

# Отчет

## Оценка качества вина Vinho Verde

Докладчик: Нурмухамедов Р.

# Содержание

- Описание задачи и используемые данные
- Основные переменные
- Распределение вин по качеству
- Используемые модели обучения
- Выводы

# Описание задачи и используемые данные

- Сложность определения качества вин ручным способом
- Использование машинного обучения для предсказания качества вин
- Предварительно подготовленные данные из Kaggle\*

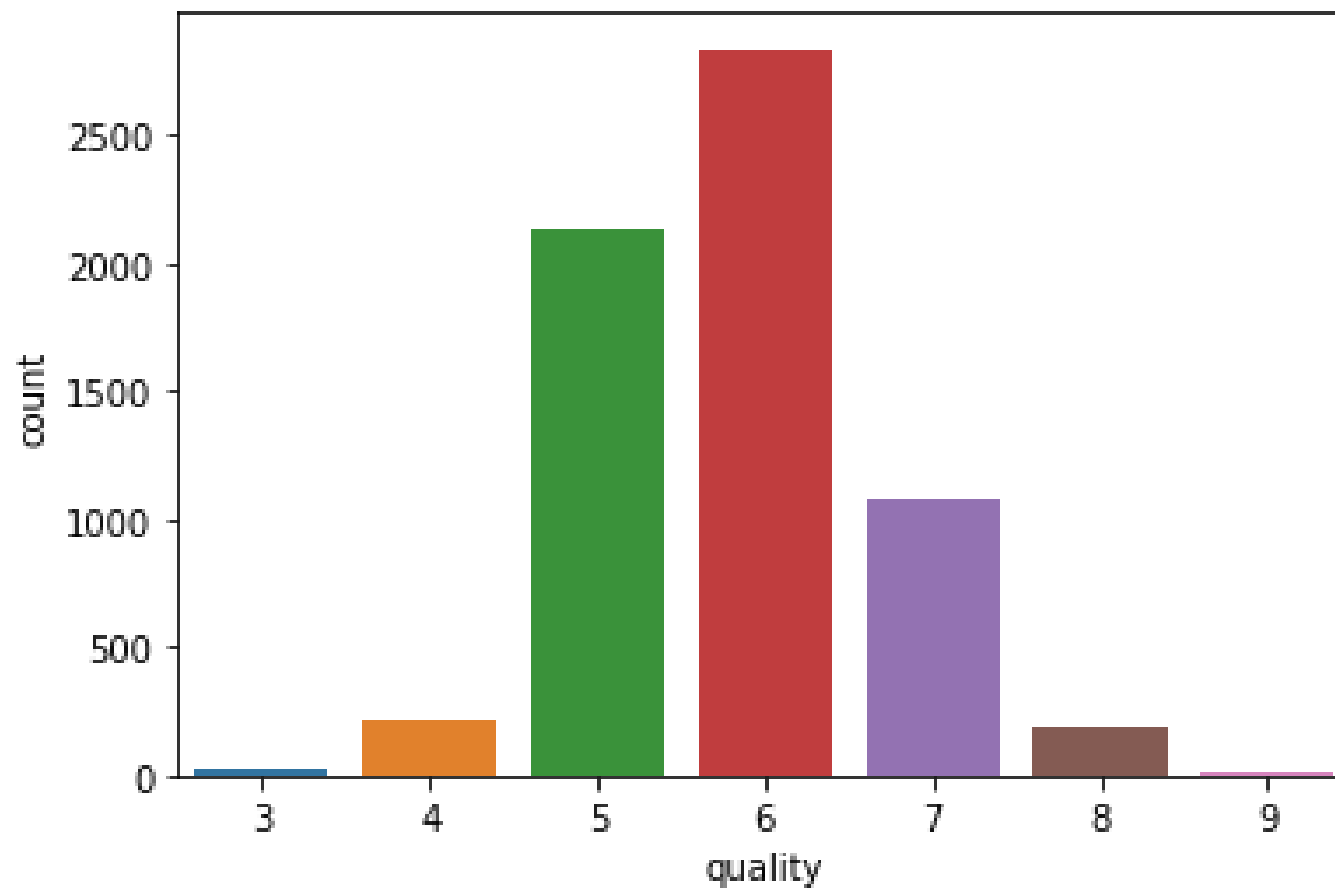
\* <https://www.kaggle.com/rajyellow46/wine-quality>

# Основные переменные\*

- фиксированная кислотность
- летучая кислотность
- лимонная кислота
- остаточный сахар
- хлориды
- свободный диоксид серы
- общий диоксид серы
- плотность
- pH
- сульфаты
- алкоголь

\* Для получения дополнительной информации прочитайте [Cortez et al., 2009]

# Распределение вин по качеству



# Используемые модели обучения

- LogisticRegression
- RandomForestClassifier
- Дефолтные настройки моделей
- Accuracy для оценки качества обучения

# Выводы

	Accuracy на train	Accuracy на test
LogisticRegression	46.3%	45.7%
RandomForestClassifier	<b>67.3%</b>	<b>50.3%</b>

Ссылка на ноутбук: [https://github.com/RuslanNT/code/blob/main/DZ\\_2\\_wine-quality-prediction.ipynb](https://github.com/RuslanNT/code/blob/main/DZ_2_wine-quality-prediction.ipynb)