

- 1) Libraries to be installed:
 - a)

```
import csv
import nltk
import re
from langdetect import detect
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')
nltk.download('wordnet')
from textblob.classifiers import NaiveBayesClassifier
from textblob import TextBlob
from nltk.corpus import wordnet as wn
from nltk import word_tokenize, pos_tag
from collections import defaultdict
from nltk.stem import WordNetLemmatizer
```
 - b) If using Command Prompt:
 - i) `python -m pip install langdetect`
 - c) If using Google Collab:
 - i) `!pip install langdetect`
- 2) Please place all the .csv files along with the code in the same folder
- 3) Run the code.py file
 - a) First Input the Dataset Name, that is one of the csv files:
 - i) sd1
 - ii) sd2
 - iii) sd1-1
 - iv) rd
 - b) input the name without-'.csv'
 - c) After that you will be asked to enter (Does the dataset have other language words:(Y/N))
 - i) Please enter : 'N'
 - (1) For the datasets sd1,sd2,sd1-1
 - (2) Because they are purely English word based
 - ii) rd.csv has some Spanish words too.
 - (1) So please enter : 'Y' in the field if you use rd.csv
- 4) DatasetNLP.csv is the main dataset used for Naive Bayes classifier for phase 1
- 5) Action_data.csv is the dataset of action verbs
- 6) Action_pairing.csv is the dataset having context sensitive pairs of verbs
- 7) Other csv's are for testing in the code,you may use any.