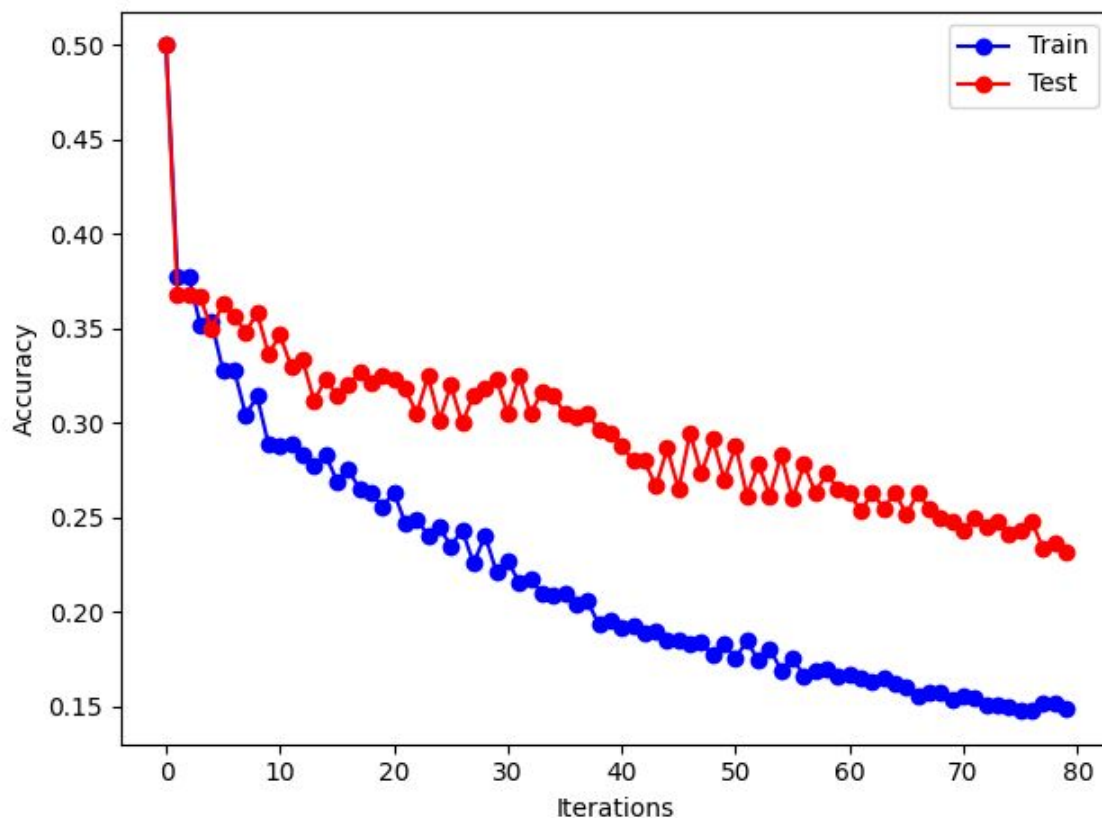


Programming Assignment

1) There training and test errors per iteration are:



2) Here are the top 10 classifiers chosen by the algorithm:

(Direction = 1 iff word implies negative review)

Iteration: #0, Dimension: 26, Theta: 1.5, Direction: 1, Word: "bad"
Iteration: #1, Dimension: 195, Theta: 1.5, Direction: -1, Word: "performances"
Iteration: #2, Dimension: 372, Theta: 1.5, Direction: 1, Word: "boring"
Iteration: #3, Dimension: 37, Theta: 2.5, Direction: -1, Word: "great"
Iteration: #4, Dimension: 88, Theta: 1.5, Direction: 1, Word: "script"
Iteration: #5, Dimension: 22, Theta: 1.5, Direction: -1, Word: "life"
Iteration: #6, Dimension: 31, Theta: 1.5, Direction: -1, Word: "many"
Iteration: #7, Dimension: 311, Theta: 1.5, Direction: 1, Word: "worst"
Iteration: #8, Dimension: 120, Theta: 1.5, Direction: -1, Word: "family"
Iteration: #9, Dimension: 76, Theta: 1.5, Direction: 1, Word: "nothing"

Three classifiers I expected are in blue (very straightforward words indicating good/bad movies).

Three less intuitive classifiers are in red, they do not directly imply a negative/positive emotion.

3) The average exponential loss per errors per iteration is:

