

Assignment-1 (Skip-gram Implementation)

Instructions for assignment-1

1. Create a cluster of “animal words”: cow, dog, bullock, cat ... (10 words)
2. Create a cluster of “bird words”: cuckoo, parrot, crow, sparrow,... (10 words)
3. Run a concordance for obtaining the neighboring words of these words (learn what a **concordancer** is) [These animal words, bird words and concordance supplied syntagmatic words form the universe of your words]
4. Train a **skip gram model** with these words.
5. Collect the word representations.
6. Ensure that “animal” words are close to another words and so are “bird” words; Inter cluster distance should be large compared to intra cluster distance; use cosine similarity.

IMP: skip gram code has to be your OWN

Note: Final assignment teams will be formed randomly for those who are left out and given by EOD.