**Задание**

Обучить векторную модель и попробовать ее методы.

Пункты:

1. Преобразовать текст в формат:

* только леммы, без знаков препинания, все в нижнем регистре
* каждое предложение на отдельной строчке

1. Обучить модель как показано в конспекте с параметрами

* размер вектора 300
* минимальное количество вхождений 5
* окно 5
* количество итераций 50
* Вопрос: сколько получилось слов в этой модели?

1. Использовать следующие функции:

* найти ближайшие 5 слов для а) любого частотного слова (союза/местоимения/частицы) б) абстрактного понятия в) имени главного героя г) прилагательных "хороший" и "плохой" (или любых других антонимов) (краткий комментарий про то, встретились ли синонимы, антонимы, близкие по распределению слова)
* найти любую семантическую пропорцию (краткий комментарий, почему Вы выбрали такую пропорцию)
* найти лишнее в списке слов (краткий комментарий про результат)

1. Попробуйте сделать визуализацию с помощью PCA (**один** из вариантов на выбор, можно предложить свой)

* попробуйте на именах героев (4+ имен, можно брать имя и фамилию героя отдельно, это даже интересно): это отражает их отношения / их характеры?
* попробуйте на каких-то названиях предметов или животных: насколько это похоже на правду?
* попробуйте на союзах, частицах, экспрессивах, многозначных словах: насколько результат соответствует Вашим ожиданиям?

1. Возьмите предложение и замените все слова (которые не являются служебными) на ближайшие по векторам