**Report of KNN**

In the attached .ipy file, we used a dataset file of diabates.csv containing about 768 rows of data which we split into four independent and dependent varaibles using the python libarary into a ratio of .3(70% used as training and 30% used as testing dataset).

The three classifiers are declared with the different values of k, in this case these are 100,200 and the third one is equal to the len of the dependent test variable which is 231.

After predication we got a confusion matrix for each classifier.

* The confusion matrix for 100 nearest values shows that there are 33 cases in which we predicted yes (they have the disease), and they do have the disease and 140 are those which we predicted no, and they don't have the disease. 11 cases are predicted yes, but they don't have the disease. (Also known as a "Type I error.").47 cases are predicted no, but they do have the disease. (Also known as a "Type II error.")
* The confusion matrix for 200 nearest values shows that there are 0 cases in which we predicted yes (they have the disease), and they do have the disease and 151 are those which we predicted no, and they don't have the disease. 0 cases are predicted yes, but they don't have the disease. (Also known as a "Type I error.").80 cases are predicted no, but they do have the disease. (Also known as a "Type II error.")
* The confusion matrix for the values equal to the length of dependent variable of test data set shows that there are 0 cases in which we predicted yes (they have the disease), and they do have the disease and 151 are thoese which we predicted no, and they don't have the disease. 0 cases are predicted yes, but they don't have the disease. (Also known as a "Type I error.").80 cases are predicted no, but they do have the disease. (Also known as a "Type II error.")

Accuracy has also been checked through a formula( which is copied from stackoverflow ☺).

Accuracy for 1st prediction is 74% and for 2nd is 65%. And accuracy for the predicted values of 3rd prediction( in which k is equal to the length of depedent variable) is also 65%.