Machine Learning Project 3

1. The breast cancer dataset is comprised of data that periodically arrive as Dr. Wolberg reports his clinical cases. There are two different classes benign and malignant. There are 30 different features including mean\_radius, mean\_texture, and perimeter\_mean. The dataset contains 570 samples of patients. The objective is to determine if the spot is benign and malignant by using mean\_radius and the mean \_texture.
2. The two classes benign and malignant and the different features have the following observations.

Benign

|  |  |  |
| --- | --- | --- |
|  | radius\_mean | texture\_mean |
| Min | 6.981 | 9.71 |
| Max | 17.85 | 33.81 |
| Mode | 13.05 | 16.84 |
| Mean | 12.15 | 17.91 |
| Median | 12.2 | 17.39 |
| STD DEV | 1.78 | 3.955 |

Malignant

|  |  |  |
| --- | --- | --- |
|  | radius\_mean | texture\_mean |
| Min | 10.95 | 10.38 |
| Max | 28.11 | 39.28 |
| Mode | 15.46 | 19.98 |
| Mean | 17.5 | 21.6 |
| Median | 17.325 | 21.46 |
| STD DEV | 3.204 | 3.779 |

1. For my study, I tested changing the hidden layer, which in the study seemed to change the accuracy of the models. The higher the number of neurons, the more likely it is to be overfitting the data. The more layers you add the more accurate the results become because you can fit more parameters into it. When I started changing the data from unscaled to scaled increased the accuracy of the findings by 2 to 5%.
2. The first layer exploration started with exploring 5 and 10 nodes in the hidden layers. The accuracy of both of my results ranges from 88% to 90%. I also explored the different alpha values on the one hidden level with Adam. Exploring the different alpha values went from 0.0001 – 1 found that the lower the learning rate of the network resulted in better results but with a slower wait time and the higher the learning rate resulted in faster results but less accurate and more likely to have undesirable divergent behavior.

My results show that with Adam the more hidden layers there are the more accurate it seems to be. Adam also seemed to be the most accurate but was closely followed by LBFGS and SGD. The logistic seems a little bit finicky with my test with all the solvers. The SGD had trouble on the first level and LBFGS had trouble on the second level. My results show that the best general solver is Adam and rulu is the activation function.

1. My conclusion the Adam solver seems to be the most accurate. The Rulu activation function seemed to keep its accuracy with the number of hidden layers thrown at it. The logistic activation function seemed not to work with certain solvers. I noticed that too many neurons in the hidden layers resulted in overfitting. Also, the higher the number of neutrons, the higher the amount of time it takes to train them. By adding more hidden layers you can add more parameters to the model making it more accurate. I also noticed that scaling the data can be helpful in many cases but not all.