КОММЕНТАРИИ КОМАНДЫ КУРСА

Антон, здравствуйте! Вами подготовлен проект, который заслуживает высокие баллы. Все пункты подробно раскрыты. Обсудим основные достоинства и рекомендации по представленным результатам.

Что сделано хорошо:  
- Информация об отелях и отзывах анализируется по отдельности, для отелей получено (создано) много новых признаков: страна и город отеля. Конечно, эффективность признаков была проверена в работе. Используются данные из внешних источников. Например, используется информация о координатах центров городов. Был выполнен подробный анализ тегов.  
- Очень подробно анализируется текстовые поля с отзывами, что для данной задачи можно назвать лучшей стратегией для улучшения качества признаков для итоговой модели. Выполнен анализ слов, которые позволяют негативно или положительно оценить контекст отзыва. Используется пакет библиотек и программ для символьной и статистической обработки естественного языка (nltk), что можно назвать best practice.  
- Выполнен отбор признаков. Анализ мультиколлинеарности (.corr) и значимости признаков (f\_classif). Результаты анализа для наилучшего восприятия представлены в виде графиков.  
- Полученная модель наилучшим образом позволяет улучшить значение MAPE для базового шаблона (baseline). Считаю, что получен отличный результат по улучшению. Проект полностью оформлен. Решение представлено на kaggle и github (2-версии). О возможных способах улучшения мы поговорим на итоговом вебинаре.

Что можно улучшить:  
- Проект очень хорошо подготовлен. Есть небольшая рекомендация. Существует такое понятие, как векторизация слов. Каждое слово может быть представлено отдельным числом. Для этого можно использовать библиотеку gensim. Использование Doc2Vec. Подобный подход позволяет нам получить множество дополнительных признаков. Подробнее об этом мы поговорим на итоговом вебинаре.

Антон, работа выполнена! Считаю, что она заслуживает максимальный балл. Обязательно продолжайте улучшать свои навыки. Удачи в следующих проектах.  
Проверку выполнил ментор Сергей Добдин. По любым вопросам, пожалуйста, пишите в общий канал slack [#03\_project-3].  
Запись итогового вебинара:  
https://us02web.zoom.us/rec/share/qoBLJFQpQHhkoaFS77dl0EkWRRvy226xr1pHO7b7Cz6cwJN8Y148vaJQ-DDCEEz6.aFGotJSRNbS59tUP?startTime=1663866000000  
Код доступа: +5s72pfL