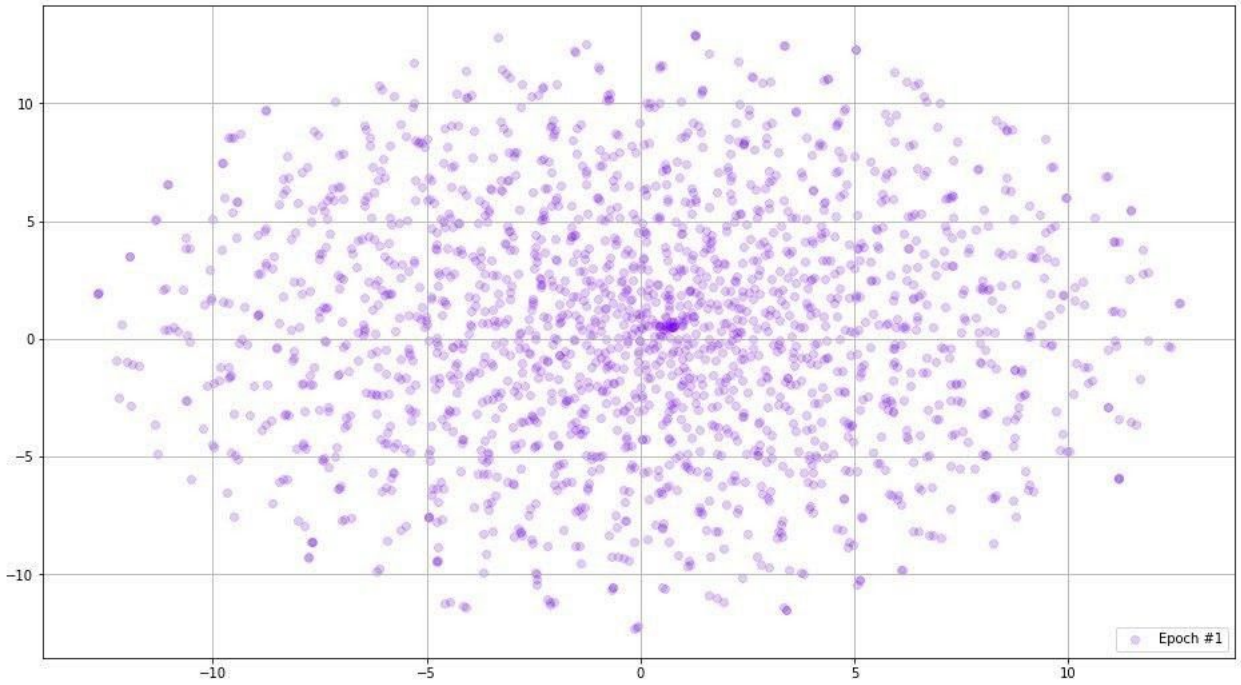
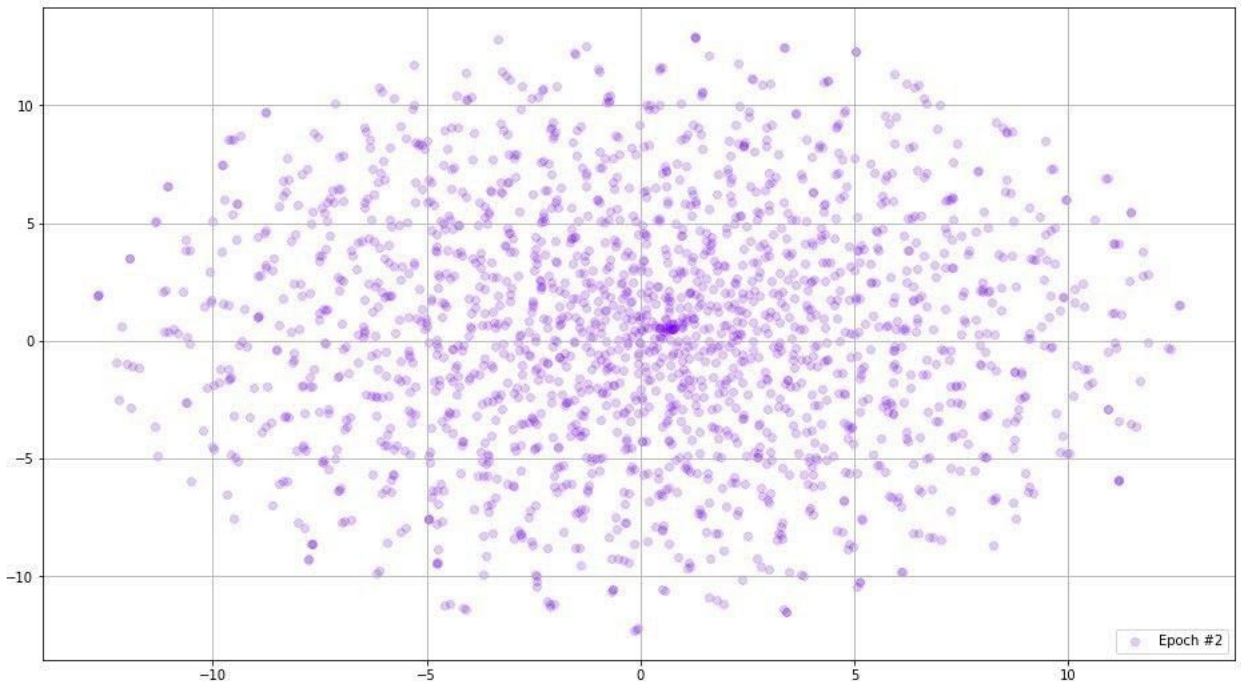
Assignment 3

Archit Singh
2017029

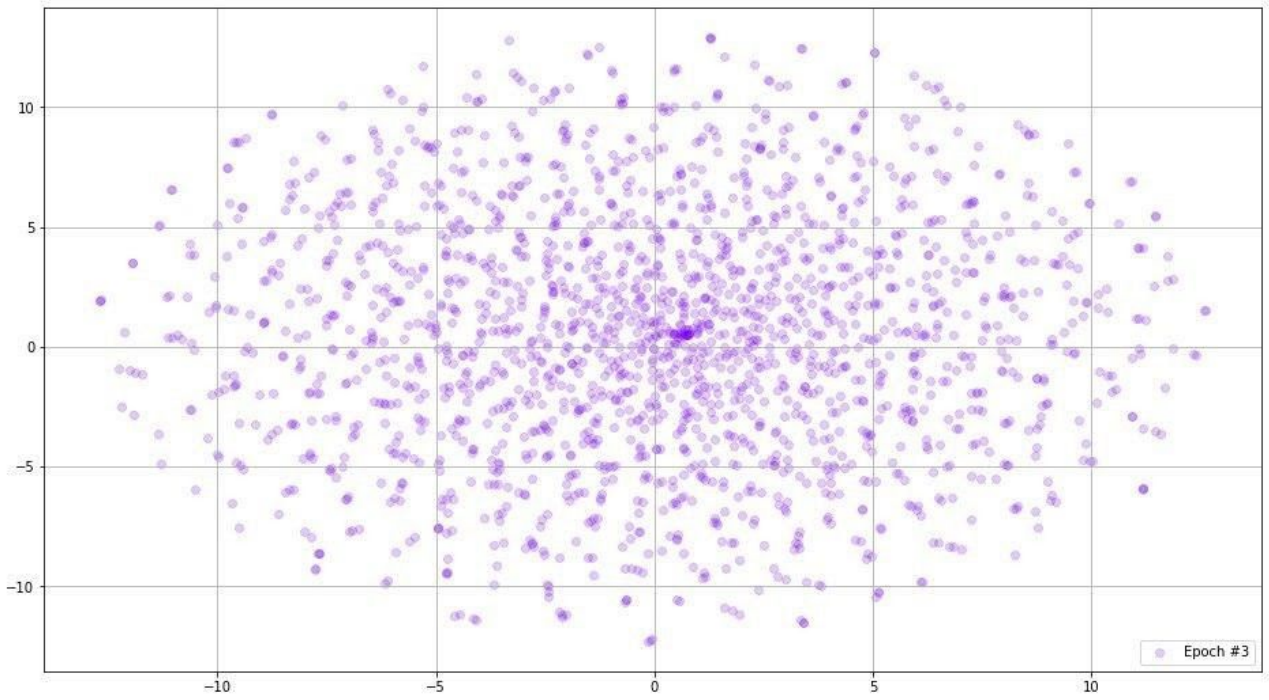
Question 1: Word2Vec





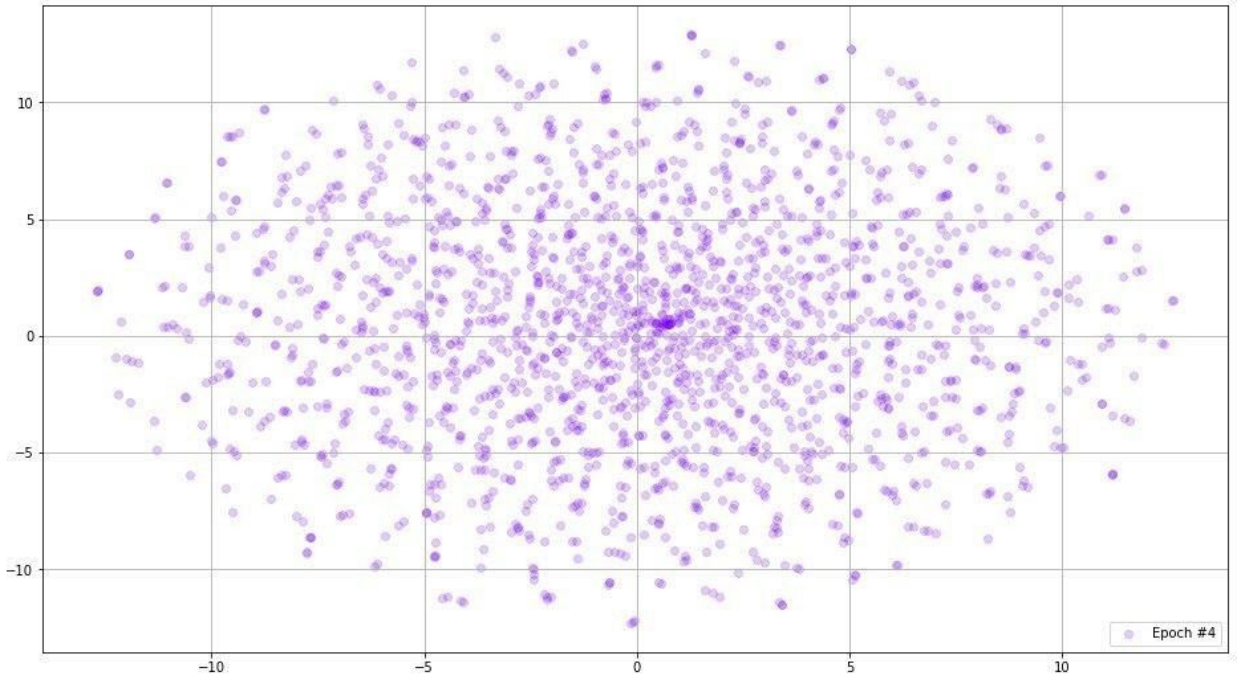
Assignment 3

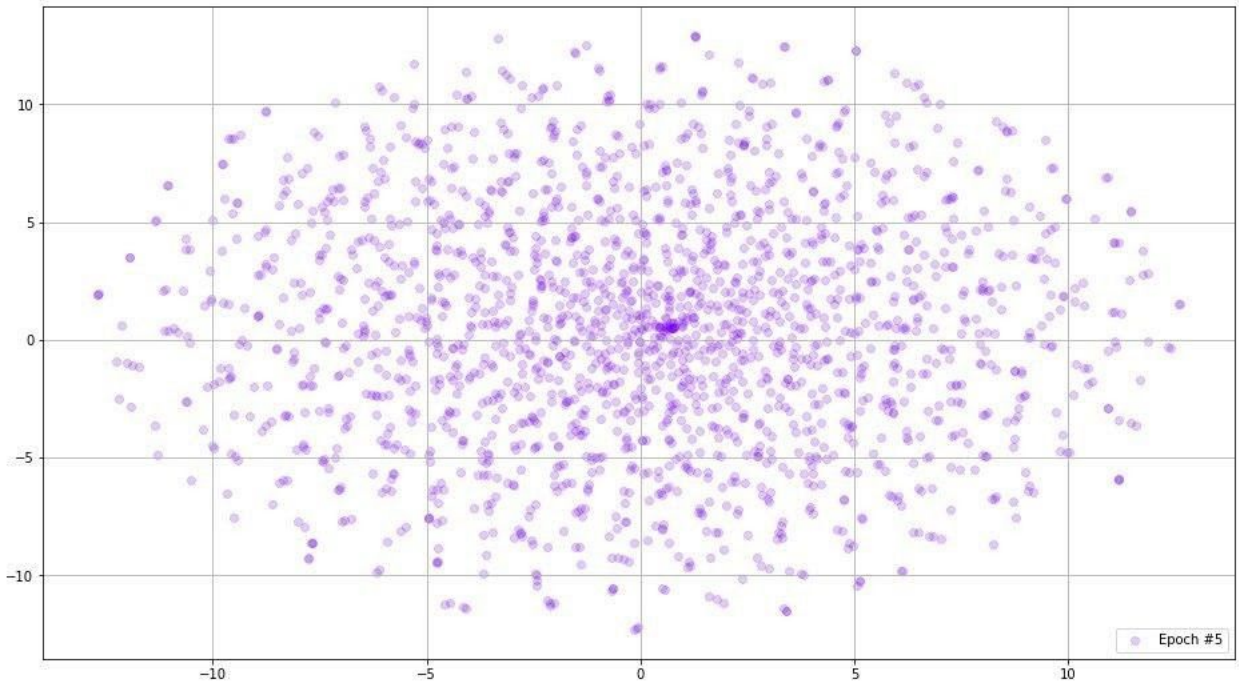
Archit Singh
2017029



Assignment 3

Archit Singh
2017029





(i)

Word2Vec is a deep learning architecture with 4 parts, namely:

1. Word Embedding, this layer is the representation of text and their similarities.
2. Deep Network, this layer takes the sequence of embedding vectors as input and converts them to a compressed representation.

Assignment 3

Archit Singh
2017029

3. The Fully Connected layer, this layer takes the deep representation from the RNN/LSTM/GRU and transforms it into the final output classes or class scores.
4. Output Layer, this layer is a simple sigmoid softmax output function.

(ii)

During the training, we observe that similar words have similar values and accordingly, made small clusters, of 3-4 data points.

Question 2:

(i)

Scores: [iterations 3 : alpha 0.75, beta 0.15, top 5 terms, top 10 documents]

1. Without relevance feedback: 0.46
2. With relevance feedback: 0.54
3. With relevance feedback and query expansion: 0.57

(ii)

The result is inline with our expectation because, the relevance feedback takes the result that is initially returned from a given query, to gather user feedback, and to use information about whether or not those results are relevant to perform a new query, as such it slightly increases the accuracy.