**Divorce Predictors Project**

Summary :

Its good to us to study what the factors that effects our lives directly and try to predict what is going to happen with us in the future , so I took the Divorce Predictors data set from UCI ML Repository and made report us the Required Listed of this Course.

The Divorce Predictors Data Set Contains 170 Participants that tell us how they acted with theirs’s spouses in 54 different specific Attributes are the Authors of this data set find it directly affecting the divorce situation between their couples. for example, how much you enjoy your holidays with your wife, or the time that you spent with your wife is it special for you or not the participants can choose between 0 and 4 how much they agree with attributes or not and as a result of that are they still married or divorced.

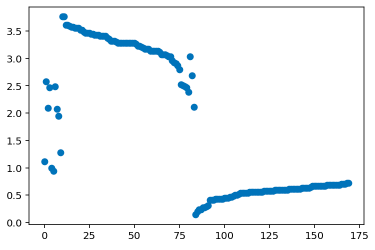
**Plan :**

The data set Contains 170 participants that describing their actions in 54 different attributes happens with their spouses, and depending on their marriage situation ; if they are still married or deforced we are practicing our codes on this target to predict and classify our new target in the future , and that why we are going to use Supervised Learning method for our Our project. The data set is Comma Separated Values CSV file downloaded from UCI ML Website, so we are going to retrieve the data from the file using Pandas library, we will clean our data if there is any outliers and missing values after we detect them by visualize them in histotrophic plots, then we are going make Removing, Assigning, Transforming, or Predicting it so we could have totally clean data and ready for evaluation .

As we can see the data set is simple evaluation took each attribute between 0 and 4 values and it’s out of missing values, and outliers so we don’t need to clean or transform any one of them ;

1. Because their isn’t an unique values so we don’t have to drop any one of the attributes .
2. Their isn’t any string values need to convert to integers.

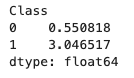
After that we group the data into 2 groups ; first is going to be the spouse that doesn’t divorced which is the Class 1 value , and the other group that the spouses divorced which is the Class 0 value .then we took the mean of the 54 Attribute of each of the 170 instance ( The Class ; the target column is outside the mean process ) to check the range of the values that ends up with been Class 0 or Class 1 .



Depending on data file that we have , all the first 85 instances took the value Class = 1,

And all the others took the value Class = 0. As we can see from plot for the mean of the data we can say that when an instance data mean takes the value between 2 – 4 so the target will be Class = 1 , and for an instance data mean takes value between 0 – 1 the target will be Class = 0 except some a few data .

As we can see below, when the mean values of all the Attributes combined is 0.550818 it’s gave us zero (0) as result and when it’s 3.046517 it’s gave us one (1) as result.



Confusion matrix applied Using KNeighbors classification module :

Result %95 accuracy .

