

```
1 import nltk
2 from nltk.corpus import wordnet
3 from nltk.stem import WordNetLemmatizer
4 from nltk import pos_tag
5 nltk.download('wordnet')
6 nltk.download('averaged_perceptron_tagger')
7 from nltk.tokenize import sent_tokenize,
   word_tokenize
8 import string
9
10 class Text:
11     def __init__(self, text):
12         self.text = text
13
14     def split_into_sent(self):
15         sentences = sent_tokenize(self.text)
16         return sentences
17
18     def __str__(self):
19         return self.text
20
21 class Sent:
22     def __init__(self, sentences):
23         self.sentences = [sentence for sentence in
   sentences]
24
25     def get_sentence(self, index):
26         if index < 0 or index >= len(self.sentences):
27             raise IndexError("Index out of range")
28         return self.sentences[index]
29
30     def __str__(self):
31         return ' '.join(self.sentences)
32
33 class LexEntry:
34     wnL = WordNetLemmatizer()
35
36     pos = {"noun", "pronoun", "verb", "adjective", "
   adverb", "preposition", "determiner", "conjunction",
37           "interjection", }
38
```

```

39     def __init__(self, sentence):
40         self.wnl = WordNetLemmatizer()
41         self.lex_entries = self.process_text(sentence
42 ).split()
43     def process_text(self, text):
44         translator = str.maketrans('', '', string.
45 punctuation)
46         cleaned_text = text.translate(translator)
47         return cleaned_text.lower()
48     def split_into_words(self):
49         return self.lex_entries
50
51     def get_word_len(self):
52         return [len(word) for word in self.
53 lex_entries]
54
55     def get_word_index(self, word):
56         try:
57             return self.lex_entries.index(word)
58         except ValueError:
59             return -1
60
61     def get_pos(self, word):
62         tag = pos_tag([word])[0][1]
63         if tag.startswith('J'):
64             return wordnet.ADJ
65         elif tag.startswith('V'):
66             return wordnet.VERB
67         elif tag.startswith('N'):
68             return wordnet.NOUN
69         elif tag.startswith('R'):
70             return wordnet.ADV
71         else:
72             return wordnet.NOUN # Default to noun if
73 not found
74
75     def lemmatize_word(self, word):
76         pos = self.get_pos(word)
77         return self.wnl.lemmatize(word, pos)

```

```
76
77 text = Text("Hello world. This is a test sentence.")
78 sentences = text.split_into_sent()
79 print(sentences)
80
81 sent = Sent(sentences)
82 print(sent.get_sentence(1))
83
84 lex_entry = LexEntry(sent.get_sentence(1))
85 print(lex_entry.split_into_words())
86 print(lex_entry.get_word_len())
87 print(lex_entry.get_word_index("test"))
88 print(lex_entry.lemmatize_word("favoring"))
```