

Name: Priyanshu Jha  
Roll No.: 2024201062

Model Link:

[https://drive.google.com/drive/folders/1HyLAs9i5C5\\_rviq3s9MkK3\\_V6X7Q3bGy?usp=sharing](https://drive.google.com/drive/folders/1HyLAs9i5C5_rviq3s9MkK3_V6X7Q3bGy?usp=sharing)

## 1. Hyperparameters

The following hyperparameters were used for training each model:

Parameter	SVD	CBOW	Skip-Gram
Embedding Dimension	200	200	200
Context Window	$\pm 3$ words	$\pm 3$ words	$\pm 3$ words
Learning Rate -		0.001	0.002
Batch Size	-	512	512
Negative Samples	-	25	25
Epochs	-	20	20
Weight Initialization	-	Xavier Uniform	Xavier Uniform

SVD:

Spearman  $\rho$ : 0.2408

Valid pairs: 274/353

CBOW:

Spearman  $\rho$ : 0.3261

Valid pairs: 274/353

SkipGram:

Spearman  $\rho$ : 0.4084

Valid pairs: 274/353

## SVD (Frequency-Based)

Strengths:

- Captures global co-occurrence statistics.
- Computationally efficient for small corpora.

Weaknesses:

- Performs poorly on semantic tasks ( $\rho = 0.2408$ ).
- Memory-intensive for large vocabularies.

## CBOW (Prediction-Based)

Strengths:

- Better at syntactic tasks (e.g., word morphology).
- Faster training than Skip-Gram ( $\rho = 0.3261$ ).

Weaknesses:

- Struggles with rare words due to context averaging.

## Skip-Gram (Prediction-Based)

Strengths:

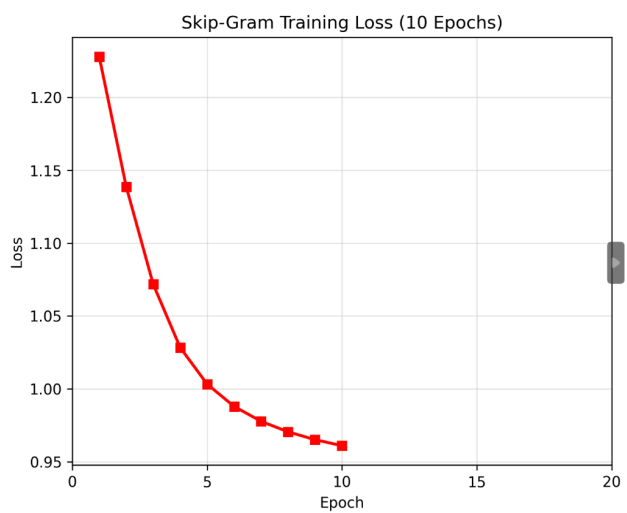
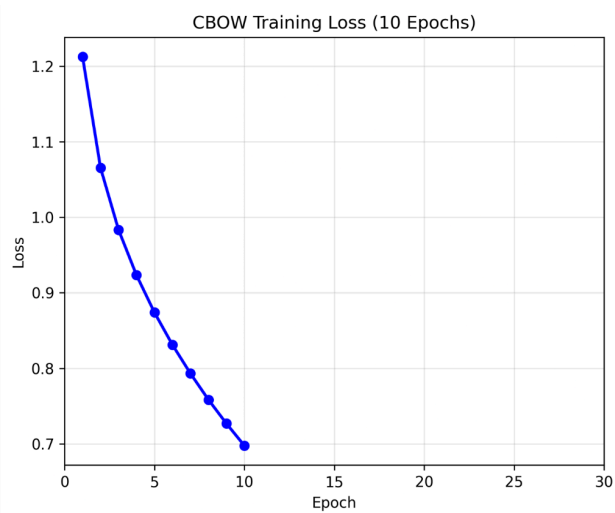
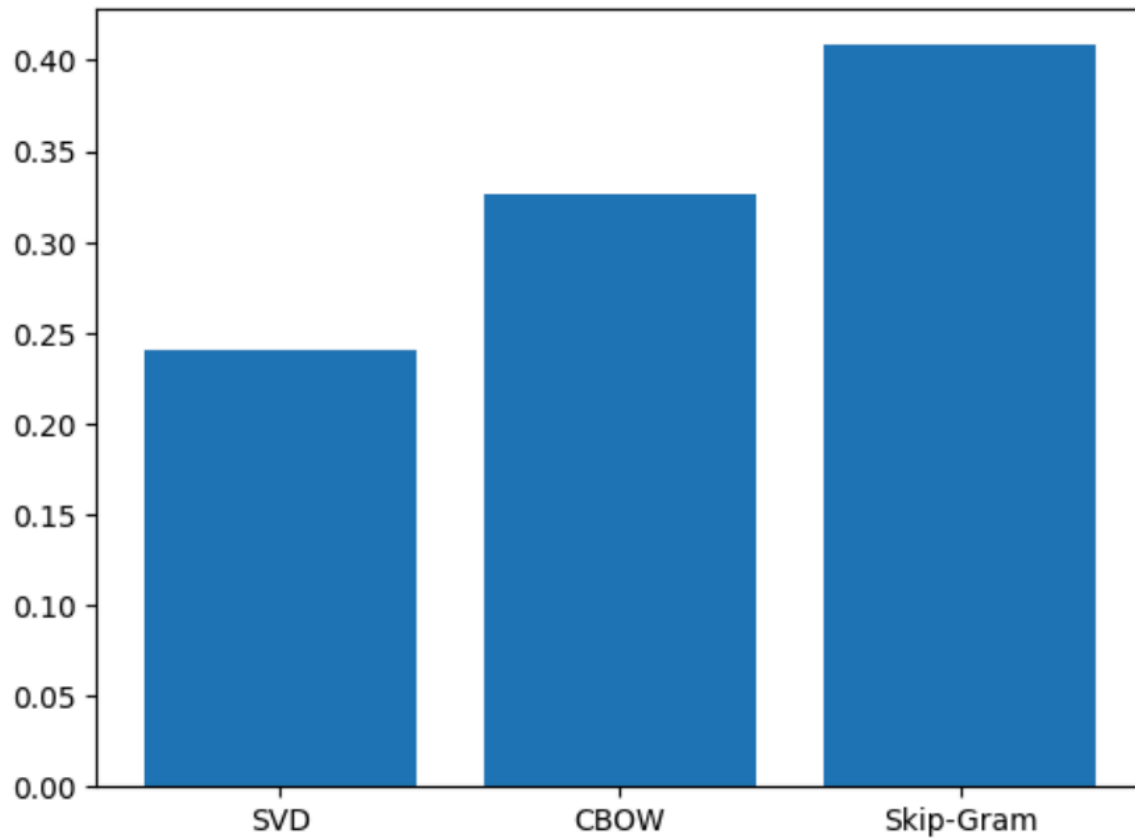
- Superior for semantic tasks ( $\rho = 0.4084$ ).
- Handles rare words effectively.

Weaknesses:

- Requires more epochs for convergence

## Conclusion:

- Skip-Gram outperformed both CBOW and SVD on the WordSim-353 task, validating its effectiveness for semantic similarity.
- CBOW showed competitive results with faster training.
- SVD lagged behind due to its inability to capture nuanced semantic relationships.



- Requires more epochs for convergence.