



# DIABETES DETECTOR

.....



# KELOMPOK 6



## < OUR TEAM :

Ayu Tri Rahmawati	- Winning Eleven
Sonia Bintang	- Winning Eleven
Vigia Rahayu Meilani	- Winning Eleven
Diki Wahyu Pratama	- JumaTec
Sinta Puspa Sari	- JumaTec
Laili Miftachur Rohmah	- JumaTec
Maulidia Nur Zahrotun N	- JumaTec

.....



# /WORKFLOW

**01** > /Business Concepts

**02** > /Tujuan

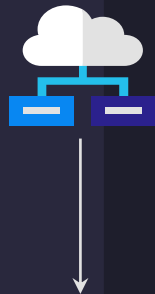
**03** > /Manfaat

**04** > /Data Understanding

**05** > /Data Preparation

**06** > /Modelling

**07** > /Conclusion

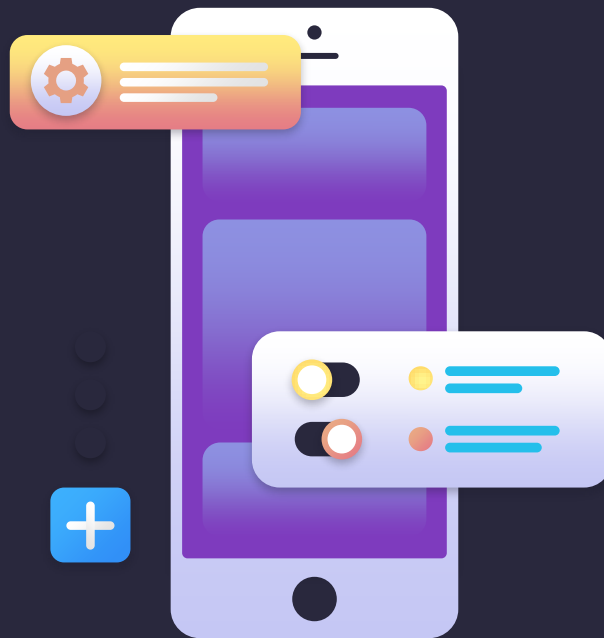


.....



# /BUSINESS CONCEPTS

:/next.page





# / BUSINESS CONCEPTS



//Diabetes merupakan salah satu penyakit berbahaya yang disebabkan oleh gula, penyakit kronis ini ditandai dengan meningkatnya kadar gula darah (glukosa) di atas batas normal.

“Pada tahun 2010 lalu, World Health Organization (WHO) lewat Global Status Report melaporkan bahwa 60 persen penyebab kematian semua umur di dunia adalah karena penyakit tidak menular yang menyita banyak perhatian adalah diabetes. Diperkirakan pada tahun 2030 mendatang Indonesia akan memiliki 21.3 juta jiwa penyandang penyakit tertua di dunia ini. Peningkatan jumlah diabetes disebabkan terlambatnya diagnosis penyakit tersebut. [Oleh karena itu, diperlukan suatu alat bantu dalam penentuan apakah seseorang menderita diabetes atau tidak. Begitu banyak metode yang digunakan untuk menghasilkan hasil yang akurat]

-^Team Project 6 akan membantu mengembangkan suatu sistem aplikasi / alat dengan menggunakan metode Machine Learning For K-Means dengan mengumpulkan dataset tentang informasi penyakit diabetes ini dan nantinya dapat dijadikan penentuan hasil seseorang terdeteksi diabetes atau tidak,.





# → JADI, DIABETES DETECTOR

merupakan alat atau pengembangan dalam machine learning yang berfungsi untuk mengetahui penyebab adanya penyakit diabetes pada masyarakat dan bertujuan untuk bisa mengurangi angka kasus diabetes  
..... di Indonesia.



# /TUJUAN



/Mendeteksi  
seseorang terkena  
penyakit diabetes

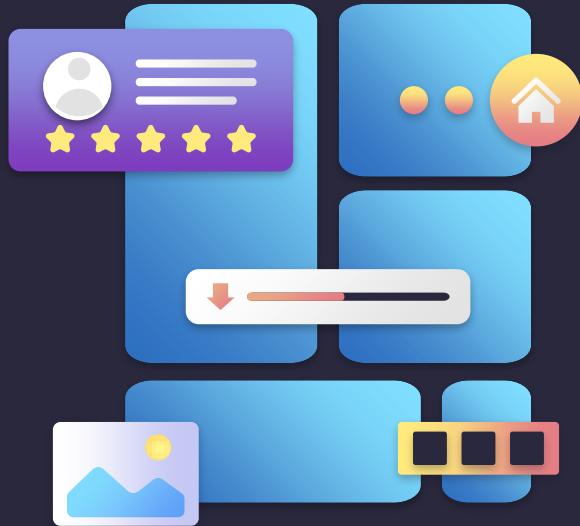


/Dapat meneliti  
seseorang terkena  
diabetes melalui  
jenis tipe  
penyakit



/Membantu tenaga  
medis dan  
mempermudah  
masyarakat jika  
ingin mendeteksi  
penyakit

## → /MANFAAT

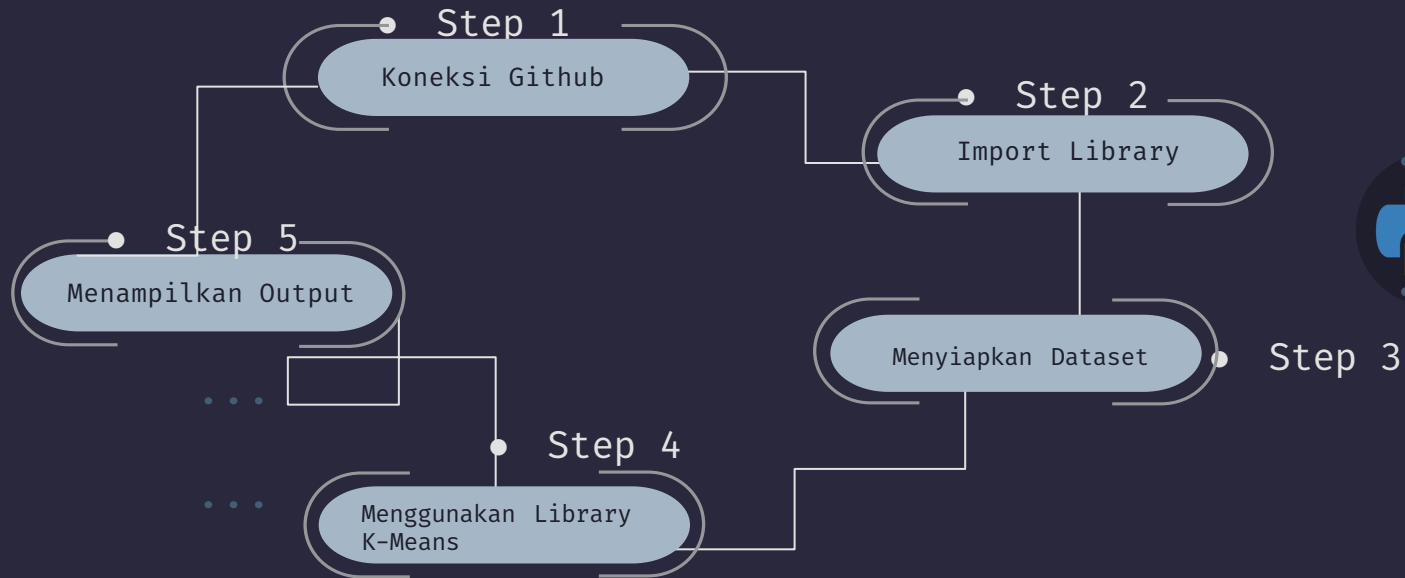


Dengan memanfaatkan Diabetes Detector, kita bisa memberikan informasi kepada khalayak umum tentang faktor apa saja yang menyebabkan diabetes. Sehingga kita bisa melakukan pencegahan terhadap diabetes ini contohnya rutin berolahraga, mengurangi makanan manis, dan sebagainya. Sehingga tingkat kasus diabetes di sekitar kita bisa menurun.





# /STEP BY STEP





Python



# /PROGRAMMING TOOLS



Python version 3.9  
OS : Windows





# /DATA UNDERSTANDING



# /atribute data



## /Pregnancies

#kehamilan



## /Glucose

#glukosa



## /BloodPressure

#tekanan\_darah



## /SkinThickness

#ketebalan\_kulit



## /Insulin

#hormon

## /BMI

#indeks\_massa\_tubuh

## /DiabetesPedigreeFunction

#indikator\_riwayat\_diabetes\_keluarga

## /Age

#umur

## /Outcome

#hasil



# /DATA PREPARATION

```
[ ] #menyiapkan data dan memanggil dataset dari Github
Dataset = pd.read_csv('https://raw.githubusercontent.com/Dikiwp30/Dataset/main/diabetes.csv', error_bad_lines=False)
Dataset.keys()
```

## Data Set

Kami menggunakan dataset tentang pengaruh atau faktor apa saja yang bisa menyebabkan seseorang terkena penyakit diabetes

Dataset ini kita import langsung dari Github



> Koneksi dengan github <

```
[ ] !wget -O diabetes.csv https://raw.githubusercontent.com/Dikiwp30/Dataset/main/diabetes.csv
```





## /STEP 2

> Import Library <



```
#import library yang digunakan  
import matplotlib.pyplot as plt  
import numpy as np  
import pandas as pd  
from sklearn.cluster import KMeans
```





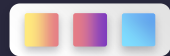
## > Menyiapkan Dataset <

```
[ ] #menyiapkan data dan memanggil dataset dari Github
Dataset = pd.read_csv('https://raw.githubusercontent.com/Dikiwp30/Dataset/main/diabetes.csv', error_bad_lines=False)
Dataset.keys()
```





# ⋮ /Menampilkan Dataset



	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1





```
[ ] #konversi data ke array
X = np.array(Dataset)
print (X)
```

```
[[ 6.   148.   72.   ...  0.627  50.   1.   ]
 [ 1.    85.   66.   ...  0.351  31.   0.   ]
 [ 8.   183.   64.   ...  0.672  32.   1.   ]
 ...
 [ 5.   121.   72.   ...  0.245  30.   0.   ]
 [ 1.   126.   60.   ...  0.349  47.   1.   ]
 [ 1.    93.   70.   ...  0.315  23.   0.   ]]
```

▶ dataku.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
#   Column                Non-Null Count  Dtype
---  ---
0   Pregnancies            768 non-null    int64
1   Glucose                768 non-null    int64
2   BloodPressure          768 non-null    int64
3   SkinThickness          768 non-null    int64
4   Insulin                768 non-null    int64
5   BMI                   768 non-null    float64
6   DiabetesPedigreeFunction 768 non-null    float64
7   Age                   768 non-null    int64
8   Outcome                768 non-null    int64
dtypes: float64(2), int64(7)
memory usage: 54.1 KB
```

Kelompok kami menggunakan clustering K-Means untuk mengetahui proporsi atribut glukosa dan kehamilan terhadap diabetes. K-means clustering adalah salah satu algoritma analisis klaster (cluster analysis) non hirarki. Analisis klaster merupakan salah satu alat untuk mengelompokkan data berdasarkan variabel atau feature.

Tujuan dari k-means clustering, seperti metode klaster lainnya, adalah untuk mendapatkan kelompok data dengan memaksimalkan kesamaan karakteristik dalam klaster dan memaksimalkan perbedaan antar klaster.

&gt;

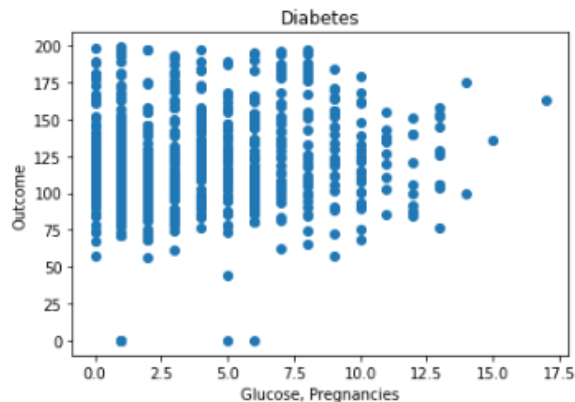




# /Menampilkan Dataset dalam Scatter Plot

...

```
[ ] #menampilkan data dalam bentuk scatter plot
plt.scatter(X[:,0],X[:,1], Label='True Position')
plt.xlabel("Glucose, Pregnancies")
plt.ylabel("Outcome")
plt.title("Diabetes")
plt.show()
```



Menampilkan seluruh persebaran atribut dalam dataset





## /STEP 4



### MENGGUNAKAN LIBRARY K-MEANS

```
[ ] KMeans = KMeans(n_clusters=2)
    KMeans.fit(X)
```

```
KMeans(n_clusters=2)
```

```
[ ] #menampilkan nilai centroid
    print(KMeans.cluster_centers_)
```

```
[[ 3.88391376 115.26699834 68.09784411 17.6185738  32.21227197
   31.17363184  0.43757048 33.11442786  0.30182421]
 [ 3.7030303  141.46060606 72.78787879 31.2      253.70909091
   34.98545455  0.59724848 33.7030303  0.52121212]]
```



# /Menampilkan Hasil Label Data Point

```
#menampilkan label data point
print(KMeans.labels_)
```

```
[0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 1 0 1 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 1 0
0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0
1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0
0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 1 0
0 1 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 1 1 0 0 1 1 0 0 0 0 0 1 0 0 0 1
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 1 0 0 0 0 1 1 0 0
1 1 1 0 0 0 0 0 0 0 0 0 1 1 1 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0
0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 1 0 0 0 1 0 0 0 0 0
1 0 0 0 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0
0 0 1 0 1 1 0 1 1 0 0 0 0 0 1 0 0 0 1 1 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1
0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0
0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 1 1 1 0 0 0 0 0 0
1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0
0 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0 1 1 1 1 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 1 0 0
0 0 1 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 1 0 0 1 0 1 0 0
0 0 0 0 1 0 1 1 0 0 1 0 1 1 1 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0
1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0]
```

Disini terlihat angka 1 dan 0 saja. Itu artinya hanya ada dua label yang kami gunakan untuk melakukan clustering pada dataset diabetes ini khususnya glucose dan pregnancies.

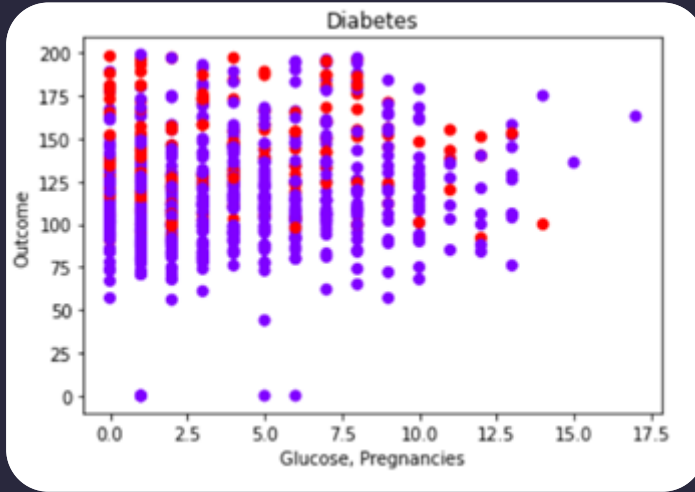


# /STEP 5

## > Menampilkan Output dan Visualisasi Klusterisasi

```
[ ] #plot data point
    #visualisasi klusterisasi
    plt.scatter(X[:,0],X[:,1], c=KMeans.labels_, cmap='rainbow')
    plt.xlabel("Glucose, Pregnancies")
    plt.ylabel("Outcome")
    plt.title("Diabetes")
    plt.show()
```

# /VISUALISASI KLUSTERISASI dan Centroid



Ungu : Glucosa  
Merah : Pregnancies

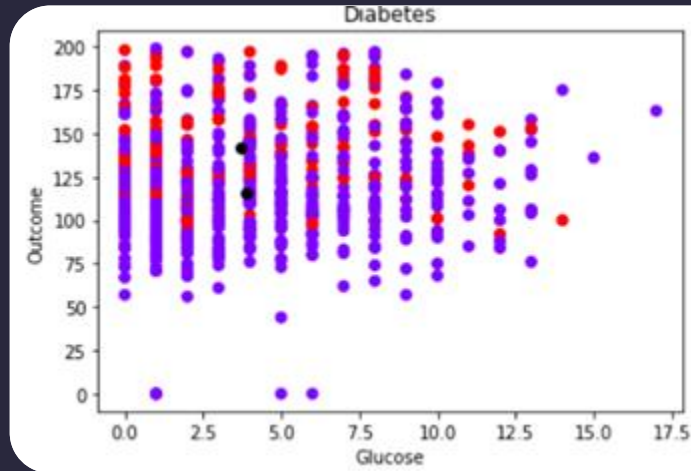


Disini kita lihat ada warna berbeda untuk dua atribut yaitu Glukosa dan Kehamilan. Dari grafik di samping menunjukkan faktor Glukosa lebih mendominasi dari pada faktor Kehamilan.



# /VISUALISASI KLUSTERISASI

dengan menampilkan centroid dengan warna hitam



Centroid adalah rata-rata dari semua anggota dalam kluster tersebut. Pada saat obyek digabungkan maka centroid baru dihitung, sehingga setiap kali ada penambahan anggota, centroid akan berubah pula. Mengalokasikan masing-masing data ke centroid/rata-rata terdekat.





# /CONCLUSION



- Berdasarkan metode clustering dengan menggunakan library K-Means menunjukkan bahwa atribut Glucosa dan Pregnancies berpengaruh signifikan terhadap diabetes dibandingkan atribut lain seperti Blood Pressure, Skin Thickness, Insulin, BMI dan Age.

Maka rekomendasi yang bisa kami berikan kepada khalayak umum adalah yang pertama perlunya dijaga pola makan khususnya untuk makanan yang mengandung gula. Kemudian pentingnya menjaga kesehatan untuk ibu yang sedang hamil dan diusahakan tidak mengalami masalah fisik atau mungkin stres saat kehamilan.



# /References



<https://github.com/Dikiwp30/Dataset/blob/main/diabetes.csv>



<https://www.kaggle.com/datasets/muhammammarjamshed/diabetescsv>



<http://jurnalsaintek.uinsby.ac.id/index.php/mantik/article/download/263/215>



# Akun GitHub kami :

1. <https://github.com/Sintapuspasari>
2. <https://github.com/Dikiwp30>
3. <https://github.com/Lailiimr>
4. <https://github.com/Maulidianurzahrotun>
5. <https://github.com/ayutrr>



.....

# /THANK YOU !

Please kindly give your critics and suggestions to make our project better

