Discussion

Dataset

- Model Training: : NLTK library (nltk.corpus.reuters)
- Word analogies dataset: https://www.kaggle.com/datasets/julianschelb/word-analogy-test
- similarity dataset: https://www.kaggle.com/datasets/julianschelb/wordsim353-crowd?resource=download

Results

Model	Window Size	Training Loss	Training Time	Syntactic Accuracy	Semantic Accuracy
Skipgram	2	13.188715	4456.45 seconds	0.0	0.0
Skipgram (NEG)	2	12.844395	113.68 seconds	0.0	0.0
Glove	2	8.381731	6627.40 seconds	0.0	0.0
Glove (Gensim)	2	-	-	0.5545	0.9387

Skip-gram and Skip-gram (NEG) showed similar training loss, but Skip-gram (NEG) trained much faster]compared to other models. GloVe achieved the lowest training loss (8.38) but required the longest training time (6627.40s).

Both Skip-gram and GloVe models achieved **0% syntactic and semantic accuracy**, indicating that the Reuters corpus size and domain were insufficient for analogy tasks. Pre-trained GloVe (Gensim) model achieved high syntactic (0.5545) and semantic (0.9387) accuracy, showing the benefit of large-scale training data.

Model	Spearman Correlation
Skipgram	0.1131

Skipgram (NEG)	0.1065
Glove	0.0993
Glove (Gensim)	0.4176

Spearman correlation values were low for all trained models (around 0.10), meaning weak correlation with human judgments. The GloVe (Gensim) model achieved a much higher correlation (0.4176), confirming its superior embedding quality.