**Лабораторная работа № 3 – Регрессионный анализ данных**

Цель лабораторной работы – получение навыков работы с методами регрессии

1. Изучить набор данных. Создать описание набора данных на русском языке. Описать признаки, используемые в наборе данных (включить полученные описания в отчёт).
2. Удалите дубликаты строк в наборе данных; приведите размер набора данных до и после данной операции;
3. Выполните масштабирование количественных признаков; Постройте диаграммы BoxPlot для признаков до и после масштабирования. Выберите способ масштабирования (например, нормализацию или стандартизацию);
4. Выполните замену категориальных признаков; выберите и обоснуйте способ замены;
5. Оцените корреляцию между признаков и удалите те признаки, которые коррелируют с наибольшим числом других;
6. Заполните пропущенные значения в данных;
7. Решите поставленную задачу регрессии в соответствии с заданием. При подборе параметров метода принятия решения (метода регрессии) используйте перекрёстную проверку (изучите возможные для изменения параметры метода регрессии). Вычислите точность решения задачи, вычислив разницу между реальным значением и предсказанным. Вычислите коэффициент корреляции (Пирсона, Спирмена) между реальным значением и предсказанным с учётом p-value.

Кодировка классификаторов:

1 – линейная регрессия без использования регуляризаторов;

2 – линейная регрессия, LASSO регуляризатор;

3 – линейная регрессия, RIDGE регуляризатор;

Варианты заданий:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Вариант | Набор данных / Классификаторы | Вариант | Набор данных / Классификаторы | Вариант | Набор данных / Классификаторы |
| 1 | 1 / 1, 3 | 2 | 2 / 1, 3 | 3 | 3 / 1, 3 |
| 4 | 1 / 1, 2 | 5 | 2 / 1, 2 | 6 | 3 / 1, 2 |
| 7 | 1 / 2, 3 | 8 | 2 / 2, 3 | 9 | 3 / 2, 3 |
| 10 | 1 / 1, 2 | 11 | 2 / 1, 2 | 12 | 3 / 1, 2 |
| 13 | 1 / 2, 3 | 14 | 2 / 2, 3 | 15 | 3 / 2, 3 |
| 16 | 4 / 1, 3 | 17 | 5 / 1, 3 | 18 | 6 / 1, 3 |
| 19 | 4 / 1, 2 | 20 | 5 / 1, 2 | 21 | 6 / 1, 2 |
| 22 | 4 / 2, 3 | 23 | 5 / 2, 3 | 24 | 6 / 2, 3 |
| 25 | 4 / 1, 3 | 26 | 5 / 1, 3 | 27 | 6 / 1, 3 |
| 28 | 4 / 1, 2 | 29 | 5 / 1, 2 | 30 | 6 / 1, 2 |
| 31 | 7 / 1, 3 | 32 | 8 / 1, 3 |  |  |