Stemming_NLP

December 10, 2023

Date : 10/12/2023

0.1 Stemming

1. Stemming is the process of reducing the word to it's root form.
2. The root form of word is known as the **'Lemma'**
3. There are two methods for stemming **1) PorterStemmer**, **2) LancasterStemmer**

1 Code for Stemming

1.1 1. Porter Stemmer method

[1]:
```
### --------------------
### Importing libraries
### --------------------

from nltk.stem import PorterStemmer, LancasterStemmer
```

[2]:
```
word1 = "chocolates", "chocolatey", "choco"

word2 = "retrieval", "retrieved", "retrieves"

word3 = "fishing", "fished", "fisher"

word4 = "argue", "argued", "argues", "arguing", "argus"
```

[3]:
```
# Initializing the Stemmer

por_stem = PorterStemmer()
```

[4]:
```
# Iterating over each word to reduce it to root form

word_list = [word1, word2, word3, word4]

for element in word_list:

    for word in element:

```

```
        stem_word = por_stem.stem(word)

        print(stem_word)

    print()
```

```
chocol
chocolatey
choco

retriev
retriev
retriev

fish
fish
fisher

argu
argu
argu
argu
argu
```

1.2 2. Lancaster Stemmer Method

[5]:
```
# Inirializing the Lancaster Stemmer

lan_stemmer = LancasterStemmer()
```

[6]:
```
# Iterating over each word to reduce it to root form

word_list = [word1, word2, word3, word4]

for element in word_list:

    for word in element:

        stem_word = lan_stemmer.stem(word)

        print(stem_word)

    print()
```

```
chocol
chocolatey
choco
```

```
retriev
retriev
retriev

fish
fish
fish

argu
argu
argu
argu
arg
```