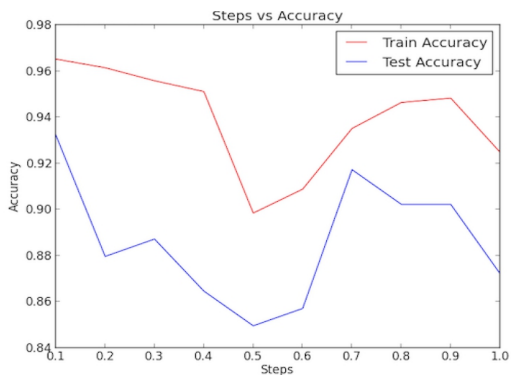
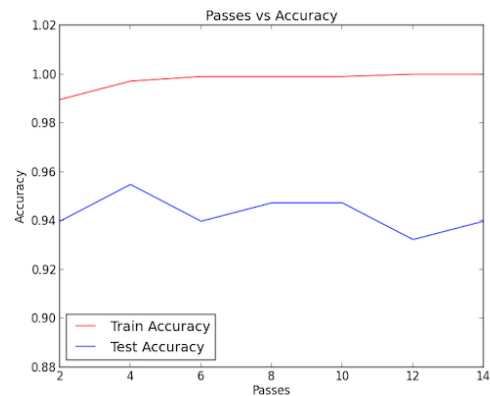


Name: Krishna Chaitanya Sripada

A. Role of Learning Rate: As we increase the step size, we see that the accuracy decreases. The peak accuracy is obtained at the value of 0.1.



B. Numbers of passes over the data is 12 to get accuracy =100%.

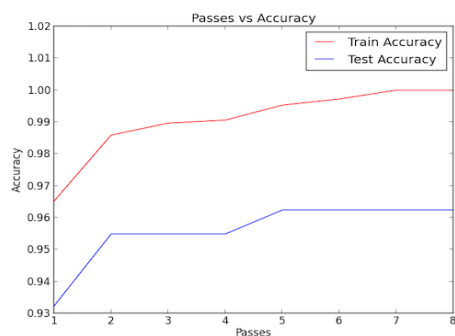


C. The words that are best predictors of the “Positive” and “Negative” class are “hit” and “hockey” respectively. To mathematically calculate these values, we have to take the absolute value of all the β values and take the index of the maximum value and find the corresponding word from the vocab.

D. The words that are poorest predictors are “everywhere, blasted, intermissions, bloody, broad, deceased, hesitate, hooked, memoriam, pitiful, racist, riel, rode, silence, tone, vintage, wrestling”. To mathematically calculate these values, we have to take the absolute value of all the β values and take the index of the values that are 0 and find the corresponding word from the vocab.

Extra Credit:

A. As the step is varied, we could see that the number of passes to reach the highest accuracy fell from 12 to 7. Also there was a considerable increase in the test data accuracy. A non-adaptive scheduling called “Search-then-coverage schedule” was used.



B. DF vector instead of X for update shows that the accuracy increases for a fewer number of passes.

