Assignment 2

Advanced Topics in High-Performance Computing (MCSC 6230G/7230G)

Faisal Qureshi

http://faculty.uoit.ca/qureshi

Computer Science, Faculty of Science

University of Ontario Institute of Technology

Due Back Nov. 1, 11:59 pm

Introduction

The goal of this assignment is to get you started with *tensorflow*.

You are asked to setup a *feed-forward neural network* in tensorflow for the purposes of binary classification. See below (Fig. 1) an example of a 2 layers, 3 perceptron feed forward network.
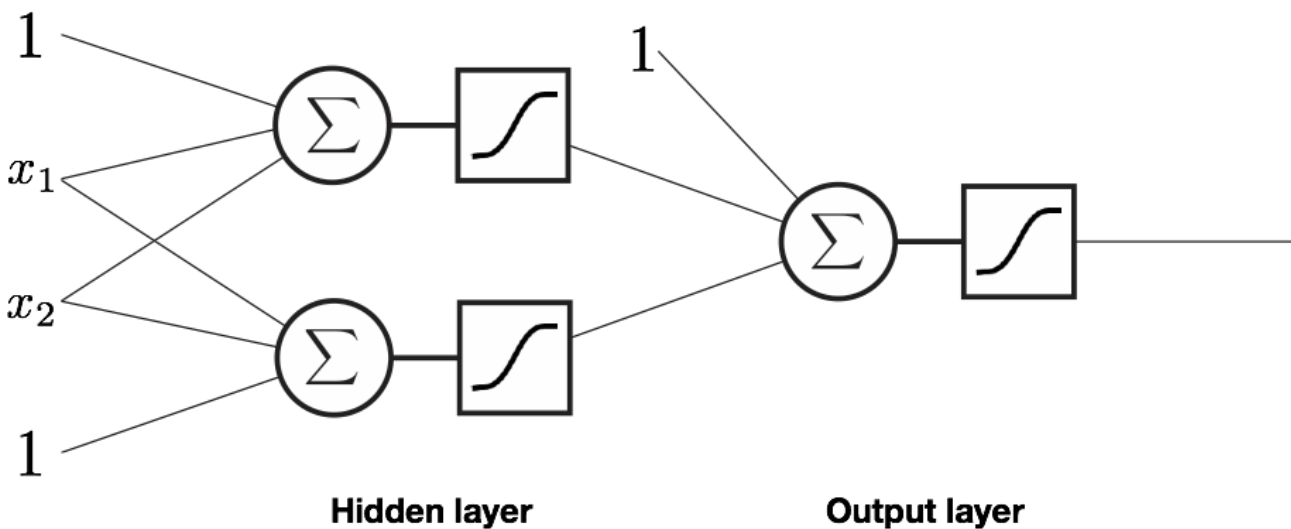

Figure 1: 2 layers and 3 perceptrons

For the sake of this assignment, we assume that the data is 2D and is generated using the following python code.

```
import numpy
import sklearn
from sklearn import datasets
import matplotlib.pyplot as plt
```

```
x, y = sklearn.datasets.make_moons(n_samples=1000, shuffle=True, noise=0.1,  random_state=0)

plt.figure()
plt.scatter(x[:,0], x[:,1], c=y)
plt.show()
```
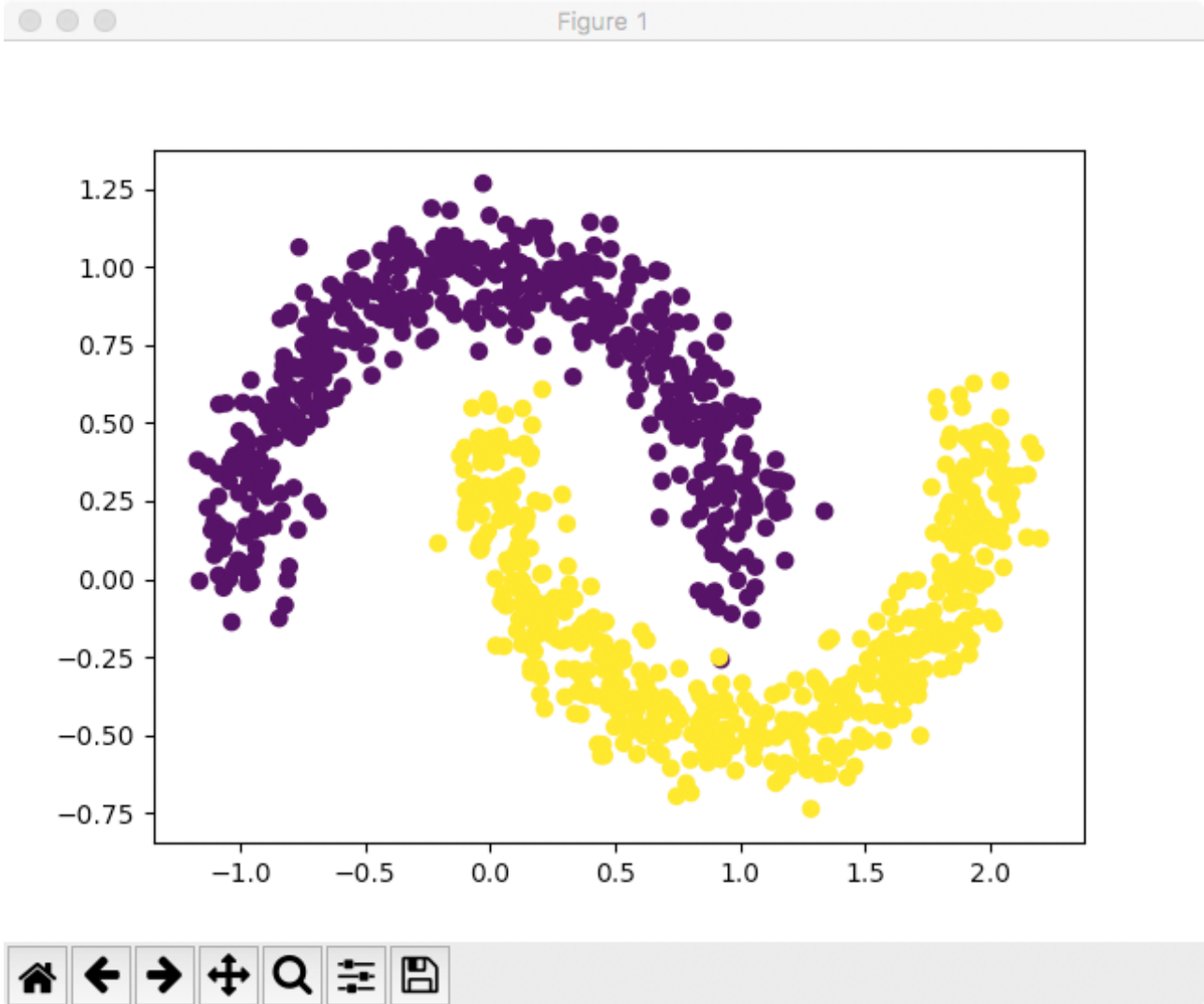
The dataset looks as seen below (Fig. 2)

Figure 2: Two moons

Tasks

1. Set up a two layer, 3 perceptron feed forward network as seen in the picture above. Train it using a percentage of the data (say 60%), and check its accuracy.

2. Now add perceptrons in the hidden layer and see if you get any improvement in accuracy. To keep things manageable increase the number of perceptrons in the hidden layer from 3 to 20 using appropriate stepping.

Hints

You should treat data with care:

- Split data into test and training sets randomly
- Try different percentages of data for training (or testing)
- Play with different number of samples
- Play with the noise paramter to create cleaner or more messy data
- If you code it right, you may be able to write a single tensorflow model that can account for different number of perceptrons in the hidden layer
- Check out sample code on course github

Submission

Submit a 2 to 4 page report outlining your findings. We are essentially trying to find the simplest network that will do a reasonable job at classifying this data. Or a determination that our feed forward network is unable to classify data of this sorts.

To support your arguments, please present ROC curves for different "classifiers". Check out http://scikit-learn.org/stable/auto_examples/model_selection/plot_roc.html for more information.

Grades will reflect the clarity of your report, the thoughtfulness that you have put into your experiments, and of course the correctness of your experiments.

Submit via Blackboard.