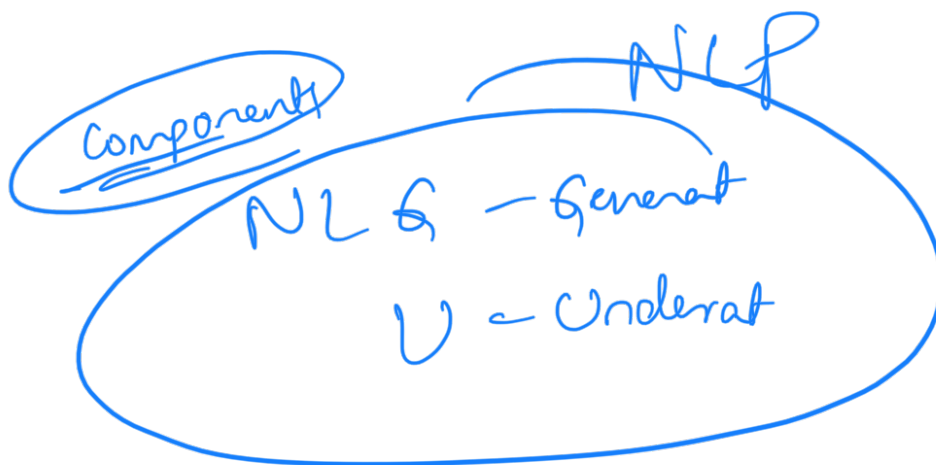


NLP

Spacy



components in NLP

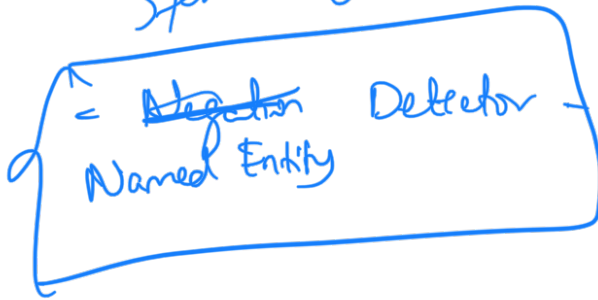
- Doc preprocessor - (cleaning
reger)

- Syntactic analysis - (parsing)

Tokenize
(words)

data

Dictionary - lemmatization part of sp J →
 - Stemming
 break text into small comp



NED
 are with
 (Noun)
 - org
 common

- Dependency parser

- Co-reference parser (- connect 1 of word to another)
 excited → emotion



Adv of NLP

- app any domain

- response in seconds

chatbot

- provides necessary info

Ex: Siri

- compute in Human language

~~Ex: Siri~~ May not good in emergency just

Lexical

break text into words

discovery → dis
→ hovers

Syntactical

meaning
grammar.

Semantic

meaning of sentence

iced coffe X
cold coffe →

pragmatic

what was said

~~①~~

Tokeniser

break big sentence to small words

vector encoding

~~②~~

Stemming

buying - buy

extract word

do leaves → leaf
leaf →

removing

prefix
& suffix

such word.

lemmatization

real meaning

leaves

leaf

③

POS tags

Part of Speech

Coding

- import NLTK library

word

- tokenize ~~word~~ ^{text}

(all words
include ',')
in sentence

- probability - Frequency Distribution

word - freq.

→ Top 10 freq. words

~~unigram~~ bigram, trigram, n-gram