HW3, Aaron Yerke

BINF 6210 / BINF 8210:  Machine Learning

Due, Tuesday, December 10, 2019

1.  (50 points)  Implement your  own neural network  having two hidden layers  in

Python.

2.  (30 points) Apply your neural network to the Wisconsin Diagnostic Breast Cancer

(WDBC) data for classifying the cancer.  Perform leave-one-out cross validation.

The data can be downloaded from the UCI (UC Irvine) Machine Learning Repos-

itory.  The webpage that describes this breast cancer data repository can be ac-

cessed by clicking on this link:

Breast Cancer Wisconsin (Diagnostic) Data Set

and then click

Data Folder

at the top of the page to go to the download site.

Or you can click the following link directly to go to the download site:

Download

The data is in file

wdbc.data

and description of the data is in file

wdbc.names

.

We (the class) collaborated to build the neural net with 2 hidden layers and implemented leave one out validation for predicting breast cancer. Since the dataset had 569 records, our leave-one-out-validation also has 569 records. These are plotted below. Our model gets an average of 92.44 percent accuracy in the cross-validation. This is really good, though it can probably be improved by manipulating the thresholds and other parameters.

****