Московский государственный технический университет им. Н.Э. Баумана  
Кафедра «Системы обработки информации и управления»

****

Лабораторная работа №6  
по дисциплине  
«Методы машинного обучения»

Выполнил:  
студент группы ИУ5-21М  
Ян Тяньци

Москва — 2022 г.

**1. Цель лабораторной работы**

Изучение методов классификации текстов..

**2. Задание**

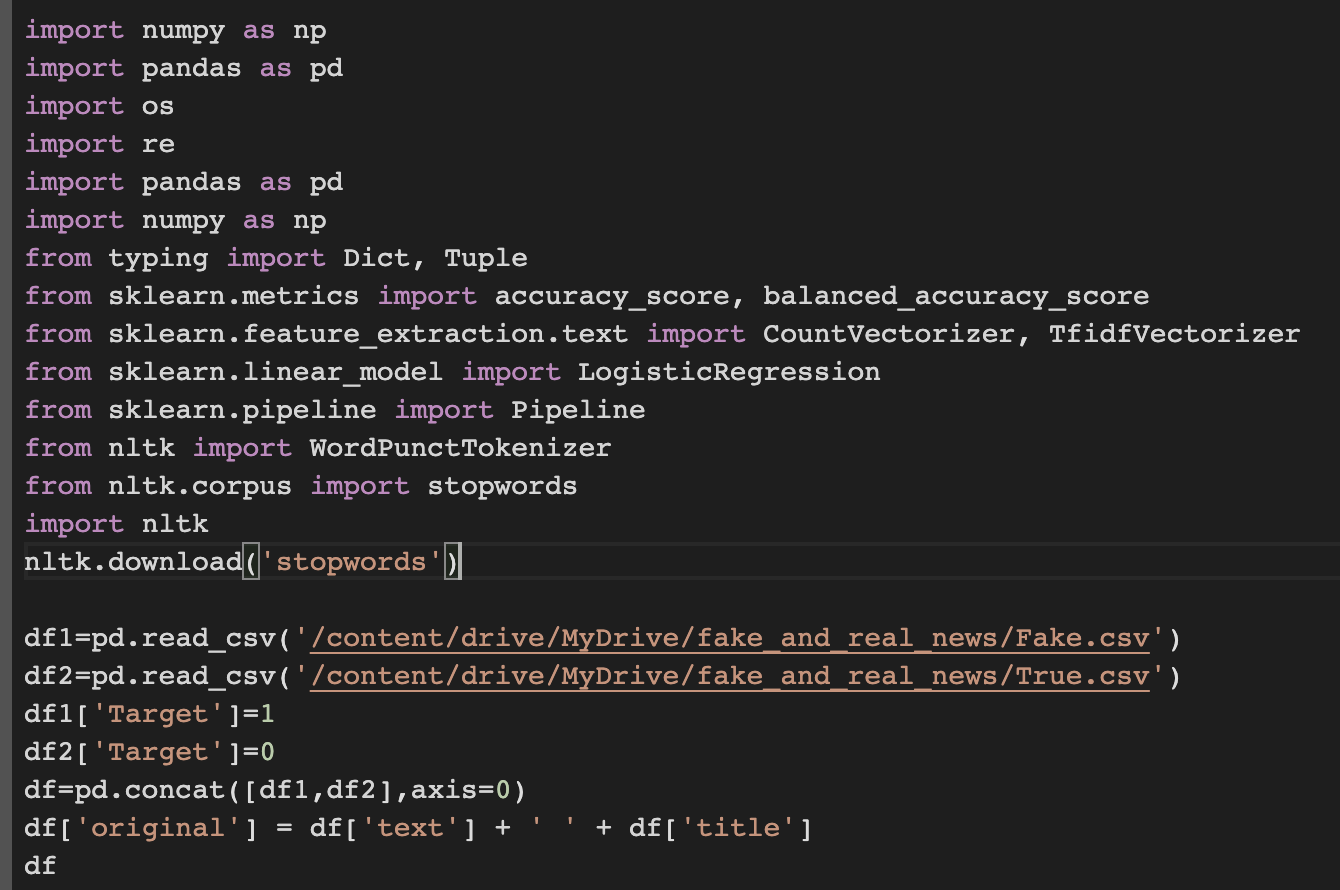
Для произвольного набора данных, предназначенного для классификации текстов, решите задачу классификации текста двумя способами:

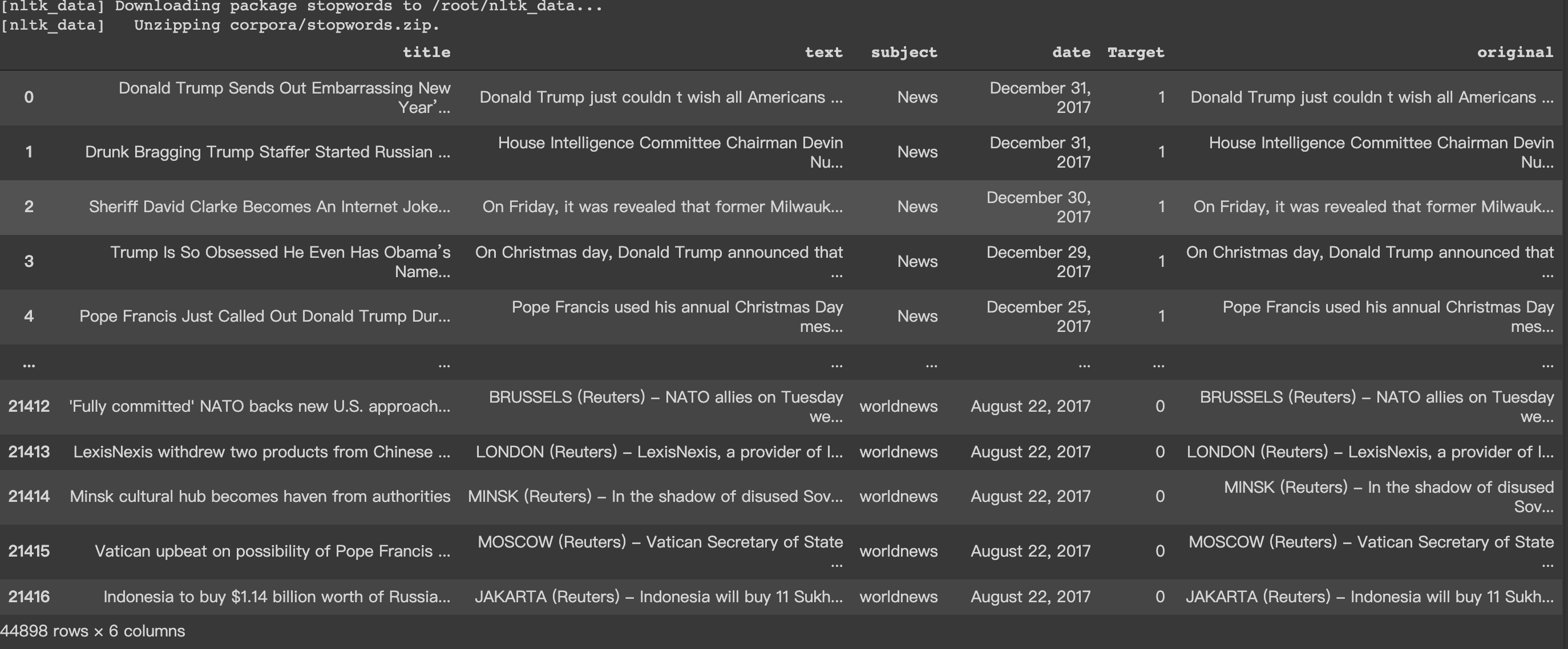
Способ 1. На основе CountVectorizer или TfidfVectorizer.

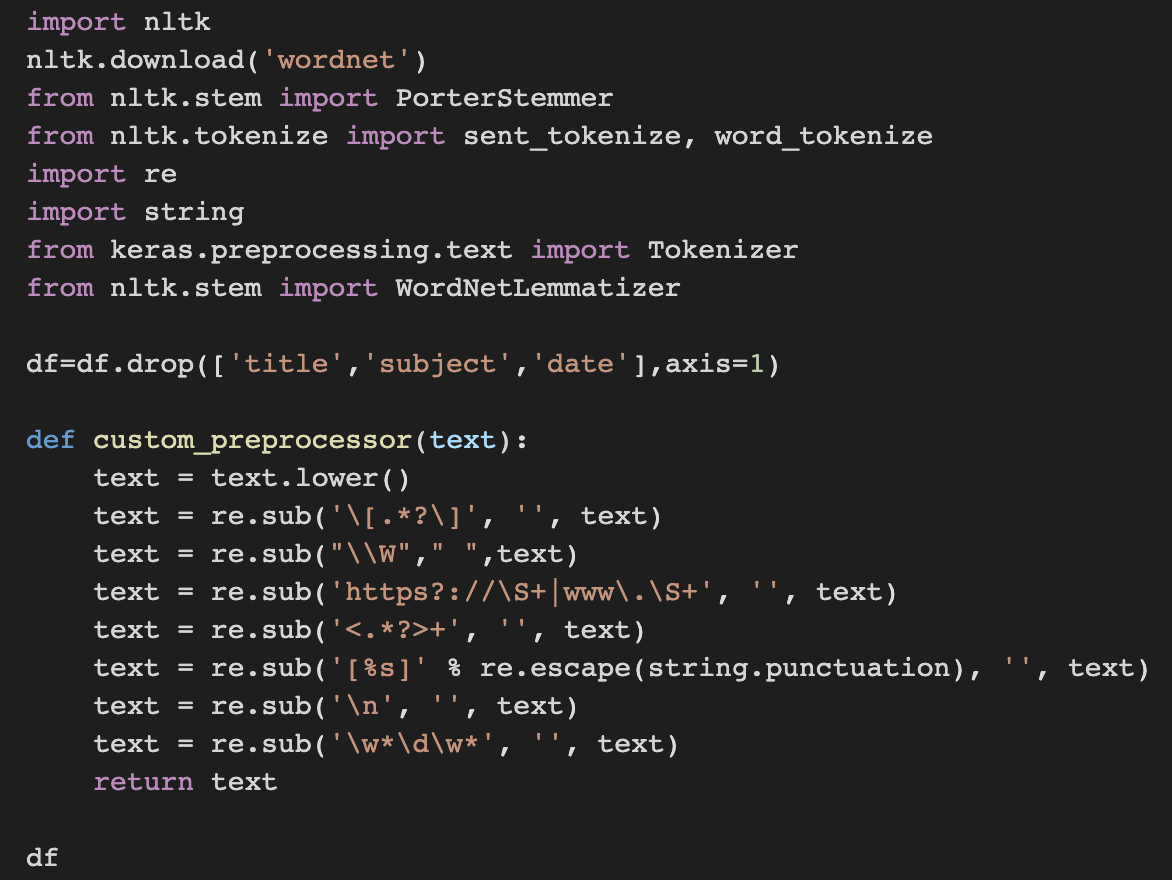
Способ 2. На основе моделей word2vec или Glove или fastText.

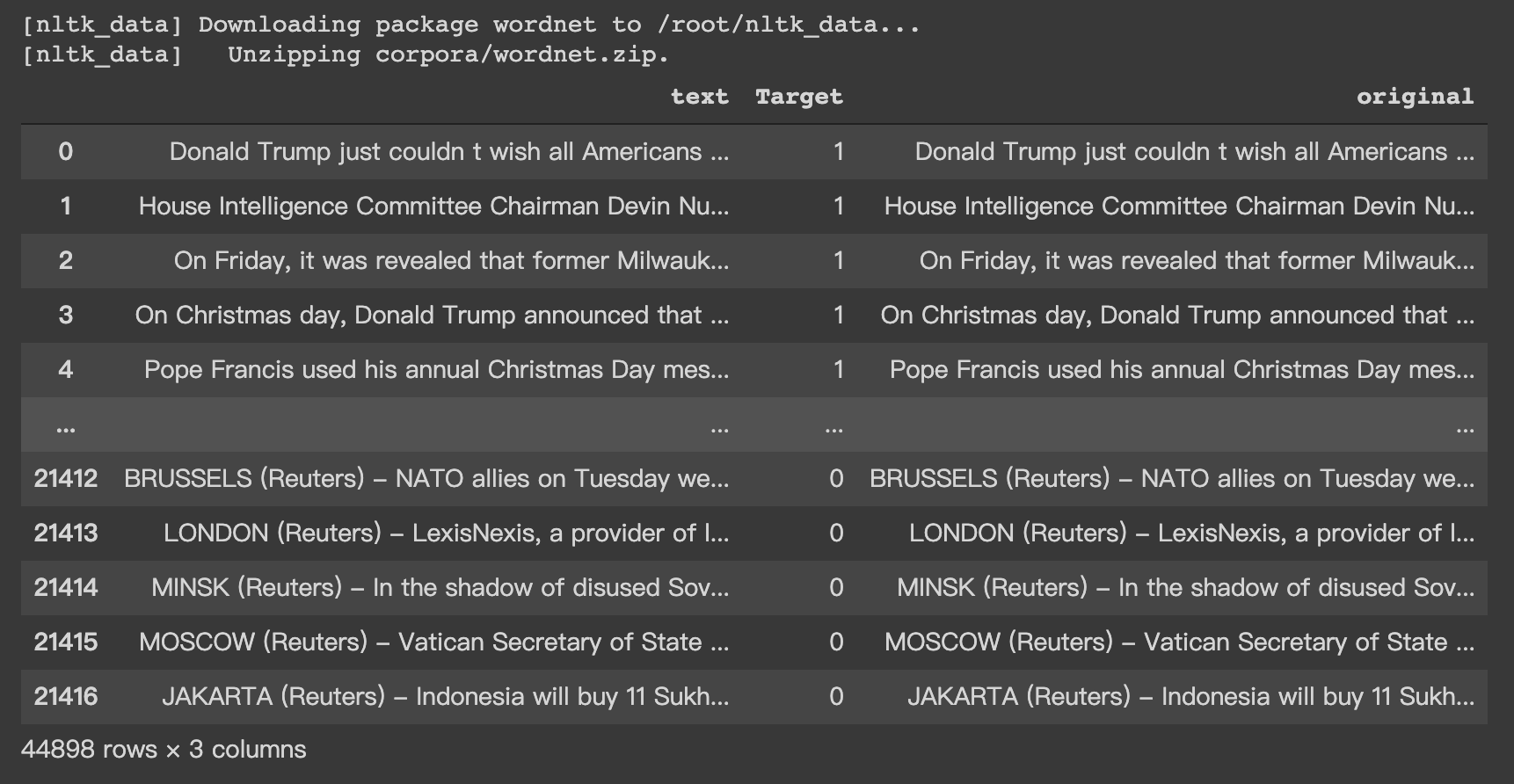
Сравните качество полученных моделей.Разбор предложения.

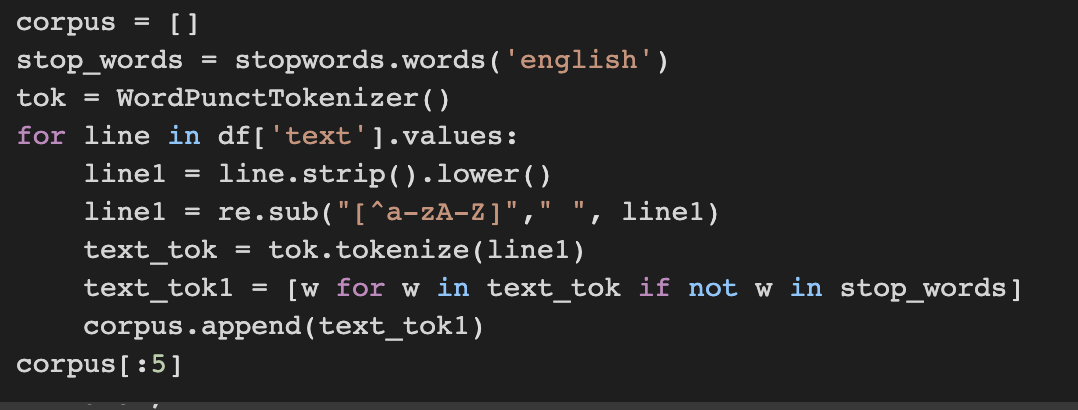
**3. Ход выполнения работы**

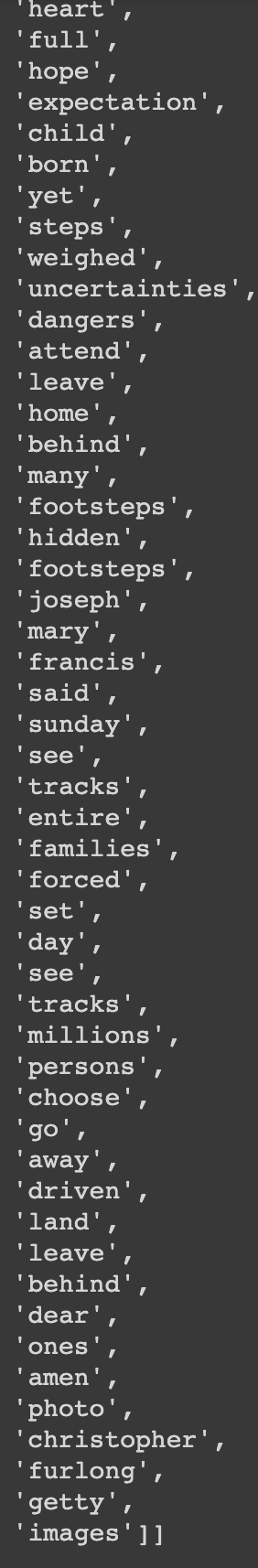


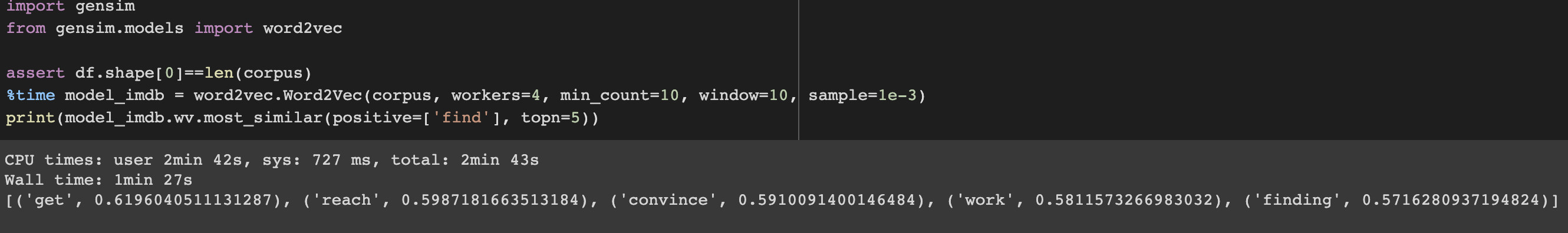


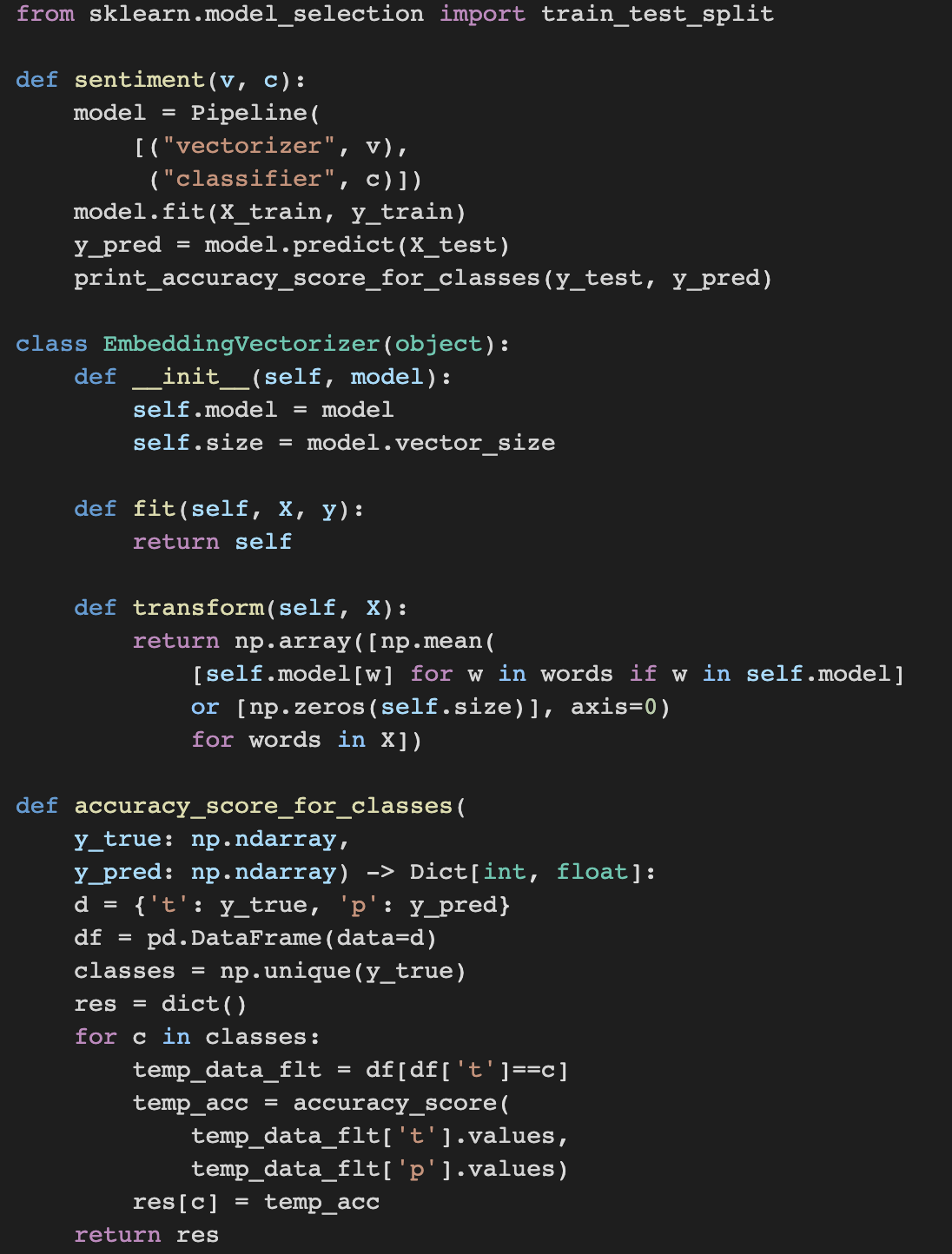


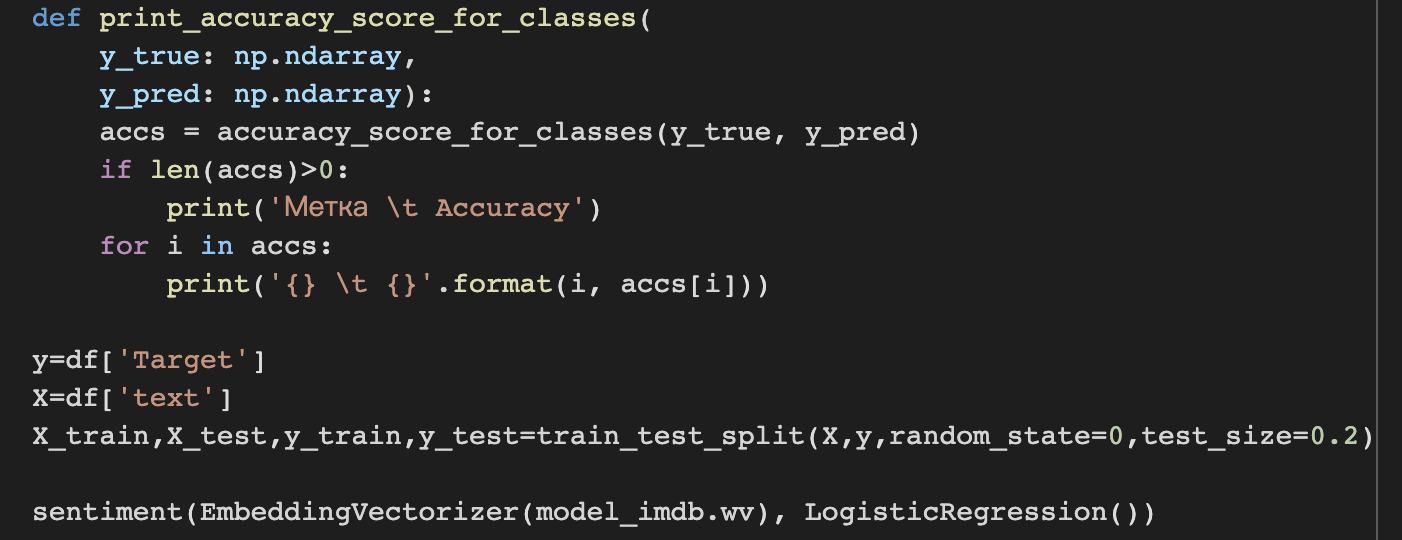


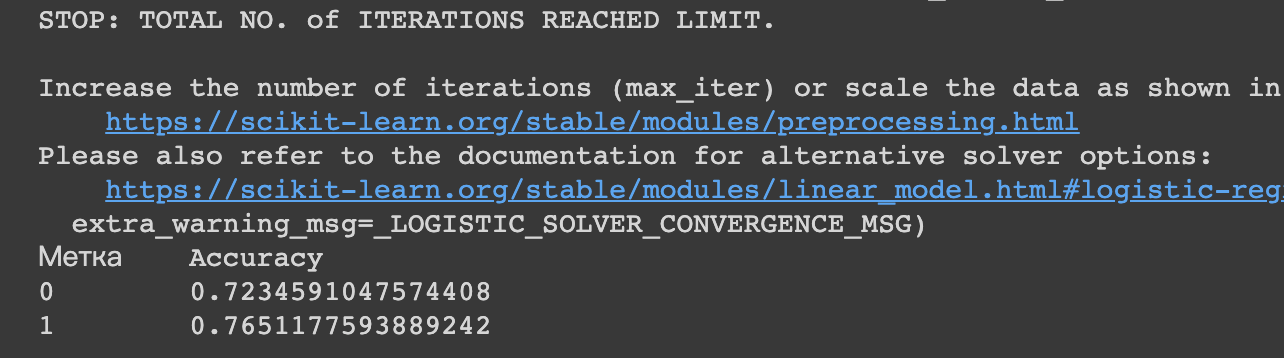


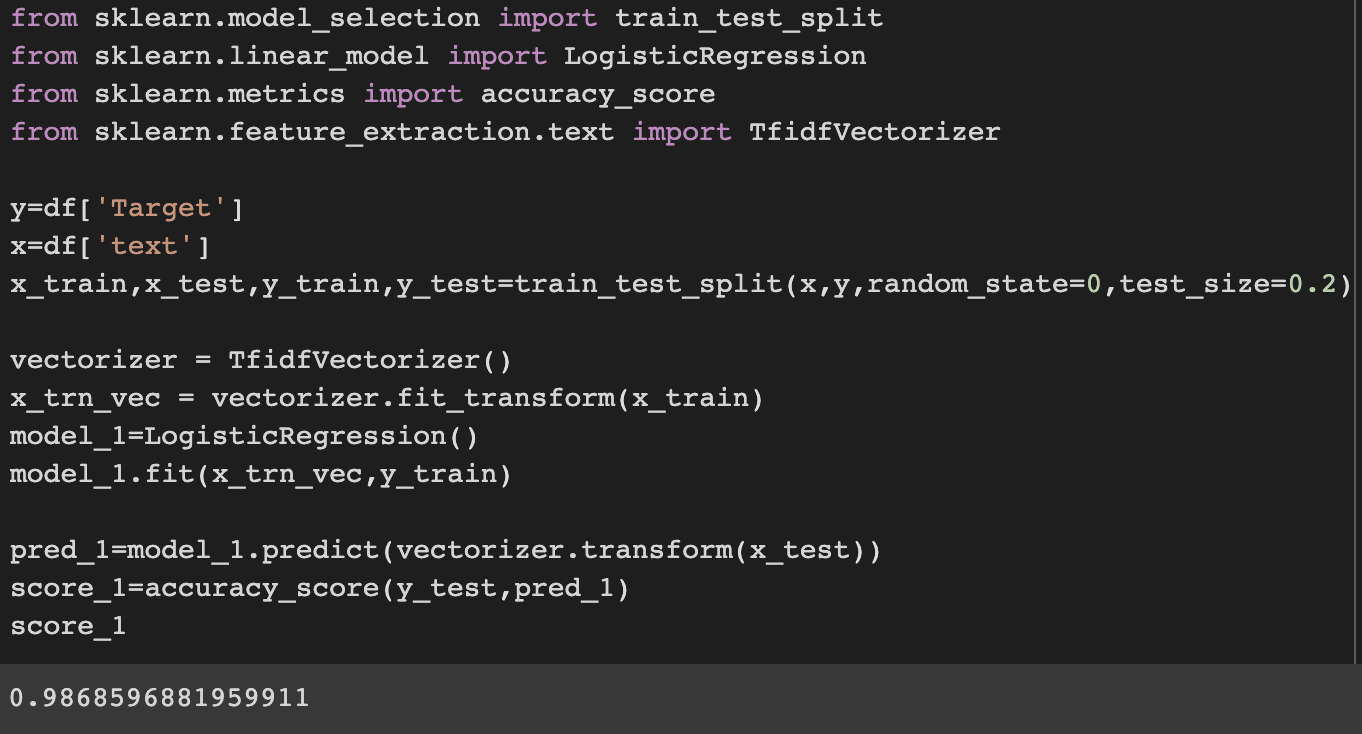












**Вывод：**

Мы видим, что при векторизации модели с использованием метода TfidfVectorizer точность использования алгоритма логистической регрессии является самой высокой. Самая высокая точность достигала 98%.

**Список литературы**

[1] Гапанюк Ю. Е. Лабораторная работа «Линейные модели, SVM и деревья решений»[Электронный ресурс] // GitHub. –– 2019. –– Режим доступа: https://github.com/ugapanyuk/ml\_course/wiki/LAB\_TREES (дата обращения: 19.04.2019).

[2] Team The IPython Development. IPython 7.3.0 Documentation [Electronic resource] //Read the Docs. — 2019. — Access mode: https://ipython.readthedocs.io/en/stable/ (online; accessed: 20.02.2019).

[3] Waskom M. seaborn 0.9.0 documentation [Electronic resource] // PyData. — 2018. —Access mode: https://seaborn.pydata.org/ (online; accessed: 20.02.2019).

[4] pandas 0.24.1 documentation [Electronic resource] // PyData. — 2019. — Access mode:http://pandas.pydata.org/pandas-docs/stable/ (online; accessed: 20.02.2019).

[5] dronio. Solar Radiation Prediction [Electronic resource] // Kaggle. — 2017. — Accessmode: https://www.kaggle.com/dronio/SolarEnergy (online; accessed: 18.02.2019).

[6] Chrétien M. Convert datetime.time to seconds [Electronic resource] // Stack Overflow.— 2017. — Access mode: https://stackoverflow.com/a/44823381 (online; accessed:20.02.2019).

[7] scikit-learn 0.20.3 documentation [Electronic resource]. — 2019. — Access mode: https://scikit-learn.org/ (online; accessed: 05.04.2019).