CMSC726: Assignment #2 — Logistic Regression

Soheil Behnezhad

September 15, 2017

1.

2.

3. Words with higher (resp. lower) biases are the better predicators for class 1 (resp. 0). Therefore by sorting the words based on their bias we can determine the best predicators for the classes.

   The first 5 words with the lowest bias are as follows: [(-1.8519, 'hockey'), (-1.2709, 'playoffs'), (-0.9959, 'pick'), (-0.9045, 'playoff'), (-0.9012, 'points')]

   The first 5 words with the highest bias are as follows: [(1.2073, 'hit'), (1.1462, 'runs'), (0.9299, 'bat'), (0.9003, 'saves'), (0.8407, 'pitching')]

4. Ok.