

CS294-1 Programming Assignment 1

Richard Hwang, David Huang

February 12, 2013

1 Our Solution

We followed advice in the assignment statement, splitting the text with the regex `"[\\s.,()!?:&\\"]+"` and indexing each word. Afterwards, we constructed a sparse matrix, with words as rows and documents as columns

2 Smoothing/Backoff

3 Refinement Effects

4 Performance