

Nama : Mohammad Zainal Fuad

NIM : A11.2021.13332

Tugas Akhir Data Mining

Url github

: <https://github.com/MZainalFuadz/UjianAkhirSemesterDM.git>

Sumber dataset

: <https://www.kaggle.com/datasets/uciml/iris>

### **1. Deskripsi eksperimen**

Tema dari eksperimen yang saya buat adalah mengklasifikasi bunga iris, yang didalamnya terdapat macam spesies bunga iris yang dapat dibedakan melalui mengklasifikasi data set menggunakan metode machine learning. Cara mengklasifikasinya dengan menggunakan entitasnya lalu direprosesing dengan menggunakan metode Naïve bayes, Decision Thre, Random Fores, KNN, LR, SVC lalu dilihat akurasi. Berdasarkan keenam metode tersebut dibandingkan mana yang memiliki kinerja terbaik berdasarkan akurasi .

### **2. Dataset yang digunakan**

Dataset yang digunakan pada penelitian ini berisi data lebar dan panjang dari sepal dan petal bunga serta spesies dari bunga tersebut yang disimpan dalam file csv dengan nama Iris.csv.

### **3. Langkah yang digunakan**

- Mengimpor library yang diperlukan: numpy untuk komputasi numerik, matplotlib.pyplot untuk plotting, dan pandas untuk manipulasi data.
- Membaca dataset dari file CSV dengan nama 'ds\_salaries.csv' dan menyimpannya .
- Mengecek dan menghilangkan data set yang memiliki nilai null atau kosong serta data yang duplikat
- Memvisualisasikan data agar terlihat lebih jelas

- Menentukan nilai x (fitur) dan y (target atau label)
- Mengubah label spesies menjadi integer menggunakan label encoder
- Melakukan scaling fitur untuk enormalisasikan nilai pada fitur
- Split data train dan testing dengan perbandingan 4 : 1
- Load model Decision Tree, Random Forest, KNN, Naïve Bayes, Logistic Regression, SVC dan menghasilkan nilai akurasi setiap model
- Evaluasi menggunakan confusion matrix.

#### **4. Hasil Eksperiment**

Hasil eksperimen yang telah saya buat sebagai berikut :

Decision Tree dengan akurasi 1.0

Random Forest dengan akurasi 0,966667

KNN dengan akurasi 0,966667

Naïve Bayes dengan akurasi 0,966667

Logistic Regression dengan akurasi 0,966667

SVC dengan akurasi 0,966667

#### **5. Kesimpulan**

Dari ke enam model yang digunakan pada penelitian ini menghasilkan nilai akurasi yang sudah sangat baik, dengan nilai sempurna pada model decision tree dan pada model lainnya menghasilkan nilai dengan selisih yang sangat kecil yaitu 0,97