

## Описание задачи

1. Обучить модель на языке Python для классификации отзывов.
2. Разработать веб-сервис на базе фреймворка Django для ввода отзыва о фильме с автоматическим присвоением рейтинга (от 1 до 10) и статуса комментария (положительный или отрицательный).

## Датасет

Датасет содержит обзоры на фильмы взятые с сайта IMDB. Тестовая и обучающая выборки содержат по 25 тысяч обзоров. Для каждого обзора предоставлены оценка и статус. Однако разделение на негативные и позитивные обзоры происходило по принципу оценка  $\leq 4$  - негативный, оценка  $\geq 7$  - позитивный. Следовательно в датасете отсутствуют нейтральные отзывы (с оценками 5 и 6).

## Подход к решению задачи.

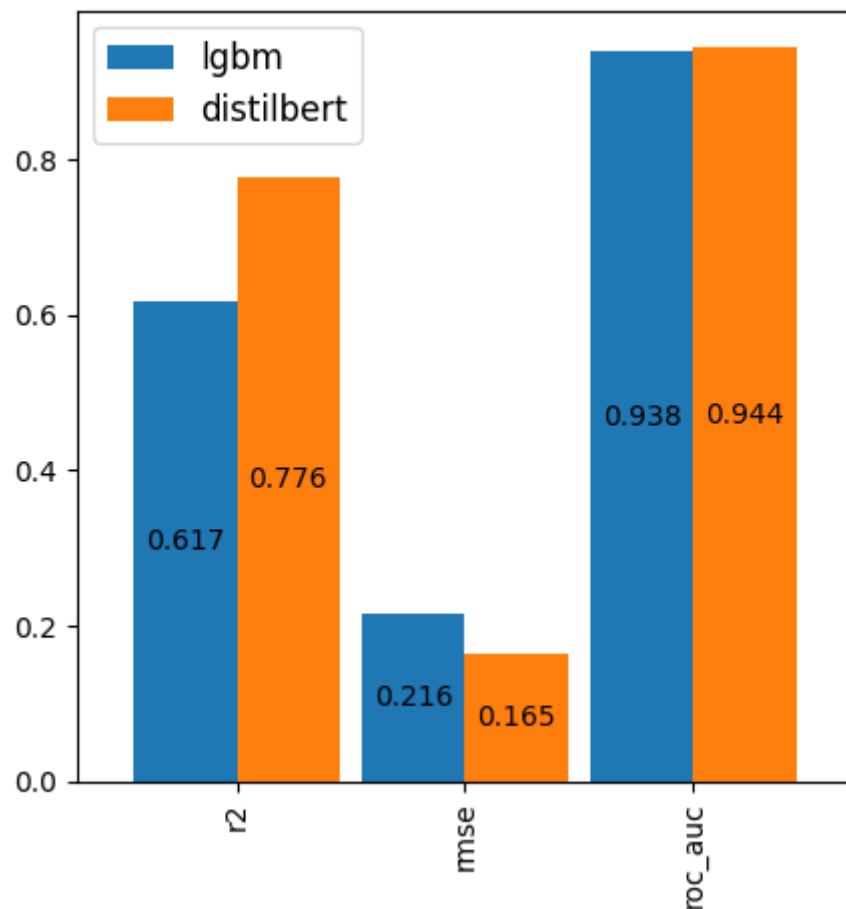
Задачу было принято решать как задачу регрессии по следующим причинам:

1. Отсутствие примеров для классов 5 и 6 (следовательно невозможность в будущем выставлять отзывам такие оценки)
2. Близость классов. В отличие от типичных классов, которые четко отделены друг от друга, в случае рейтинга разница между соседне-стоящими значениями незначительна, соответственно гораздо лучше, если модель будет это учитывать.

Для решения задачи было выбрано два метода - bag of words lgbm model и дообучение верхних слоев модели distilbert. Данные модели были выбраны с учетом ограничений располагаемых вычислительных возможностей (Так как прототип сервиса, использующего данные модели, должен быть размещен в открытом доступе, а серверы с гпу стоят дорого)

обучение моделей представлено в jupyter ноутбуках в гит репозитории.

## Результаты

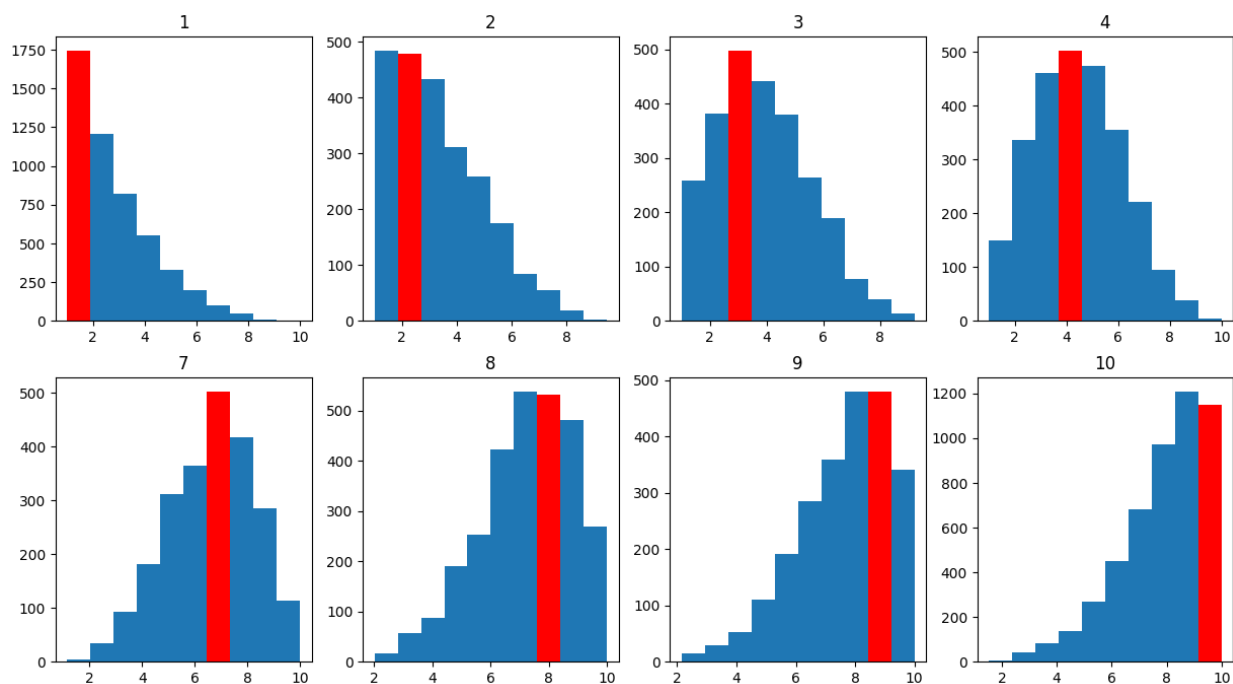


Сравнение двух моделей на тестовой выборке

Из рисунка выше видно четкое превосходство distilbert над BoW моделью. Нейросетевая модель гораздо четче попадает в нужный рейтинг о чем свидетельствует заметно больший коэффициент детерминации и заметно меньший rmse, означающий что в среднем мы отклоняемся от нужного рейтинга где-то на 1.5 звезды. Roc\_auc рассчитан для задачи sentiment analysis (обзоры с рейтингом меньше трешхолда считаются негативными), как можно заметить обе модели показывают примерно одинаковые результаты.

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.86      | 0.86   | 0.86     | 12500   |
| 1            | 0.86      | 0.86   | 0.86     | 12500   |
| accuracy     |           |        | 0.86     | 25000   |
| macro avg    | 0.86      | 0.86   | 0.86     | 25000   |
| weighted avg | 0.86      | 0.86   | 0.86     | 25000   |

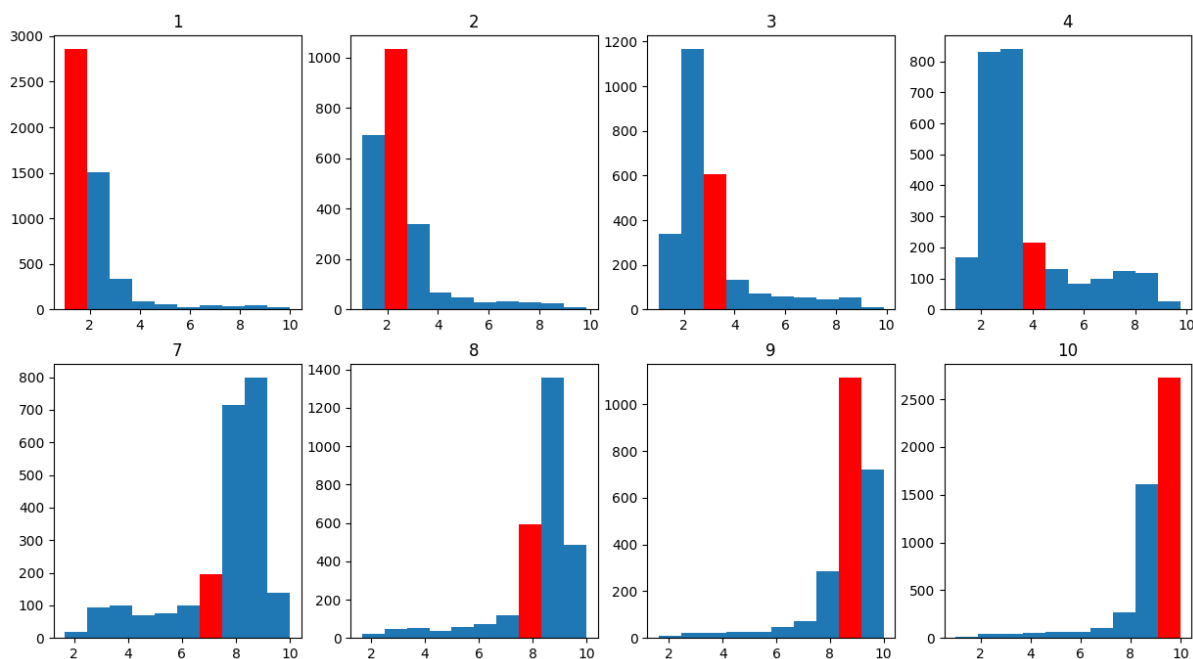
Результаты sentiment analysis для LGBM



Распределение предсказанных рейтингов по оценкам для LGBM

```
rating: 1, preds mean: 2.8, std: 1.57
rating: 2, preds mean: 3.33, std: 1.663
rating: 3, preds mean: 3.877, std: 1.645
rating: 4, preds mean: 4.436, std: 1.689
rating: 7, preds mean: 6.61, std: 1.63
rating: 8, preds mean: 7.231, std: 1.592
rating: 9, preds mean: 7.613, std: 1.56
rating: 10, preds mean: 7.887, std: 1.573
```

Как видно по распределению предсказанных рейтингов моделирование задачи как регрессии позволяет сконцентрировать большинство предсказаний около нужных нам рейтингов, по мере удаления рейтинга от нужного количество ошибочных предсказаний уменьшается. Однако чем лучше будет работать модель, тем острее должны быть пики в полученных распределениях, что и можно четко пронаблюдать на следующем рисунке. BoW модель не может справиться с неизвестными словами, которые не присутствовали на обучении, а также не способна учитывать контекст, как и порядок слов в предложениях.



Распределение предсказанных рейтингов по оценкам для DISTILBERT

```
rating: 1, preds mean: 2.0920000076293945, std: 1.305999994277954
rating: 2, preds mean: 2.5350000858306885, std: 1.3519999980926514
rating: 3, preds mean: 3.0250000953674316, std: 1.534000039100647
rating: 4, preds mean: 3.6670000553131104, std: 1.878000020980835
rating: 7, preds mean: 7.558000087738037, std: 1.7259999513626099
rating: 8, preds mean: 8.222000122070312, std: 1.4329999685287476
rating: 9, preds mean: 8.586999893188477, std: 1.2300000190734863
rating: 10, preds mean: 8.831000328063965, std: 1.3359999656677246
```

Как видно сеть более четко отделяет негативные и позитивные примеры, но при этом имеет склонность завышать позитивные оценки и занижать негативные, что четко видно на пограничных оценках таких как 4 и 7.

## Возможные улучшения

Так как подход к оцениванию фильмов у каждого человека индивидуален, то не для всех людей оценка 5-6 является позитивной, так же как и оценка 4 порой может нести позитивный настрой. Поэтому было бы правильнее использовать отдельную модель для решения задачи sentiment analysis в купе с моделью регрессии рейтинга.

Так же использовании более крупных моделей могло бы привести к улучшению работы сервиса, однако из-за ограниченных возможностей используемого vps на данный момент пришлось ограничиться относительно небольшими моделями.

## Прототип Веб Сервиса

Был разработан простой прототип сервиса с использованием фреймворка flask. Flask был выбран по причине моего предыдущего знакомства с ним. Примеры работы сервиса приведены ниже.

Сервис развернут на крайне слабом vps без наличия gpu, поэтому его быстродействие сильно ограничено.

Ссылка на сервис <http://82.146.60.164:5000/>

Ссылка на исходный код проекта, а также ноутбуки с обучением модели  
[https://github.com/daniltirsk/review\\_rating\\_prediction](https://github.com/daniltirsk/review_rating_prediction)

I am a newbie to the source material, but the first 3 episodes of this show blew me away! And, now having seen the whole first season, my rating remains even more solidly a 10! :-)

Groundbreaking/beautiful animation, awesome environments, spectacular cinematography, cool musical choice, compelling characters, great voice casting/acting, and some interesting Sci-Fi/Fantasy lore! :-)

YOU HAVE TO SEE THIS SHOW!!! :-)

Submit

Method: lgbm | Rating: 9 | Sentiment: Positive

Method: bert | Rating: 10 | Sentiment: Positive

Every scene tries to build up as if something amazing is just about to happen but then nothing ever happens. Plus a real mess of storylines and characters everywhere. Boring as crap, not for me.

If you want some good straight to the point animation just watch "Dark Fury".

Submit

Method: lgbm | Rating: 4 | Sentiment: Negative

Method: bert | Rating: 2 | Sentiment: Negative