MCA Assignment 3

Sudhanshu Mohan
2017110

Q1:
tldr:
Algo, is to take 3 layers of NN then feed it a word to predict its neighbourhood word, then remove the output layer and then when we input a word the output will now would be word embedding of the input word

The CBOW model:

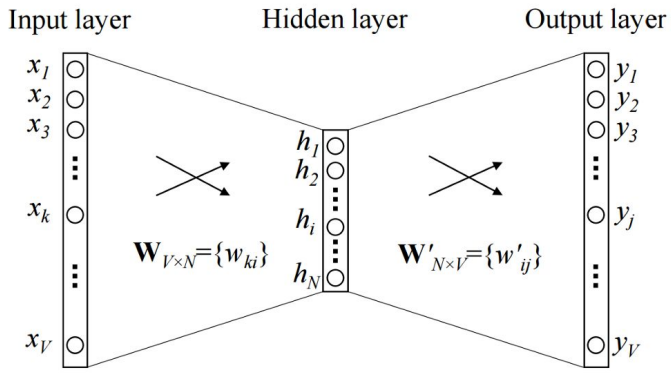
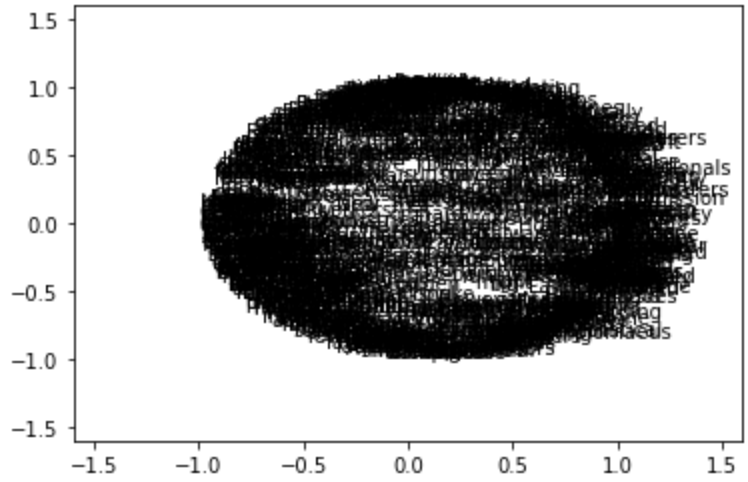


Figure 1: A simple CBOW model with only one word in the context

Output of tSNE dim reduced to 3



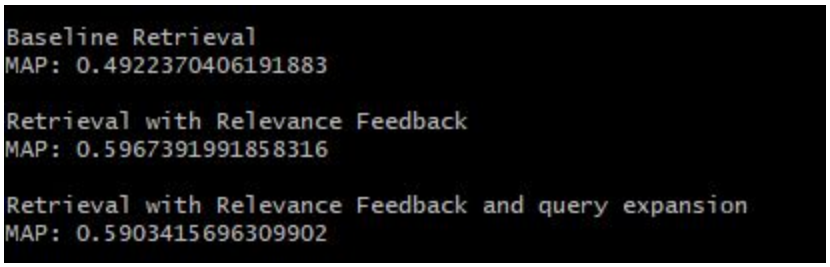
Q2 (Used python3)
Original

- Baseline Retrieval MAP: 0.4922370406191883
- Retrieval with Relevance Feedback MAP: 0.4922370406191883
- Retrieval with Relevance Feedback and query expansion MAP: 0.4922370406191883

After relevant changes to implement relevance feedback and relevance feedback query expansion:

Over 3 iterations:

- Baseline Retrieval MAP: 0.4922370406191883
- Retrieval with Relevance Feedback MAP: 0.5967391991858316
- Retrieval with Relevance Feedback and query expansion MAP: 0.5903415696309902



As expected the MAP score after implementing relevance feedback increased as now we are focussed on the relevant documents and penalised the irrelevant documents

For Relevance Feedback:

beta = 0.4

alpha = 0.6

For Relevance Feedback and query expansion

beta = 0.2

number_of_terms = 9

alpha = 0.8