Задание 1. Сравнительный Анализ моделей

Материалы:

* Размеченный Датасет (balanced\_dataset.csv)
* Колаб с LSTM: https://colab.research.google.com/drive/13BJN-eWshcKx5yA7Gb5Bu75n-0kMGhXi?usp=sharing
* Колаб с остальными моделями

Что нужно:

* Выявить метрики на векторах glove, fasttext(<http://vectors.nlpl.eu/repository/>), tfidf с помощью классических методов регрессии (полный список в from sklearn.utils import all\_estimators)
* Поэкспериментировать с кол-вом слоев и их размером в LSTM и выявить наилучшие метрики, попробовать комбинировать bidirectional и unidirectional слои
* Записать получившиеся лучшие значения метрик и параметры использованных моделей в таблицу

Пример:A black and white text with numbers

Description automatically generated

Задание 2. Сделать сайт обертку для пайплайна модели

Функции:

* Загрузка текстовых файлов (drag n drop)
* Вывод предсказаний при нажатии на кнопку (по нажатию текст разбивается на пары предложений, и они скармливаются берту)

Пример:

A screenshot of a computer

Description automatically generated

A screenshot of a question answer

Description automatically generated