

# Отчет о проделанной работе

---

1. Я предобработал данные, создал и обучил нейронную сеть для решения задачи классификации. (IMDBModelTraining.ipynb)
2. Создал single-page приложение на базе Django
3. Приобрел VPS, развернул приложение на удаленном сервере (<http://45.9.41.27:8000/>)
4. Предоставил отчет о работе модели на тестовых данных (IMDBTestAnalysis.ipynb)

## Приобретенные навыки

---

1. Научился использовать вычислительные мощности платформы Kaggle для обучения нейросетей и использовать обученные модели локально/в приложении
2. В сжатые сроки изучил фреймворк Django, смог поднять свое первое приложение.
3. Научился работе с VPS-сервером через терминал Bash

## Что хотелось бы улучшить

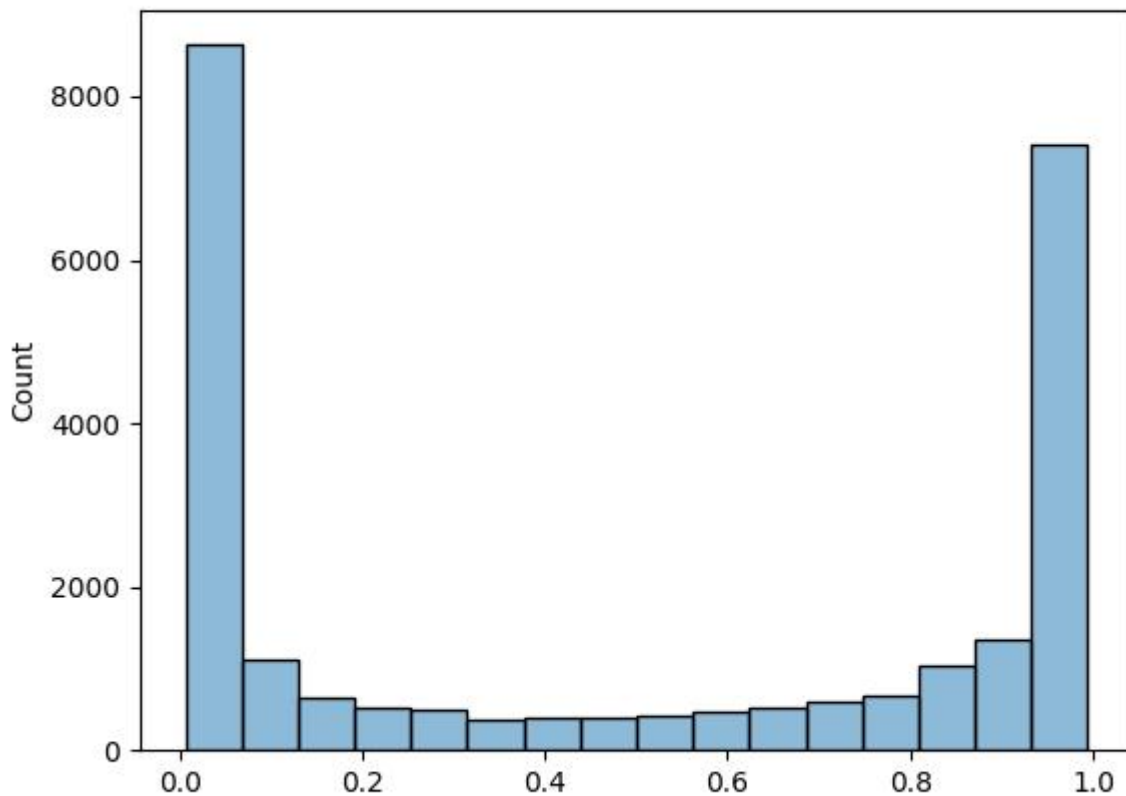
---

### 1. Точность работы модели

Точность 86% - это хорошо, но на данном датасете можно выжать и более высокую точность. Это можно было бы сделать, увеличив размерность эмбединга, прибавив количество слоев в нейронной сети, добавив несколько дополнительных модулей LSTM. Можно и вовсе использовать модели из семейства BERT. Однако, в таком случае, для обучения и работы нейросети понадобится гораздо больше ресурсов.

### 2. Уточнить шкалу от 1 до 10

В тренировочном датасете каждому отзыву сопоставлена отметка - positive или negative. Однако, в задании требовалось, чтобы модель ранжировала отзывы по шкале от 1 до 10. Выбрав функцию активации на последнем слое сигмоидой, мы можем интерпретировать результат работы модели как уверенность НС в том, что данный отзыв - положительный. Результат всегда будет принадлежать отрезку  $[0,1]$ , следовательно, домножив результат на 10 и округлив, мы можем получить оценку по десятибалльной шкале. Тем не менее, модель стремится дать ответ как можно ближе к 0 или 1, поэтому оценки часто бросаются из крайности в крайность.



Чтобы исправить такое поведение, нужно каждому отзыву в тренировочной выборке присваивать именно оценку от 1 до 10, а не только бинарный категориальный признак. Но, скорее всего, при таком подходе понадобится больше данных для адекватного обучения модели.

### 3. Работа над приложением

Белая страничка с полем для ввода и кнопкой - это не верх мечтаний. Главная страница приложения была создана исключительно для проверки работоспособности, и конечно, хотелось бы её улучшить. Все же, я считаю, что такими делами должен заниматься frontend-разработчик, поэтому я не стал наводить красоту в таком базовом проекте.

К backend тоже могут быть вопросы: приложение работает через команду `runserver`, что, мягко скажем, не является рекомендуемым способом деплоить проект, но надеюсь, что для тестового варианта такой путь сойдет.