



SISTEM PREDIKSI DINI DIABETES

Tiear Rafa Komara
S1 Data Sains
1305210092

Widha Dwiyanti
S1 Data Sains
1305213019

Wildan Aufa Rafid
S1 Data Sains
1305213022

DR. Warih Maharani S. T., M. T.
Dosen Pengampu MataKuliah

Latar Belakang

Pola makan dan gaya hidup manusia cenderung beralih ke arah yang **tidak sehat**, ditandai dengan konsumsi makanan cepat saji yang **tinggi lemak, gula, dan garam**.



Tujuan Proyek

Bertujuan untuk mendeteksi apakah seseorang berpotensi terkena penyakit diabetes.



Dataset

"Diabetes for Beginners" Data diperoleh dari kaggle, berisi 9 kolom dan terdiri dari 769 baris data.

Pregnancies	Glucose	Blood Pressure	Skin Thickness	Insulin	BMI	Diabetes Pedigree Function	Age	Outcome
6	148	72	35	0	33.6	0.6	50	1
1	85	66	29	0	26.6	0.351	31	0
3	78	50	32	88	31	0.248	26	1



Method

• Random Forest

Kombinasi masing-masing tree dari decision tree yang kemudian digabungkan menjadi satu model.

• K-Means

Algoritma yang bertujuan untuk membagi sebuah set data ke dalam kelompok (clusters).

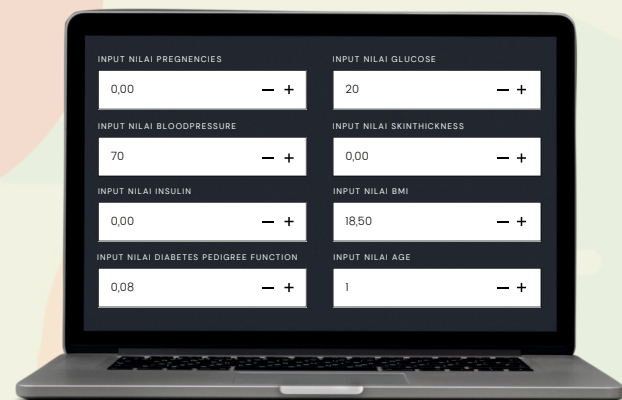
• Visualisasi

Representasi data atau informasi dalam bentuk visual, seperti grafik, diagram, peta, dan grafik interaktif.



Hasil

User menginput sendiri nilai - nilai yang dibutuhkan untuk setiap fitur dan memiliki batasan batasan yang relevan.



Kemudian user bisa memilih salah satu dari 3 fitur utama yaitu **prediksi** menggunakan Random Forest, **pengelompokan (clustering)** menggunakan K-Means, dan **visualisasi**.



Kesimpulan dan Evaluasi

Sistem ini dapat memprediksi apakah seseorang tersebut dapat dikatakan diabetes atau tidak berdasarkan 8 faktor. Sistem ini memiliki akurasi sebesar 81,870%.

Referensi

<https://www.kaggle.com/code/melikedilekci/diabetes-dataset-for-beginners>
<https://www.sciencedirect.com/science/article/pii/S1877050922021858>

