How to quickly check and count English grammar errors in Python? [closed]

Asked 1 year, 7 months ago   Active 1 year, 7 months ago   Viewed 2k times

0

I'm looking for a fast way to count grammar errors in Python. For example, I need something along these lines:

```
matches = grammar_checker.check('You is awesome!')
assert len(matches) == 1
```

★

1

Speed is much more an issue than accuracy. I could live with a few false positives or negatives as a trade-off for fast checks. I am talking here about problem sizes on the order of 100k documents (with roughly 5-10k characters) in only a few minutes.

I came across LanguageTool in Java and its Python wrapper language_check. However, these are unfortunately too slow for my purposes. Any other suggestions or ideas? Thanks!

`python` `python-3.x` `performance` `nlp`

edited Feb 20 '18 at 21:42   asked Feb 20 '18 at 21:38

                                                  SmCaterpillar
                                                  3,047   2   25   50

---

closed as off-topic by Fred Larson, juanpa.arrivillaga, Chris_Rands, Thierry Lathuille, AdrianHHH Feb 20 '18 at 22:08

This question appears to be off-topic. The users who voted to close gave this specific reason:

- "Questions asking us to **recommend or find a book, tool, software library, tutorial or other off-site resource** are off-topic for Stack Overflow as they tend to attract opinionated answers and spam. Instead, describe the problem and what has been done so far to solve it." – Fred Larson, juanpa.arrivillaga, Chris_Rands, Thierry Lathuille, AdrianHHH

If this question can be reworded to fit the rules in the help center, please edit the question.

On hold, really? This question is much less specific (without the efficiency constraint) and has over 20 votes. Why is that one not on hold? – SmCaterpillar Feb 21 '18 at 7:23

---

2 Answers

6

⌄

For more information, check out https://pypi.python.org/pypi/grammar-check/1.3.1

```python
import grammar_check
tool = grammar_check.LanguageTool('en-GB')
text = 'This are bad.'
matches = tool.check(text)
len(matches)
>>>1

grammar_check.correct(text, matches)
>>>'These are bad'
```

edited Feb 20 '18 at 22:00   answered Feb 20 '18 at 21:54

Orhan Solak
578   1   4   20

---

1   I'll look into the NLTK stuff thanks. Nonetheless, I cannot use `grammar_check` because it's just a fork of `language_check`. Therefore, also uses the Java LanguageTool which is too slow for my purposes. Still could be that I can configure the LanguageTool server better that it internally uses multiprocessing, but I haven't found a way to do that, yet. – SmCaterpillar   Feb 21 '18 at 7:22

⌃

0

⌄

You can try Grammarly, it is one of the most popular speller checkers. Probably, they can provide some kind of an API. (I am not an expert in Python frameworks but I guess you should try this if you need to just check grammar as fast as possible.

answered Feb 20 '18 at 21:45

MAP Inc.
7   6

---

1   I don't think there is an API for Grammarly, but I feel like this would be even slower because you would have to make requests to their server with large amounts of text. And even as I'm typing this, Grammarly is taking a few seconds for each mistake in this text box. – whackamadoodle3000   Feb 20 '18 at 21:51 ✎