

Nama : Firman Ramdhani

NIM : 1301164052

Kelas : IF 40-11

LAPORAN TUPRO 3 AI

ALGORITMA *K-NEAREST NEIGHBOR* (KNN)

1. Definisi Algoritma KNN

Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran diproyeksikan k ruang dimensi banyak, dimana masing-masing dimensi memoresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi data pembelajaran. Sebuah titik pada ruangan ini ditandai dengan kelas c , jika kelas c merupakan klasifikasi yang paling banyak ditemui pada k bauh tetangga terdekat titik tersebut

Algoritma K-NN secara umum.

- a. Menentukan jumlah tetangga paling dekat (K)
- b. Menghitung kuatdat jarak Euclid dari objek terhadap data sample yang ada
- c. Mengurutkan objek kedalam kelompok yang memiliki jarak terkecil
- d. Kumpulkan kategori dari klasifikasi nearest neighbor
- e. Dengan kategori nearest neighbor yang paling banyak, maka dapat diprediksikan nilai query instance yang telah dihitung

2. Source Code

a. Function dalam program

```
import math
import sys
import csv
import collections

#menghitung jarak antara dua data
def jarak(ax,ay):
    a = map(lambda x,y: (x - y)**2, ax, ay)
    return math.sqrt(sum(a))

#membaca data train dan data test
def loadData():
    dataHasil1=[]
    dataHasil2=[]
    with open('DataTrain_Tugas3_AI.csv') as dataTrain:
        reader1 = csv.reader(dataTrain, delimiter=',')
        next(reader1)
        for row in reader1:
            dataHasil1.append(list(map(lambda x: float(x), row[1:])))
    with open('DataTest_Tugas3_AI.csv') as dataTest:
        reader2 = csv.reader(dataTest, delimiter=',')
        next(reader2)
        for row in reader2:
            dataHasil2.append(list(map(lambda x: float(x), row[1:6])))
    return dataHasil1,dataHasil2

#menyimpan data
def dataHasil(hasil):
    with open('TebakanTugas3.csv', mode='w') as dHasil:
        writer = csv.writer(dHasil, delimiter=',', quotechar='"', quoting=csv.QUOTE_MINIMAL)
        i = 1
        for n in hasil:
            writer.writerow([n])
            i += 1
```

Sebelum dijalankan di main program saya membuat 3 function yaitu:

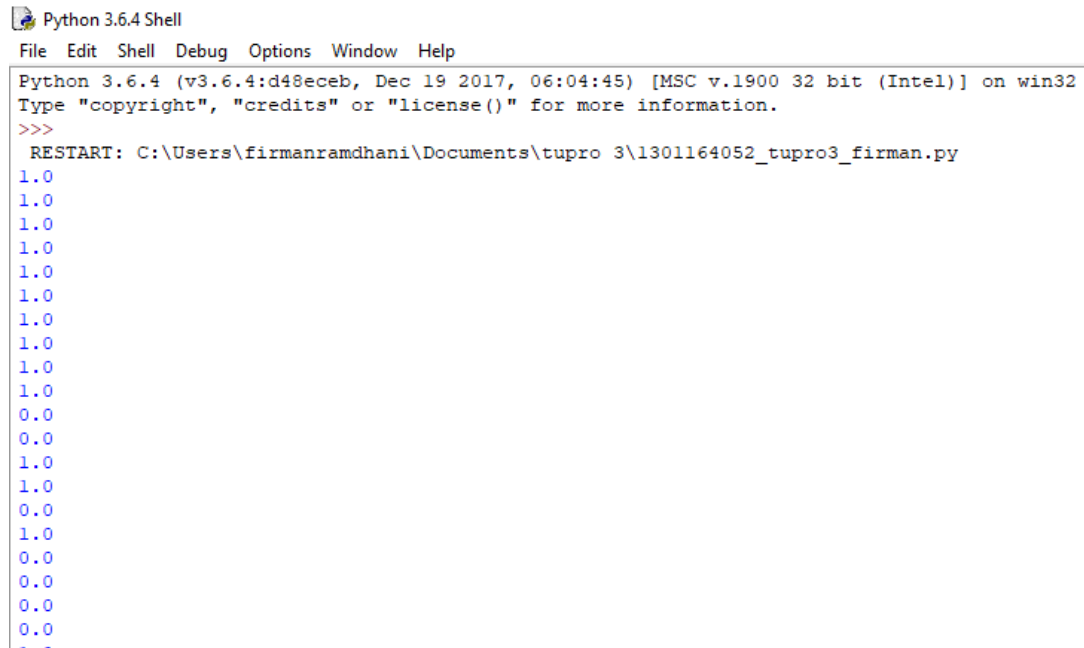
- Function jarak untuk menghitung jarak antara data yang ada di file DataTest_Tugas3_AI.csv t dengan data yang ada di file DataTrain_Tugas3_AI.csv.
- Function loadData untuk membaca file DataTest_Tugas3_AI.csv dan DataTrain_Tugas3_AI.csv kedalam program bahasa python
- Fuction datahasil adalah untuk membuat file baru yaitu TebakanTugas3.csv yaitu

b. Main Program

```
#main program
k = 7
dataTrain, dataTest = loadData()
hasil = []
for test in dataTest:
    list_distance = []
    for train in dataTrain:
        list_distance.append([jarak(test, train[:5]), train[5]])
    list_distance.sort(key=(lambda x:x[0]))
    nearest_distance = list_distance[:k]
    distance, class_type = zip(*nearest_distance)
    counter = collections.Counter(class_type)
    result = counter.most_common(1)[0][0]
    print(result)
    hasil.append(int(result))
dataHasil(hasil)
```

Pada main program saya menginputkan nilai 7 pada variable k, berarti jumlah tetangga terdekat.

c. Contoh Hasil Running



```
Python 3.6.4 Shell
File Edit Shell Debug Options Window Help
Python 3.6.4 (v3.6.4:d48eceb, Dec 19 2017, 06:04:45) [MSC v.1900 32 bit (Intel)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
RESTART: C:\Users\firmanramdhani\Documents\tupro 3\1301164052_tupro3_firman.py
1.0
1.0
1.0
1.0
1.0
1.0
1.0
1.0
1.0
1.0
0.0
0.0
1.0
1.0
0.0
1.0
0.0
0.0
0.0
0.0
0.0
~ ~
```