Birla Institute of Technology and Science, Pilani, Hyderabad Campus

2nd Semester 2017-2018 (CS F407: Artificial Intelligence)

Assignment 4 (Max Marks:15) Date of Submission:23.04.2018

A consistent increase in customer dropout was observed by an Indian Bank. The bank has its branches in Pilani, Goa and Hyderabad. The reasons of the dropout were unknown and the bank authorities wanted to find out whether the current customer base will stay or leave the bank. The bank collected data of customers who exited and those who are still active customers. This survey held information of 10000 customers which included those who exited and who are still active customers at the bank.

The dataset holds the following features:-

1. UID- Contains the Unique Customer ID assigned to individual working in the MNC.
2. Customer_name- Describes the Name of each customer.
3. CreditBonus- Credit points gained over using credit card for various transations made online.
4. City-The city office to which individually an customer works for.
5. Gender-Male or Female
6. Age
7. Tenure- Total time they have been associated with the bank.
8. Balance- Depicts the current balance that they have in their accounts.
9. Availed Products- Services of the bank that has been availed by any individual customer. The products or services maybe usage of Internet Banking Facility, Mobile Banking, Loans,etc.
10. CreditCardPresent- Information whether the customer has availed a credit card or not.
11. ActiveOnline- Gives the information whether the customer has transacted on any online sites by Internet Banking,Credit Card,Debit Card,etc or not in the last few months.
12. Current Salary- The salary deposits of individual customers registered in their respective accounts.
13. Status-Depicts whether the customer continued to work or resigned from his/her post. (ground truth)

The dataset contains 10000 customer data based on certain attributes which can be used in the creation of a MultiLayer Perceptron using Back Propagation Algorithm. It can be trained to efficiently classify the data into two sets:exit and stay. This should be able to predict whether a customer might stay with the bank or leave it in future.

Some questions that can be observed by the MNC in this scenario are:

a. What are the features that can be necessarily taken to handle such a classification task ? Give valid reasons.
b. What can be the accuracy of an MLP built on this data ? Use 5 fold cross validation and 10 fold cross validation on the data.
c. What will be the activation function that can be used with the MLP in this scenario ? Give your reasons in favour of the choice you made.
d. What are the assumptions made for incorporation of number of hidden layer and number of nodes in the hidden layers ?

The assignment has to be created using Tensorflow.The assignment has to be made using Python 3.5 or 3.6. The code has to be executed on systems available in IPC 5.

References:

Tensorflow- https://www.tensorflow.org/

(Dataset attached)

Mode of Submission:

Form your own groups of three only. Tar the entire source and executable files with your id as the tar file name (e.g. f20140055.tar) and send it to the mail id- p20150408@hyderabad.bits-pilani.ac.in. Include a readme.txt with group details in your tar file. Submit only one file per group.

(Date given:16.04.2018)