

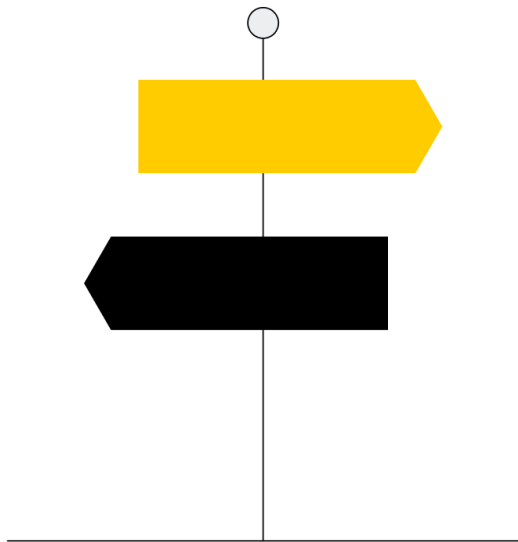
Fake reviews detection app

Студкемп Яндекс × ВШЭ

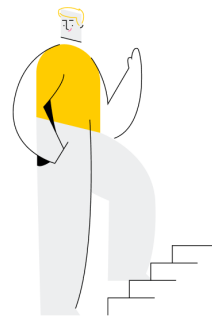
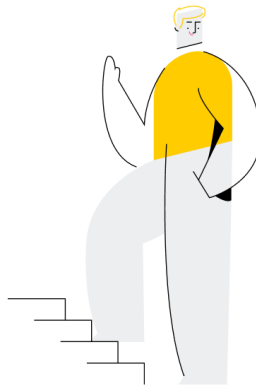
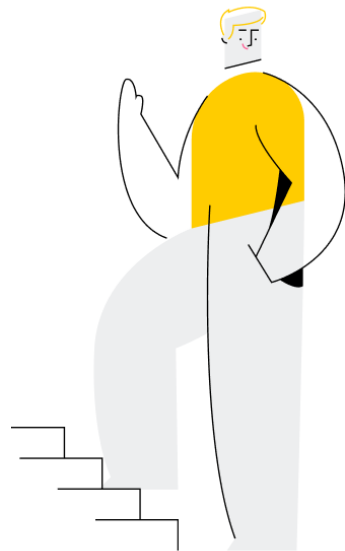
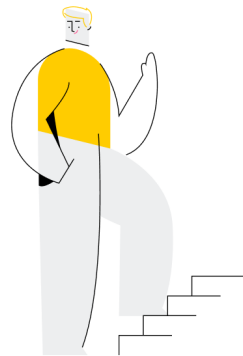


Александр Прошичев
Станислав Ивашков

Михаил Сурков
Ярослав Присталов



Это мы.





Александр Прошичев,
ЛЭТИ



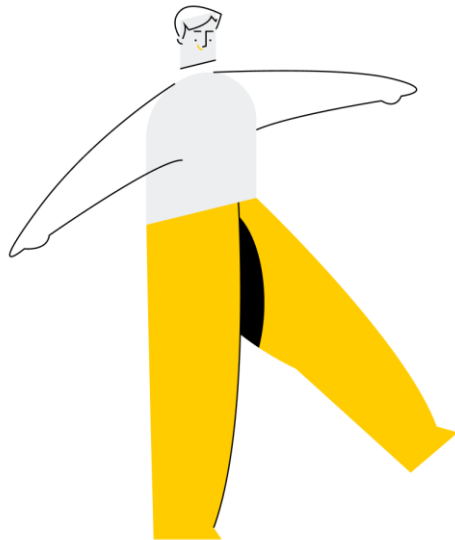
Станислав Ивашков,
МИЭТ, CV-intern, MIET



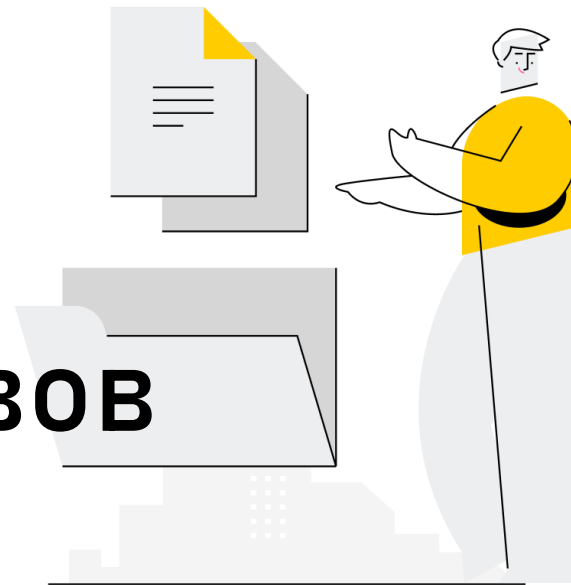
Михаил Сурков,
физфак МГУ

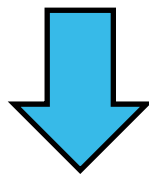


Ярослав Присталов,
НИУ ВШЭ



Получение реальных отзывов





kaggle



Генерация фейковых отзывов

"Какой бы" + `emotion[rd.randint(0, 2)]` + "отзыв ты бы оставил? Сделай так, чтобы отзыв был не больше, чем в " + `str(min(rd.randint(1,10), rd.randint(1, 10)))` + ". Ответ дай в виде json-файла."

"Представь, что ты покупатель в интернет магазине со следующими характеристиками: " + `costumers[rd.randint(0, 29)]` + ". Приобретаешь следующую вещь: " + `products[rd.randint(0, 203)]` + ". Купил ты это вещь для " + `dest[rd.randint(0, 32)]` + ". Можешь использовать не всю предоставленную о человеке информацию. Постарайся сделать естественный отзыв. Выбирай или формальный, или неформальный стиль общения в Интернете, но не нужно приводить в пример оба варианта. Возможно использовать не часто сленг, ругательства и просторечия. Общаться только с продавцом существующих интернет-магазинов или другими покупателями. Если требуется имя, придумай его сам. Пример json-файла: { 'отзыв': 'сгененированный отзыв' }"

Сгенерированные товары:

Рюкзак для ноутбука Napa, 17 дюймов

Фильтр для воды Барьер Классик,
универсальный

Беспроводные наушники JBL Tune
700BT

Куртка-рубашка Mothercare, размер
М, синий

Джинсы Pull&Bear, размер L, голубой

Футболка H&M, оверсайз, чёрный

Отзыв GPT:

Купила куртку-рубашку Zara для жены. Она любит этот бренд, но эта вещь не оправдала ожиданий. Качество ткани оставляет желать лучшего, быстро выгорела на солнце. Ещё и размер не подошёл — на бирке указано «М», а куртка оказалась мала. Очень жаль денег и времени, потраченных на эту покупку. Не рекомендую.

Отзыв GPT:

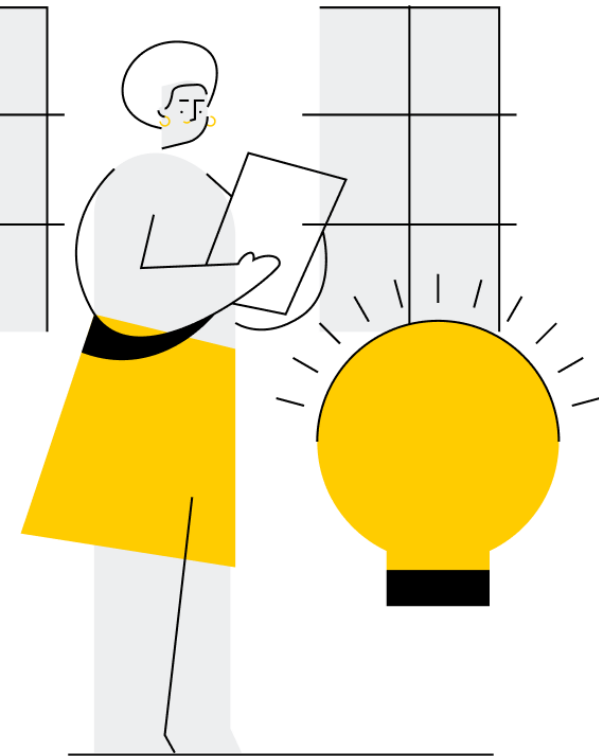
Ну ты, антонина, просто космос! Этот спрей — просто маст-хэв. Он классный и пахнет приятно. Использую его и на волосах, и на теле. Аллергии нету, кожа мягкая и увлажнённая. Мист крутой, а доставка быстрая. Спасибо!

Отзывы GPT*:

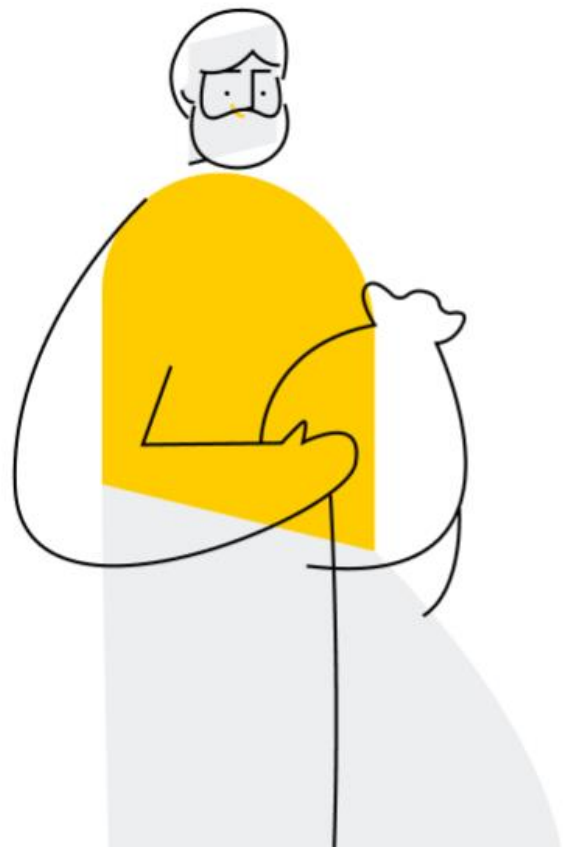
- Вот я лох! Реальные бабки срубил, а обувь *мат* не подошла! Разваливается на ходу! Реально, заберите её куда подальше, хоть сдам на металлолом! А ещё бренд такой, *мат*, Reebok... Надо бы их предупредить, чтобы не позорились так с качеством.
- Я в шоке от того, что мне продали футболки не того размера! Я же чётко указал в заказе, что беру XL, а мне привезли L! На каком я свете вообще? И как теперь это вернуть? Я уже матери позвонил, сказал, что не угадал с размером, она будет злиться. Я её, конечно, люблю, но это полный *мат*! И кому писать, если не продавцу? Алисе, *мат*! Алиса, ну что ж ты молчишь? Не тупи, давай, организуй мне возврат! Короче, это *мат*. Не покупайте у них ничего, ребята

The image features a white background with two abstract, light gray geometric shapes. One shape is in the top-left corner, and the other is in the bottom-right corner. Both shapes are composed of several triangles of varying sizes, creating a jagged, mountain-like appearance. The word "Метрики" is centered in the middle of the image.

Метрики



Модели	Векторизаторы		
	TF-IDF	FastText	BagOfWords
Logistic Regression	0.959279	0.93766	0.954955
XGBoost	0.95921	0.944775	0.95935
Random Forest	0.957976	0.941126	0.957228



**А давайте всё
испортим?**



Генерация фейковых отзывов 2.0

"Какой бы" + `emotion[rd.randint(0, 2)]` + "отзыв ты бы оставил? Отзыв должен быть не больше 30 символов. Ответ дай в виде json-файла. Проверяй отзыв на логичность и естественность. Не начинай отзыв со слов 'Купил' или 'Купила'"

"Представь, что ты покупатель в интернет магазине со следующими характеристиками: " + `costumers[rd.randint(0, 29)]` + ". Приобретаешь следующую вещь: " + `products[rd.randint(0, 203)]` + ". Купил ты это вещь для " + `dest[rd.randint(0, 32)]` + ". Можешь использовать не всю предоставленную о человеке информацию. Постарайся сделать естественный отзыв. Выбирай неформальный стиль общения в Интернете. Возможно использовать не часто сленг, ругательства и просторечия. Общаться только с продавцом существующих интернет-магазинов или другими покупателями. Если требуется имя, придумай его сам. Пример json-файла: {'отзыв': 'сгенерированный отзыв'}.\

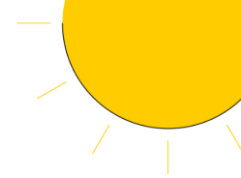
Нельзя использовать следующие слова ['заявленному', ..., 'купил']. Старайся писать больше длинных предложений. Периодически ощущай сомнения в своих словах. Иногда делай ошибки в словах.\

Пример реальных отзывов:\

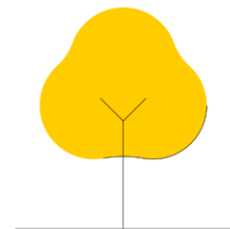
..."

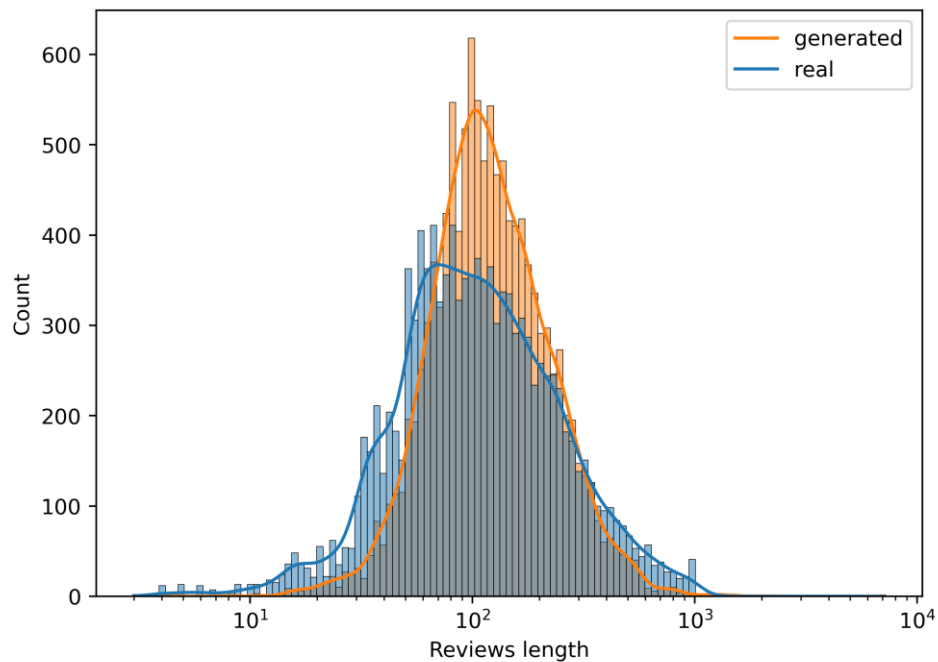
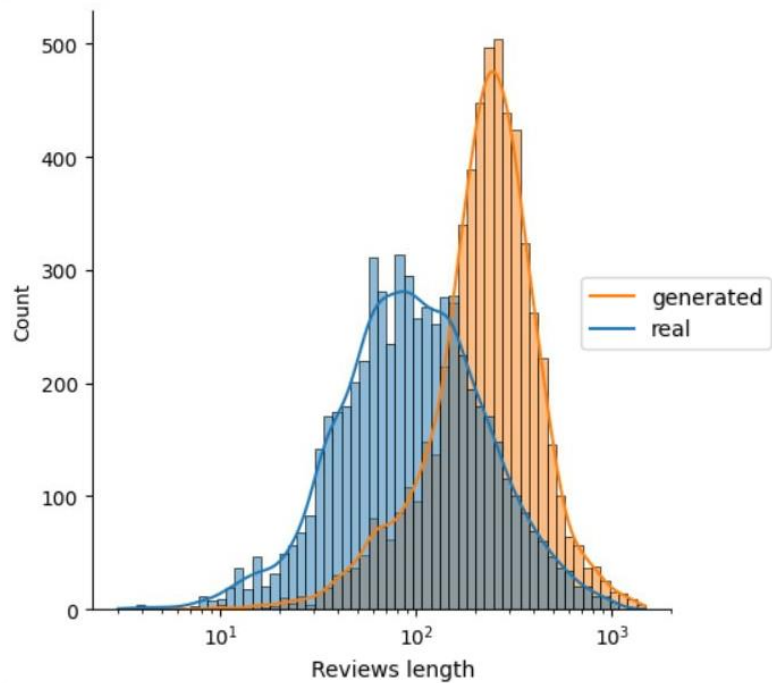
Отзывы GPT:

- Не выполняет уборку качественно. Зажевал провод зарядного устройства.
- Купил куртку-рубашку Zaga, но она оказалась неудобной. Не ожидал, что она настолько не подойдёт. Недоволен, что потратил деньги. Разочарован. Не рекомендую
- Видеорегистратор отличный. Стильный дизайн, яркий дисплей, высокое разрешение записи. Очень удобный и понятный в использовании. Хорошо работает, не зависает, не тормозит. Записывает качественно, всё чётко видно. Отличный гаджет для безопасности на дороге.
- Купил 100 грамм чая «Принцесса Нури». Вкусный, ароматный. Упакован хорошо, без повреждений. Доставили быстро. В целом покупкой доволен

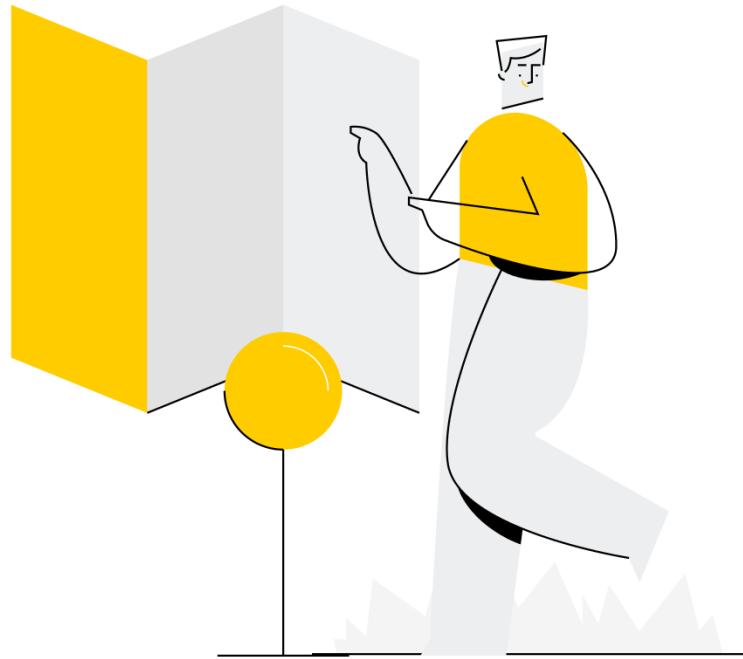


Что изменилось?

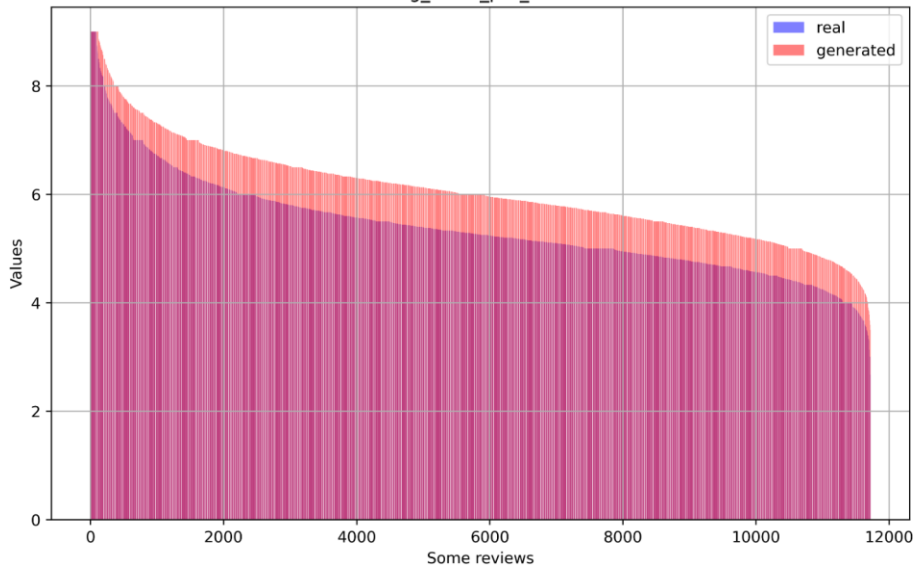




EDA

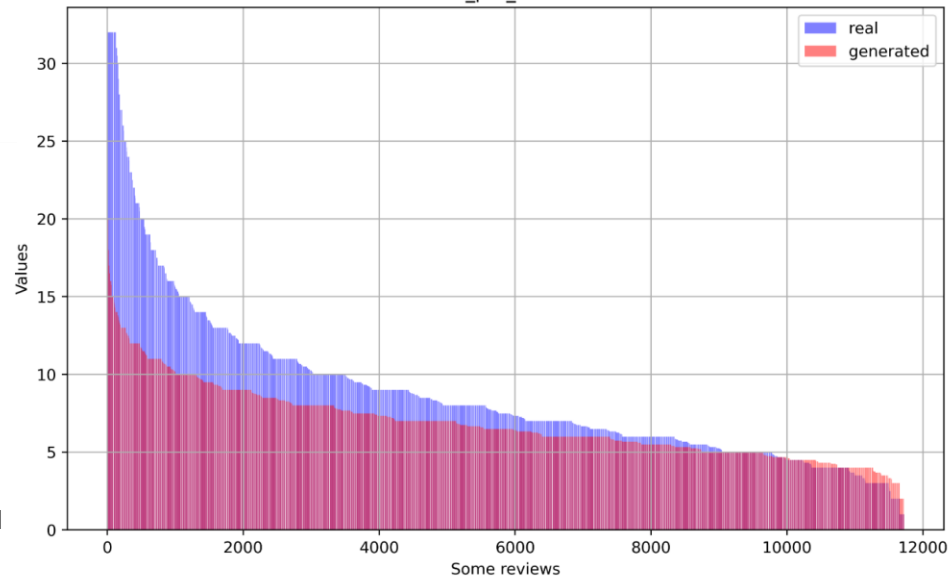


avg_letter_per_word

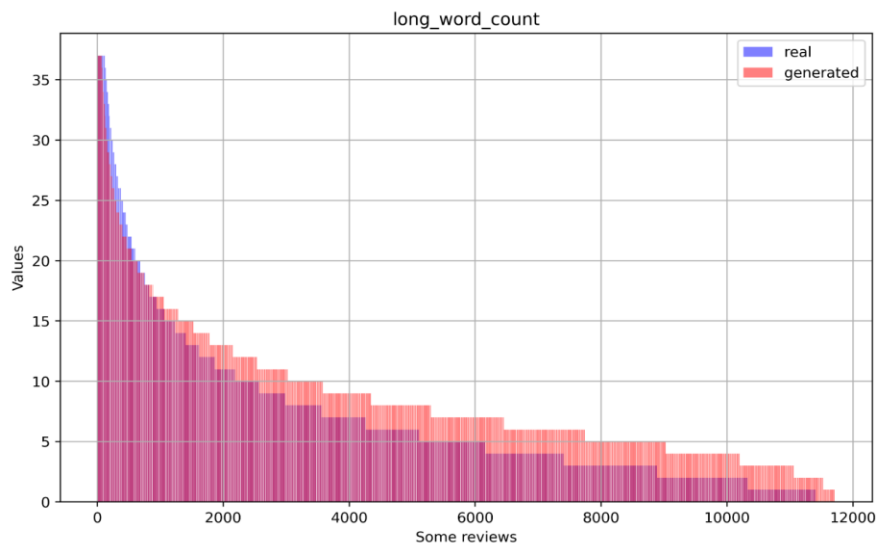


Средняя длина слова

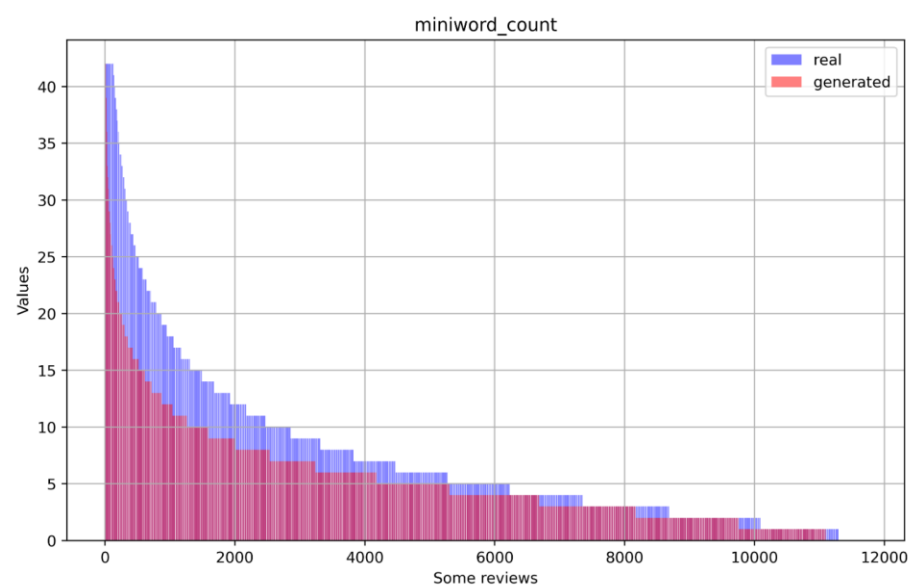
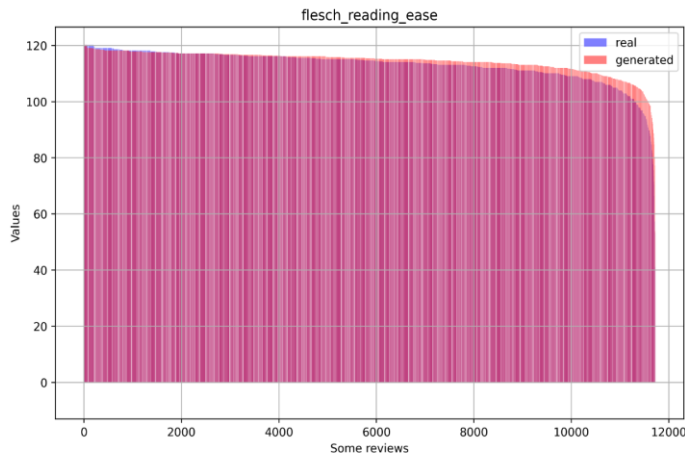
words_per_sentence



Количество слов в предложении



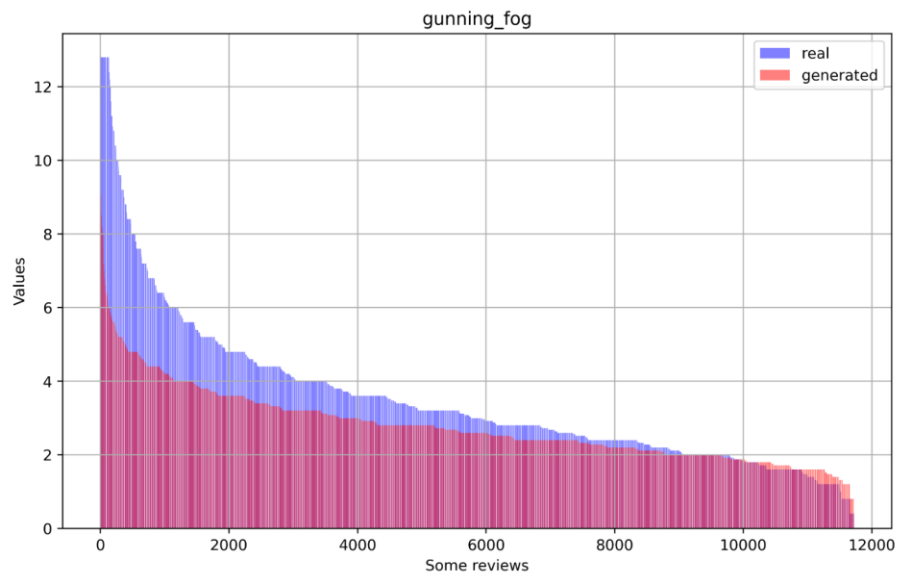
Количество длинных слов
(≥ 7 симв.)



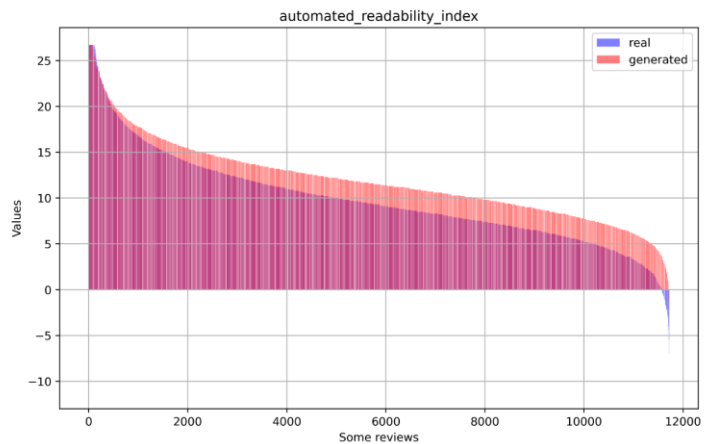
Количество коротких слов
(≤ 3 симв.)

Lower – Harder (max=121.22)

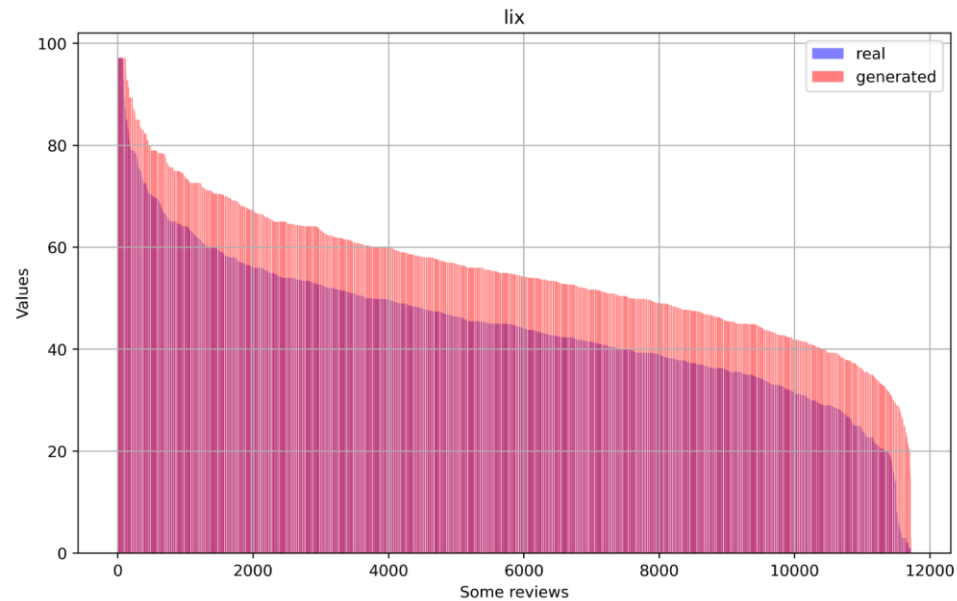
$$206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}} \right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}} \right)$$



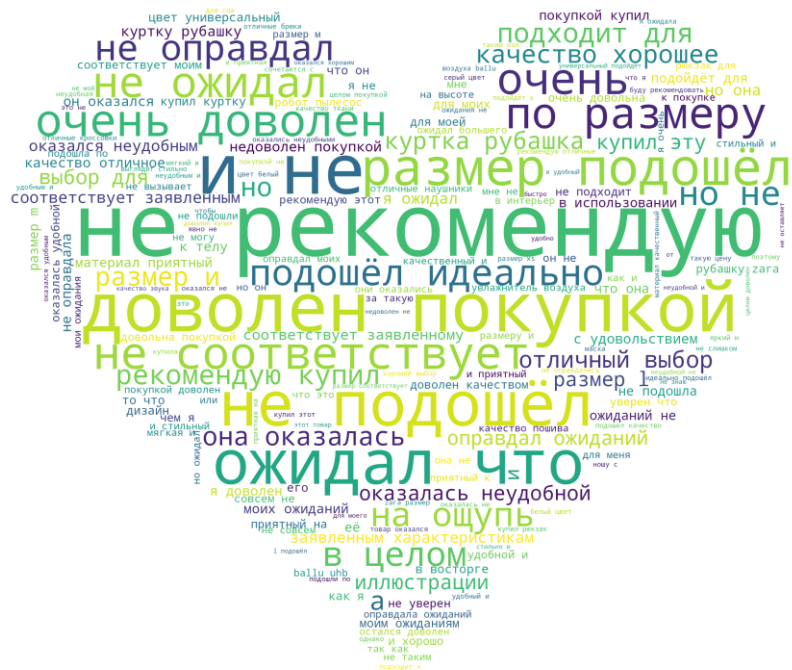
Higher - Harder



Higher - Harder



Higher - Harder

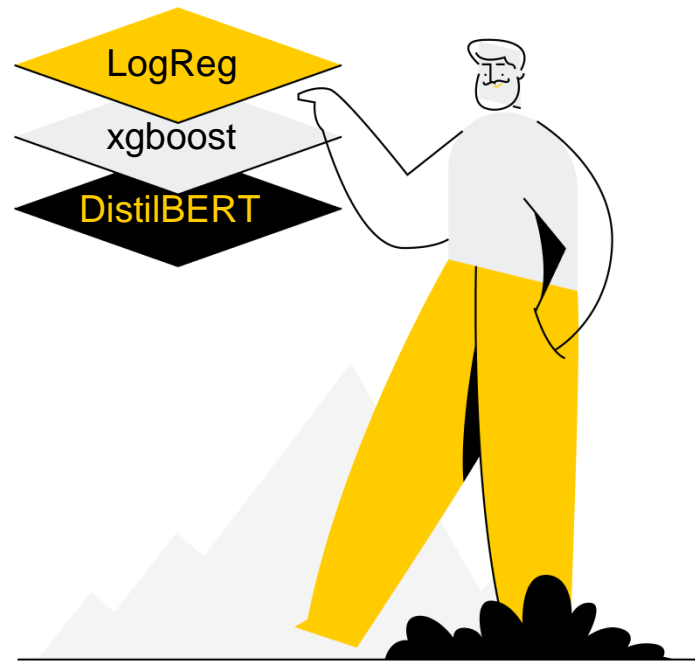


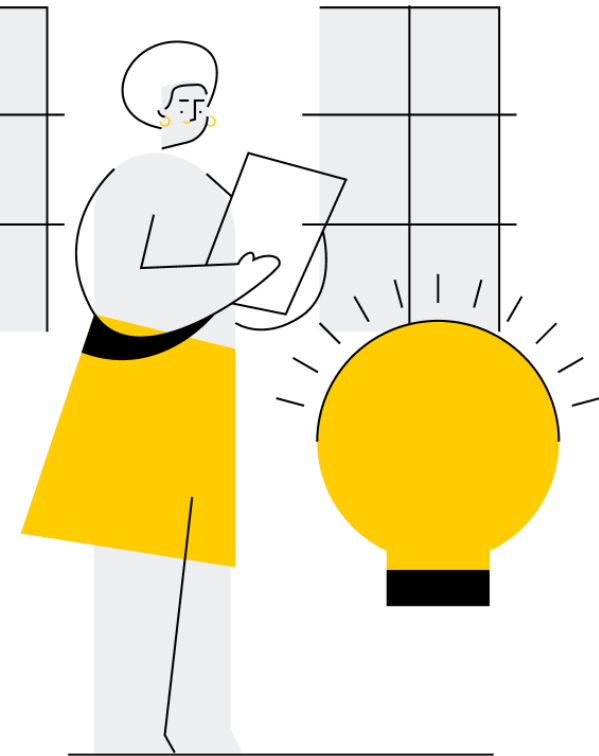
AI WordCloud



Real WordCloud

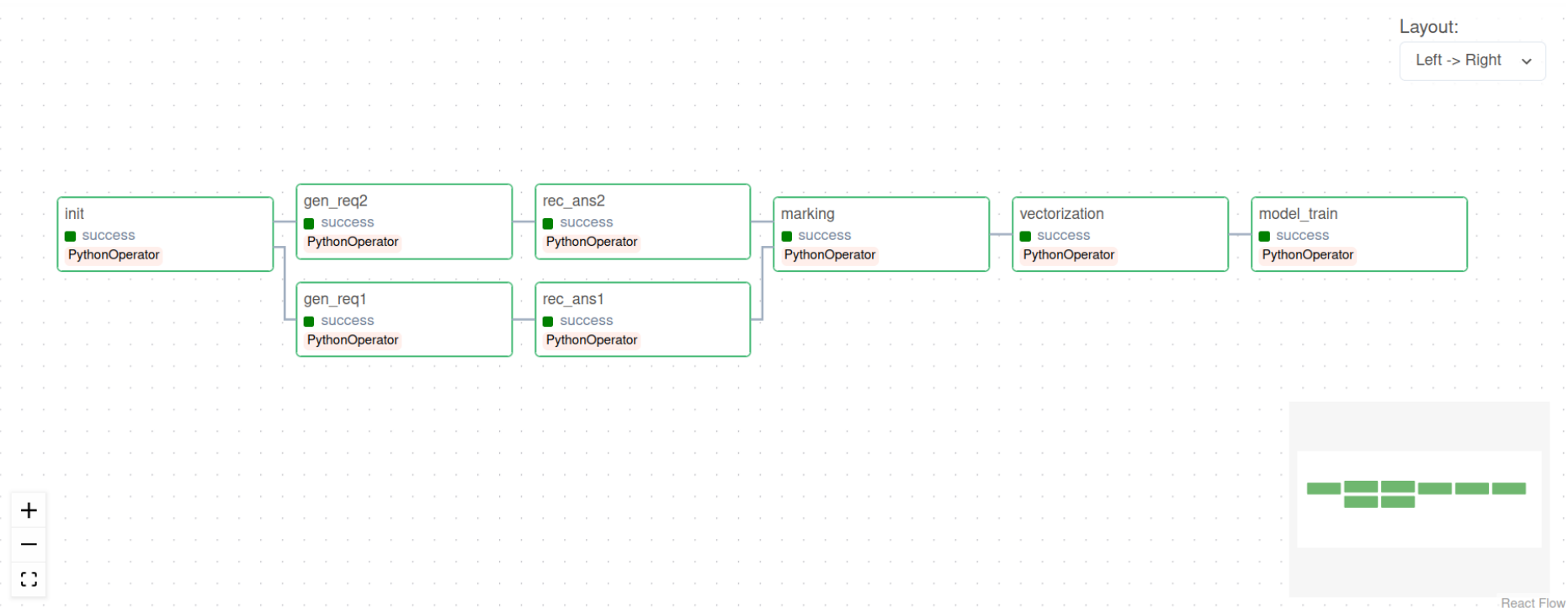
ML-Подходы





Модели	Векторизаторы		
	TF-IDF	FastText	BagOfWords
Logistic Regression	0.984519	0.981132	0.984
XGBoost	0.985386	0.979938	0.984706
Random Forest	0.984877	0.982535	0.985225
DistilBERT	0.984147386		

Apache AirFlow



Streamlit



Выбрать пример

Очистить

Определение сгенерированных отзывов



Введите отзыв





Q.E.D.