

Определение языка

Даны три файла с разными языками в папке с заданием.

1. Откройте файлы.
2. Объедините все записи из файлов в один датасет.
3. Реализуйте свой классификатор или используйте любые вам известные стандартные библиотеки для определения языка каждой записи.
4. Добейтесь точности классификации **0.85** на KFold с $k=10$.
5. * Определите, что это за языки. Опишите, как вы определили язык.
6. Задание присылайте в ipython notebook.

Полезные библиотеки

codecs, unicodedata, base64