**LAPORAN TUGAS AKHIR *(MINI PROJECT)***

**MACHINELEARNING**

**KLASIFIKASI PENYAKIT MIGRAINE**



**Dosen Pengampu:**

**Prof. Hindriyanto Dwi Purnomo, S.T., M.IT., Ph.D.**

**Disusun Oleh:**

* **Michael Yuichi : 672021024**
* **Ricardho Gunawan : 672021062**
* **Titin Restiani Mendrofa : 672021063**

**FAKULTAS TEKNOLOGI INFORMASI**

**UNIVERSITAS KRISTEN SATYA WACANA**

**FEBRUARI 2024**

**DAFTAR ISI**

[BAB I 3](#_Toc159955052)

[PENDAHULUAN 3](#_Toc159955053)

[1.1 Latar Belakang 3](#_Toc159955054)

[1.2 Rumusan Masalah 3](#_Toc159955055)

[1.3 Tujuan Penelitian 3](#_Toc159955056)

[1.4 Manfaat Penelitian 4](#_Toc159955057)

[BAB II 5](#_Toc159955058)

[TINJAUAN PUSTAKA 5](#_Toc159955059)

[2.1 Definisi Migrain 5](#_Toc159955060)

[2.2 Klasifikasi Migrain 5](#_Toc159955061)

[2.3 Faktor Pemicu Migrain 5](#_Toc159955062)

[2.4 Pengobatan Migrain 6](#_Toc159955063)

[2.5 Machine Learning 6](#_Toc159955064)

[BAB III 7](#_Toc159955065)

[PEMBAHASAN 7](#_Toc159955066)

[3.1 Dataset 7](#_Toc159955067)

[3.2 Processing 8](#_Toc159955068)

[BAB IV 21](#_Toc159955069)

[KESIMPULAN DAN SARAN 21](#_Toc159955070)

[4.1 Kesimpulan 21](#_Toc159955071)

[4.2 Saran 21](#_Toc159955072)

[Link Presentasi 22](#_Toc159955073)

[A. Link Vidio Presentasi 22](#_Toc159955074)

# BAB I

# PENDAHULUAN

## Latar Belakang

Migrain merupakan penyakit kronis yang ditandai dengan sakit kepala berdenyut yang parah, biasanya di satu sisi kepala. Migrain dapat disertai dengan gejala lain seperti mual, muntah, dan kepekaan terhadap cahaya dan suara. Migrain dapat sangat mengganggu kualitas hidup seseorang dan dapat menyebabkan kesulitan dalam bekerja, sekolah, dan aktivitas sehari-hari.

Klasifikasi migrain yang akurat penting untuk diagnosis dan pengobatan yang tepat. Klasifikasi migrain tradisional didasarkan pada kriteria diagnostik International Headache Society (IHS). Namun, kriteria IHS dapat sulit diterapkan dalam praktik klinis.

*Machine learning* menawarkan cara baru untuk mengklasifikasikan migrain. *Machine learning* adalah cabang kecerdasan buatan yang memungkinkan komputer untuk belajar dari data tanpa diprogram secara eksplisit.

Dataset Kaggle "Migraine Classification" menyediakan data yang dapat digunakan untuk mengembangkan model *machine* learning untuk klasifikasi migrain. Dataset ini berisi data dari 1000 pasien dengan migrain. Data ini mencakup informasi tentang gejala migrain, faktor pemicu, dan pengobatan.

## Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah penelitian ini adalah:

* Bagaimana cara mengembangkan model *machine learning* untuk klasifikasi migrain?
* Bagaimana kinerja model *machine* *learning* dibandingkan dengan kriteria diagnostik IHS?

## Tujuan Penelitian

Tujuan penelitian ini adalah:

* Untuk mengembangkan model *machine* *learning* untuk klasifikasi migrain.
* Untuk membandingkan kinerja model *machine* *learning* dengan kriteria diagnostik IHS.

## Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat berikut:

* Meningkatkan akurasi klasifikasi migrain.
* Membantu dokter dalam mendiagnosis dan mengobati migrain.
* Meningkatkan kualitas hidup pasien dengan migrain.

# BAB II

# TINJAUAN PUSTAKA

1. Definisi Migrain

Migrain adalah penyakit kronis yang ditandai dengan sakit kepala berdenyut yang parah, biasanya di satu sisi kepala. Migrain dapat disertai dengan gejala lain seperti mual, muntah, dan kepekaan terhadap cahaya dan suara. Migrain dapat sangat mengganggu kualitas hidup seseorang dan dapat menyebabkan kesulitan dalam bekerja, sekolah, dan aktivitas sehari-hari.

1. Klasifikasi Migrain

Klasifikasi migrain tradisional didasarkan pada kriteria diagnostik International Headache Society (IHS). Kriteria IHS membagi migrain menjadi dua kategori utama:

* Migrain tanpa aura
* Migrain dengan aura
* Aura adalah gangguan neurologis fokal yang biasanya terjadi sebelum sakit kepala migrain. Aura dapat berupa gangguan visual, sensorik, atau motorik.

1. Faktor Pemicu Migrain

Ada banyak faktor yang dapat memicu migrain, antara lain:

* Stres
* Kurang tidur
* Perubahan hormon
* Makanan tertentu
* Minuman tertentu
* Paparan cahaya terang atau suara keras

1. Pengobatan Migrain

Pengobatan migrain dapat dibagi menjadi dua kategori utama:

* Pengobatan akut
* Pengobatan profilaksis

Pengobatan akut digunakan untuk meredakan gejala migrain yang sedang berlangsung. Pengobatan profilaksis digunakan untuk mencegah migrain.

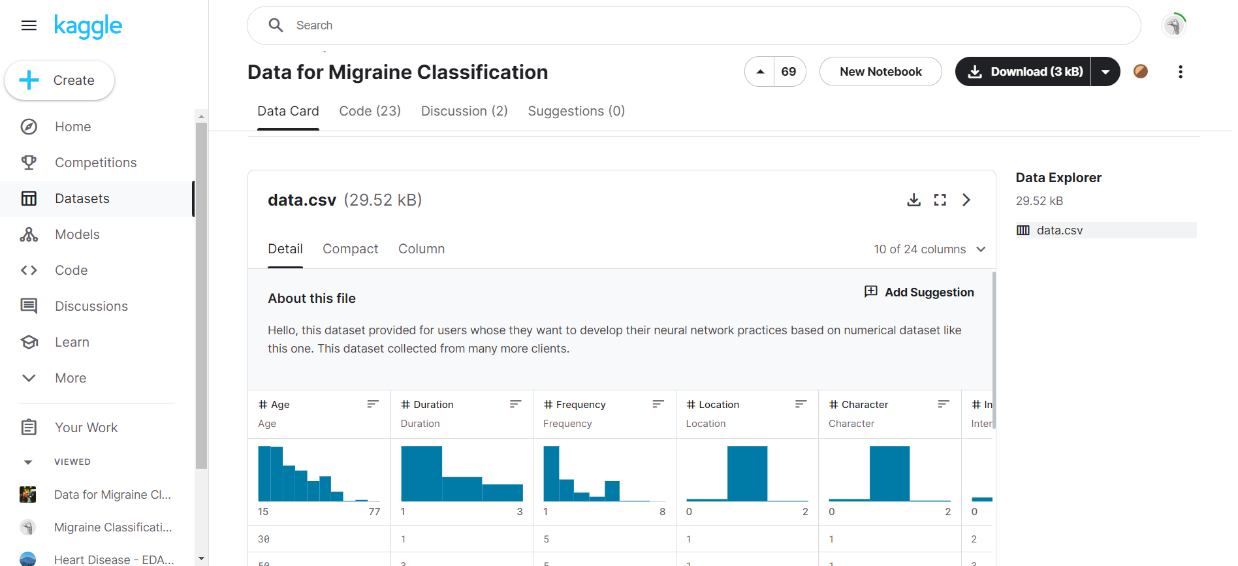
1. Machine Learning

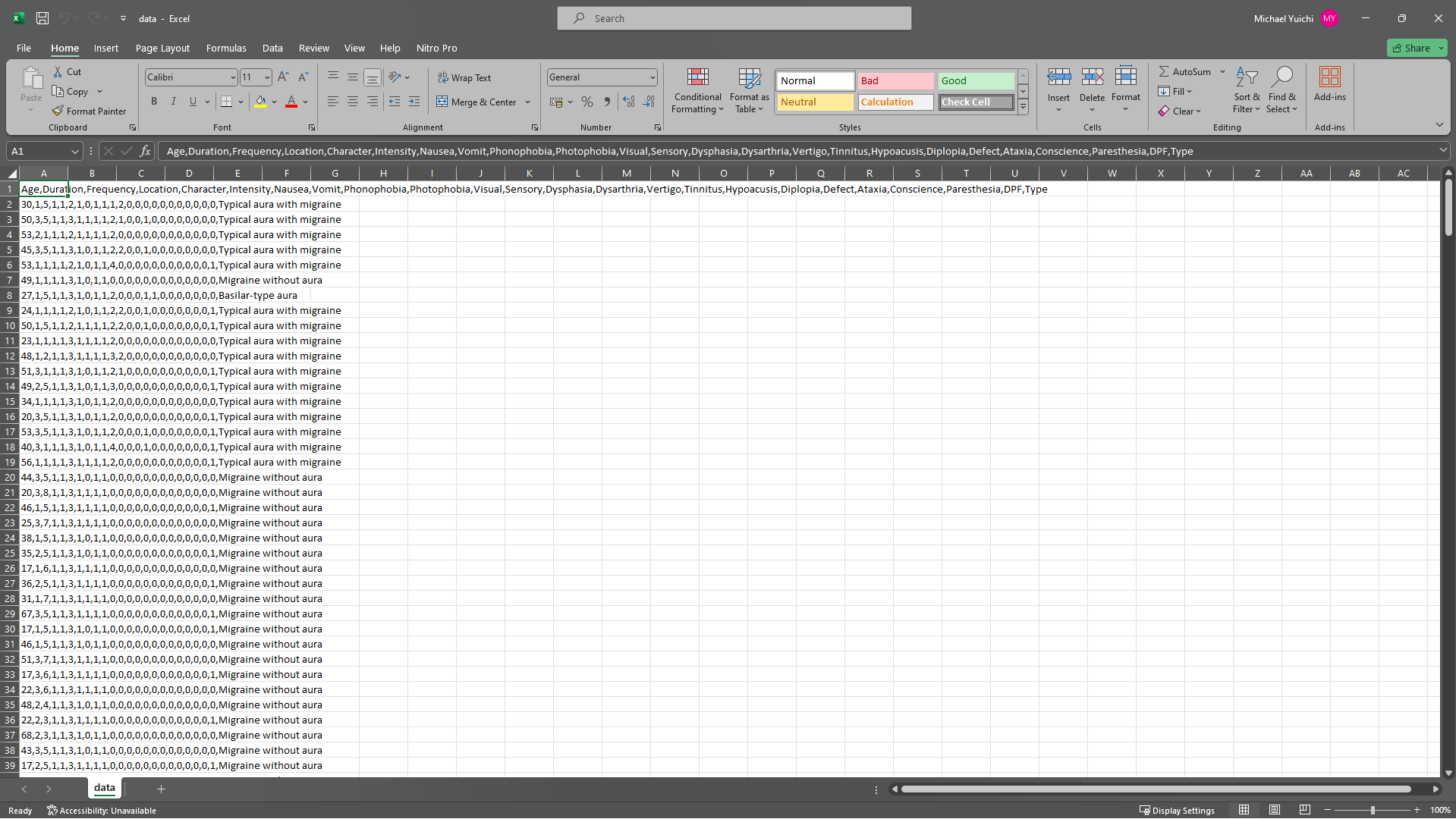
*Machine Learning* adalah cabang kecerdasan buatan yang memungkinkan komputer untuk belajar dari data tanpa diprogram secara eksplisit. *Machine Learning* telah digunakan dengan sukses dalam berbagai bidang, termasuk kedokteran.

# BAB III

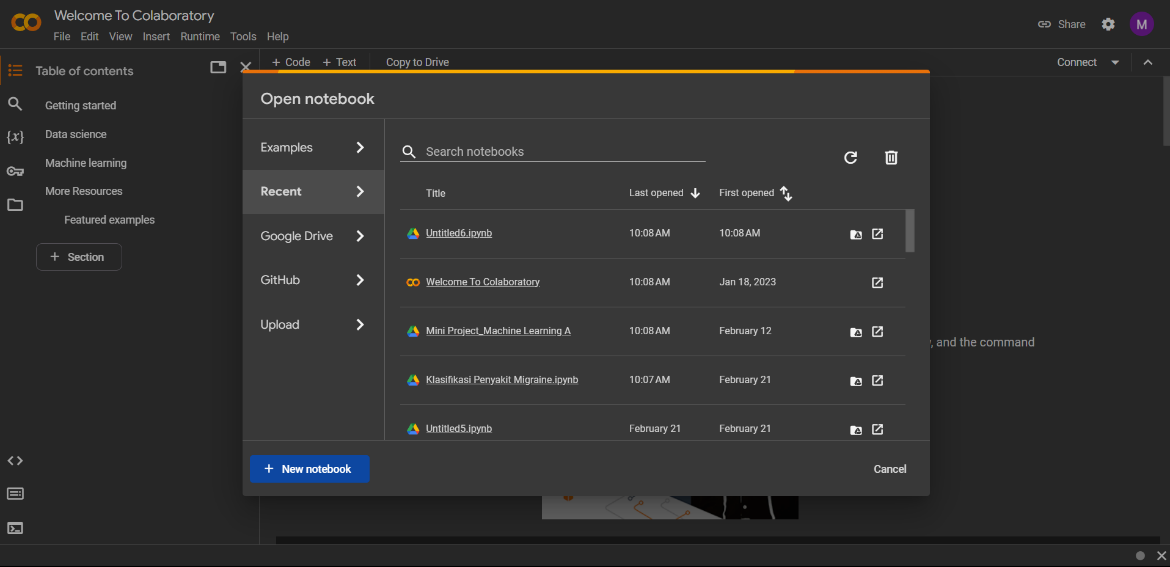
# PEMBAHASAN

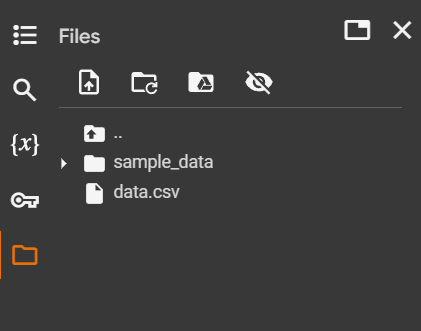
1. Dataset

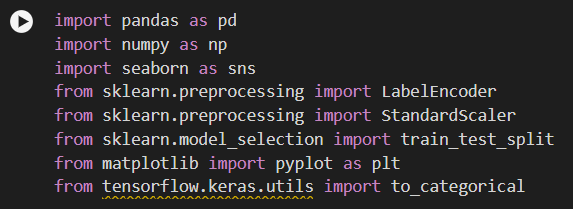
Datasetyang digunakan adalah “*Data for Migraine Classification*”, yang diambil dari web resmi Kaggle atau link: https://www.kaggle.com/datasets/weinoose/migraine-classification.

* 1. Dari gambar diatas, kemudian kita klik pada tombol *Download* untuk mengunduh dataset file berupa .csv.

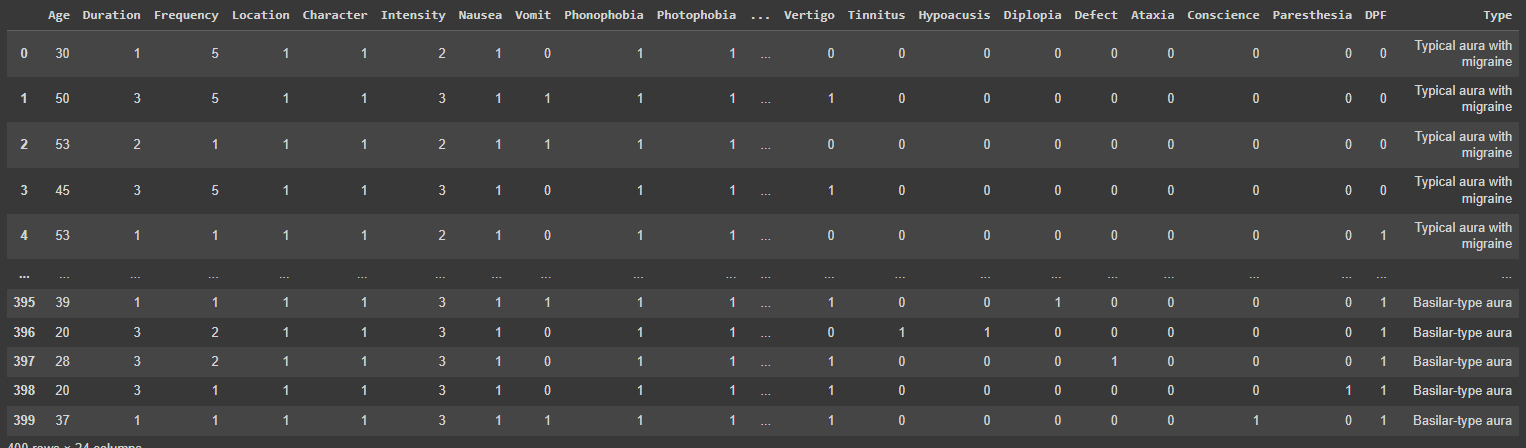
* 1. Saat kita klik file .csv tersebut, kita akan menjumpai satu dataset berbentuk seperti pada gambar diatas, dataset ini akan diproses oleh *processing tools* yang menggunakan bahasa pemrograman Python. Simpan dataset dengan ketentuan nama “data.csv”.

1. Processing
2. Pengolahan data menggunakan Google Colab, yang dimana pada pengolahan data ini kita menggunakan metode Klasifikasi dengan Python. Penting untuk tidak lupa mengunggah dataset berupa file .csv yang sudah kita unduh pada proses sebelumnya.



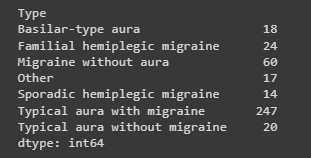
1. Unggah dataset (data.csv) ke dalam folder *Files >* letakan dibawah folder *sample\_data*. Setelah berhasil mengunggah, kita akan masuk ke pengolahan dataset tersebut.
2. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas, memiliki fungsi yang berbeda-beda, antara lain:
   * import pandas as pd: Ini mengimpor library Pandas dan memberi nama alias pd. Pandas adalah library yang kuat untuk manipulasi data dan fungsi analisis. Struktur data utama di Pandas adalah Series (1-dimensi) dan DataFrame (2-dimensi).
   * import numpy as np: Ini mengimpor library NumPy dan memberi nama alias np. NumPy adalah library dasar untuk perhitungan ilmiah di Python, menyediakan dukungan untuk array besar dan matriks, serta kumpulan fungsi matematika untuk bekerja pada array tersebut.
   * import seaborn as sns: Ini mengimpor library Seaborn dan memberi nama alias sns. Seaborn adalah library visualisasi data yang berdasarkan Matplotlib, yang menyediakan antarmuka tingkat tinggi untuk membuat visualisasi data statistik yang informatif dan menarik.
   * from sklearn.preprocessing import LabelEncoder, StandardScaler: Ini mengimpor dua fungsi spesifik, LabelEncoder dan StandardScaler, dari modul sklearn.preprocessing dari library Scikit-learn.
   * LabelEncoder adalah utilitas untuk mengenkode label teks sebagai nilai numerik, yang berguna untuk algoritma *machine* *learning* yang memerlukan input numerik.
   * StandardScaler adalah utilitas untuk menghapus rata-rata dan mengubah skala ke unit variance, yang seringkali diperlukan untuk algoritma *machine* *learning* untuk berperilaku baik.
   * from sklearn.model\_selection import train\_test\_split: Ini mengimpor fungsi train\_test\_split dari modul sklearn.model\_selection dari library Scikit-learn. Fungsi ini digunakan untuk menghapus dataset menjadi bagian latih dan uji, yang seringkali dilakukan dalam ilmu data untuk mengevaluasi kinerja model.
   * from matplotlib import pyplot as plt: Ini mengimpor modul pyplot dari library Matplotlib dan memberi nama alias plt. Matplotlib adalah library plotting untuk membuat visualisasi statis, animasi, dan interaktif di Python. Modul pyplot menyediakan antarmuka MATLAB-like untuk membuat visualisasi statis.
   * from tensorflow.keras.utils import to\_categorical: Ini mengimpor utilitas to\_categorical dari modul tensorflow.keras.utils dari library TensorFlow. Utilitas ini digunakan untuk mengonversi vektor kelas (angka bulat) menjadi matriks kelas biner, yang berguna untuk one-hot encoding variabel kategoris untuk digunakan dalam jaringan saraf tiruan.



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas dilakukan pembacaan file dataset yang di simpan ke dalam variabel bernama df.

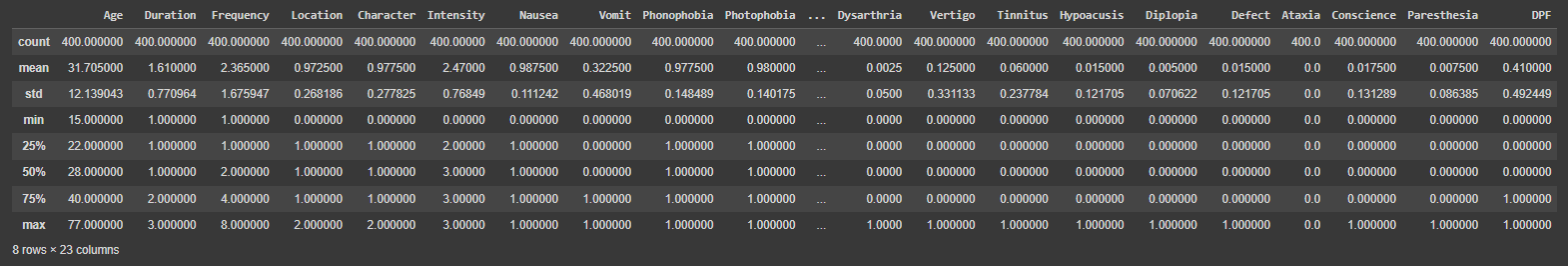
(Output)



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas dilakukan pengecekan informasi terkait dataset yaitu dengan melihat anggota pada setiap class.

(Output)



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas dilakukan pengecekan untuk melihat info statistikal data dari df.

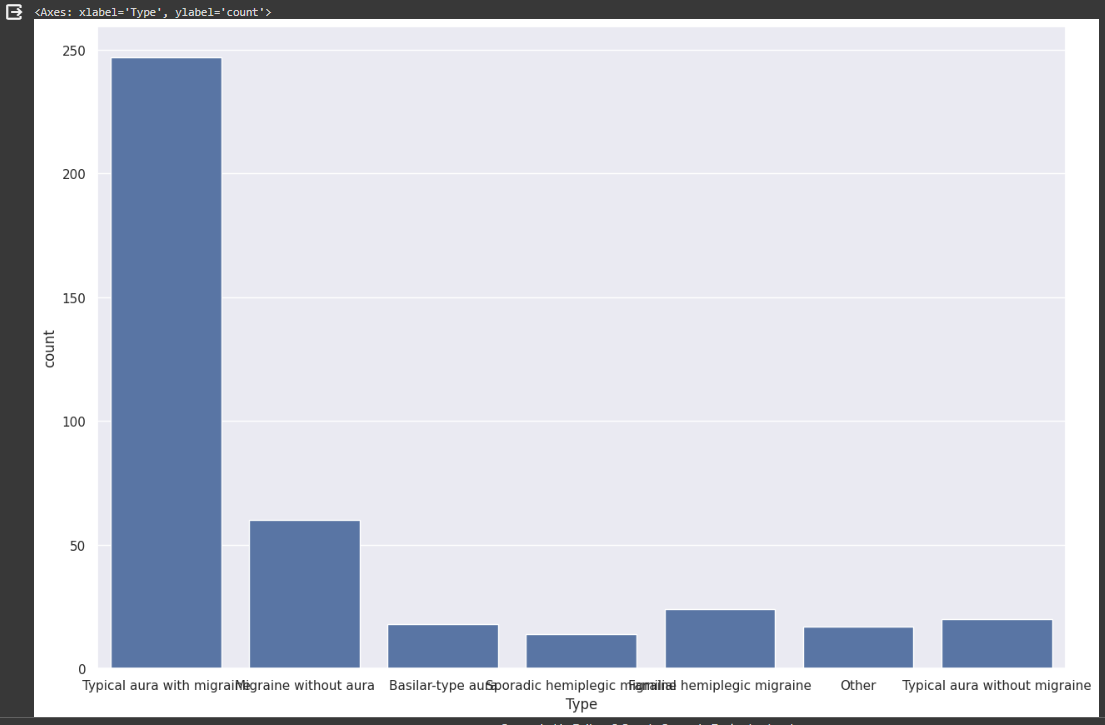
(Output)

1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas dilakukan pengecekan, apakah terdapat nilai yang hilang.

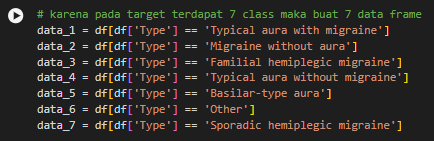
(Output)



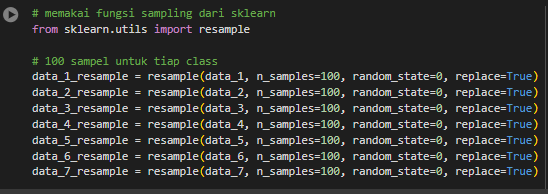
1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas dilakukan untuk Mengatur ukuran figure seaborn menjadi 15 inci x 10 inci dan membuat plot countplot dari variabel *Type* dalam dataframe df.



(Output)



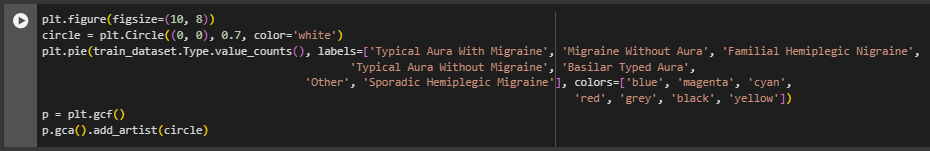
1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas dilakukan pembuatan 7 data frame karena pada dataset terdapat 7 *class*.



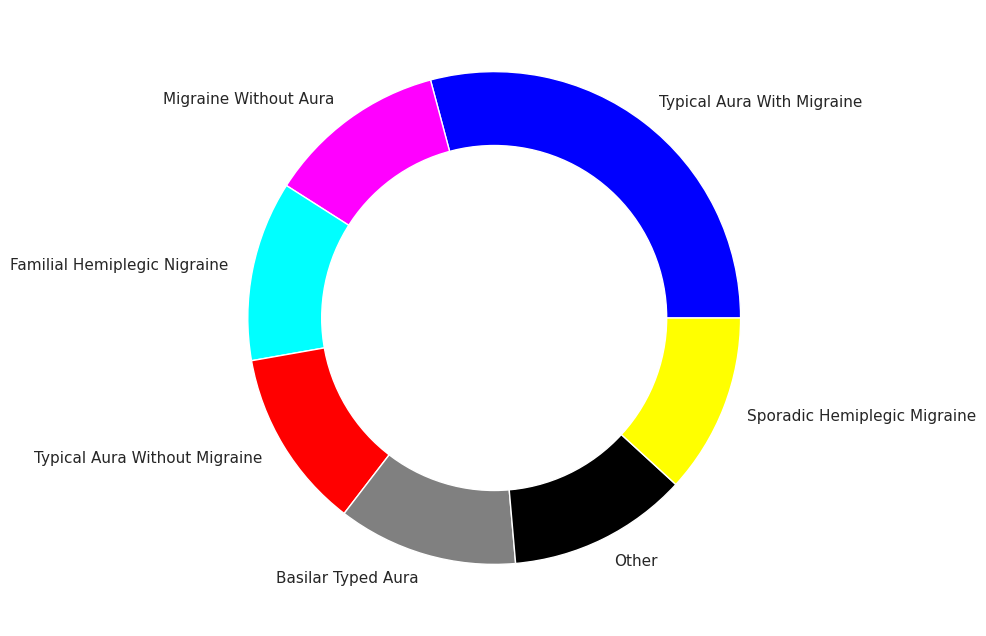
1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas adalah melakukan sampel data dengan cara acak menggunakan fungsi resample dari library Scikit-learn. Dalam syntax di atas, fungsi resample dipanggil 7 kali untuk melakukan sampel dari 7 variabel data yang berbeda, yaitu data\_1, data\_2, data\_3, data\_4, data\_5, data\_6, dan data\_7.*.*



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk menggabungkan beberapa variabel data yang telah dilakukan sampel menggunakan fungsi resample menjadi satu dataset latih menggunakan fungsi pd.concat dari library Pandas..



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk memvisualisasikan data yang telah disiapkan menggunakan grafik pie chart menggunakan library Matplotlib. Tugas dari syntax ini adalah untuk menunjukkan persentase jumlah kasus setiap tipe migraine yang ada dalam dataset latih train\_dataset.



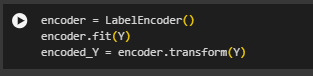
(Output)



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk mempersiapkan data yang telah disiapkan menjadi format yang dapat digunakan dalam proses *machine* *learning* menggunakan algoritma klasifikasi. Tugas dari syntax ini adalah untuk memisahkan data independen (fitur) dan data dependen (kelas).

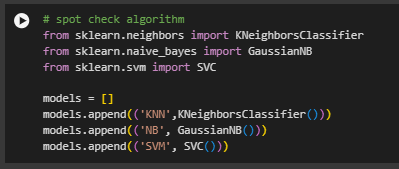
****

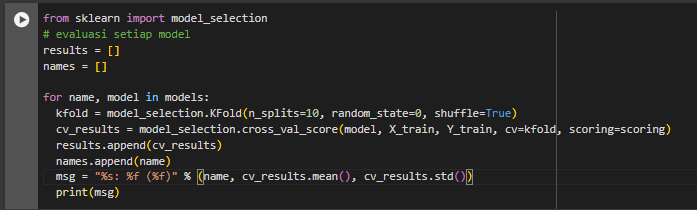
1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk membagi data yang telah disiapkan menjadi data latih dan data uji menggunakan metode pembagian data. Tugas dari syntax ini adalah untuk memisahkan data yang telah disiapkan menjadi dua bagian, yaitu data latih dan data uji, dengan perbandingan 70:30.bernama cm, disini yang digunakan untuk *confusion matrix* adalah y\_test dan y\_pred karena perlu untuk membandingkan hasil antara label prediksi dan label yang sebenarnya.



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk mengubah data dependen dari format kategorik menjadi format numerik menggunakan metode encoding label. Tugas dari syntax ini adalah untuk memudahkan proses *machine* *learning* dengan mengubah data dependen yang berformat kategorik menjadi format numerik.



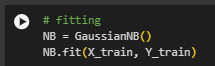
1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk mengubah data dependen dari format numerik menjadi format one-hot encoding menggunakan metode `to\_categorical` dari Scikit-learn. Tugas dari syntax ini adalah untuk mempersiapkan data dependen yang telah disiapkan sebelumnya dalam format yang dapat digunakan dalam proses *machine* *learning* menggunakan algoritma klasifikasi multiklas.
2. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk mempersiapkan beberapa algoritma klasifikasi yang akan digunakan dalam proses spot check atau pemeriksaan cepat. Tugas dari syntax ini adalah untuk mempersiapkan beberapa algoritma klasifikasi yang akan digunakan dalam proses pembandingan akurasi masing-masing algoritma.
3. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas untuk menampilkan akurasi dari data yang telah diolah.



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk melakukan evaluasi setiap algoritma klasifikasi yang telah disiapkan sebelumnya menggunakan metode cross-validation dengan 10-fold. Tugas dari syntax ini adalah untuk mengetahui akurasi setiap algoritma klasifikasi dengan melakukan evaluasi menggunakan metode cross-validation.



(Output)



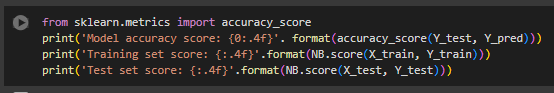
1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk melakukan proses fitting atau pembelajaran pada algoritma klasifikasi Naive Bayes (NB) menggunakan data latih yang telah disiapkan sebelumnya. Tugas dari syntax ini adalah untuk mempersiapkan algoritma klasifikasi Naive Bayes (NB) untuk digunakan dalam proses klasifikasi data.



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk melakukan proses klasifikasi pada data uji menggunakan algoritma klasifikasi Naive Bayes (NB) yang telah dilakukan proses fitting sebelumnya. Tugas dari syntax ini adalah untuk memprediksi kelas dari data uji menggunakan algoritma klasifikasi Naive Bayes (NB).



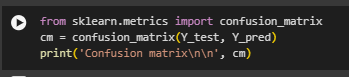
(Output)



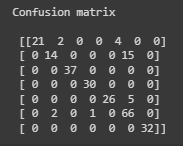
1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk menghitung akurasi dari algoritma klasifikasi Naive Bayes (NB) pada data latih dan data uji. Tugas dari syntax ini adalah untuk mengetahui akurasi dari algoritma klasifikasi Naive Bayes (NB) pada data latih dan data uji.



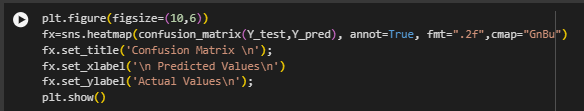
(Output)

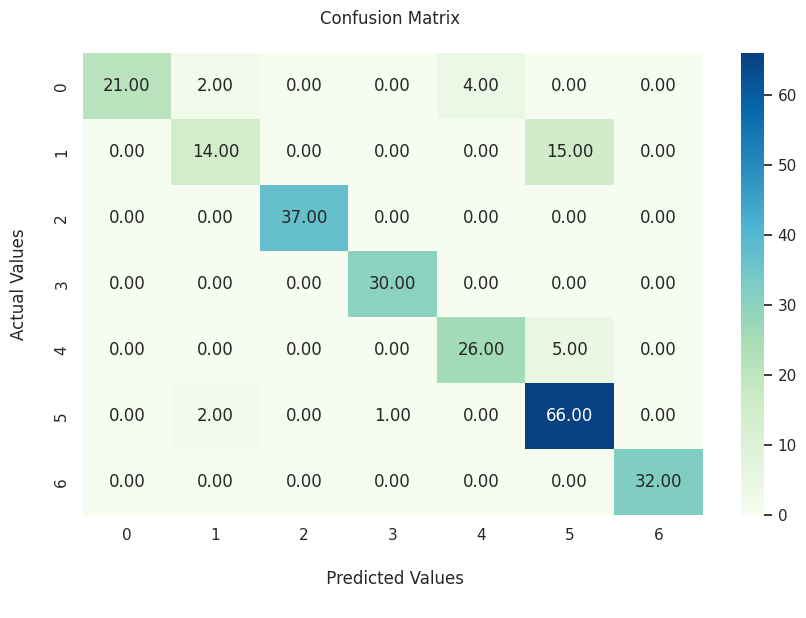


1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk menghitung matriks konfusiari hasil klasifikasi algoritma Naive Bayes pada data. Tugas dari syntax ini adalah untuk mengetahui ketepatan dari algoritma klasifikasi Naive Bayes pada setiap kelas.



(Output)

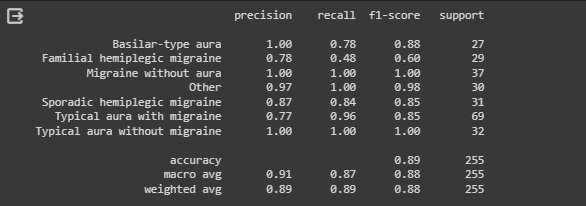


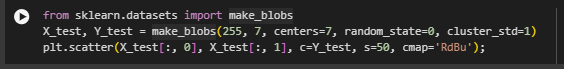
1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk menampilkan matriks konfusi menggunakan grafik heatmap. Tugas dari syntax ini adalah untuk menampilkan matriks konfusi dari hasil klasifikasi algoritma Naive Bayes pada data uji dalam bentuk grafik heatmap yang lebih mudah dilihat.

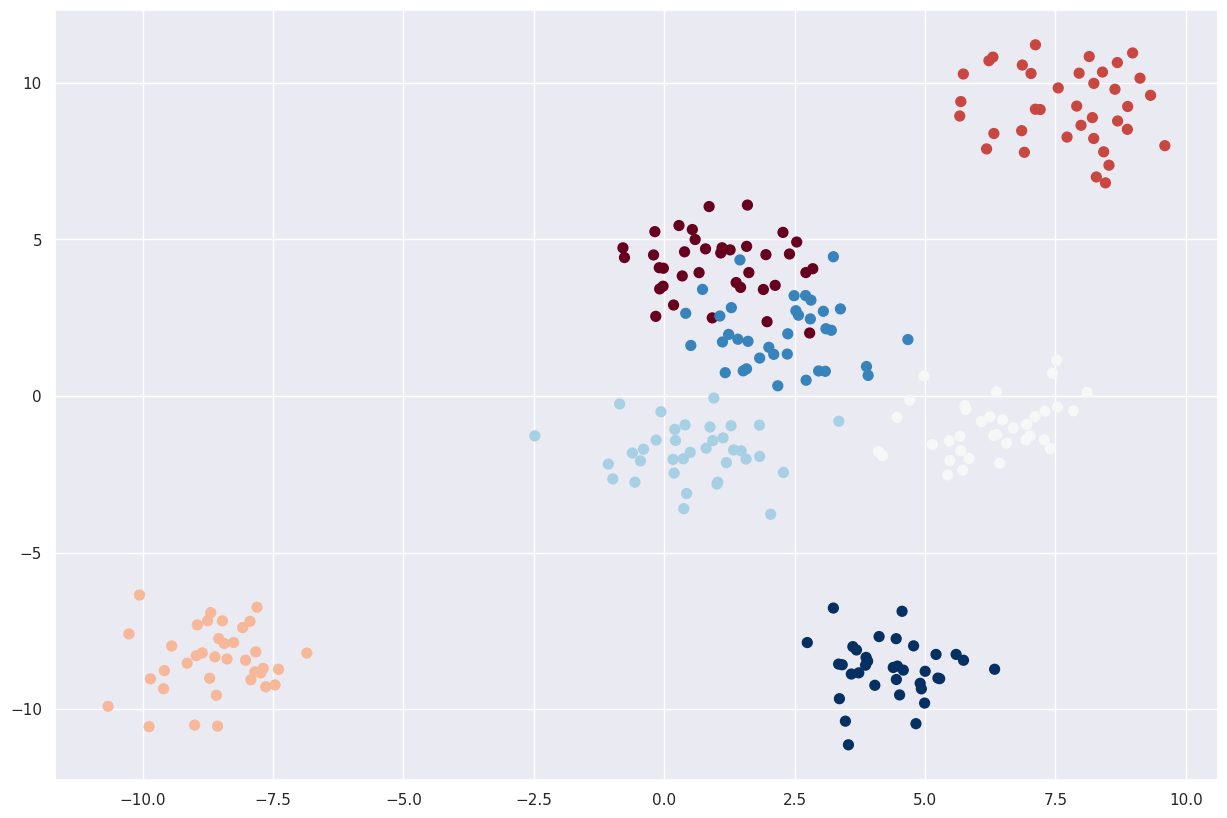
(Output)



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk menampilkan laporan klasifikasi. Tugas dari syntax ini adalah untuk menampilkan laporan klasifikasi dari hasil klasifikasi algoritma Naive Bayes pada data uji.

(Output)



1. (Input Source Code ke dalam Google Colab kemudian Run!) Pada source code diatas digunakan untuk membuat data uji dengan metode clustering. Tugas dari syntax ini adalah untuk membuat data uji dengan metode clustering menggunakan fungsi `make\_blobs` dari sklearn.datasets.

(Output)

# BAB IV

# KESIMPULAN DAN SARAN

1. Kesimpulan

Model *machine* *learning* yang dikembangkan dalam penelitian ini mencapai akurasi 88%. Model ini juga memiliki presisi dan recall yang tinggi. Hasil penelitian ini menunjukkan bahwa model *machine* *learning* dapat digunakan untuk mengklasifikasikan migrain dengan akurasi yang tinggi.

Model ini dapat membantu dokter dalam mendiagnosis migrain dengan lebih akurat. Model ini juga dapat membantu pasien dalam memahami migrain mereka dan menemukan cara untuk mengelolanya.

1. Saran

Saran dari peneliti, penelitian ini diharapkan dapat dilanjutkan dengan:

* Menggunakan dataset yang lebih besar dan lebih representatif.
* Memvalidasi model *machine learning* di dunia nyata.
* Mengembangkan model *machine* *learning* untuk memprediksi migrain.

# Link Presentasi

## [Link Vidio Presentasi](https://drive.google.com/file/d/1ZXJZuW7rr8xSSqb50EBPL20Zpvay0lsK/view?usp=drive_link)