```
!pip install nltk spacy
```

```
Requirement already satisfied: nltk in /usr/local/lib/python3.12/dist-package
Requirement already satisfied: spacy in /usr/local/lib/python3.12/dist-packag
Requirement already satisfied: click in /usr/local/lib/python3.12/dist-packag
Requirement already satisfied: joblib in /usr/local/lib/python3.12/dist-packa
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.12/d
Requirement already satisfied: tqdm in /usr/local/lib/python3.12/dist-package
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/py
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python
Requirement already satisfied: thinc<8.4.0,>=8.3.4 in /usr/local/lib/python3.
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/pyth
Requirement already satisfied: weasel<0.5.0,>=0.4.2 in /usr/local/lib/python3
Requirement already satisfied: typer-slim<1.0.0,>=0.3.0 in /usr/local/lib/pyt
Requirement already satisfied: numpy>=1.19.0 in /usr/local/lib/python3.12/dis
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/pyth
Requirement already satisfied: pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4 in /usr/l
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packa
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-p
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/d
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/pytho
Requirement already satisfied: pydantic-core==2.41.4 in /usr/local/lib/python
Requirement already satisfied: typing-extensions>=4.14.1 in /usr/local/lib/py
Requirement already satisfied: typing-inspection>=0.4.2 in /usr/local/lib/pyt
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/pyt
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.1
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.1
Requirement already satisfied: blis<1.4.0,>=1.3.0 in /usr/local/lib/python3.1
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/pyt
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/p
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/pyt
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/d
Requirement already satisfied: wrap in /usr/local/lib/python3.12/dist-packag
```

```
import nltk
import spacy
```

```
with open('nlp_text_dataset.csv', 'r', encoding='utf-8') as file:
    text = file.read()

print(text)
```

```
id,text,label
1,The weather is sunny and perfect for a picnic.,weather
2,Government announced a new policy for renewable energy.,news
```

```
3,I love learning Natural Language Processing.,education
4,The movie was boring and too long.,reviews
5,AI models are transforming healthcare systems.,technology
6,Heavy rains caused traffic jams across the city.,weather
7,The new mall opened with many international brands.,news
8,Students participated in a coding hackathon.,education
9,The restaurant had delicious food but slow service.,reviews
10,Researchers introduced a new deep learning architecture.,technology
11,Cloudy skies are expected tomorrow.,weather
12,Parliament discussed the national budget today.,news
13,Teachers are introducing online learning tools.,education
14,The hotel room was clean and comfortable.,reviews
15,Robotics is changing manufacturing industries.,technology
16,Strong winds damaged several trees.,weather
17,Local elections are scheduled next month.,news
18,Students submitted assignments through the portal.,education
19,The mobile app had too many bugs.,reviews
20,Blockchain applications are increasing rapidly.,technology
21,It rained heavily last night.,weather
22,The city hosted a cultural festival.,news
23,The professor explained transformers clearly.,education
24,The laptop battery drains very quickly.,reviews
25,Self-driving cars are being tested worldwide.,technology
26,The temperature dropped significantly this week.,weather
27,A new hospital was inaugurated yesterday.,news
28,Students prepared presentations on AI ethics.,education
29,The headphones broke after one week.,reviews
30,Cybersecurity attacks are becoming common.,technology
31,Morning fog reduced visibility on highways.,weather
32,Scientists discovered a new species of bird.,news
33,The online course helped me understand NLP basics.,education
34,The service at the bank was very slow.,reviews
35,Smartwatches can monitor heart rate and steps.,technology
36,The heat wave affected many regions.,weather
37,International summit focused on climate change.,news
38,Students learned tokenization and stemming.,education
39,The phone camera quality is excellent.,reviews
40,Quantum computing may solve complex problems.,technology
41,Snowfall covered the mountains overnight.,weather
42,The mayor launched a city-cleaning mission.,news
43,The workshop explained text classification.,education
44,The smartwatch screen cracked easily.,reviews
45,Startups are building AI-powered assistants.,technology
46,Humidity levels were high throughout the day.,weather
47,A new airport terminal became operational.,news
48,Students practiced summarizing long documents.,education
49,The coffee shop had great ambiance but was costly.,reviews
50,Edge computing reduces latency in smart devices.,technology
```

```python
words = text.split()
word_count = len(words)
print("Number of words in dataset:", word_count)
```

```
Number of words in dataset: 314
```

```
lowercase_text = text.lower()
print(lowercase_text)
```

```
id,text,label
1,the weather is sunny and perfect for a picnic.,weather
2,government announced a new policy for renewable energy.,news
3,i love learning natural language processing.,education
4,the movie was boring and too long.,reviews
5,ai models are transforming healthcare systems.,technology
6,heavy rains caused traffic jams across the city.,weather
7,the new mall opened with many international brands.,news
8,students participated in a coding hackathon.,education
9,the restaurant had delicious food but slow service.,reviews
10,researchers introduced a new deep learning architecture.,technology
11,cloudy skies are expected tomorrow.,weather
12,parliament discussed the national budget today.,news
13,teachers are introducing online learning tools.,education
14,the hotel room was clean and comfortable.,reviews
15,robotics is changing manufacturing industries.,technology
16,strong winds damaged several trees.,weather
17,local elections are scheduled next month.,news
18,students submitted assignments through the portal.,education
19,the mobile app had too many bugs.,reviews
20,blockchain applications are increasing rapidly.,technology
21,it rained heavily last night.,weather
22,the city hosted a cultural festival.,news
23,the professor explained transformers clearly.,education
24,the laptop battery drains very quickly.,reviews
25,self-driving cars are being tested worldwide.,technology
26,the temperature dropped significantly this week.,weather
27,a new hospital was inaugurated yesterday.,news
28,students prepared presentations on ai ethics.,education
29,the headphones broke after one week.,reviews
30,cybersecurity attacks are becoming common.,technology
31,morning fog reduced visibility on highways.,weather
32,scientists discovered a new species of bird.,news
33,the online course helped me understand nlp basics.,education
34,the service at the bank was very slow.,reviews
35,smartwatches can monitor heart rate and steps.,technology
36,the heat wave affected many regions.,weather
37,international summit focused on climate change.,news
38,students learned tokenization and stemming.,education
39,the phone camera quality is excellent.,reviews
40,quantum computing may solve complex problems.,technology
41,snowfall covered the mountains overnight.,weather
42,the mayor launched a city-cleaning mission.,news
43,the workshop explained text classification.,education
44,the smartwatch screen cracked easily.,reviews
45,startups are building ai-powered assistants.,technology
46,humidity levels were high throughout the day.,weather
47,a new airport terminal became operational.,news
48,students practiced summarizing long documents.,education
49,the coffee shop had great ambiance but was costly.,reviews
50,edge computing reduces latency in smart devices.,technology
```

```
uppercase_text = text.upper()
print(uppercase_text)
```

```
ID,TEXT,LABEL
1,THE WEATHER IS SUNNY AND PERFECT FOR A PICNIC.,WEATHER
2,GOVERNMENT ANNOUNCED A NEW POLICY FOR RENEWABLE ENERGY.,NEWS
3,I LOVE LEARNING NATURAL LANGUAGE PROCESSING.,EDUCATION
4,THE MOVIE WAS BORING AND TOO LONG.,REVIEWS
5,AI MODELS ARE TRANSFORMING HEALTHCARE SYSTEMS.,TECHNOLOGY
6,HEAVY RAINS CAUSED TRAFFIC JAMS ACROSS THE CITY.,WEATHER
7,THE NEW MALL OPENED WITH MANY INTERNATIONAL BRANDS.,NEWS
8,STUDENTS PARTICIPATED IN A CODING HACKATHON.,EDUCATION
9,THE RESTAURANT HAD DELICIOUS FOOD BUT SLOW SERVICE.,REVIEWS
10,RESEARCHERS INTRODUCED A NEW DEEP LEARNING ARCHITECTURE.,TECHNOLOGY
11,CLOUDY SKIES ARE EXPECTED TOMORROW.,WEATHER
12,PARLIAMENT DISCUSSED THE NATIONAL BUDGET TODAY.,NEWS
13,TEACHERS ARE INTRODUCING ONLINE LEARNING TOOLS.,EDUCATION
14,THE HOTEL ROOM WAS CLEAN AND COMFORTABLE.,REVIEWS
15,ROBOTICS IS CHANGING MANUFACTURING INDUSTRIES.,TECHNOLOGY
16,STRONG WINDS DAMAGED SEVERAL TREES.,WEATHER
17,LOCAL ELECTIONS ARE SCHEDULED NEXT MONTH.,NEWS
18,STUDENTS SUBMITTED ASSIGNMENTS THROUGH THE PORTAL.,EDUCATION
19,THE MOBILE APP HAD TOO MANY BUGS.,REVIEWS
20,BLOCKCHAIN APPLICATIONS ARE INCREASING RAPIDLY.,TECHNOLOGY
21,IT RAINED HEAVILY LAST NIGHT.,WEATHER
22,THE CITY HOSTED A CULTURAL FESTIVAL.,NEWS
23,THE PROFESSOR EXPLAINED TRANSFORMERS CLEARLY.,EDUCATION
24,THE LAPTOP BATTERY DRAINS VERY QUICKLY.,REVIEWS
25,SELF-DRIVING CARS ARE BEING TESTED WORLDWIDE.,TECHNOLOGY
26,THE TEMPERATURE DROPPED SIGNIFICANTLY THIS WEEK.,WEATHER
27,A NEW HOSPITAL WAS INAUGURATED YESTERDAY.,NEWS
28,STUDENTS PREPARED PRESENTATIONS ON AI ETHICS.,EDUCATION
29,THE HEADPHONES BROKE AFTER ONE WEEK.,REVIEWS
30,CYBERSECURITY ATTACKS ARE BECOMING COMMON.,TECHNOLOGY
31,MORNING FOG REDUCED VISIBILITY ON HIGHWAYS.,WEATHER
32,SCIENTISTS DISCOVERED A NEW SPECIES OF BIRD.,NEWS
33,THE ONLINE COURSE HELPED ME UNDERSTAND NLP BASICS.,EDUCATION
34,THE SERVICE AT THE BANK WAS VERY SLOW.,REVIEWS
35,SMARTWATCHES CAN MONITOR HEART RATE AND STEPS.,TECHNOLOGY
36,THE HEAT WAVE AFFECTED MANY REGIONS.,WEATHER
37,INTERNATIONAL SUMMIT FOCUSED ON CLIMATE CHANGE.,NEWS
38,STUDENTS LEARNED TOKENIZATION AND STEMMING.,EDUCATION
39,THE PHONE CAMERA QUALITY IS EXCELLENT.,REVIEWS
40,QUANTUM COMPUTING MAY SOLVE COMPLEX PROBLEMS.,TECHNOLOGY
41,SNOWFALL COVERED THE MOUNTAINS OVERNIGHT.,WEATHER
42,THE MAYOR LAUNCHED A CITY-CLEANING MISSION.,NEWS
43,THE WORKSHOP EXPLAINED TEXT CLASSIFICATION.,EDUCATION
44,THE SMARTWATCH SCREEN CRACKED EASILY.,REVIEWS
45,STARTUPS ARE BUILDING AI-POWERED ASSISTANTS.,TECHNOLOGY
46,HUMIDITY LEVELS WERE HIGH THROUGHOUT THE DAY.,WEATHER
47,A NEW AIRPORT TERMINAL BECAME OPERATIONAL.,NEWS
48,STUDENTS PRACTICED SUMMARIZING LONG DOCUMENTS.,EDUCATION
49,THE COFFEE SHOP HAD GREAT AMBIANCE BUT WAS COSTLY.,REVIEWS
50,EDGE COMPUTING REDUCES LATENCY IN SMART DEVICES.,TECHNOLOGY
```

```python
from nltk.tokenize import word_tokenize
nltk.download('punkt')
nltk.download('punkt_tab')
```

```
tokens = word_tokenize(text)
print("Tokens:", tokens)
```

```
, 'wave', 'affected', 'many', 'regions.', ',', 'weather', '37', ',', 'Internat
```

```
from nltk.stem import PorterStemmer

stemmer = PorterStemmer()
stemmed_words = [stemmer.stem(word) for word in tokens] # Changed 'filtered_

print("Stemmed Words:", stemmed_words)
```

```
Stemmed Words: ['id', ',', 'text', ',', 'label', '1', ',', 'the', 'weather',
```

```
import spacy

nlp = spacy.load("en_core_web_sm")

doc = nlp(text)
lemmatized_words = [token.lemma_ for token in doc]

print("Lemmatized Words:", lemmatized_words)
```

```
Lemmatized Words: ['id', ',', 'text', ',', 'label', '\n', '1,the', 'weather',
```

```
text = "Natural Language Processing is interesting"

words = text.split()        # Split the text into words
word_count = len(words)     # Count total number of words

print(word_count)           # Display the word count
```

```
5
```

```
from nltk.corpus import stopwords         # Import stopwords
nltk.download('stopwords') # Download the stopwords corpus

stop_words = set(stopwords.words('english'))  # Load English stopwords

filtered_words = [word for word in tokens      # Loop through tokens
                  if word.lower() not in stop_words]  # Remove stopwords

print(filtered_words)                          # Print filtered words
```

```
['id', ',', 'text', ',', 'label', '1', ',', 'weather', 'sunny', 'perfect', 'p
[nltk_data] Downloading package stopwords to /root/nltk_data...
```

```
[nltk_data]    Unzipping corpora/stopwords.zip.
```