Московский государственный технический университет

им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Системы обработки информации и управления»



**Рубежный контроль № 1**

**По курсу «методы машинного обучения в АСОИУ»**

**Выполнил:**

студент ИУ5-24М

Ширшов А.С.

**Проверил:**

Гапанюк Ю.Е.

Подпись:

29.02.2024

Москва, 2024

**Задание**

* Для набора данных проведите нормализацию для одного (произвольного) числового признака с использованием функции "обратная зависимость - 1 / X".
* Для набора данных проведите процедуру отбора признаков (feature selection). Используйте метод обертывания (wrapper method), алгоритм полного перебора (exhaustive feature selection).
* Для произвольной колонки данных построить график "Скрипичная диаграмма (violin plot)".

**Ход работы**

Для выполнения работы возьмем набор хорошо известный тестовый набор данных про вино.

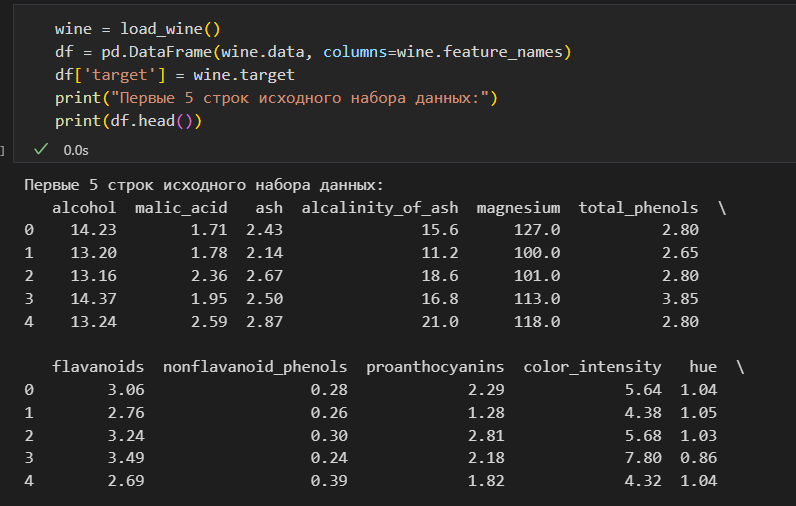


Рисунок 1 - Вывод набора данных

Построение обратной зависимости для поля - “Алкоголь”

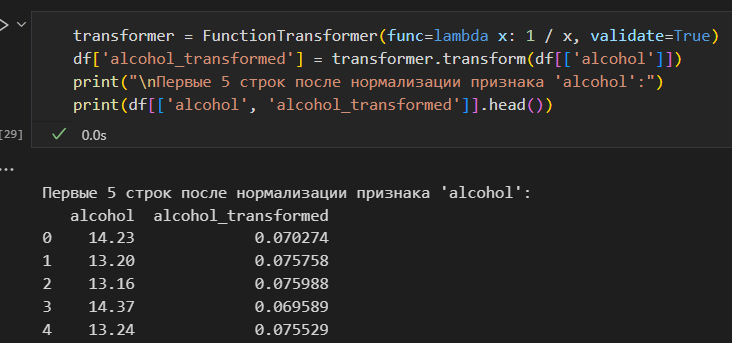


Рисунок 2 - Нормализация числового признака при помощи обратной зависимости

Для набора данных проведём процедуру отбора признаков (feature selection). Алгоритм полного перебора (exhaustive feature selection)

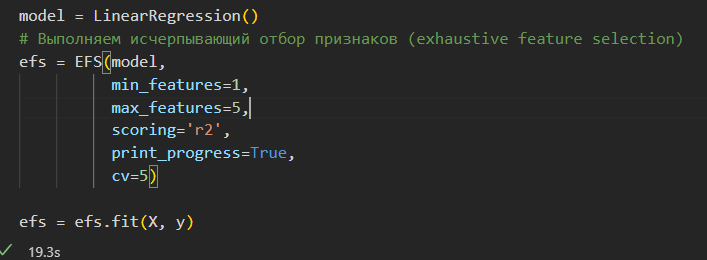


Рисунок 3 - Использование EFS

Ниже представлены итоговые полученные признаки.

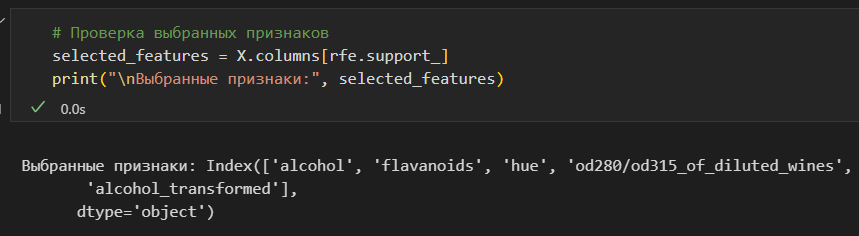


Рисунок 4 - Итоговые полученные признаки

Также выполним небольшое доп. задание. Построим violin plot для поля alcohol.



Рисунок 5 - График violin plot

**Вывод**

В ходе выполнения данного задания были проведены шаги по предварительной обработке и анализу набора данных, такие как нормализация числовых признаков, отбор наиболее значимых признаков и визуализация данных с помощью графиков. Они могут значительно улучшить качество и точность моделей машинного обучения, которые будут использоваться для анализа данного набора данных в будущем.