Justin Park
112184933
park1996
CMSC389A

Practical 1

Analysis

1. **What is the role of the learning rate?**
   The learning rate determines the rate of change for the bias and weights. A larger learning rate will cause the bias and weights to increase/decrease more quickly and a smaller learning rate will cause the bias and weights to increase/decrease less quickly.

2. **How many passes over the data do you need to complete?**
   The logistic regression accuracy is above 60 even at 0 epochs. The models are trained on 80% of the 768 examples which is 614 examples.

3. **What features are the best predictors of each class? How (mathematically) did you find them?**
   The best predictors are 2-hour serum insulin, plasma glucose concentration, and diastolic blood pressure. I found these by finding which 3 features had the greatest weights.

4. **What features are the poorest predictors of each class? How (mathematically) did you find them?**
   The poorest predictors are diabetes pedigree function, number of times pregnant, and body mass index. I found these by finding which 3 features had the smallest weights.

5. **What is an advantage of the perceptron algorithm compared to logistic regression?**
   The perceptron algorithm is simple to understand and easy to train.

Extra Credit

1. **Explain why the normalize_dataframe function would be useful.**
   This would be useful so that the models will be trained in a way that is independent of the relative scales of the features.

2. **The effect of using a schedule to update the learning rate.**
   When I train the models using 1000 epochs, I am able to achieve accuracies of 72.089 and 67.532 for the logistic regression model and perceptron model respectively. However, when I use a decay rate of 0.001, I am able to achieve accuracies of 74.675 and 74.026.