

Laporan Tugas k-NN Sistem Cerdas

Nama: M. Faris Riyan

Kelas: IT-41-03

NIM: 1303174079

Studi Kasus

Diberikan sebuah data latih Train_Tugas2.csv dengan data tes Test_Tugas2.csv. Menggunakan dua data tersebut carilah kelas dari data tes menggunakan metode k-NN dan buatlah *output* data dari hasil tersebut yang dibuat menjadi file csv dengan nama Predict_Tugas2.csv.

Analisis

Data Train_Tugas2.csv adalah sebuah data latih yang digunakan sebagai parameter untuk menentukan kelas pada data tes Test_Tugas2.csv. Tiap data tes akan dihitung dengan Tiap data Train menggunakan suatu rumus untuk menentukan kelas pada data tes. Hasil kelas dari perhitungan tersebut akan disimpan kedalam file Predict_Tugas2.csv.

Strategi Penyelesaian Masalah

Pertama hitung distance dari data train dengan data tes menggunakan Euclidean distance yaitu

$$\begin{aligned}d(\mathbf{p}, \mathbf{q}) &= d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2} \\&= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}.\end{aligned}$$

Dimana n adalah banyak data, q sebagai data pada data Train dan p sebagai data pada data Test. Lalu dicari nilai k menggunakan akar kuadrat dari banyak data pada train. Setelah itu urutkan data train berdasarkan distance dari yang terkecil hingga terbesar. Ambil data train sebanyak k . Tentukan kelas data tes berdasarkan frekuensi kelas terbanyak yang ada di data train sebanyak k . Simpan hasil kedalam file baru.

Fungsi Program

Program terdiri dari beberapa fungsi, yaitu:

1. `hitung_distance(Train, Test)` dengan masukan berupa data Train_Tugas2.csv dan data Test_Tugas2.csv. Fungsi ini menghitung distance, dan mengoutputkan hasil tersebut .
2. `hitung_K(Train)` untuk mencari nilai dari k dimana rumusnya adalah akar kuadrat dari banyak data train.
3. `hitung_class(Data)` dengan masukan berupa data yang telah dicari distancenya yang diurutkan dari yang terkecil. Fungsi akan memeriksa satu persatu class dan menghitung frekuensi kelas. Lalu akan mengembalikan class yang paling sering muncul.
4. `To2array(Data)` merupakan fungsi untuk merubah array satu dimensi menjadi array dua dimensi
5. `CombineArray(Array1, Array2)` merupakan fungsi yang menggabungkan array1 dengan array2