task_1

```
import nltk
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger_eng')
nltk.download('punkt_tab')
sentence = "Students are studying Natural Language Processing for the upcoming Semester examinations"
tokens = nltk.word_tokenize(sentence)
pos_tags = nltk.pos_tag(tokens)
print(pos_tags)
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package averaged_perceptron_tagger_eng to
[nltk_data]     /root/nltk_data...
[nltk_data]   Unzipping taggers/averaged_perceptron_tagger_eng.zip.
[nltk_data] Downloading package punkt_tab to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt_tab.zip.
[('Students', 'NNS'), ('are', 'VBP'), ('studying', 'VBG'), ('Natural', 'NNP'), ('Language', 'NNP'), ('Processing', 'NNP'), (
```

```
import spacy
nlp = spacy.load("en_core_web_sm")

doc = nlp("Students are studying Natural Language Processing for the upcoming Semester examinations")
for token in doc:
    print(token.text, token.pos_)
```

```
Students NOUN
are AUX
studying VERB
Natural PROPN
Language PROPN
Processing NOUN
for ADP
the DET
upcoming ADJ
Semester PROPN
examinations NOUN
```

```
import spacy

nlp = spacy.load("en_core_web_sm")
doc = nlp("Apple is looking at buying a startup in India.")

for token in doc:
    print(token.text, token.pos_, token.tag_)
```

```
Apple PROPN NNP
is AUX VBZ
looking VERB VBG
at ADP IN
buying VERB VBG
a DET DT
startup NOUN NN
in ADP IN
India PROPN NNP
. PUNCT .
```

```
import spacy
from collections import Counter

nlp = spacy.load("en_core_web_sm")

text = "The new features of the moble are awesome!, after updating 😍 #AI #MachineLearning"
doc = nlp(text)

nouns = []
verbs = []

for token in doc:
    if token.pos_ in ["NOUN", "PROPN"]:
        nouns.append(token.text)
    elif token.pos_ == "VERB":
        verbs.append(token.text)

noun_freq = Counter(nouns)
verb_freq = Counter(verbs)
```

```
print("Noun Frequency:", noun_freq)
print("Verb Frequency:", verb_freq)
```

```
Noun Frequency: Counter({'features': 1, 'moble': 1, '😍': 1, 'AI': 1, 'MachineLearning': 1})
Verb Frequency: Counter({'updating': 1})
```

TASK_2

```
import nltk
SRUniversity="""The SR University campus is located in Ananthasagar village of Hasanparthy Mandal in Warangal, Telangana, I
It is in 150 acres, with both separate hostel facilities for boys and girls.
There is a huge central library along with Indias largest Technology Business Incubator (TBI) in tier 2 cities."""

pos_tags=nltk.pos_tag(nltk.word_tokenize(SRUniversity))
print(pos_tags)
```

```
[('The', 'DT'), ('SR', 'NNP'), ('University', 'NNP'), ('campus', 'NN'), ('is', 'VBZ'), ('located', 'VBN'), ('in', 'IN'), ('A
```

```
import spacy
nlp = spacy.load("en_core_web_sm")

doc = nlp("The SR University campus is located in Ananthasagar village of Hasanparthy Mandal in Warangal, Telangana, India.
for token in doc:
    print(token.text, token.pos_)
```

```
The DET
SR PROPN
University PROPN
campus NOUN
is AUX
located VERB
in ADP
Ananthasagar PROPN
village NOUN
of ADP
Hasanparthy PROPN
Mandal PROPN
in ADP
Warangal PROPN
, PUNCT
Telangana PROPN
, PUNCT
India PROPN
. PUNCT
It PRON
is AUX
in ADP
150 NUM
acres NOUN
, PUNCT
with ADP
both DET
separate ADJ
hostel NOUN
facilities NOUN
for ADP
boys NOUN
and CCONJ
girls NOUN
. PUNCT
There PRON
is VERB
a DET
huge ADJ
central ADJ
library NOUN
along ADP
with ADP
Indias PROPN
largest ADJ
Technology PROPN
Business PROPN
Incubator PROPN
( PUNCT
TBI PROPN
) PUNCT
in ADP
tier NOUN
2 NUM
cities NOUN
. PUNCT
```

```
nltk.download('averaged_perceptron_tagger')
```

```
nltk.download('averaged_perceptron_tagger')
from nltk.tokenize import word_tokenize
words = word_tokenize(SRUniversity)
nltk.pos_tag(words)
```

```
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     /root/nltk_data...
[nltk_data]   Unzipping taggers/averaged_perceptron_tagger.zip.
[('The', 'DT'),
 ('SR', 'NNP'),
 ('University', 'NNP'),
 ('campus', 'NN'),
 ('is', 'VBZ'),
 ('located', 'VBN'),
 ('in', 'IN'),
 ('Ananthasagar', 'NNP'),
 ('village', 'NN'),
 ('of', 'IN'),
 ('Hasanparthy', 'NNP'),
 ('Mandal', 'NNP'),
 ('in', 'IN'),
 ('Warangal', 'NNP'),
 (',', ','),
 ('Telangana', 'NNP'),
 (',', ','),
 ('India', 'NNP'),
 ('.', '.'),
 ('It', 'PRP'),
 ('is', 'VBZ'),
 ('in', 'IN'),
 ('150', 'CD'),
 ('acres', 'NNS'),
 (',', ','),
 ('with', 'IN'),
 ('both', 'DT'),
 ('separate', 'JJ'),
 ('hostel', 'NN'),
 ('facilities', 'NNS'),
 ('for', 'IN'),
 ('boys', 'NNS'),
 ('and', 'CC'),
 ('girls', 'NNS'),
 ('.', '.'),
 ('There', 'EX'),
 ('is', 'VBZ'),
 ('a', 'DT'),
 ('huge', 'JJ'),
 ('central', 'JJ'),
 ('library', 'NN'),
 ('along', 'IN'),
 ('with', 'IN'),
 ('Indias', 'NNP'),
 ('largest', 'JJS'),
 ('Technology', 'NN'),
 ('Business', 'NNP'),
 ('Incubator', 'NNP'),
 ('(', '('),
 ('TBI', 'NNP'),
 (')', ')'),
 ('in', 'IN'),
 ('tier', '$'),
 ('2', 'CD'),
 ('cities', 'NNS'),
```

```
nltk.download('tagsets')
nltk.download('tagsets_json')
nltk.help.upenn_tagset()
```

```
    man baby diddle hush sonuvabitch ...
VB: verb, base form
    ask assemble assess assign assume atone attention avoid bake balkanize
    bank begin behold believe bend benefit bevel beware bless boil bomb
    boost brace break bring broil brush build ...
VBD: verb, past tense
    dipped pleaded swiped regummed soaked tidied convened halted registered
    cushioned exacted snubbed strode aimed adopted belied figgered
    speculated wore appreciated contemplated ...
VBG: verb, present participle or gerund
    telegraphing stirring focusing angering judging stalling lactating
    hankerin' alleging veering capping approaching traveling besieging
    encrypting interrupting erasing wincing ...
VBN: verb, past participle
    multihulled dilapidated aerosolized chaired languished panelized used
    experimented flourished imitated reunifed factored condensed sheared
    unsettled primed dubbed desired ...
VBP: verb, present tense, not 3rd person singular
    predominate wrap resort sue twist spill cure lengthen brush terminate
    appear tend stray glisten obtain comprise detest tease attract
    emphasize mold postpone sever return wag ...
VBZ: verb, present tense, 3rd person singular
    bases reconstructs marks mixes displeases seals carps weaves snatches
    slumps stretches authorizes smolders pictures emerges stockpiles
    seduces fizzes uses bolsters slaps speaks pleads ...
WDT: WH-determiner
    that what whatever which whichever
WP: WH-pronoun
    that what whatever whatsoever which who whom whosoever
WP$: WH-pronoun, possessive
    whose
WRB: Wh-adverb
    how however whence whenever where whereby whereever wherein whereof why
``: opening quotation mark
    ` ``
[nltk_data]   Unzipping help/tagsets_json.zip.
```

TASK_3 canvas

```
!pip install nltk spacy
!python -m spacy download en_core_web_sm
```

```
Requirement already satisfied: nltk in /usr/local/lib/python3.12/dist-packages (3.9.1)
Requirement already satisfied: spacy in /usr/local/lib/python3.12/dist-packages (3.8.11)
Requirement already satisfied: click in /usr/local/lib/python3.12/dist-packages (from nltk) (8.3.1)
Requirement already satisfied: joblib in /usr/local/lib/python3.12/dist-packages (from nltk) (1.5.3)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.12/dist-packages (from nltk) (2025.11.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.12/dist-packages (from nltk) (4.67.1)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.12/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (1.0.15)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.0.13)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.12/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: thinc<8.4.0,>=8.3.4 in /usr/local/lib/python3.12/dist-packages (from spacy) (8.3.10)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.12/dist-packages (from spacy) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.5.2)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.4.2 in /usr/local/lib/python3.12/dist-packages (from spacy) (0.4.3)
Requirement already satisfied: typer-slim<1.0.0,>=0.3.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (0.21.1)
Requirement already satisfied: numpy>=1.19.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.0.2)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.32.4)
Requirement already satisfied: pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.12/dist-packages (from spacy)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from spacy) (3.1.6)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from spacy) (75.2.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (25.0)
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!=1.8.1
Requirement already satisfied: pydantic-core==2.41.4 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!=1.8.1,
Requirement already satisfied: typing-extensions>=4.14.1 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!=1.
Requirement already satisfied: typing-inspection>=0.4.2 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!=1.8
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spacy)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->
Requirement already satisfied: blis<1.4.0,>=1.3.0 in /usr/local/lib/python3.12/dist-packages (from thinc<8.4.0,>=8.3.4->spac
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.12/dist-packages (from thinc<8.4.0,>=8.3.4
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/python3.12/dist-packages (from weasel<0.5.0,>=0.
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/python3.12/dist-packages (from weasel<0.5.0,>=0.4.
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->spacy) (3.0.3)
Requirement already satisfied: wrapt in /usr/local/lib/python3.12/dist-packages (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0
Collecting en-core-web-sm==3.8.0
  Downloading https://github.com/explosion/spacy-models/releases/download/en_core_web_sm-3.8.0/en_core_web_sm-3.8.0-py3-none
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 12.8/12.8 MB 67.8 MB/s eta 0:00:00
✓ Download and installation successful
You can now load the package via spacy.load('en_core_web_sm')
⚠ Restart to reload dependencies
If you are in a Jupyter or Colab notebook, you may need to restart Python in
order to load all the package's dependencies. You can do this by selecting the
```

```
'Restart kernel' or 'Restart runtime' option.
```

```
import nltk
import spacy
from nltk.tokenize import TweetTokenizer
from collections import Counter

nltk.download('averaged_perceptron_tagger')
```

```
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     /root/nltk_data...
[nltk_data]   Package averaged_perceptron_tagger is already up-to-
[nltk_data]       date!
True
```

```
from google.colab import files
files.download("/content/json-to-text-converter.txt")
```

```
import nltk
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger_eng')
sentence = "Students are learning Natural Language Processing"
tokens = nltk.word_tokenize(sentence)
pos_tags = nltk.pos_tag(tokens)
print(pos_tags)
```

```
[('Students', 'NNS'), ('are', 'VBP'), ('learning', 'VBG'), ('Natural', 'NNP'), ('Language', 'NNP'), ('Processing', 'NNP')]
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger_eng to
[nltk_data]     /root/nltk_data...
[nltk_data]   Package averaged_perceptron_tagger_eng is already up-to-
[nltk_data]       date!
```

```
file_name = file_path = "/content/json-to-text-converter.txt"

with open(file_name, 'r', encoding='utf-8') as f:
    text = f.read()

print("Input Text:\n")
print(text)
```

```
anguage. It combines computational linguistics—rule-based modeling of human language—with statistical, machine learning, and
```

```
from nltk.tokenize import word_tokenize

tokens = word_tokenize(text)

nltk_pos = nltk.pos_tag(tokens)

print("NLTK POS Tagging:\n")
nltk_pos
```

```
NLTK POS Tagging:

[('Natural', 'JJ'),
 ('Language', 'NNP'),
 ('Processing', 'NNP'),
 ('(', '('),
 ('NLP', 'NNP'),
 (')', ')'),
 ('is', 'VBZ'),
 ('a', 'DT'),
 ('field', 'NN'),
 ('of', 'IN'),
 ('artificial', 'JJ'),
 ('intelligence', 'NN'),
 ('that', 'WDT'),
 ('enables', 'VBZ'),
 ('computers', 'NNS'),
 ('to', 'TO'),
 ('understand', 'VB'),
 (',', ','),
 ('interpret', 'VB'),
 (',', ','),
```

```
 ('and', 'CC'),
 ('generate', 'VB'),
 ('human', 'JJ'),
 ('language', 'NN'),
 ('.', '.'),
 ('It', 'PRP'),
 ('combines', 'VBZ'),
 ('computational', 'JJ'),
 ('linguistics—rule-based', 'JJ'),
 ('modeling', 'NN'),
 ('of', 'IN'),
 ('human', 'JJ'),
 ('language—with', 'JJ'),
 ('statistical', 'JJ'),
 (',', ','),
 ('machine', 'NN'),
 ('learning', 'NN'),
 (',', ','),
 ('and', 'CC'),
 ('deep', 'JJ'),
 ('learning', 'NN'),
 ('models', 'NNS'),
 ('.', '.')]
```

```
nlp = spacy.load("en_core_web_sm")

doc = nlp(text)

spacy_pos = [(token.text, token.pos_) for token in doc]

print("spaCy POS Tagging:\n")
spacy_pos
```

```
spaCy POS Tagging:

[('Natural', 'PROPN'),
 ('Language', 'PROPN'),
 ('Processing', 'PROPN'),
 ('(', 'PUNCT'),
 ('NLP', 'PROPN'),
 (')', 'PUNCT'),
 ('is', 'AUX'),
 ('a', 'DET'),
 ('field', 'NOUN'),
 ('of', 'ADP'),
 ('artificial', 'ADJ'),
 ('intelligence', 'NOUN'),
 ('that', 'PRON'),
 ('enables', 'VERB'),
 ('computers', 'NOUN'),
 ('to', 'PART'),
 ('understand', 'VERB'),
 (',', 'PUNCT'),
 ('interpret', 'VERB'),
 (',', 'PUNCT'),
 ('and', 'CCONJ'),
 ('generate', 'VERB'),
 ('human', 'ADJ'),
 ('language', 'NOUN'),
 ('.', 'PUNCT'),
 ('It', 'PRON'),
 ('combines', 'VERB'),
 ('computational', 'ADJ'),
 ('linguistics', 'NOUN'),
 ('—', 'PUNCT'),
 ('rule', 'NOUN'),
 ('-', 'PUNCT'),
 ('based', 'VERB'),
 ('modeling', 'NOUN'),
 ('of', 'ADP'),
 ('human', 'ADJ'),
 ('language', 'NOUN'),
 ('—', 'PUNCT'),
 ('with', 'ADP'),
 ('statistical', 'ADJ'),
 (',', 'PUNCT'),
 ('machine', 'NOUN'),
 ('learning', 'NOUN'),
 (',', 'PUNCT'),
 ('and', 'CCONJ'),
 ('deep', 'ADJ'),
 ('learning', 'NOUN'),
 ('models', 'NOUN'),
 ('.', 'PUNCT'),
 ('\n', 'SPACE')]
```