

Q) Stemming for a given sentence.

A) \Rightarrow stemming is a process of extracting a root word.

\Rightarrow for example: 'jumping', 'jumps' & 'jumped' are stemmed into jump.

\Rightarrow stemming helps us in standardizing words to their base stem regardless of their pronunciation.

This helps us to classify or cluster the text.

\Rightarrow Search engines use these techniques extensively to give better & more accurate results irrespective of the word form.

\Rightarrow The nltk. package has several implementations for stemmers.

- ① using Porter stemmer.
- ② using Lancaster stemmer.
- ③ using snowball stemmer.
- ④ using Regexp stemmer.

PROGRAM :

```
from nltk.stem import PorterStemmer
```

```
st = PorterStemmer()
```

```
text = ['where did he learn to dance like that?',  
        'His eyes were dancing with humor.',  
        'She shook her head and danced away',  
        'Aditi is an excellent dancer.']
```

```
output = []
```

```
for sentence in text:
```

```
    output.append(" ".join([st.stem(i) for  
                             i in sentence.split()]))
```

```
for item in output:
```

```
    print(item)
```

```
print("-" * 50)
```

```
print(st.stem('jumping'), st.stem('jumps'), st.stem('jumped'))
```

Expected output :

where did he learn to dance like that?

his eyes were dancing with humor.

She shook her head and danced away.

Aditi was an excellent dancer.

jump jump jump.

Output

where did he learn to dance like that?

his eyes were dancing with humor.

She shook her head and danced away.

Aditi was an excellent dancer.

jump jump jump.

OUTPUT SCREENSHOT



The screenshot shows a Google Colab notebook interface. At the top, the notebook is titled '160118737001.ipynb'. The menu bar includes 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help', with a status message 'All changes saved'. On the right, there are buttons for 'Comment', 'Share', and a settings icon. Below the menu bar, the notebook is in 'Editing' mode, with RAM and Disk usage indicators. The code cell contains the following Python code:

```
from nltk.stem import PorterStemmer

st = PorterStemmer()
text = ['Where did he learn to dance like that?',
        'His eyes were dancing with humor.',
        'She shook her head and danced away',
        'Aditi was an excellent dancer.']

output = []
for sentence in text:
    output.append(" ".join([st.stem(i) for i in sentence.split()]))

for item in output:
    print(item)

print("-" * 50)
print(st.stem('jumping'), st.stem('jumps'), st.stem('jumped'))
```

The output of the code is displayed below the code cell:

```
where did he learn to danc like that?
hi eye were danc with humor.
she shook her head and danc away
aditi wa an excel dancer.
-----
jump jump jump
```

Link to colab file: <https://colab.research.google.com/drive/1AU-HwDJpNldi21umz9vOfXTQIUqKps40?usp=sharing>