# HW-3 Report

Ques 1) Result :- Loss = 0.587

Effectively, the algorithm works in the following manner:
1. Initially, two pairs are made one of (target word, real context word) and other of (target word, negative context word) and both of them are sent as input.
2.   This input is then passed through the embedding layer which gives the embeddings of the words.
3.   The dot product of these 2300-length vectors is calculated to obtain the similarity.
4. Finally, the similarity score obtained above is sent to a sigmoid layer which gives us an output of 1 or 0 (1 for a true context word, 0 for a negative sample). This is matched with the ground truth label given to the real context word.
5. The errors that are generated while training are back-propagated and this updates the embedding layer. This enables that words which actually share meanings and contexts have vectors which return high similarity score.
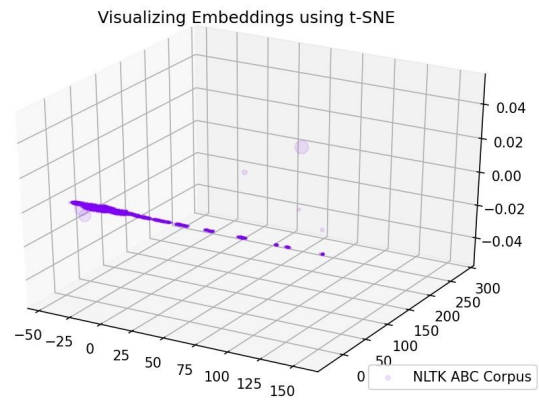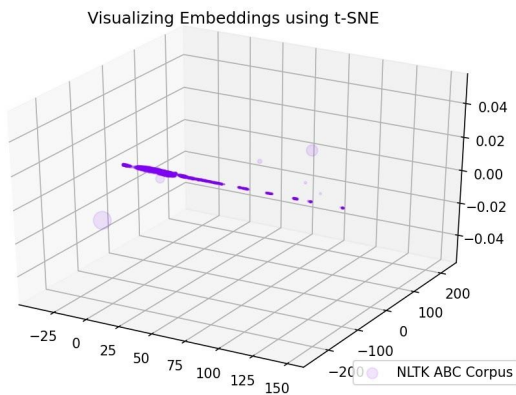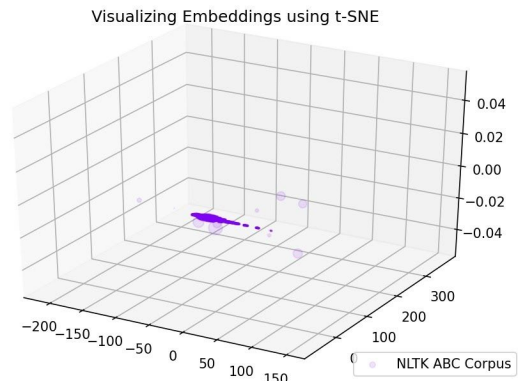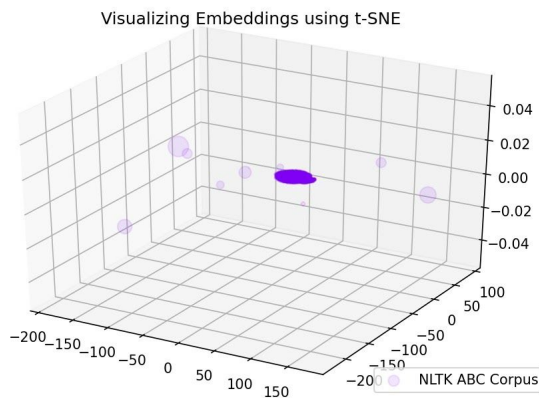
At the beginning of the algorithm, as we can clearly see the words are present in clusters/groups. And after every epoch of 10-sec words are being separated from the clusters.

Ques2) **Results**
Baseline Retrieval
MAP: 0.49176786896815833

Retrieval with Relevance Feedback (α = 0.8 β = 0.6)
MAP: 0.5829868326225287

Retrieval with Relevance Feedback and query expansion (α = 0.8 β = 0.6)
MAP: 0.6258735323494148

Yes, the changes in the results are as expected. As the results are increased upon including the extra information every time.