Ahmed Almutawa

HW2 Analysis

What is the role of the learning rate?

The learning rate decides how fast we can converge to the optimal solution, too small and it will take way too long and if it's too large we will miss the optimal solution entirely.

How many passes over the data do you need to complete?

Testing up to 4 passes, we notice that the test accuracy plateaus from the very first pass at around 93-94%. 1 pass should be enough, lest we risk overfitting (eegads!).

What words are the best predictors of each class? How (mathematically) did you find them?

BB: ['ball', 'BIAS_CONSTANT', 'baseball', 'peak', 'starters', 'pitching', 'saves', 'bat', 'runs', 'hit']

Hockey: ['hockey', 'playoffs', 'pick', 'playoff', 'points', 'period', 'pp', 'golchowy', 'goals', 'next']

I got them mathematically by sorting the beta values in ascending and descending order and grabbing the first 10 values. The highest values (positive) were the best Baseball predictors, the lowest values (negative) were the best Hockey predictors.

What words are the poorest predictors of classes? How (mathematically) did you find them?

Worst: ['racist', 'bloody', 'blasted', 'hooked', 'intermissions', 'rode', 'silence', 'vintage', 'pitiful', 'everywhere']
These were the absolute beta values closest to zero. Since we had a range from negative to positive, zeros were right in the middle and hence the worst predictors. I sorted ascending order with the absolute values of the betas.