STOP WORD REMOVAL

Import nltk

nltk.download("stopwords")

from nltk.corpus import stopwords

Sentence = nltk.sent_tokenize (Paragraph)

For i in range (len (Sentence)):

words = nltk.word_tokenize (Sentences [i])

newwords = [word for word in words if not in stopwords.words ('
English')

Sentences [i] = ' '. join (newwords)

* This will be Used mainly in the text classification problems

POS (Parts of Speach Tagging)

→ nouns, Adjetive, Verbs ; Adverbs
  └ Propernans

Import nltk;

words = nltk. word_tokenize (Paragraph)

tagged_words = nltk.pos_tag (words)

→ Returns a list
  └ Each word will a tuple
  [('Thank', 'nnp')]

word_tags = [ ]

for tw in tagged_words:

word_tags.append (tw[0]+ "_" + tw[1])

tagged_paragraph = ' '.join (words_tags)

CC	Coordinating, Conjunction	(and, or, but)
CD	Cardinal numbers	(five, three, 13-)
DT	Determiner	(The, a, these)
EX	Existential there	(There were six boys)
FW	Foreign word	(mais)
IN	Preposition or subordinating, Conjunction (of, on before unless)	
JJ	Adjective (nice, Easy)	
JJR	Adjective, Comparative. (nicer, easier)	
JJS	Adjective, superlative (nicest, easiest)	
LS	List item marker	
MD	Modal (may, should)	
NN	Noun, singular or mass (tiger, chairs, laughter)	
NNS	Noun, plural (tiger, chairs, insects)	
NNP	Proper noun, singular (Germany, God, Alice)	
NNPS	Proper noun, plural (we met two christmases ago)	
PDT	Predeterminer (both his children)	
POS	Possesive Ending ('s)	
PRP	Personal, pronoun (me, you, it)	
PRP$	Possesive pronoun (me, your, own)	
RB	Adverb (Extremely, loudy, hard)	
RBR	Adverb, Comparative (better)	
RBS	Adverb, superlative (best)	
RP	adverb, article (about, off, up)	
SYM	Symbol (%.)	
TO	infinitival to (what to do?)	
UH	Interjection (oh, oops, gosh)	
VB	Verb, bare form (think)	
VBD	Verb, past tense	
VBG	Verb, present participle	

VBN - Verb, past participle

VBP: Verb, non 3rd person singular present

VBZ: Verb, 3rd person singular present

WDT wh-determiner

WD wh-pronoun

WP$ Possesive wh-pronoun

WRB wh-adverb