Владислав, здравствуйте! Работа выполнена и заслуживает высокой оценки! Особо хочется отметить создание большого количества полезных функций для обработки данных. Далее немного подробнее.

Что сделано хорошо:  
- Безусловным достоинством работы является подготовка данных и создание функций, позволяющих реализовать новые признаки (feature engineering). Особенностью реализации функций является соблюдение стандартов PEP. Функции в обязательном порядке содержат подробный комментарий. Правильное комментирование является важным навыком, т.к. подробное описание всегда упрощает понимание кода. Обязательно не забывайте об этом в будущем.  
- В работе подробно анализируется отзывы, что для данной задачи можно назвать лучшей стратегией. Выполнен анализ слов, которые позволяют негативно или положительно оценить контекст отзыва. Отмечу, что негативные и положительные отзывы анализируются по отдельности. Используется пакет библиотек и программ для символьной и статистической обработки естественного языка (nltk), что можно назвать best practice для данной задачи.  
- Используется кодировка признака с использование LabelEncoder() и BinaryEncoder(). Выполнен анализ и значимости признаков (f\_classif). Можно было дополнительно проверить влияние нормализации и стандартизации на результаты улучшения модели.  
- Полученная модель позволяет достаточно хорошо улучшить значение MAPE для базового шаблона (baseline). Многие очень часто стараются существенно улучшить этот показатель, что является небольшой ошибкой. Основная задача проекта в данном случае связана с предобработкой данных. Проанализировать и улучшить именно алгоритм машинного обучения вам предстоит в следующих проектах. Хотя ещё раз скажу, что в вашем проекте результат получился достойный.

Что можно улучшить:  
- Владислав, проект нужно сохранить на github в своём аккаунте. Это нужно не только для анализа решений ментора, но и для подготовки будущего портфолио. Обязательно добавьте проект в свой репозиторий и настройте его согласно всем правилам. Есть способ трансляции блокнота kaggle на github: https://www.kaggle.com/product-feedback/295170  
- Лучшей стратегией для данной задачи является обработка отзывов. Можно было использовать специальные библиотеки, которые позволяют получить много дополнительной информации. Существует такое понятие, как векторизация слов. Каждое слово может быть представлено отдельным числом. Для этого можно использовать библиотеку genism, класс Doc2Vec. В итоге, получите множество новых полезных признаков, которые созданы на основе отзывов. Подробнее об этом поговорим на итоговом вебинаре. Если есть желание разобраться в этом самостоятельно, то рекомендую вот [Б. Бенгфорт, Р. Билбро, Т. Охеда Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка. СПб.: Питер, 2019. 368 с.].

Считаю, что работа выполнена может быть оценена максимальным баллом! Некоторые недостатки скомпенсированы очень грамотно выполненным блоком по созданию новых признаков. Всё-таки не забудьте разместить проект на github. Искренне желаю удачи в будущих блоках и проектах!  
Проверку выполнил ментор Сергей Добдин. По любым вопросам, пожалуйста, пишите в общий канал slack [#03\_project-3].  
Буду ждать вас на итоговый созвон-вебинар по текущему проекту 2 июня в четверг в 20:00 (Мск). На нём мы разберём основные ошибки группы и посмотрим эталонное решение. До встречи на вебинаре!