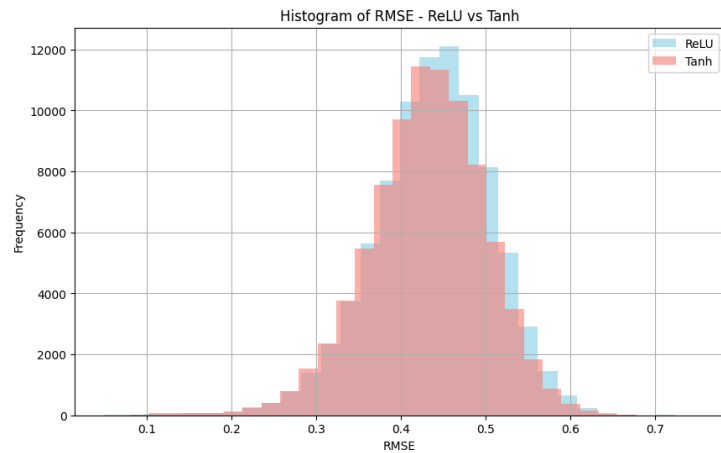


1 Word Prediction with Word2Vec and LSTM (50 points)

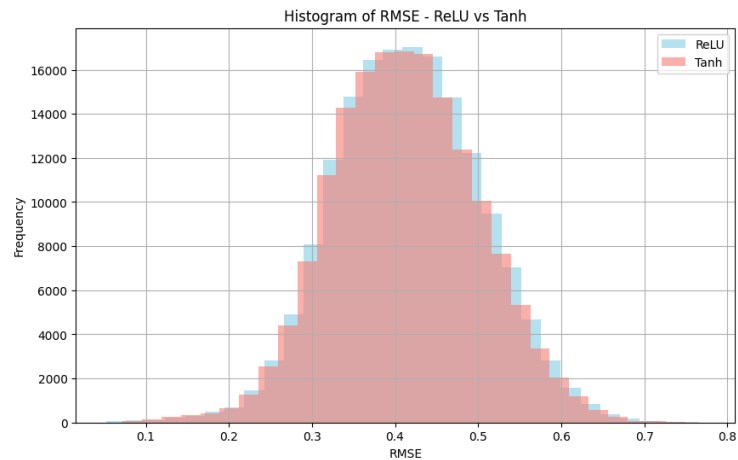
3.3) Mean RMSE Values (Stopwords Removed)

Histogram of RMSE When Removing Stopwords (ReLU vs Tanh) - [Image 1]



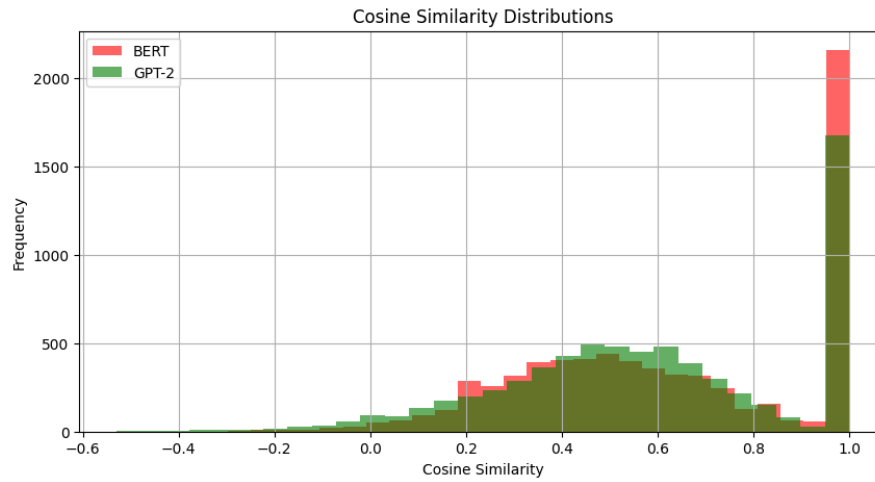
4) Not Removing Stopwords

Histogram of RMSE When Not Removing Stopwords (ReLU vs Tanh) - [Image 2]



2 Word Prediction with Transformers - (30 points)

Cosine Similarity Between BERT and GPT-2 - [Image 3]



3 Topic Modeling - Multinomial PMM Model and LDA(20 points)

Part B - [Image 4]

