

# Advanced NLP Assgn 1

Raghav Saboo (2018101065)

---

## CBOW with full Softmax

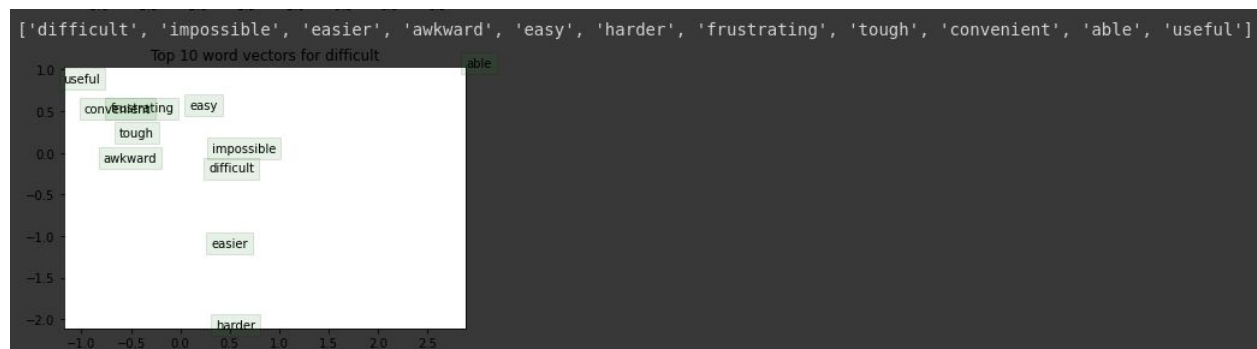
Parameters and Hyperparameters:

- Embedding Size - 50 Dimensions
- Window Size - 5
- Learning Rate - 0.03 (initially)
- Dataset - 50k reviews
- 2 epochs
- Full Softmax

**Display the top-10 word vectors for 5 different words (a combination of nouns, verbs, adjectives etc) using the above pre-trained model.**

```
Words = ["camera", "ipad", "difficult", "wall", "table"]
```







**What are the top 10 closest words for the word ‘camera’ in the embeddings generated by your program. Compare them against the pre-trained word2vec embeddings.**

```
['camera', 'bag', 'tripod', 'lens', 'lense', 'dslr', 'camcorder', 'film', 'sensor', 'slr', 'body']
```

**Total 11 words in which the first is the word itself and after that remaining 10 words**

Gensim (excluding the word “camera” itself) (in descending order)

**[‘cameras’, ‘cam’, ‘rebel’, ‘camcorder’, ‘digicam’, ‘lens’, ‘nikon’, ‘lense’, ‘canon’, ‘slr’]**

## Skip-gram with full Softmax

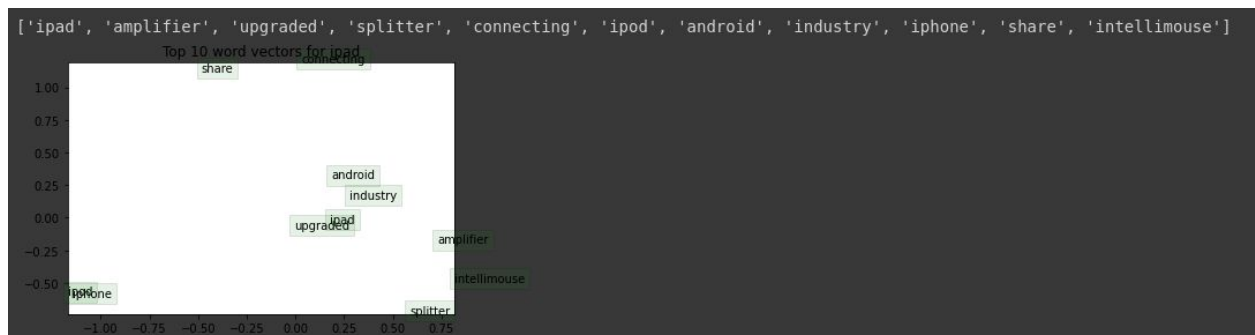
Parameters and Hyperparameters:

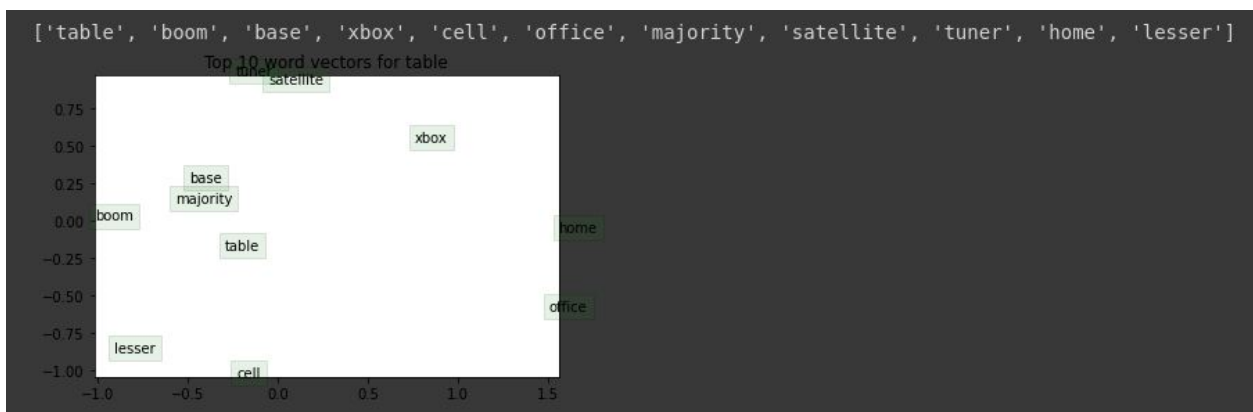
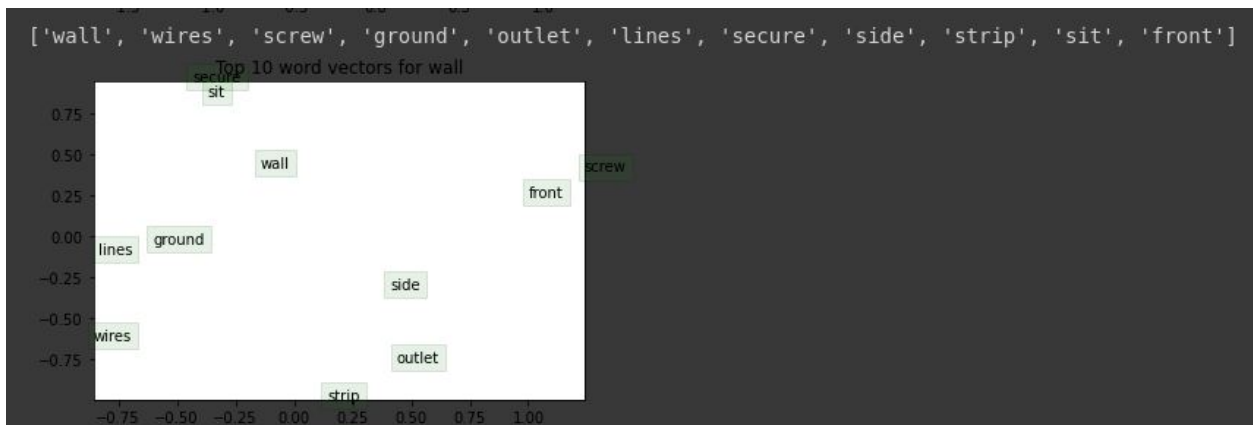
- Embedding Size - 50 Dimensions
- Window Size - 5

- Learning Rate - 0.03 (initially)
- Dataset - 50k reviews
- 2 epochs
- Full Softmax
- Negative sampling

**Display the top-10 word vectors for 5 different words (a combination of nouns, verbs, adjectives etc) using the above pre-trained model.**

```
Words = ["camera", "ipad", "difficult", "wall", "table"]
```





**What are the top 10 closest words for the word ‘camera’ in the embeddings generated by your program. Compare them against the pre-trained word2vec embeddings.**



---

Gensim (excluding the word “camera” itself) (in descending order)

**['cameras', 'camcorder', 'rebel', 'cam', 'slr', 'cannon', 'lens', 'monopod', 'filter', 'zoom']**

**Link for my Embeddings:**

**[https://drive.google.com/drive/folders/1\\_ouH9uPCd8oOI8lkIOoP1DgKL2iteM7c?usp=sharing](https://drive.google.com/drive/folders/1_ouH9uPCd8oOI8lkIOoP1DgKL2iteM7c?usp=sharing)**