**9. Предсказание** **расходов на здравоохранение**

Предсказать стоимость медицинского обслуживания для пациента на основе данных о различных факторах. (insurance.csv)

*Данные:  
В датасете представлены следующие признаки:*

1. Возраст (age): возраст человека в годах.

2. Пол (sex): пол человека (мужчина или женщина).

3. Индекс массы тела (bmi): показатель, вычисляемый на основе массы тела и роста человека. Позволяет оценить наличие лишнего или недостатка веса.

4. Количество детей (children): количество детей у человека.

5. Курение (smoker): является ли человек курильщиком (да или нет).

6. Регион (region): регион проживания человека.

7. Медицинские расходы (charges): медицинские расходы человека в долларах.

1. Импортировать необходимые библиотеки и загрузить данные из файла insurance.csv.

2. Провести предварительный анализ данных: ознакомиться с размерностью датасета, типами данных и наличием пропусков.

3. Очистить данные от выбросов, если таковые имеются.

4. Преобразовать категориальную переменную "пол" в числовую форму, например, с помощью метода Label Encoding.

5. Преобразовать категориальную переменную "регион" в числовую форму, например, с помощью метода One-Hot Encoding.

6. Разделить данные на тренировочный и тестовый наборы в соотношении, например, 70:30.

7. Осуществить шкалирование числовых признаков для более эффективного обучения модели.

Пункты визуализации данных:

8. Визуализировать распределение целевой переменной - медицинских расходов.

9. Визуализировать связь между каждым числовым признаком и целевой переменной, использовав диаграммы рассеяния или ящики с усами.

Пункты построения регрессии:

10. Проанализировать корреляцию между признаками и целевой переменной.

11. Выделить наиболее значимые признаки для построения регрессии.

12. Построить модель регрессии, используя выбранные признаки.

13. Оценить точность модели с помощью метрик, таких как средняя абсолютная ошибка (MAE) или коэффициент детерминации (R^2).

14. Провести кросс-валидацию модели для более надежной оценки ее результативности.

15. Проверить модель на тестовом наборе данных и оценить ее предсказательную способность.