

Importing of libraries and vocabulary

Text

Description automatically generated

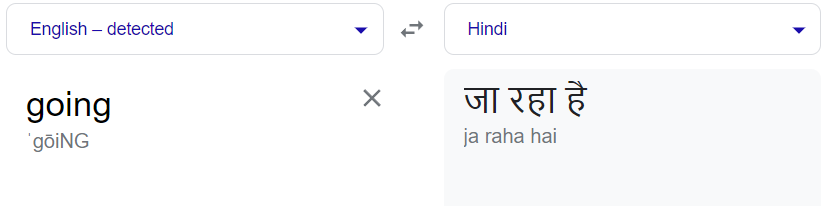
Translations of each word in the vocab to hindi using googleTrans

Text

Description automatically generated

Saving translated vocabulary in different file

**WE CREATED THE NEW HINDI VOCABULARY FROM THE EXISTING ENGLISH WORDS BUT GOOGLE TRANSLATIONS CREATED MULTIPLE WORDS FOR SINGLE ENGLISH WORD LIKE :**

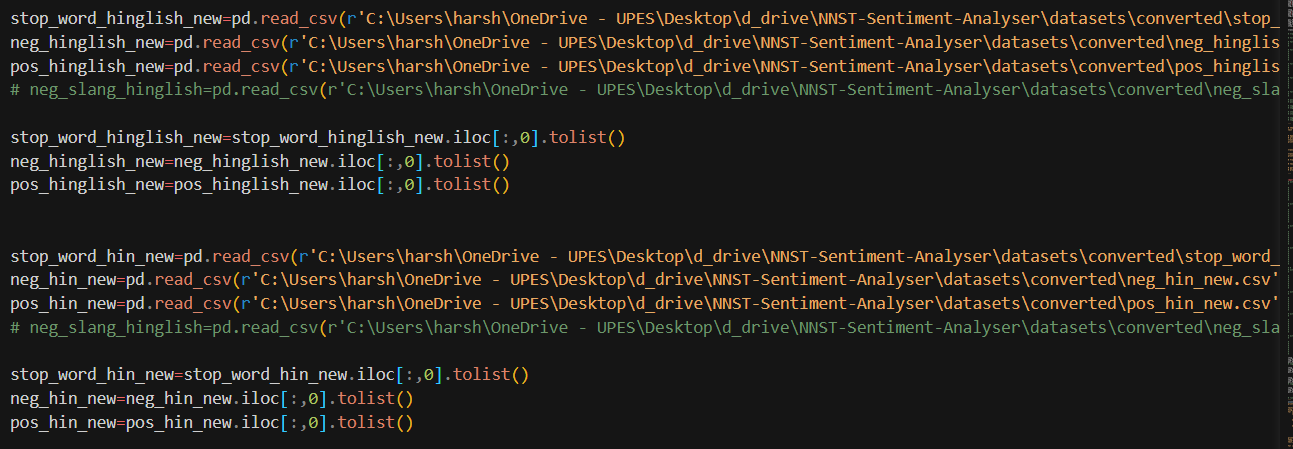
****

**So handle this :**

Text

Description automatically generated

New vocabulary has been created by adding new line to every translated sentence which also increased in improving the vocabulary dataset.



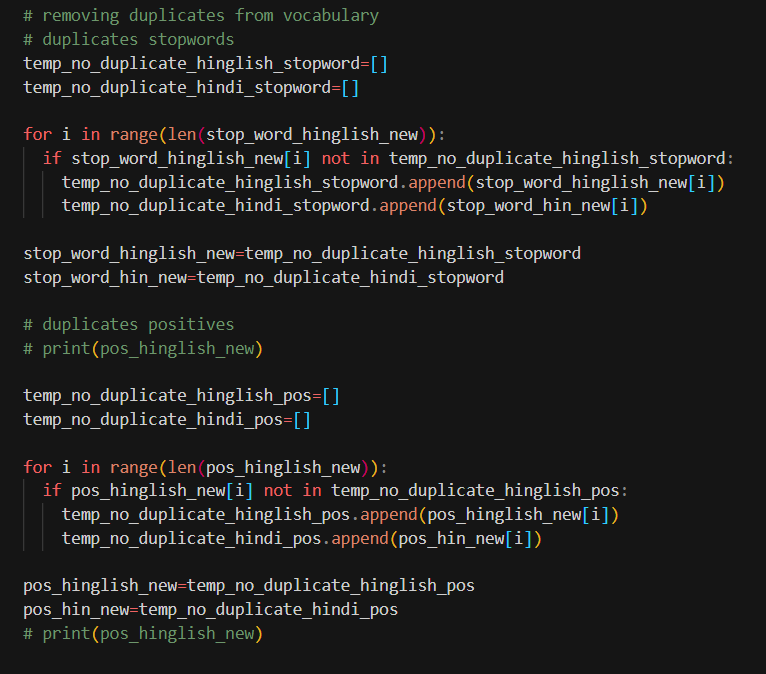
Now reading the newly created vocabulary with having one word per row

Text

Description automatically generated

Since translations could have multiple times same stopwords in them like the word

है would be very common and so are needed to be removed



Duplicates have been removed

Text

Description automatically generated

Saved to new list

Text

Description automatically generated

Importing Minor\_methods and generating adjacency list of all the possible erroneous words for hinglish dataset.