

# School of Computer Science Engineering and Technology

Course- B. Tech  
Course Code- CSET346

Year- 2022  
Date: 20-09-2022

Type- Elective  
Course Name: Natural language  
processing  
Semester- odd  
Batch- ALL

## Lab Assignment 05 – Text Mining and Classification

### CO-Mapping

Exp. No.	Name	CO1	CO2	CO3
06	Apply NLP to real life data	✓	✓	✓

**Objective:** The main motivation is to generate a review summary and develop recommendation system.

1. Collect the data from the following link

<https://github.com/subhamyadav580/Amazon-Reviews-Predictions-of-Cell-Phones-and-Accessories>

The data contains file which is a subset of amazon reviews from the cell phone and accessories.

1. Implement word2vec using genism library using the above data.
  - a. Train the model to find the similar word.
  - b. Find out the cosine similarity of all similar words.
2. Now you take headline of your mother tong newspaper convert with google API feed to your model and check the accuracy.
  - a. Apply the above two words steps for training your data.
  - a. Include UNK token in your corpus to deal with unknown words.
3. Create a recommendation system on the basis of reviews of the customers for the product.