**

COMP814 – Text Mining

Skip-gram Lab

# Objective

1. To modify an existing Skip-gram model and run it with your own data.

# Task

1. You will need the sample code on skip-gram from the lab.
2. Your task is to write your own method to replace “*remove\_stop\_words*” to only consider all types of nouns and verbs when computing the vector space model.
3. Test it with the existing data and then extend the data set by adding other sentences and verify that the model gives the ”correct” output.
4. Write a method that outputs the nearest word for a given word. Use this method to output the nearest word for all given words in your dataset. (Hint: You can use Euclidean distance)