**Machine Learning Assignment 1**

**Neural Networks**

Neural networks are a set of algorithms, modeled loosely after the human brain, that are designed to recognize patterns. Neural networks help us cluster and classify data.

**Neural Network Architecture for Binary Classification**

* The neural network we used has a three layer architecture. There are two layers of 5 nodes each and one output node.
* The last node uses the sigmoid activation function that will squeeze all the values between 0 and 1 into the form of a sigmoid curve.
* The other two layers use ReLU (Rectified Linear Units) as the activation function.
* Biases are initialized with zeros.
* Weights are initialized by He Normal (He-et-al) Initialization.

It draws samples from a truncated normal distribution centered on 0 with stddev = √(2 / fan\_in) where fan\_in is the number of incoming connections coming into a given layer from the previous layer’s output. This is proven to result in faster and more efficient gradient descent when ReLU activation is used.

Mini-batch gradient descent has been used that seeks to find a balance between the robustness of stochastic gradient descent and the efficiency of batch gradient descent.

**Results**

Number of epochs = 1000

Mini-batch size = 10

Learning rate = 0.005

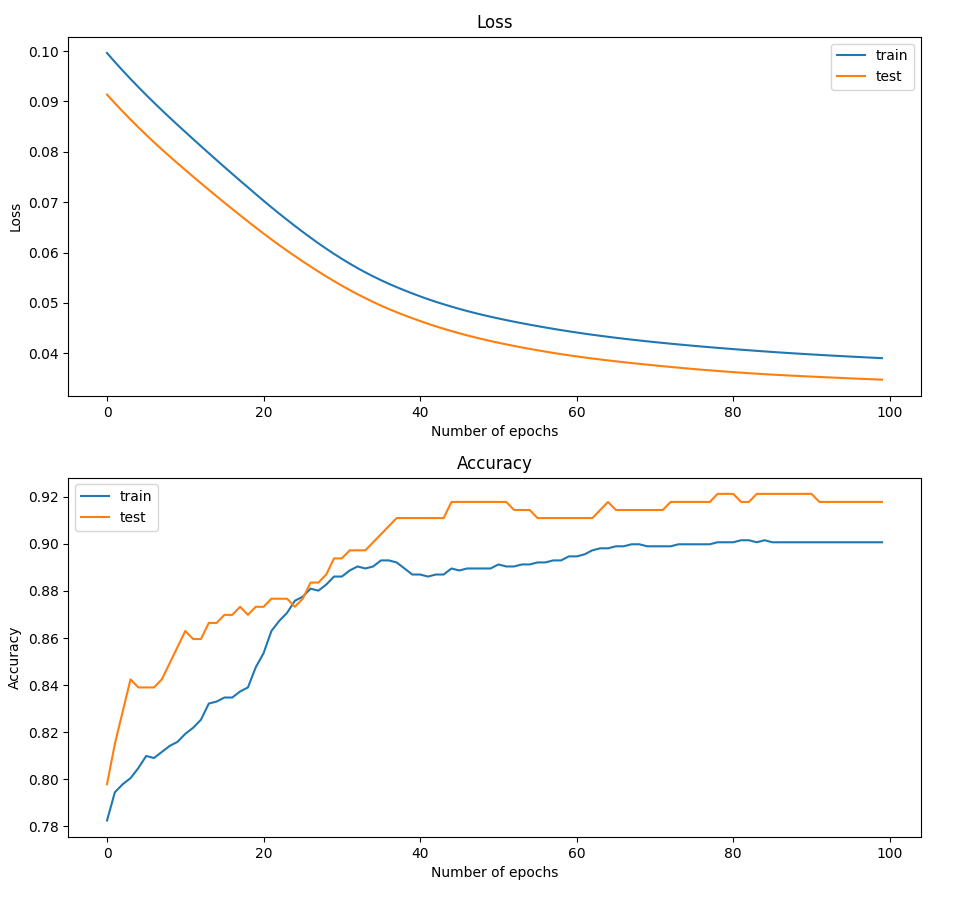
Running neural network 10 times

**Accuracy = 0.905 ± 0.014**

**F-score = 0.906 ± 0.013**

**Loss and Accuracy Plots:**

It is observed that loss function decreases and accuracy increases gradually with number of epochs for both training and testing data.



100 epochs, mini batch size = 10, learning rate = 0.01

Accuracy = 268 / 292 = 0.918