

Оценка важности признаков

Цель

изучить особенности и важность признаков в модели многомерной линейной регрессии, использовать разные методы отбора признаков и оценить их влияние на качество модели.

Что нужно сделать:

00 Выберите набор данных для анализа и выполните следующие шаги:

- a)** Обучите модель многомерной линейной регрессии и извлеките коэффициенты регрессии.
- b)** Оцените важность признаков, основываясь на величине коэффициентов.
Обратите внимание на знаки коэффициентов.
- c)** Реализуйте как минимум три метода отбора признаков:
 - Удаление признаков с низкой важностью (например, по критерию p-value).
 - Метод рекурсивного исключения признаков (RFE).
 - Использование регрессии с Lasso (в случае использования L1-регуляризации).
- d)** Сравните качество моделей до и после отбора признаков. Используйте метрики R^2 и RMSE.
- e)** Напишите вывод о том, какие признаки оказались наиболее важными, и как отбор признаков повлиял на качество модели.

Критерии оценивания

K1 Обучена модель многомерной линейной регрессии и корректно извлечены коэффициенты регрессии. **1 балл**

K2 Проведена оценка важности признаков, основываясь на величине и знаках коэффициентов. **1 балл**

K3 Реализованы как минимум три метода отбора признаков: удаление по p-value, RFE и регрессия с Lasso. **1 балл**

K4 Сравнено качество моделей до и после отбора признаков с использованием метрик R^2 и RMSE. **1 балл**

K5 Написан вывод о наиболее важных признаках и о влиянии отбора признаков на качество модели. **1 балла**

K6 Полное документирование процесса с кодами, результатами и выводами. **1 балл**

Максимальное количество баллов

5 баллов

Минимальное количество баллов чтобы преподаватель смог зачесть вашу работу

2 балла