

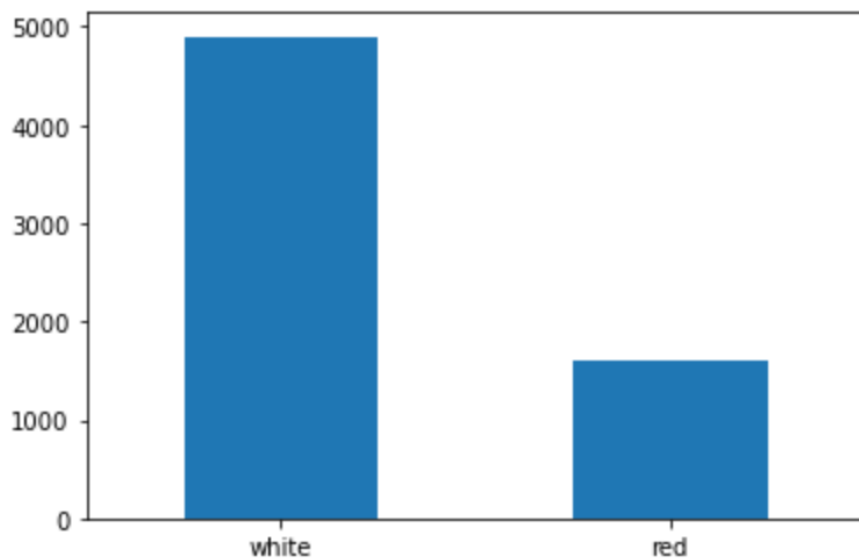
Определение качества вина по его объективным характеристикам

Целью исследования являлась проверка гипотезы о том, что по объективным характеристикам вина можно предсказать с высокой точностью оценку качества, которую оно получит от экспертов.

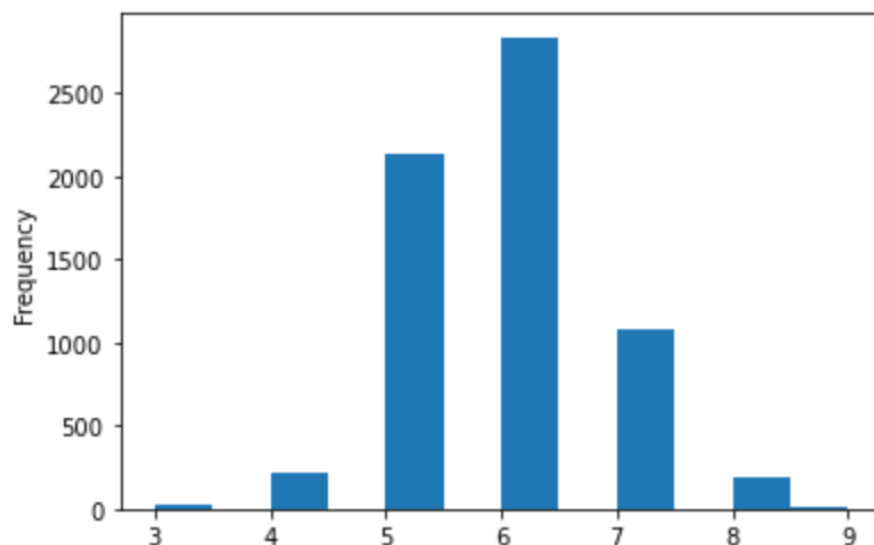
Исходные данные

В качестве данных для исследования были взяты результаты лабораторных анализов белых и красных вин “Vinho Verde”, а также их оценки по 12-балльной шкале.

Вина по классам распределены неравномерно, информации о красных винах значительно меньше, чем о белых



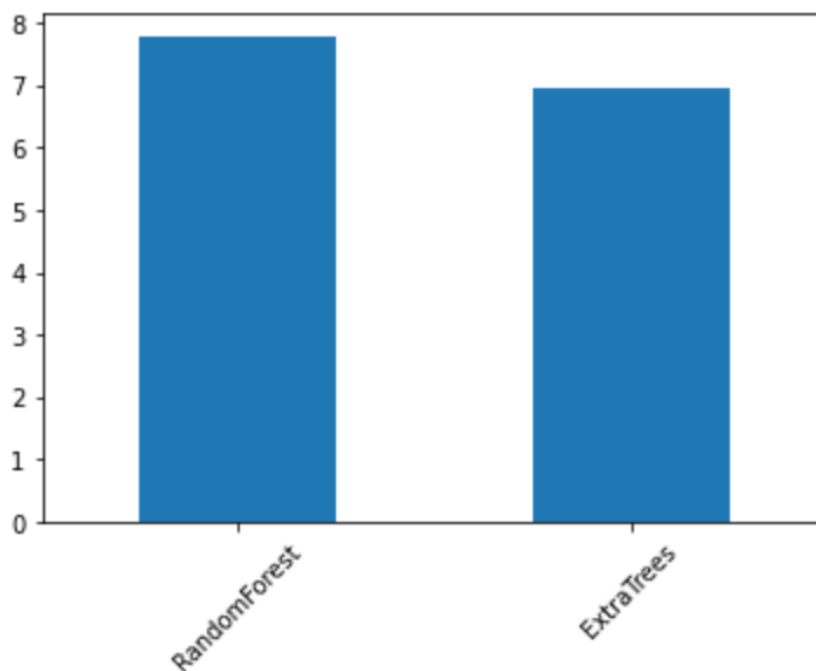
Видно, что эксперты избегают минимальных и максимальных оценок, большая часть оценок сосредоточена вокруг значений 5, 6 и 7



Данные были качественно собраны и, за редким исключением, покрывали весь спектр собираемых метрик. Информация о винах с неполными данными была удалена для чистоты эксперимента.

Результаты исследования

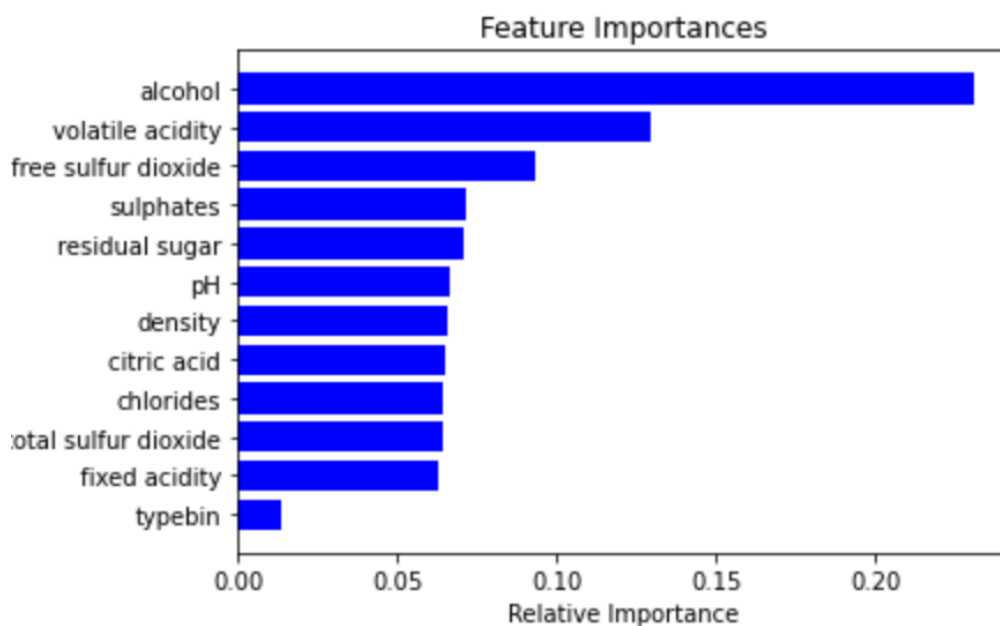
Первоначально выбранная модель на базе случайного леса дала невысокую точность, которую явно можно было улучшить. Необходимое улучшение дала модель «Extra Trees»



Учитывая ошибку менее 7% можно с уверенностью утверждать, что существует чёткая взаимосвязь между лабораторными метриками вина и его экспертной оценкой

One more thing...

Изучение важности разных параметров модели наводит на мысль, что определяющим фактором качества вина, с огромным отрывом от других параметров, является его крепость



Выпьем же за это! 😊