**Figure.docx**

Figure 3.1: A two-dimensional t-SNE projection of word vectors using GloVe-6B with 300-dimensional word embeddings. The relative positions of these word vectors in this two-dimensional vector space reflect the semantic meanings captured in the GloVe embeddings. Here, words of similar types are neighbours, which are clustered in the dotted ellipses. Notice how ‘apple’ lies between the cluster of ‘food’ and the ‘devices.’ Also, note that the axes values are arbitrary t-SNE components, and only the relative positions are meaningful.

Figure 3.2: An illustration of the skip-Gram model to predict the context words ‘I’ and ‘in’ given the input centre word ‘life’. The centre word is converted to a one-hot encoder vector of dimension |V |, and the output final layers are the one-hot encoding vectors for the two predicted outputs. The model learns the weight matrices, P and Q, which are used as the embedding. Each row of the P matrix is a d-dimensional embedding.

Figure 3.3: Illustration of the CBOW model - the context words {‘I’ and ‘in’} is used for predicting centre word ‘live’.