien yang diinputkan. Dengan bantuan algoritma-algoritma yang sesuai, website ini akan memproses data pasien dan menghasilkan rekomendasi terkait obat-obat yang sesuai dengan kebutuhan pasien, berdasarkan karakteristik kesehatan mereka.

Keunggulan dari website ini adalah kemampuannya untuk memberikan rekomendasi obat yang lebih personal dan disesuaikan dengan data kesehatan individu pasien, yang akan membantu dalam meningkatkan efektivitas pengobatan dan mengurangi risiko penggunaan obat yang tidak sesuai.

Website ini juga bisa digunakan sebagai alat bantu untuk pendekatan awal dalam proses pengambilan keputusan medis, yang dapat memudahkan dokter dalam mengambil keputusan terkait dengan pengobatan pasien mereka. Dengan demikian, website ini akan membantu dalam mengoptimalkan

pengobatan pasien dan meningkatkan kualitas pelayanan kesehatan secara keseluruhan.

## II. METODE

Pada penelitian ini ada beberapa metode yang digunakan guna melaksanakan penelitian secara baik dan efisien:

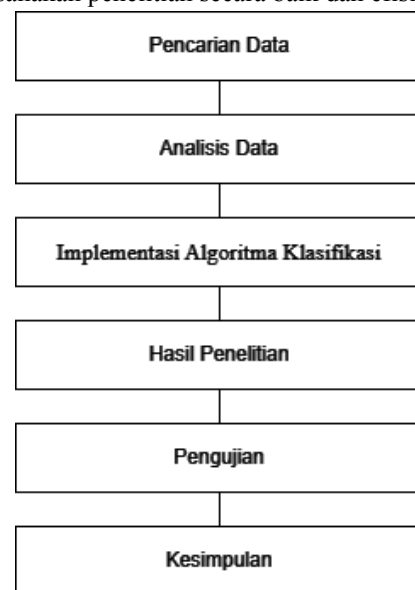


Diagram. 1 Aliran metode penelitian

### • Pencarian Data

Data yang digunakan pada penelitian ini merupakan data yang berasal dari *website Kaggle* dengan nama dataset Klasifikasi Obat atau *Drug Classification*. Data ini memiliki fitur dan target atau hasil didalamnya, adapun isi fitur pada data tersebut yaitu:

1. Age : Age of patient (Umur pasien)
2. Sex: Gender of patient (kelamin pasien)
3. BP: Blood pressure of patient (tekanan darah pasien)
4. Cholesterol: Cholesterol of patient (kadar kolesterol pasien)
5. Na\_to\_K: Sodium to Potassium Ratio in Blood (rasio sodium to potassion di darah)

Dan target atau hasil yang didapat yaitu berupa:

1. Drug: Drug Type (Tipe obat).

### • Analisis Data

Setelah mendapatkan data, langkah selanjutnya adalah melakukan analisis terhadap data dengan cara

menerapkan *pre-processing* pada data agar data sesuai ketika dilakukan proses *data mining*.

- Implementasi Algoritma Klasifikasi  
Implementasi Algoritma Klasifikasi dilakukan ketika dataset sudah melalui *pre-processing*. Algoritma yang digunakan pada saat proses data mining ini yakni algoritma Decision Tree, Naive Bayes, K-Nearest Neighbors, Support Vector Machine, Reinforcement Learning, dan Artificial Neural Network
- Hasil Penelitian  
Dataset yang telah melalui proses data mining dengan berbagai algoritma akhirnya akan dipilih salah satu yang menghasilkan akurasi paling baik, pada penelitian ini digunakan algoritma Naive Bayes dikarenakan algoritma ini menunjukkan hasil yang paling baik. Dataset dengan algoritma ini kemudian kita simpan untuk ekstensi pada website pemilihan obat nanti.
- Pengujian  
Pengujian yang dilakukan merupakan pengujian langsung pada website pemilihan obat. pengujian berjalan dengan cara menginputkan berbagai macam data yang ada dan melihat apakah output yang keluar sesuai dengan data yang ada atau tidak.
- Kesimpulan  
Dengan adanya website ini diharapkan kedepannya dapat mempermudah tenaga kesehatan dalam melakukan identifikasi terhadap obat yang diperlukan guna menanggulangi masalah yang ada pada pasien.

### III. HASIL DAN PEMBAHASAN

Hasil dan pembahasan dari penelitian ini adalah :

#### A. Analisa Data

Dataset ini merupakan data yang berisi prediksi kemungkinan terjadinya kolesterol. Analisis ini berdasarkan berbagai fitur, kemudian prediksi dari analisis tersebut akan memberitahu obat manakah yang cocok sesuai dengan yang dialami pasien. Berikut ini informasi kolom dari Klasifikasi obat yang sesuai dengan peneliti buat :

1. Age : Age of patient (Umur pasien)
2. Sex: Gender of patient (kelamin pasien)
3. BP: Blood pressure of patient (tekanan darah pasien)
4. Cholesterol: Cholesterol of patient (kadar kolesterol pasien)
5. Na\_to\_K: Sodium to Potassium Ratio in Blood (rasio sodium to potassium di darah)
6. Drug: Drug Type (Tipe obat)

Dari dataset tersebut selanjutnya akan dilakukan perbandingan dengan algoritma Decision Tree, Naive Bayes, K-Nearest Neighbors, Support Vector Machine, Reinforcement Learning, dan Artificial Neural Network untuk

mendapatkan model terbaik dengan menggunakan bantuan Google Colab.

Hasil dari perbandingan setiap algoritma dapat dilihat sebagai berikut :

#### 1. Decision Tree

```
Fitting 30 folds for each of 27 candidates, totalling 810 fits
Best Parameters (Accuracy): {'criterion': 'gini', 'max_depth': 3}
Best Score (Accuracy): 0.8716666666666665
Best Score (F1 Macro): 0.8224910813636713
```

Gbr. 1 Hasil evaluasi menggunakan Decision Tree

Hasil yang didapat dari algoritma DT adalah 0,871 dan 0,822

#### 2. Naive Bayes

```
Fitting 30 folds for each of 100 candidates, totalling 3000 fits
Best Parameters (Accuracy): {'var_smoothing': 0.0006579332246575676}
Best Score (Accuracy): 0.9666666666666667
Best Score (F1 Macro): 0.9455176442637742
```

Gbr. 2 Hasil evaluasi menggunakan Naive Bayes

Hasil yang didapat dari algoritma NB adalah 0,966 dan 0,945

#### 3. K-Nearest Neighbors

```
Best Parameters (Accuracy): {'n_neighbors': 14, 'weights': 'distance'}
Best Score (Accuracy): 0.7300000000000001
Best Score (F1 Macro): 0.5513459173459173
```

Gbr. 3 Hasil evaluasi menggunakan K-Nearest Neighbors

Hasil yang didapat dari algoritma KNN adalah 0,730 dan 0,551

#### 4. Support Vector Machine

```
Best Parameters (Accuracy): {'C': 0.7430147299885189}
Best Score (Accuracy): 79.83333333333334
Best Score (F1 Macro): 0.6426165802460055
```

Gbr. 4 Hasil evaluasi menggunakan Support Vector Machine

Hasil yang didapat dari algoritma SVM adalah 0,798 dan 0,642

#### 5. Reinforcement Learning

```
Best Parameters (Accuracy): {'C': 21118.24384789908}
Best Score (Accuracy): 96.66666666666666
Best Score (F1 Macro): 0.9435390846969796
```

Gbr. 5 Hasil evaluasi menggunakan Reinforcement Learning

Hasil yang didapat dari algoritma RL adalah 0,966 dan 0,943

#### 6. Artificial Neural Network

```
Best Parameters (Accuracy): {'alpha': 0.01, 'hidden_layer_sizes': (100, 50, 25)}
Best Score (Accuracy): 77.50000000000001
Best Score (F1 Macro): 0.6311707944618162
```

Gbr. 6 Hasil evaluasi menggunakan Artificial Neural Network

Hasil yang didapat dari algoritma ANN adalah 0,775 dan 0,631

Dari hasil pemrosesan data menggunakan algoritma Decision Tree, Naive Bayes, K-Nearest Neighbors, Support Vector Machine, Reinforcement Learning, dan Artificial Neural Network didapatkan model terbaik adalah model NB (naive bayes) sehingga model tersebut yang ditentukan untuk membangun website.

## B. Hasil

Setelah menemukan algoritma terbaik, data disimpan dan dimasukkan dalam rancang bangun website pada *Visual Studi Code*. Berikut adalah *source code* yang digunakan dalam pembuatan website pemilihan obat:

### 1. Library dan load data

```
1 import numpy as np
2 import pickle
3 import streamlit as st
4 from PIL import Image
5
6 # Load the saved model
7 loaded_model = pickle.load(open('ML_NB_DRUG3.pkl','rb'))
8
```

Gbr. 7 Code Library dan load data

Kode diatas merupakan bagian library serta bagian dimana data dari data mining sebelumnya kita masukan sebagai landasan data pada website pemilihan obat.

### 2. Form input data

```
11 def main():
12
13     #Give a title
14     img1 = Image.open("logo2.png")
15     img1 = img1.resize((500, 300))
16     st.image(img1)
17
18     # To get the input data from the user
19     age = st.number_input("Umur Anda",value=0)
20
21     sex_display = ('female', 'Male')
22     sex_option = list(range(len(sex_display)))
23     sex=st.selectbox("Gender Anda", sex_option, format_func=lambda x: sex_display[x])
24
25     bp_display = ('Low', 'Normal', 'High')
26     bp_option= list(range(len(bp_display)))
27     bp=st.selectbox("Tekanan Darah Anda", bp_option, format_func=lambda x: bp_display[x])
28
29     cholesterol_display = ('Normal', 'High',)
30     cholesterol_option= list(range(len(cholesterol_display)))
31     chol=st.selectbox("You selected:", cholesterol_option, format_func=lambda x: cholesterol_displ
32
33     na_to_k = st.number_input("Masukan Rasio Sodium to Potassium dalam Darah",value=0)
34
35     #Code for prediction
36     drug = ''
37
```

Gbr. 8 Code Form input data

Pada bagian input form data digunakan kode seperti gambar diatas untuk melakukan input data, dimana pada bagian input data diberi berbagai macam objek yang sesuai dengan fitur pada dataset yang ada serta memisahkan target dari fitur.

### 3. Output target obat

```
35 #Code for prediction
36 drug = ''
37
38 # Create a button for Prediction
39
40 if st.button("Drug Test Result"):
41     Features = [[age, sex, bp, chol, na_to_k]]
42     prediction = loaded_model.predict(Features)
43     lc = [str(i) for i in prediction]
44     ans = int("".join(lc))
45     if ans == 0:
46         st.success(
47             "DrugV"
48         )
49     elif ans == 1:
50         st.success(
51             "DrugA"
52         )
53     elif ans == 2:
54         st.success(
55             "DrugA"
56         )
57     elif ans == 3:
58         st.success(
59             "DrugB"
60         )
61     else:
62         st.success(
63             "DrugC"
64         )
65
66 main()
```

Gbr. 9 Code Output target obat

Untuk bagian output target obat digunakan metode *if else* dimana bila inputan data tidak sesuai dengan hasil yang pertama maka hasil yang muncul adalah hasil yang kedua ataupun hasil yang lain.

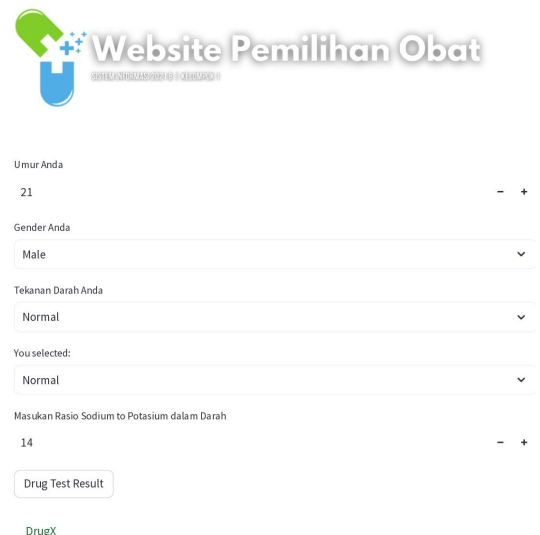
### 4. Running code

```
PS C:\TUGAS KULIAH/SIST 5/DATA MINING/UTS> streamlit run uts.py
You can now view your Streamlit app in your browser.
Local URL: http://localhost:8501
Network URL: http://10.60.1.21:8501
```

Gbr. 10 Code Output target obat

Ketika ingin melakukan running terhadap program, dapat diketikkan pada terminal yang ada berupa “streamlit run uts.py” maka secara otomatis website akan terbuka dan klasifikasi dapat dilakukan.

### 5. Tampilan website serta input dan output data



Gbr. 11 Tampilan website dan output hasil

Tampilan pada website dibuat sesuai dengan data yang dimiliki agar dapat memunculkan output target yang sesuai dengan data yang ada.

#### IV. KESIMPULAN

Pada website klasifikasi obat terdapat beberapa fitur yaitu usia, jenis kelamin, blood pressure, kolesterol, dan rasio sodium to potassium di darah yang dapat membantu dalam menentukan jenis obat yang cocok dari fitur yang telah diberikan.

Pengolahan data pada penelitian ini menghasilkan algoritma naive bayes sebagai algoritma yang paling sesuai dengan dataset yang ada, sehingga menjadikan algoritma naive bayes sebagai algoritma yang digunakan untuk mengelola sekaligus mengklasifikasi untuk menghasilkan data yang bagus untuk website pemilihan obat.

Website ini juga akan memudahkan tenaga kesehatan yang tidak perlu menghabiskan banyak waktu untuk menentukan obat secara manual dan tidak menghabiskan banyak kertas untuk menulisnya, karena dapat dibantu dengan website yang telah kami buat.

#### DAFTAR PUSTAKA

- [1] Akbar, N. T. (2022). Klasifikasi Obat. Retrieved from Kaggle: <https://www.kaggle.com/code/galerinyanopal/klasifikasi-obat>
- [2] Zhang, H., Ding, L., Zou, Y., Hu, S. Q., Huang, H. G., Kong, W. B., & Zhang, J. (2016). Predicting drug-induced liver injury in human with Naïve Bayes classifier approach. *Journal of computer-aided molecular design*, 30, 889-898.
- [3] Rajalakshmi, R., & Aravindan, C. (2011, April). Naive Bayes approach for website classification. In *International Conference on Advances in Information Technology and Mobile Communication* (pp. 323-326). Berlin, Heidelberg: Springer Berlin Heidelberg.
- [4] Shmueli, G., Bruce, P. C., Gedeck, P., & Patel, N. R. (2019). *Data mining for business analytics: concepts, techniques and applications in Python*. John Wiley & Sons.
- [5] Wahyuni, S. (2018, March). Implementation of Data Mining to Analyze Drug Cases Using C4.5 Decision Tree. In *Journal of Physics: Conference Series* (Vol. 970, No. 1, p. 012030). IOP Publishing.
- [6] Hananto, A. L., Assiroj, P., Priyatna, B., Fauzi, A., Rahman, A. Y., & Hilabi, S. S. (2021, June). Analysis of drug data mining with clustering technique using K-Means algorithm. In *Journal of Physics: Conference Series* (Vol. 1908, No. 1, p. 012024). IOP Publishing.