

Nihar Ranjan Murudi CSE 2 (210273)

Printing the input sentences line by line

```
import nltk
nltk.download('punkt')
from nltk.tokenize import sent_tokenize
file_path = "/content/drive/MyDrive/nihar.txt"
with open(file_path, 'r') as file:
    text = file.read()
sentences = sent_tokenize(text)
for i, sentence in enumerate(sentences):
    print(f"Sentence {i+1}: {sentence}")
```

☞ Sentence 1: The modern-day computer has become an important part of our daily life.
 Sentence 2: Also, their usage has increased much fold during the last decade.
 Sentence 3: Nowadays, they use the computer in every office whether private or government.
 Sentence 4: Mankind is using computers for over many decades now.
 Sentence 5: Also, they are used in many fields like agriculture, designing, machinery making, defense and many more.
 Sentence 6: Above all, they have revolutionized the whole world.
 Sentence 7: essay on computer

History of Computers

It is very difficult to find the exact origin of computers.

Sentence 8: But according to some experts computer exists at the time of world war-II.

Sentence 9: Also, at that time they were used for keeping data.

Sentence 10: But, it was for only government use and not for public use.

Sentence 11: Above all, in the beginning, the computer was a very large and heavy machine.

Sentence 12: Working of a Computer

The computer runs on a three-step cycle namely input, process, and output.

Sentence 13: Also, the computer follows this cycle in every process it was asked to do.

Sentence 14: In simple words, the process can be explained in this way.

Sentence 15: The data which we feed into the computer is input, the work CPU do is process and the result which the computer give is out

Sentence 16: Components and Types of Computer

The simple computer basically consists of CPU, monitor, mouse, and keyboard.

Sentence 17: Also, there are hundreds of other computer parts that can be attached to it.

Sentence 18: These other parts include a printer, laser pen, scanner, etc.

Sentence 19: The computer is categorized into many different types like supercomputers, mainframes, personal computers (desktop), PDAs,

Sentence 20: The mobile phone is also a type of computer because it fulfills all the criteria of being a computer.

Sentence 21: Get the huge list of more than 500 Essay Topics and Ideas

Uses of Computer in Various Fields

As the usage of computer increased it became a necessity for almost every field to use computers for their operations.

Sentence 22: Also, they have made working and sorting things easier.

Sentence 23: Below we are mentioning some of the important fields that use a computer in their daily operation.

Sentence 24: Medical Field

They use computers to diagnose diseases, run tests and for finding the cure for deadly diseases.

Sentence 25: Also, they are able to find a cure for many diseases because of computers.

Sentence 26: Research

Whether it's scientific research, space research or any social research computers help in all of them.

Sentence 27: Also, due to them, we are able to keep a check on the environment, space, and society.

Sentence 28: Space research helped us to explore the galaxies.

Sentence 29: While scientific research has helped us to locate resources and various other useful resources from the earth.

Sentence 30: Defense

For any country, his defence is most important for the safety and security of its people.

Sentence 31: Also, computer in this field helps the country's security agencies to detect a threat which can be harmful in the future.

Sentence 32: Above all the defense industry use them to keep surveillance on our enemy.

Sentence 33: Threats from a Computer

Computers have become a necessity also, they have become a threat too.

Sentence 34: This is due to hackers who steal your private data and leak them on internet.

Sentence 35: Also, anyone can access this data.

Sentence 36: Apart from that, there are other threats like viruses, spams, bug and many other problems.

[nltk_data] Downloading package punkt to /root/nltk_data...

[nltk_data] Package punkt is already up-to-date!

#Tokenization

```
from nltk.tokenize import word_tokenize, sent_tokenize
file_path = "/content/drive/MyDrive/nihar.txt"
with open(file_path, 'r') as file:
    text = file.read()
sentences = sent_tokenize(text)
tokenized_words = [word_tokenize(sentence) for sentence in sentences]
for i, sentence in enumerate(tokenized_words):
    print(f"Sentence {i+1}: {sentence}")
```

```

nt      Se ence 1: ['The', 'modern-day', 'computer', 'has', 'become', 'an', 'important', 'part', 'of', 'our', 'daily', 'life', '.']
t       Sen ence 2: ['Also', ',', 'their', 'usage', 'has', 'increased', 'much', 'fold', 'during', 'the', 'last', 'decade', '.']
t       Sen ence 3: ['Nowadays', ',', 'they', 'use', 'the', 'computer', 'in', 'every', 'office', 'whether', 'private', 'on', 'government', '.']
t       Sen ence 4: ['Mankind', 'is', 'using', 'computers', 'for', 'over', 'many', 'decades', 'now', '.']
        Sentence 5: ['Also', ',', 'they', 'are', 'used', 'in', 'many', 'fields', 'like', 'agriculture', ',', 'designing', ',', 'machinery', '.']
        Sentence 6: ['Above', 'all', ',', 'they', 'have', 'revolutionized', 'the', 'whole', 'world', '.']
        Sentence 7: ['essay', 'on', 'computer', 'History', 'of', 'Computers', 'It', 'is', 'very', 'difficult', 'to', 'find', 'the', 'exact', '.']
        Sentence 8: ['But', 'according', 'to', 'some', 'experts', 'computer', 'exists', 'at', 'the', 'time', 'of', 'world', 'war-II', '.']
        Sentence 9: ['Also', ',', 'at', 'that', 'time', 'they', 'were', 'used', 'for', 'keeping', 'data', '.']
        Sentence 10: ['But', ',', 'it', 'was', 'for', 'only', 'government', 'use', 'and', 'not', 'for', 'public', 'use', '.']
        Sentence 11: ['Above', 'all', ',', 'in', 'the', 'beginning', ',', 'the', 'computer', 'was', 'a', 'very', 'large', 'and', 'heavy', '.']
        Sentence 12: ['Working', 'of', 'a', 'Computer', 'The', 'computer', 'runs', 'on', 'a', 'three-step', 'cycle', 'namely', 'input', ',', 'process', ',', 'output', '.']
        Sentence 13: ['Also', ',', 'the', 'computer', 'follows', 'this', 'cycle', 'in', 'every', 'process', 'it', 'was', 'asked', 'to', 'perform', '.']
        Sentence 14: ['In', 'simple', 'words', ',', 'the', 'process', 'can', 'be', 'explained', 'in', 'this', 'way', '.']
        Sentence 15: ['The', 'data', 'which', 'we', 'feed', 'into', 'the', 'computer', 'is', 'input', ',', 'the', 'work', 'CPU', 'do', 'it', 'all', '.']
        Sentence 16: ['Components', 'and', 'Types', 'of', 'Computer', 'The', 'simple', 'computer', 'basically', 'consists', 'of', 'CPU', 'monitor', 'mouse', '.']
        Sentence 17: ['Also', ',', 'there', 'are', 'hundreds', 'of', 'other', 'computer', 'parts', 'that', 'can', 'be', 'attached', 'to', 'the', 'mainframe', '.']
        Sentence 18: ['These', 'other', 'parts', 'include', 'a', 'printer', ',', 'laser', 'pen', ',', 'scanner', ',', 'etc', '.']
        Sentence 19: ['The', 'computer', 'is', 'categorized', 'into', 'many', 'different', 'types', 'like', 'supercomputers', ',', 'mainframes', ',', 'personal', 'computers', '.']
        Sentence 20: ['The', 'mobile', 'phone', 'is', 'also', 'a', 'type', 'of', 'computer', 'because', 'it', 'fulfills', 'all', 'the', 'criteria', 'of', 'a', 'computer', '.']
        Sentence 21: ['Get', 'the', 'huge', 'list', 'of', 'more', 'than', '500', 'Essay', 'Topics', 'and', 'Ideas', 'Uses', 'of', 'Computer', 'Applications', '.']
        Sentence 22: ['Also', ',', 'they', 'have', 'made', 'working', 'and', 'sorting', 'things', 'easier', '.']
        Sentence 23: ['Below', 'we', 'are', 'mentioning', 'some', 'of', 'the', 'important', 'fields', 'that', 'use', 'a', 'computer', 'in', 'their', 'operation', '.']
        Sentence 24: ['Medical', 'Field', 'They', 'use', 'computers', 'to', 'diagnose', 'diseases', ',', 'run', 'tests', 'and', 'for', 'finding', 'cure', 'deadly', 'diseases', '.']
        Sentence 25: ['Also', ',', 'they', 'are', 'able', 'to', 'find', 'a', 'cure', 'for', 'many', 'diseases', 'because', 'of', 'computer', 'technology', '.']
        Sentence 26: ['Research', 'Whether', 'it', 'is', 'scientific', 'research', ',', 'space', 'research', 'or', 'any', 'social', 'research', 'all', 'require', 'the', 'use', 'of', 'computers', '.']
        Sentence 27: ['Also', ',', 'due', 'to', 'them', ',', 'we', 'are', 'able', 'to', 'keep', 'a', 'check', 'on', 'the', 'environment', 'and', 'resources', '.']
        Sentence 28: ['Space', 'research', 'helped', 'us', 'to', 'explore', 'the', 'galaxies', '.']
        Sentence 29: ['While', 'scientific', 'research', 'has', 'helped', 'us', 'to', 'locate', 'resources', 'and', 'various', 'other', 'things', 'in', 'the', 'universe', '.']
        Sentence 30: ['Defense', 'For', 'any', 'country', ',', 'his', 'defense', 'is', 'most', 'important', 'for', 'the', 'safety', 'and', 'security', 'of', 'the', 'country', '.']
        Sentence 31: ['Also', ',', 'computer', 'in', 'this', 'field', 'helps', 'the', 'country', ',', 's', 'security', 'agencies', 'to', 'perform', 'their', 'tasks', '.']
        Sentence 32: ['Above', 'all', 'the', 'defense', 'industry', 'use', 'them', 'to', 'keep', 'surveillance', 'on', 'our', 'enemy', 'and', 'detect', 'any', 'threats', '.']
        Sentence 33: ['Threats', 'from', 'a', 'Computer', 'Computers', 'have', 'become', 'a', 'necessity', 'also', ',', 'they', 'have', 'made', 'it', 'easier', 'to', 'communicate', 'and', 'share', 'information', '.']
        Sentence 34: ['This', 'is', 'due', 'to', 'hackers', 'who', 'steal', 'your', 'private', 'data', 'and', 'leak', 'them', 'on', 'the', 'internet', '.']
        Sentence 35: ['Also', ',', 'anyone', 'can', 'access', 'this', 'data', '.']
        Sentence 36: ['Apart', 'from', 'that', ',', 'there', 'are', 'other', 'threats', 'like', 'viruses', ',', 'spams', ',', 'bug', 'and', 'malware', '.']

```

```

#Stopwords
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize, sent_tokenize
nltk.download('stopwords')
file_path = "/content/drive/MyDrive/nihar.txt"
with open(file_path, 'r') as file:
    text = file.read()
sentences = sent_tokenize(text)
stop_words = set(stopwords.words('english'))
tokenized_words_without_stopwords = []
for sentence in sentences:
    words = word_tokenize(sentence)
    filtered_words = [word for word in words if word.lower() not in stop_words]
    tokenized_words_without_stopwords.append(filtered_words)
for i, sentence in enumerate(tokenized_words_without_stopwords):
    print(f"Sentence {i+1}: {sentence}")

Sentence 1: ['modern-day', 'computer', 'become', 'important', 'part', 'daily', 'life', '.']
Sentence 2: ['Also', ',', 'usage', 'increased', 'much', 'fold', 'last', 'decade', '.']
Sentence 3: ['Nowadays', ',', 'use', 'computer', 'every', 'office', 'whether', 'private', 'government', '.']
Sentence 4: ['Mankind', 'using', 'computers', 'many', 'decades', '.']
Sentence 5: ['Also', ',', 'used', 'many', 'fields', 'like', 'agriculture', ',', 'designing', ',', 'machinery', 'making', ',', 'development', '.']
Sentence 6: ['revolutionized', 'whole', 'world', '.']
Sentence 7: ['essay', 'computer', 'History', 'Computers', 'difficult', 'find', 'exact', 'origin', 'computers', '.']
Sentence 8: ['according', 'experts', 'computer', 'exists', 'time', 'world', 'war-II', '.']
Sentence 9: ['Also', ',', 'time', 'used', 'keeping', 'data', '.']
Sentence 10: ['government', 'use', 'public', 'use', '.']
Sentence 11: ['beginning', ',', 'computer', 'large', 'heavy', 'machine', '.']
Sentence 12: ['Working', 'Computer', 'computer', 'runs', 'three-step', 'cycle', 'namely', 'input', ',', 'process', ',', 'output', '.']
Sentence 13: ['Also', ',', 'computer', 'follows', 'cycle', 'every', 'process', 'asked', '.']
Sentence 14: ['simple', 'words', ',', 'process', 'explained', 'way', '.']
Sentence 15: ['data', 'feed', 'computer', 'input', ',', 'work', 'CPU', 'process', 'result', 'computer', 'give', 'output', '.']
Sentence 16: ['Components', 'Types', 'Computer', 'simple', 'computer', 'basically', 'consists', 'CPU', ',', 'monitor', ',', 'mouse', '.']
Sentence 17: ['Also', ',', 'hundreds', 'computer', 'parts', 'attached', '.']
Sentence 18: ['parts', 'include', 'printer', ',', 'laser', 'pen', ',', 'scanner', ',', 'etc', '.']
Sentence 19: ['computer', 'categorized', 'many', 'different', 'types', 'like', 'supercomputers', ',', 'mainframes', ',', 'personal', 'computers', '.']
Sentence 20: ['mobile', 'phone', 'also', 'type', 'computer', 'fulfills', 'criteria', 'computer', '.']
Sentence 21: ['Get', 'huge', 'list', '500', 'Essay', 'Topics', 'Ideas', 'Uses', 'Computer', 'Various', 'Fields', 'usage', 'computer', 'applications', '.']
Sentence 22: ['Also', ',', 'made', 'working', 'sorting', 'things', 'easier', '.']
Sentence 23: ['mentioning', 'important', 'fields', 'use', 'computer', 'daily', 'operation', '.']
Sentence 24: ['Medical', 'Field', 'use', 'computers', 'diagnose', 'diseases', ',', 'run', 'tests', 'finding', 'cure', 'deadly', 'diseases', '.']

```

```

nt
t
t
t
Sentence 25: ['Also', ',', 'able', 'find', 'cure', 'many', 'diseases', 'computers', '.']
Sentence 26: ['Research', 'Whether', '', 'scientific', 'research', ',', 'space', 'research', 'social', 'research', 'computers', 'help',
Sentence 27: ['Also', ',', 'due', ',', 'able', 'keep', 'check', 'environment', ',', 'space', ',', 'society', '.']
Sentence 28: ['Space', 'research', 'helped', 'us', 'explore', 'galaxies', '.']
Sentence 29: ['scientific', 'research', 'helped', 'us', 'locate', 'resources', 'various', 'useful', 'resources', 'earth', '.']
Sentence 30: ['Defense', 'country', ',', 'defence', 'important', 'safety', 'security', 'people', '.']
Sentence 31: ['Also', ',', 'computer', 'field', 'helps', 'country', '', 'security', 'agencies', 'detect', 'threat', 'harmful', 'future'
Sentence 32: ['defense', 'industry', 'use', 'keep', 'surveillance', 'enemy', '.']
Sentence 33: ['Threats', 'Computer', 'Computers', 'become', 'necessity', 'also', ',', 'become', 'threat', '.']
Sentence 34: ['due', 'hackers', 'steal', 'private', 'data', 'leak', 'internet', '.']
Sentence 35: ['Also', ',', 'anyone', 'access', 'data', '.']
Sentence 36: ['Apart', ',', 'threats', 'like', 'viruses', ',', 'spams', ',', 'bug', 'many', 'problems', '.']
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Package stopwords is already up-to-date!

```

```

import nltk
from nltk.tokenize import word_tokenize
from nltk.probability import FreqDist
import matplotlib.pyplot as plt
from nltk.corpus import stopwords
import string
from wordcloud import WordCloud

nltk.download('punkt')
nltk.download('stopwords')

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

file_path = "/content/drive/MyDrive/nihar.txt"
with open(file_path, 'r') as file:
    text = file.read()

words = word_tokenize(text)
words = [word.lower() for word in words if word.isalnum()]
words = [word for word in words if word not in stop_words]
freq_dist = FreqDist(words)

# Print word count
print("Word count:")
for word, frequency in freq_dist.most_common():
    print(f"{word}: {frequency}")

# Plot frequency distribution
freq_dist.plot(30, cumulative=False)
plt.show()

# Generate word cloud
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(' '.join(words))

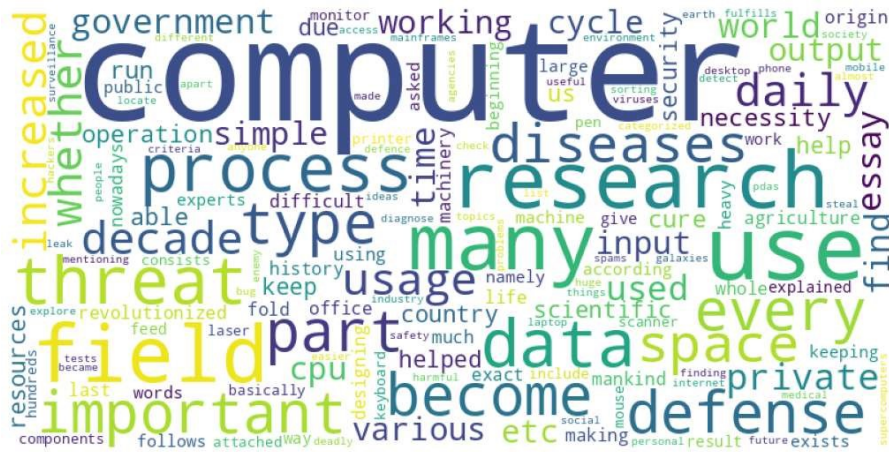
# Display the generated word cloud
plt.figure(figsize=(10, 8), facecolor=None)
plt.imshow(wordcloud)
plt.axis("off")
plt.tight_layout(pad=0)

plt.show()

```

neavy: 1
machine: 1
runs: 1
namely: 1
follows: 1
asked: 1
words: 1
explained: 1
way: 1
feed: 1
work: 1
result: 1
give: 1
components: 1
basically: 1
consists: 1
monitor: 1
mouse: 1
keyboard: 1
hundreds: 1
attached: 1
include: 1
printer: 1
laser: 1
pen: 1
scanner: 1
categorized: 1
different: 1
supercomputers: 1
mainframes: 1
personal: 1
desktop: 1
pdas: 1
laptop: 1
mobile: 1
phone: 1
type: 1
fulfills: 1
criteria: 1
get: 1
huge: 1
list: 1
500: 1
topics: 1
ideas: 1
uses: 1
became: 1
almost: 1
operations: 1
made: 1
sorting: 1
things: 1
easier: 1
mentioning: 1
operation: 1
medical: 1
diagnose: 1
run: 1
tests: 1
finding: 1
deadly: 1
social: 1
help: 1
check: 1
environment: 1
society: 1
explore: 1
galaxies: 1
locate: 1
useful: 1
earth: 1
defence: 1
safety: 1
people: 1
helps: 1
agencies: 1
detect: 1
harmful: 1
future: 1
industry: 1
surveillance: 1
enemy: 1
hackers: 1
steal: 1
leak: 1

Sample	Count
computer	21
also	12
computers	9
use	7
many	6
research	6
data	4
process	4
become	3
important	3
every	3
fields	3
like	3
defense	3
field	3
diseases	3
space	3
daily	3
usage	2
increased	2
whether	2
private	2
government	2
used	2
world	2
essay	2
find	2
time	2
working	2
cycle	2



```
#Converting to Lowercase
file_path = "/content/drive/MyDrive/nihar.txt"
with open(file_path, 'r') as file:
    text = file.read()
lowercase_text = text.lower()
print(lowercase_text)
```

the modern-day computer has become an important part of our daily life. also, their usage has increased much fold during the last decade
essay on computer

history of computers
it is very difficult to find the exact origin of computers. but according to some experts computer exists at the time of world war-ii. a

working of a computer
the computer runs on a three-step cycle namely input, process, and output. also, the computer follows this cycle in every process it was

6

components and types of computer
the simple computer basically consists of cpu, monitor, mouse, and keyboard. also, there are hundreds of other computer parts that can b

the computer is categorized into many different types like supercomputers, mainframes, personal computers (desktop), pdas, laptop, etc.

get the huge list of more than 500 essay topics and ideas

uses of computer in various fields
as the usage of computer increased it became a necessity for almost every field to use computers for their operations. also, they have m

medical field
they use computers to diagnose diseases, run tests and for finding the cure for deadly diseases. also, they are able to find a cure for

research
whether it's scientific research, space research or any social research computers help in all of them. also, due to them, we are able to

defense
for any country, his defence is most important for the safety and security of its people. also, computer in this field helps the country

threats from a computer
computers have become a necessity also, they have become a threat too. this is due to hackers who steal your private data and leak them