Московский государственный технический университет им. Н.Э. Баумана  
Кафедра «Системы обработки информации и управления»

****

Рубежный контроль №1  
по дисциплине  
«Методы машинного обучения»

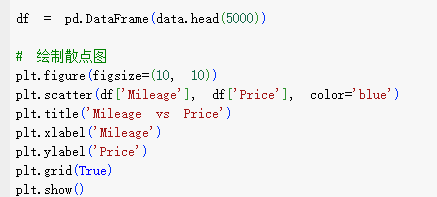
Выполнил:  
студент группы ИУ5И-21М  
Ван Чжэн

Москва — 2024 г.

**Вариант:№.16**

Я выбрал набор данных-«продажа подержанных автомобилей».

**Дополнительные требования по группам:**  для пары произвольных колонок данных построить график "Диаграмма рассеяния".



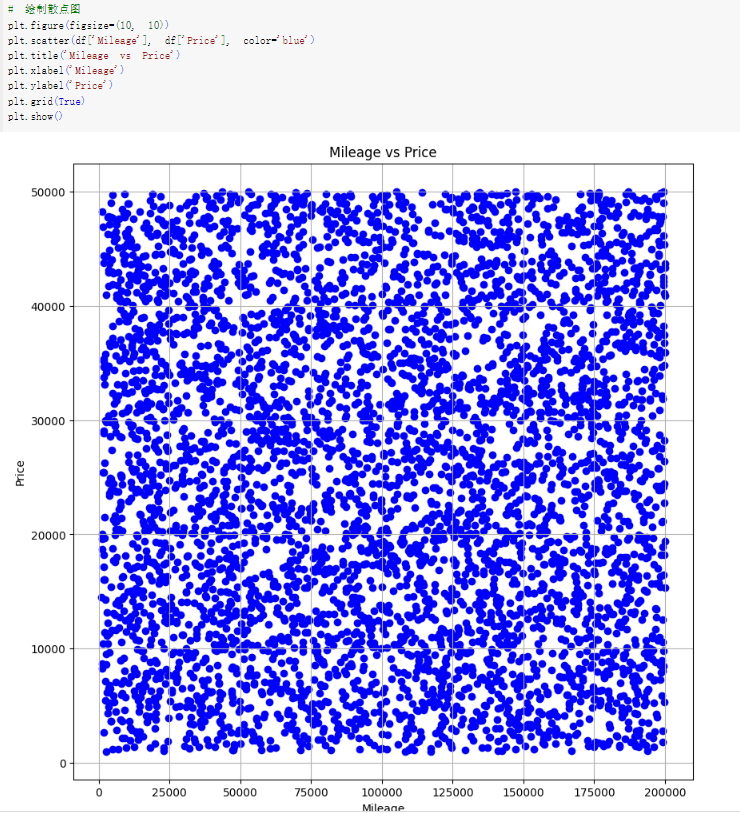


Рис1. - Диаграмма рассеяния

**Задача1: №.16.**

Для набора данных проведите нормализацию для одного (произвольного) числового признака с использованием преобразования Бокса-Кокса (Box-Cox transformation).

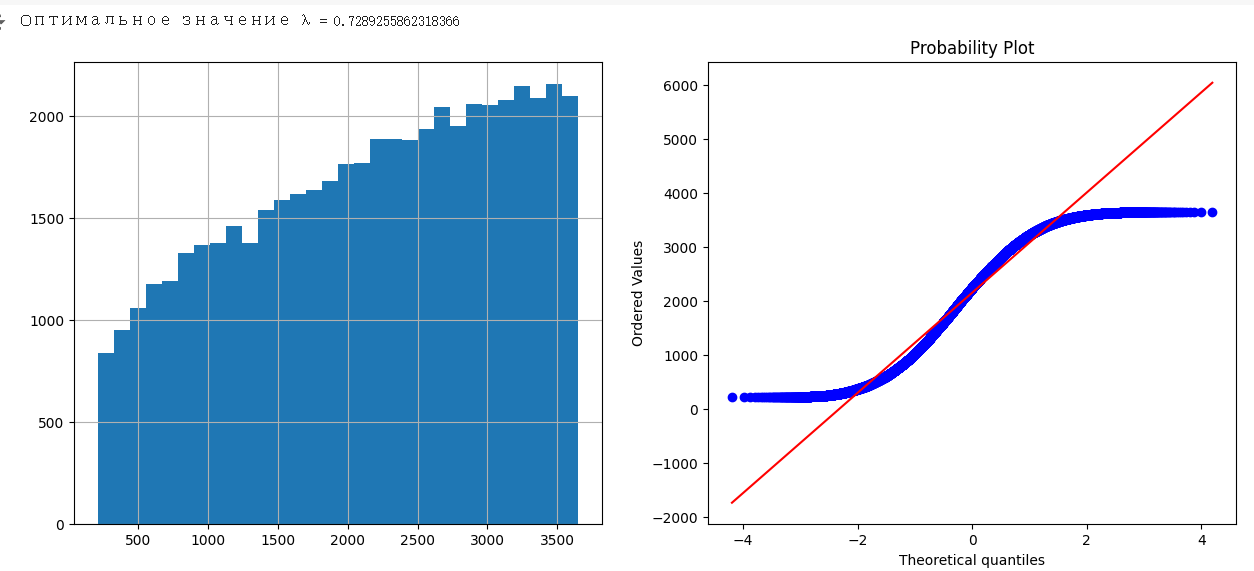
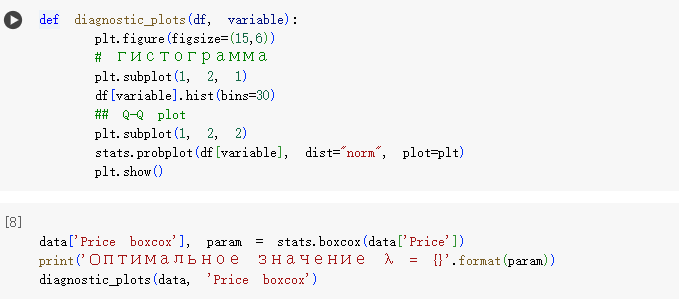


Рис2. - проведите нормализацию для «Price»

**Задача2: №.36.**

Для набора данных проведите процедуру отбора признаков (feature selection). Используйте класс SelectKBest для 5 лучших признаков, и метод, основанный на взаимной информации.

Мы сначала создаем DataFrame из предоставленных данных, затем преобразуем категориальные переменные в числовые с помощью кодирования One-Hot. После этого мы разделяем данные на матрицу признаков (X) и целевую переменную (y). Далее мы используем класс SelectKBest с методом взаимной информации (mutual\_info\_regression) для отбора 5 лучших признаков. Наконец, мы выводим список отобранных признаков.

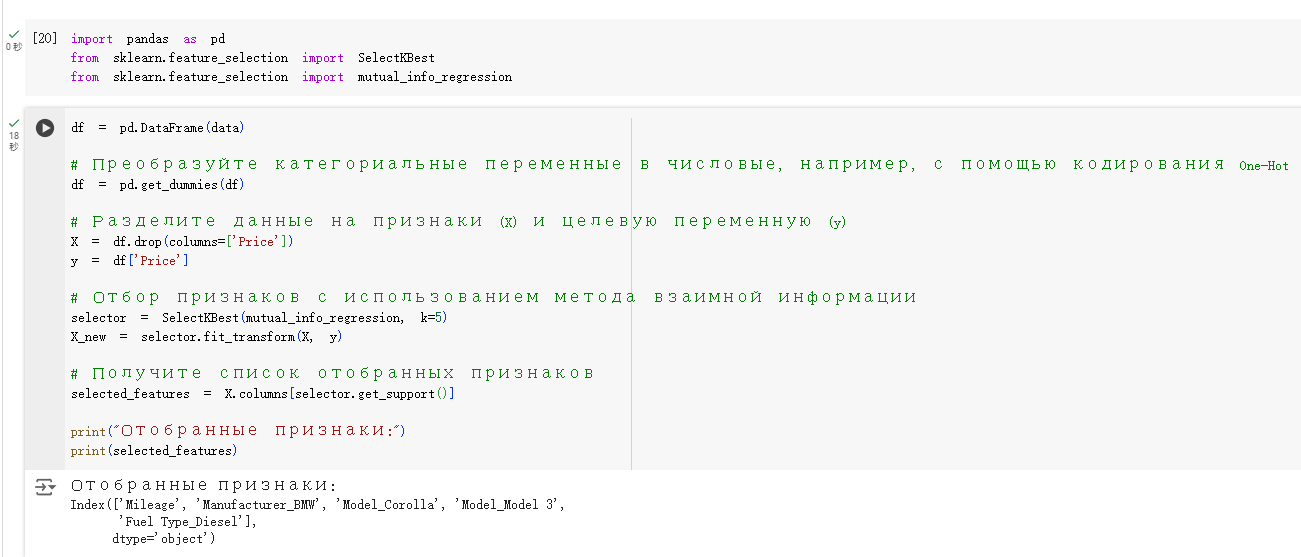


Рис3. - проведите процедуру отбора признаков