Instructions

Review the [German Credit DataSet (Links to an external site.)](https://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data)) (https://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data)) in the attachment. It has 1,000 observations.

You can also [download data here in the excel format. Preview the document](https://njit.instructure.com/courses/12618/files/1053572/download?wrap=1)

Train, test and validate a neural network with the first 980 observations and however many neurons in the hidden layer as you like.  The raw data needs to be preprocessed to be used by your ML tools/modules.    We cleaned up the data.   You can download the cleaned-up data here.    Here is an example that you can use it for your ML project.

Take a look at the data and remove a few attributes that you think do not help to determine the creditworthiness of a customer.

The last column is whether a customer is actually "good" or "bad" (i.e., their credit rating). See if you can improve the accuracy by changing various parameters, such as the number of neurons, and the number of layers. After you train, test your holdout 20 samples and report the results using the method below.

If your predictions are correct (good or bad) for each example, that counts as 0. If your prediction for a good customer is "bad" add 1 to your total. If your prediction for a bad customer is "good", add 5 to your total. The lower the total the better your neural network. Use your Group Discussion forum to either set up times and ways to meet or to discuss the assignment.

Evaluation criteria: the score should be as small as possible.

Language to be used: Python,  Notebook.

 What you need to do.

 1) Partition the data into training (980 data points) and holdout (last 20 data points) datasets.

 2) Get the evaluation score of the 20 holdout data.

Submission of project python notebook and project report (your project report should be embedded in the notebook).  You need to make it appear to be as professional as possible).  Here is what you need to provide for your project report:

* A) Summary of your project
* B) Discuss Data Mining Methods that are appropriate for your project problem.
* C) Business Understanding, what motivates you to do the project.
* D) Data Understanding, describe your understanding of the credit data.
* E) How do you prepare data for your chosen machine learning models?
* F) Testing and Evaluations  (show all your screen plots and performance results).
* G) Conclusions

Here are the [German-Credit-Card-To-Be-Filled-Example.ipynb](https://njit.instructure.com/courses/12618/files/1053611/download?wrap=1) (incomplete) and [dataPreview the document](https://njit.instructure.com/courses/12618/files/1053617/download?wrap=1)txt file, you need to fill in the necessary components.