

LAPORAN HASIL KLASIFIKASI PENYAKIT DIABETES

Nama: Firda Muflif Fauzi
NIM: 231011402305
Prodi: Teknik Informatika FT UNPAM

1. Deskripsi Dataset

Dataset yang digunakan merupakan data medis pasien untuk mendeteksi kemungkinan terkena diabetes. Terdapat beberapa fitur seperti jumlah kehamilan, kadar glukosa, tekanan darah, BMI, dan usia. Target variabel (Outcome) bernilai 0 untuk tidak diabetes dan 1 untuk diabetes.

2. Metodologi

Langkah-langkah yang dilakukan meliputi preprocessing data (normalisasi dan pembagian data), kemudian dilakukan pelatihan model menggunakan dua algoritma klasifikasi yaitu Logistic Regression dan K-Nearest Neighbors (KNN). Evaluasi dilakukan menggunakan metrik akurasi dan confusion matrix.

3. Hasil dan Evaluasi

| Model | Akurasi |
|---------------------|---------|
| Logistic Regression | 0.53 |
| KNN | 0.50 |

Hasil pengujian menunjukkan bahwa Logistic Regression memiliki akurasi 0.53, sedangkan KNN memiliki akurasi 0.50.

4. Kesimpulan

Model Logistic Regression menunjukkan performa yang lebih baik dan stabil dibandingkan KNN pada dataset ini. Oleh karena itu, model ini direkomendasikan untuk deteksi awal kemungkinan penyakit diabetes berdasarkan data medis pasien.

5. Referensi

Dataset: Pima Indians Diabetes Dataset (Kaggle / UCI Machine Learning Repository).