**Nama : Alfian Yulianto**

**NIM : 1301178160**

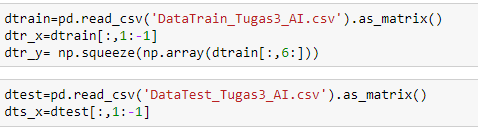
**Kelas : IF-40-GAB05**

**K-NEAREST NEIGHBOR(KNN)**

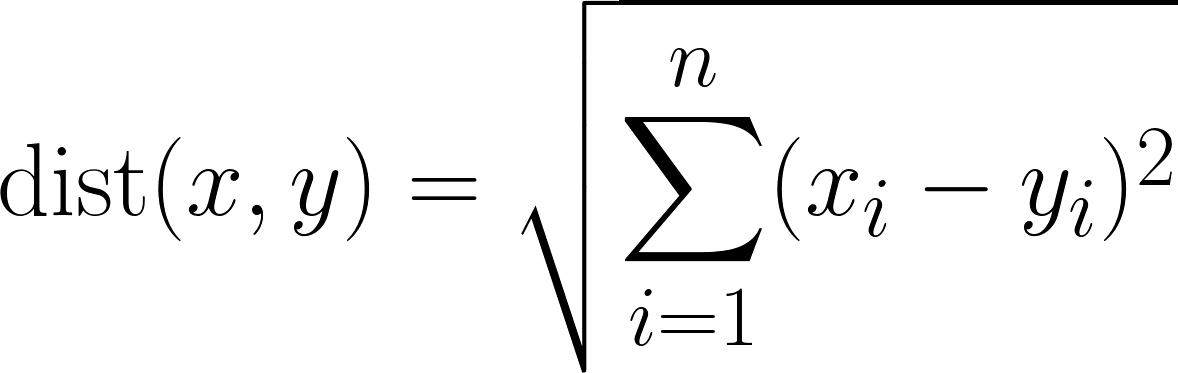
K-Nearest Neighbor (KNN) adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari *query instance* yang baru diklasifikan berdasarkan mayoritas dari kategori pada KNN. Tujuan dari algoritma ini adalah mengklasifikasikan obyek baru bedasarkan atribut dan training sample.

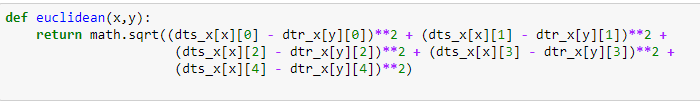
Pada tugas 3 ini diberikan 200 data test, 800 data training yang memiliki 5 atribut yaitu (X1,X2,X3,X4 dan X5) serta memiliki 4 kelas yaitu (0,1,2,3) yang merupakn keluaran/*output* dari hasil pencocokan dari 200 data test dg data train yg ada sehingga diperolehlah nilai kelas-kelas tersebut, berikut penjelasan langkah-langkahnya :

1. Load data training dan data test yang telah diberikan, disini data training dan data test diload dalam bentuk matrix.



1. Melakukan penghtingungan jarak tetangga terdekat dengan rumus *Euclidean.*



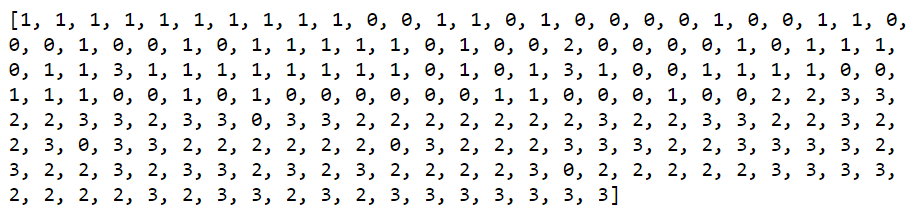


1. Perhitungan *Euclidean distance* akan dilakukan dalam perulangan dan disimpan dalam array yang bernama best\_k[]
2. Melakukan proses *sorting* pada array distance[] dari yang terkecil atau *ascending* kemudian disimpan sebagai nilai Y nya.
3. Menentukan nilai K, untuk mendapatkan nilai K yang bagus maka digunakanlah validasi data yaitu disini menggunakan *K-folds cross validation* yang dimana merupakan salah satu pendekatan untuk mengevaluasi model classifier dengan cara membagi data yang sudah memiliki kelas ke dalam K bagian. K yang sering digunakan adalah 10, jadi K=10.



1. Hasil dari voting dengan K=10 disimpan dalam array bernama best\_k.

Berikut hasil Running terbaik yang diperoleh :



Hasil array terbaik diatas meninterpretasikan klasifikasi dari 200 data test terhadap 800 data teraining yang ada.