capstone_summary

December 5, 2019

[34]:
```
import sumy
from sumy.parsers.plaintext import PlaintextParser
from sumy.nlp.tokenizers import Tokenizer
from sumy.summarizers.lex_rank import LexRankSummarizer
```

[35]:
```
document1="""Text summarization refers to the technique of shortening long␣
 ↪pieces of text. The intention is to create a coherent and fluent summary␣
 ↪having only the main points outlined in the document. Automatic text␣
 ↪summarization is a common problem in machine learning and natural language␣
 ↪processing (NLP)."""
```

[36]:
```
document1
```

[36]:
```
'Text summarization refers to the technique of shortening long pieces of text.
The intention is to create a coherent and fluent summary having only the main
points outlined in the document. Automatic text summarization is a common
problem in machine learning and natural language processing (NLP).'
```

[37]:
```
parser = PlaintextParser.from_string(document1 ,Tokenizer("english"))
```

[38]:
```
parser
```

[38]:
```
<sumy.parsers.plaintext.PlaintextParser at 0x7f10a4cf2eb8>
```

[40]:
```
from sumy.summarizers.lex_rank import LexRankSummarizer
```

[41]:
```
summarizer = LexRankSummarizer()
```

[55]:
```
summary=summarizer(parser.document,2)
```

[56]:
```
for sentence in summary:
    print(sentence)
```

```
Text summarization refers to the technique of shortening long pieces of text.
The intention is to create a coherent and fluent summary having only the main
points outlined in the document.
```

[46]:
```
from sumy.summarizers.luhn import LuhnSummarizer
luhn_summarizer = LuhnSummarizer()
```

[63]:
```
l_summary=luhn_summarizer(parser.document,1)
for sentence in l_summary:
```

```
    print(sentence)
```

```
Text summarization refers to the technique of shortening long pieces of text.
```

[64]:
```
from sumy.summarizers.lsa import LsaSummarizer
lsa_summarizer = LsaSummarizer()
lsa_summary=lsa_summarizer(parser.document,1)
for sentence in lsa_summary:
    print(sentence)
```

```
Automatic text summarization is a common problem in machine learning and natural
language processing (NLP).
```

[52]:
```
from sumy.nlp.stemmers import Stemmer
from sumy.utils import get_stop_words
```

[53]:
```
summarizer_lsa2 = LsaSummarizer()
summarizer_lsa2 = LsaSummarizer(Stemmer("english"))
summarizer_lsa2.stop_words = get_stop_words("english")
```

[54]:
```
summary2 = summarizer_lsa2(parser.document ,1)
print(summary2)
```

```
(<Sentence: Automatic text summarization is a common problem in machine learning
and natural language processing (NLP).>,)
```

[ ]:
```

```