Lab Assignment 6

(KNN Algorithm)

**Name:** Sourabh Barala

**Course:** M.Sc. Data Science

**Year:** 1st

**Reg. No.:** 23MSD7044

**Subject:** Machine Learning and its Applications

**Implementation:**

**cal\_dist:**

This funtion calculates distance of a data point with all data points

**input:**

point: one data point

train\_para: all other data points

train\_label: labels of ecach data points in train\_para (1- dimensional)

**output:**

DataFrame with distances from a given point to all points, with labels

**example:**

temp\_cls.cal\_dist(point=X\_train.iloc[0],

train\_para=X\_train,

train\_label= y\_train.values.ravel())

output:

dist class

0 0.000000 1

1 38.449169 1

2 271.548671 1

3 109.262951 1

4 160.529340 1

... ... ...

421 1252.334390 0

422 574.286256 0

423 1519.622517 0

424 1552.695208 0

425 184.725959 1

**get\_closest\_class:**

This function takes DataFrame given by cal\_dist function ,sorts the DataFrame according to diatance then selects top 'k' smallest distances and returns label with max count.

**input:**

dist\_df: DataFrame given by cal\_dist function

**output:**

label with max count

**example:**

temp=temp\_cls.cal\_dist(point=X\_train.iloc[0],

train\_para=X\_train,

train\_label=y\_train.values.ravel())

temp\_cls.get\_closest\_k\_classes(temp)

output:

1

**predict**

This function predicts class of each point and appends in a list

**input:**

test\_para: points to be classified

**output:**

list with predicted classes for each point

**example:**

temp\_cls.fit(X\_train,y\_train.values.ravel())

pred=temp\_cls.predict(X\_test)

pred=np.array(pred)

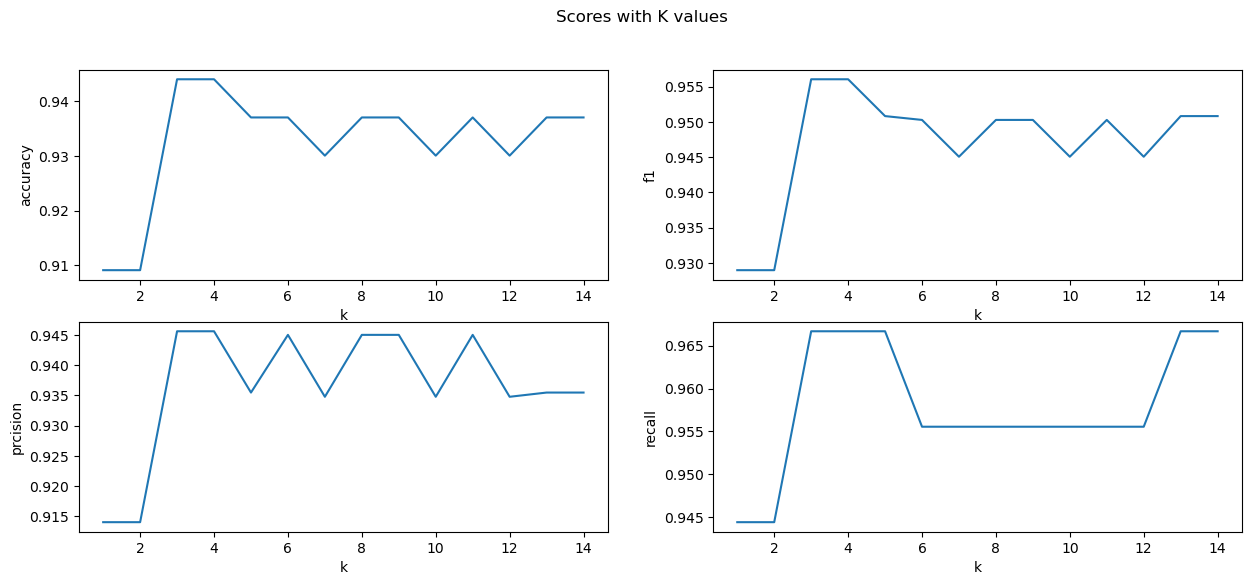
output:

array([1, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, ....])

**Elbow Method**

accuracy\_score, f1\_score, precision\_score, recall\_score are calculated with value of K in range 1 to 14.

Observation:



K=3,4,5 gives best scores