

ДЗ 16. Класифікація тексту

usetextblob.py

Спробував використати

```
nlp.add_pipe('spacytextblob')
```

Досить оптимальний інструмент

Результат виконання:

Polarity:

0.10972222222222222

Subjectivity:

0.5593434343434344

Sentiment assessments:

[(['wonderful'], 1.0, 1.0, None), (['little'], -0.1875, 0.5, None), (['very'], 0.2, 0.3, None), (['very'], 0.2, 0.3, None), (['entire'], 0.0, 0.625, None), (['extremely'], -0.125, 1.0, None), (['not', 'only'], -0.0, 1.0, None), (['down', '!'], -0.19444444444444448, 0.2888888888888889, None), (['seamless'], 0.1, 0.1, None), (['not', 'only'], -0.0, 1.0, None), (['worth'], 0.3, 0.1, None), (['terrificly'], 0.0, 1.0, None), (['masterful'], 1.0, 1.0, None), (['great'], 0.8, 0.75, None), (['really'], 0.2, 0.2, None), (['little'], -0.1875, 0.5, None), (['traditional'], 0.0, 0.75, None), (['solid'], 0.0, 0.1, None), (['particularly'], 0.16666666666666666, 0.3333333333333333, None), (['particularly'], 0.16666666666666666, 0.3333333333333333, None), (['flat'], -0.025, 0.125, None), (['terribly'], -1.0, 1.0, None)]

result

positive

Значення polarity та subjectivity відповідає класифікації positive

HW16.ipynb

Навчання моделі за допомогою класифікатора `MultinomialNB()`, `spacy` був використаний для стандартизації тексту. Вдалось досягти Accuracy: 0.86

classification.py

Була використаний `LogisticRegression()` для класифікації `Spacy` був використаний для стандартизації тексту. Вдалось досягти Accuracy: 0.90

Accuracy: 0.90

	precision	recall	f1-score	support
negative	0.91	0.89	0.90	4961
positive	0.89	0.91	0.90	5039
accuracy			0.90	10000
macro avg	0.90	0.90	0.90	10000
weighted avg	0.90	0.90	0.90	10000