**LOGISTIC REGRESSION HOMEWORK**

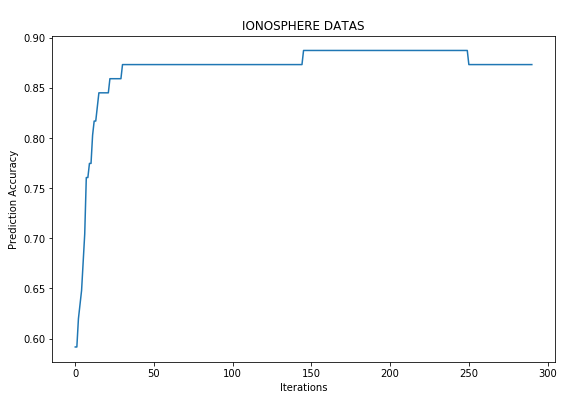
Data was added to a dataFrame and shuffled. For seperating data as train and test, ’rate’variable was created. With int(((len(df2))\*80)/100) operation, rate was determined. And with the help of this rate, our data was divided into 80 train and 20 test.

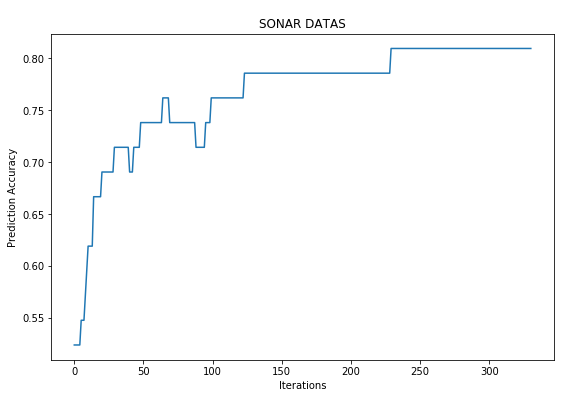
As we saw in the code, after creating test and train datas, we create a model with using LinearRegression class. With ‘fit’ method in this class, our model is being created. We have two nested for. The reason of this, we want to use mini-batch gradient descent. Our aim is processing data little by little. In inner for, we do this with ‘batch\_size’ variable. ‘batc\_size’ was set to 32. Because, choosing the multiples of 2 has always helped us in computer science. Generally, batch\_size is set to between 32-512. In this experiment we have small dataset. So, 32 will be sufficient.

We use sigmoid function to get our hypothesis. Then we apply gradient descent to get better theta vector and so on to get better prediction accuracies.

To stopping iterations, log\_likelihood is used. With this function, we can understand where to stop. The change of our loss will decrease after a point. The point i choosed corresponds 0.001. If (total loss of one iteration – total loss of previous iteration) is smaller than 0.001, model fitting will be finished.

We can’t check test error to stop iterations. Because, although our hypothesis changes, test results may not change, and that would be pretty deceptive.





**Ahmet Haşim Delibaş**