

Модель классификации комментариев по работе УК



самолет

Работу выполнил студент групп ИСП-22:

Хусаинов Марат

Преподаватель:

Коновалов Игорь Васильевич



О кейсе

Постановщик задачи:

Компания «Самолет»

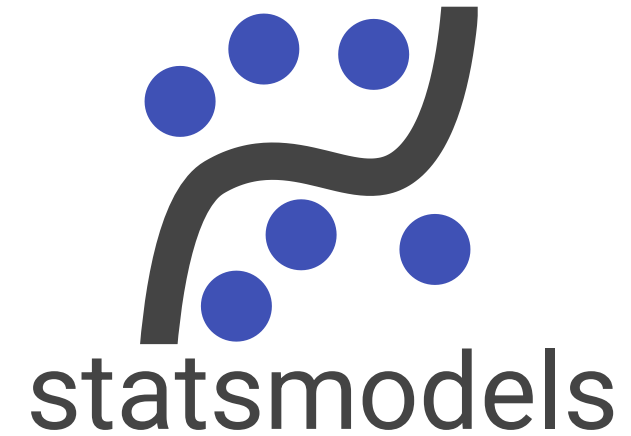
Название кейсового задания:

Классификация позитивных
комментариев

Цель: Разработать модель классификации
комментариев по работе УК на категории



Стек технологий



самолет



Этапы работы

1

Разбор
технического
задания

2

Подготовка данных

3

Построение
моделей

4

Изучение
и сравнение
полученных
решений

5

Понять, что тз
понято не так и
изменений не
терпит

6

Обучение,
настройка и
оптимизация
итоговой модели

7

Получение
и интерпретация
итогового
результата

8

Создание
презентации



Обработка и подготовка данных



Удаление дублей

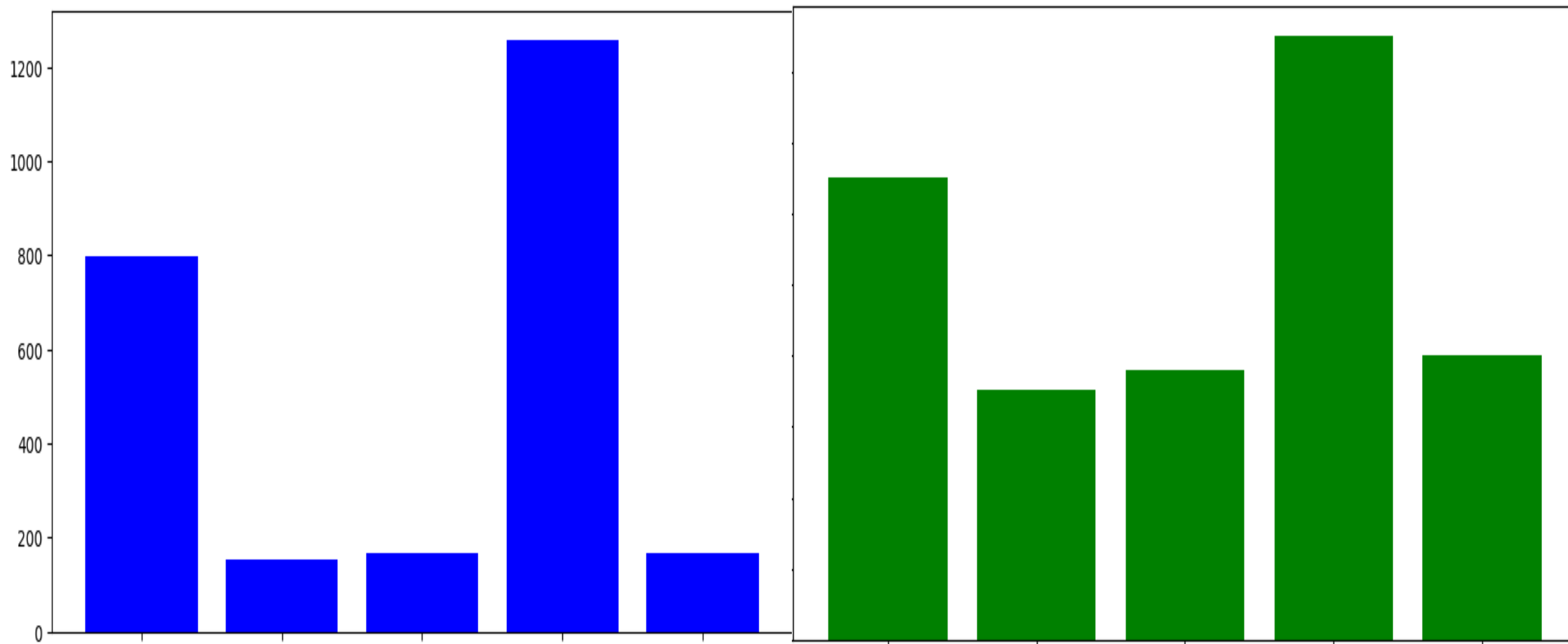


Разметка по категориям



Попытка сбалансировать классы

Использование nlpaug для генерации комментариев недостающих классов



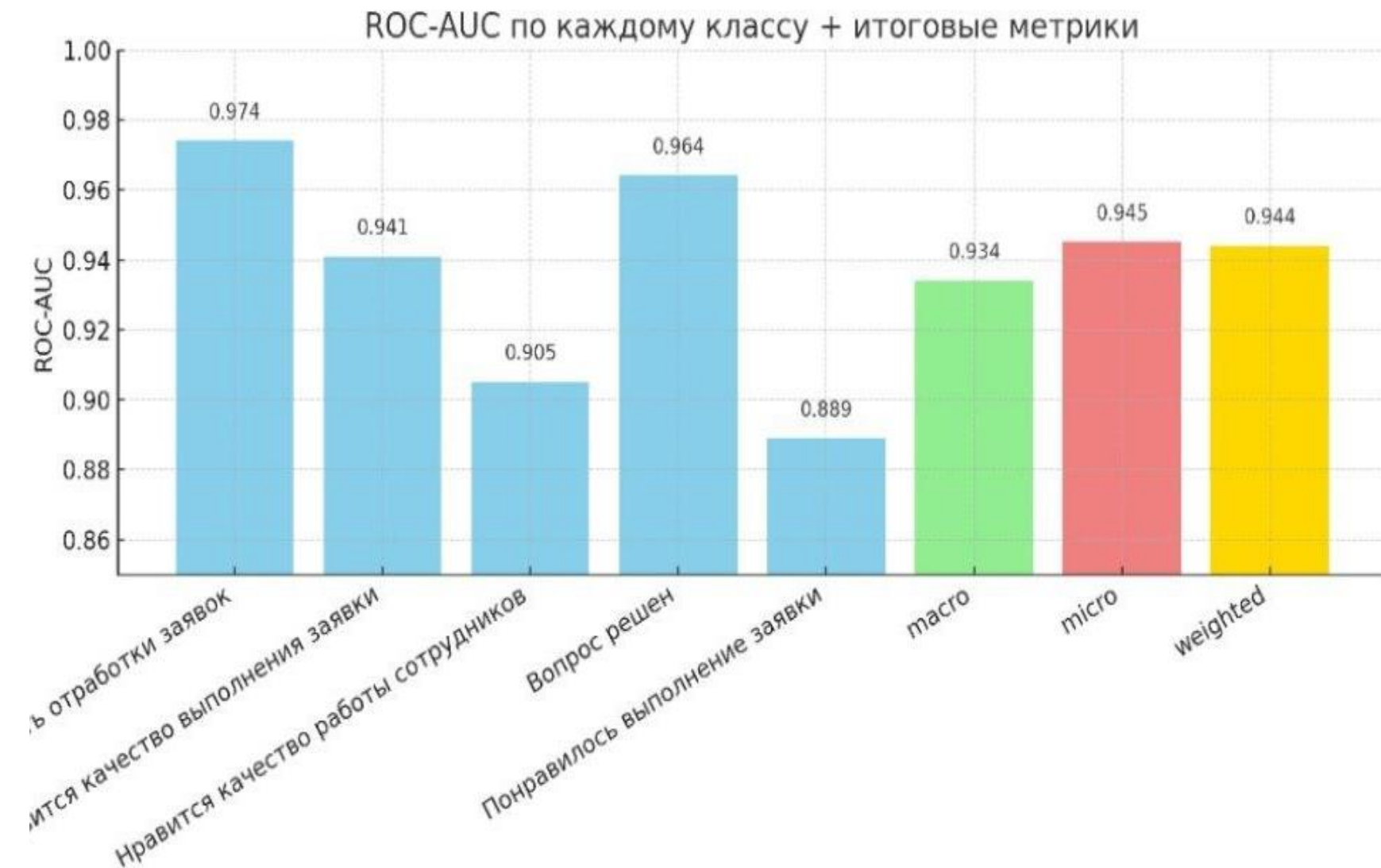
самолет

Выбор модели

Выбранная модель: bert-base-multilingual-cased для задач мультиклассовой классификации, с размером батча 8



Оценка качества модели



ROC-AUC по каждому классу:

скорость отработки:
0.974

Нравится качество
выполнения: 0.941

Нравится качество
работы сотрудников:
0.905

Вопрос решен: 0.964

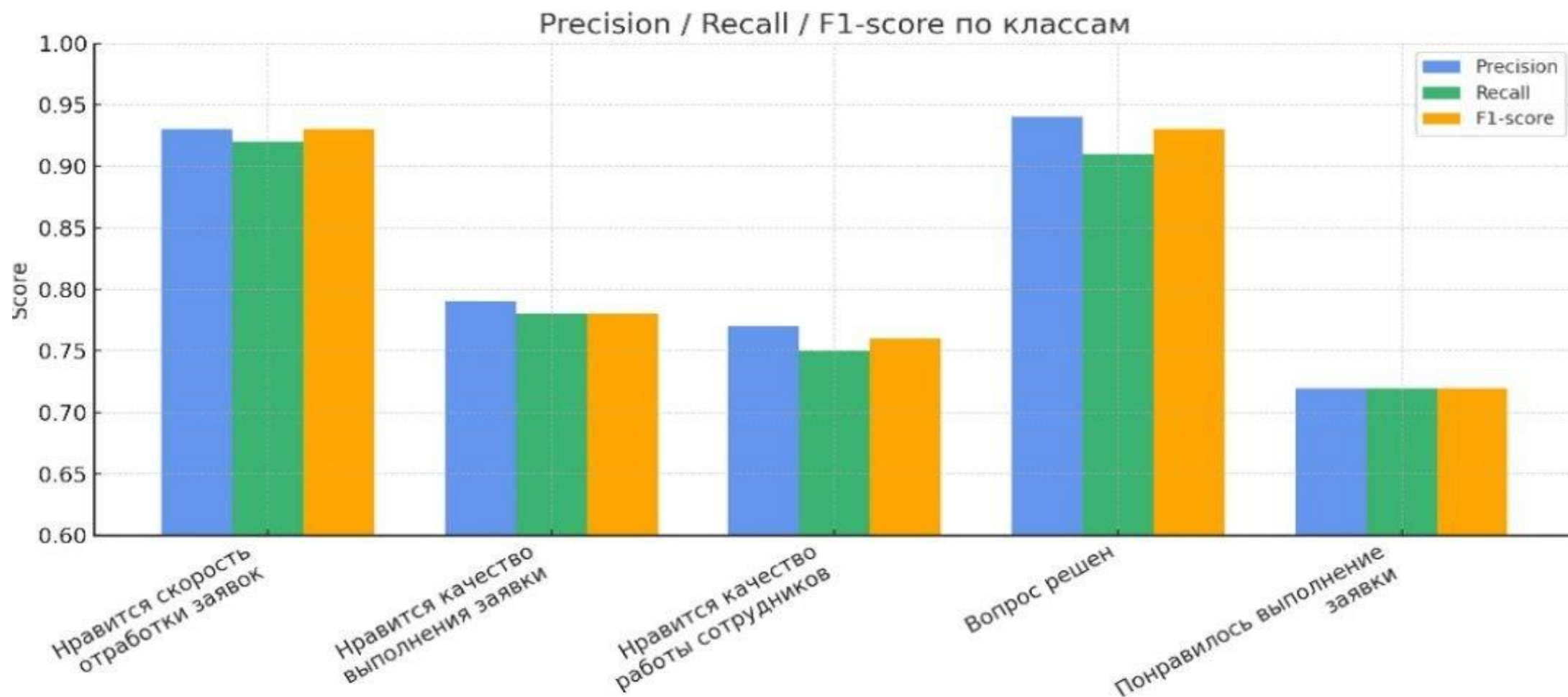
Понравилось
выполнение заявки:
0.889

ROC-AUC (macro): 0.934

ROC-AUC (micro): 0.945

ROC-AUC (weighted):
0.944

Оценка качества моделей



Вывод по проделанной работе

- Получилось модифицировать датасет для лучших результатов
- Получена работоспособная модель работающая четко по ТЗ и выдающая неплохие метрики
- Все классно сделано я считаю (поставьте 5ку пожалуйста)

