

Лабораторная работа 4.

Тема: Сравнительный анализ методов регрессии.

Цель: Сформировать комплексное понимание различных методов регрессионного анализа и выработать навыки осознанного выбора моделей в зависимости от характеристик данных и решаемой задачи.

Задание

1. Подготовка данных

1.1. Выберите датасет подходящий для решения задачи регрессии с сайта Kaggle.com.

Датасеты для примера: [Egyptian Real Estate Listings](#), [Walmart Sales](#), [Student Performance](#), [Car Prices Dataset](#) и др.

1.2. Загрузите данные и выполните предобработку:

Обработайте пропущенные значения.

Закодируйте категориальные признаки (Label Encoding).

Разделите данные на обучающую и тестовую выборки.

2. Базовые модели

2.1. Обучите три базовые модели LinearRegression, Lasso и ElasticNet.

3. Оценка качества моделей на тестовой выборке

3.1. Рассчитайте MSE, RMSE, MAE и R^2 для каждой модели

3.2. Постройте графики "предсказанные vs фактические значения"

3.3. Сравните коэффициенты моделей.

4. Подбор гиперпараметров и кросс-валидация

4.1. Реализуйте подбор гиперпараметров для Lasso и ElasticNet.

4.2. Проведите кросс-валидацию для всех моделей.

4.3. Ответьте на вопросы:

- Какие оптимальные гиперпараметры для Lasso и ElasticNet?
- Улучшило ли качество подбор гиперпараметров?

5. Отчёт

Оформите отчет в электронном виде, приложив ссылку на Jupyter Notebook/ Google Colab, где код сопровождается краткими выводами по каждому шагу, электронный вид отчёта загрузите в формате pdf в ЭИОС.