|  |
| --- |
| https://lh6.googleusercontent.com/QcftzNtI05T0Y6fjdSh1Rr2rt8oqZ1IvnLvbn1jLJ7CCyteVir3k-xBLv4SL1wAgWJsRhmmJSR0UW-RP63_GQenE4vVWv05BRoZTsmIcBccVTnfxwmsnNMvjg599x9SqZd8E3dkd |
| МИНОБРНАУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение высшего образования  **«МИРЭА** - **Российский технологический университет»**  **РТУ МИРЭА** |

Институт информационных технологий (ИТ)

Кафедра прикладной математики (ПМ)

|  |  |
| --- | --- |
| **ОТЧЕТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ №5** | |
| **по дисциплине** | |
| «Технологии и инструментарий анализа больших данных» | |
| Выполнил студент группы ИКБО-02-20 | Злобин А.О. |
|  |  |
| Принял | Горячев А.А. |

Москва 2023

# Задания

1. Найти данные для классификации. Данные в группе повторяться не

должны. Предобработать данные, если это необходимо.

2. Изобразить гистограмму, которая показывает баланс классов. Сделать

выводы.

3. Разбить выборку на тренировочную и тестовую. Тренировочная для

обучения модели, тестовая для проверки ее качества.

4. Применить алгоритмы классификации: логистическая регрессия, SVM,

KNN. Построить матрицу ошибок по результатам работы моделей

(использовать confusion\_matrix из sklearn.metrics).

5. Сравнить результаты классификации, используя accuracy, precision,

recall и f1-меру (можно использовать classification\_report из

sklearn.metrics). Сделать выводы.

6. Оформить отчет о проделанной работе

# Выполнение заданий

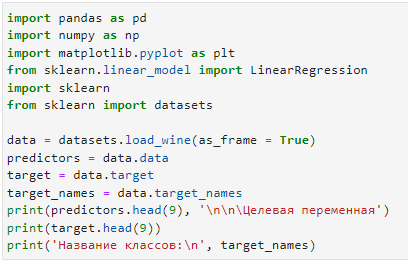
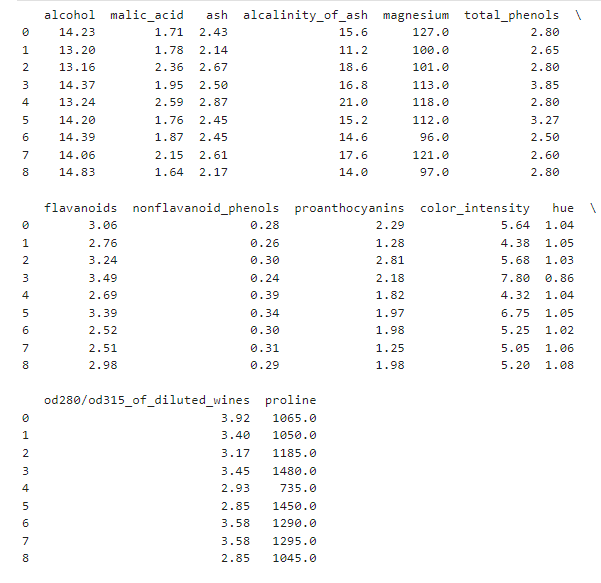


Рисунок 2.1 – Задание 1



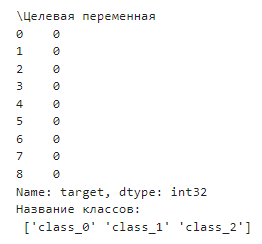


Рисунок 2.2 – Результат задания 1



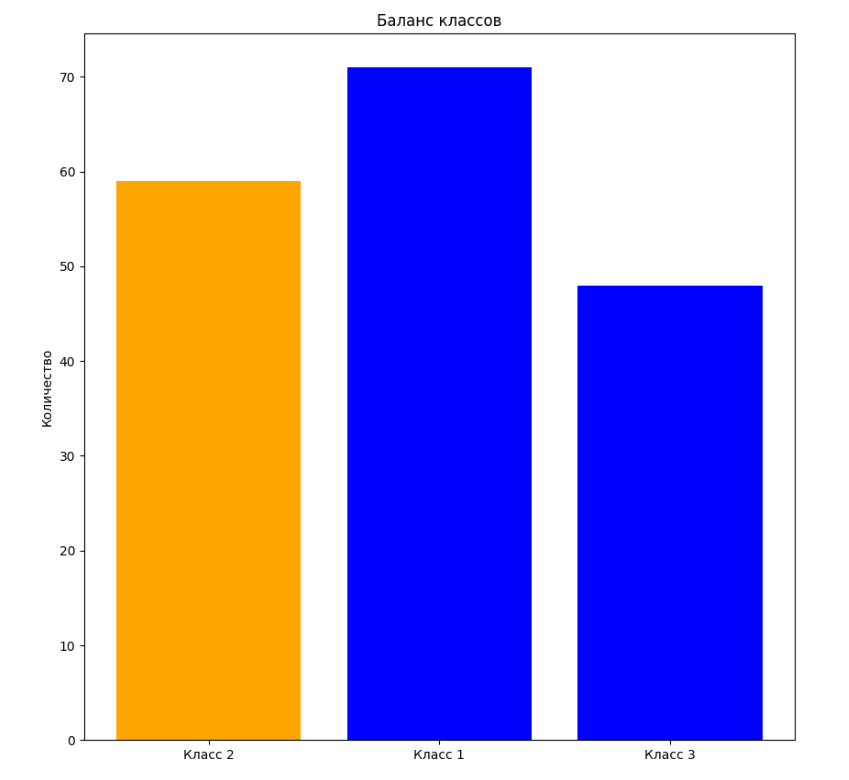


Рисунок 2.3 – Задание 2

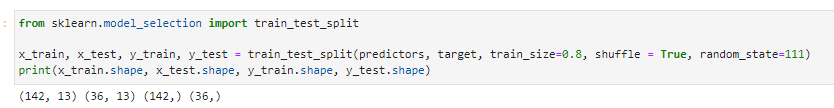


Рисунок 2.4 – Задание 3

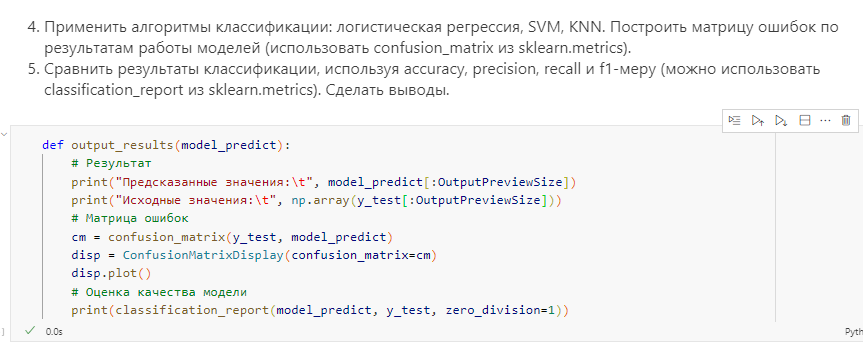
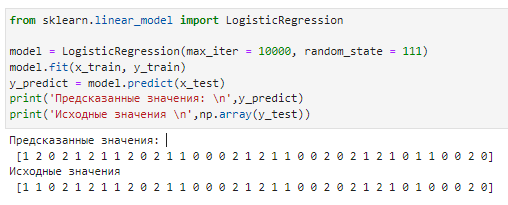


Рисунок 2.5 – Задания 4-5



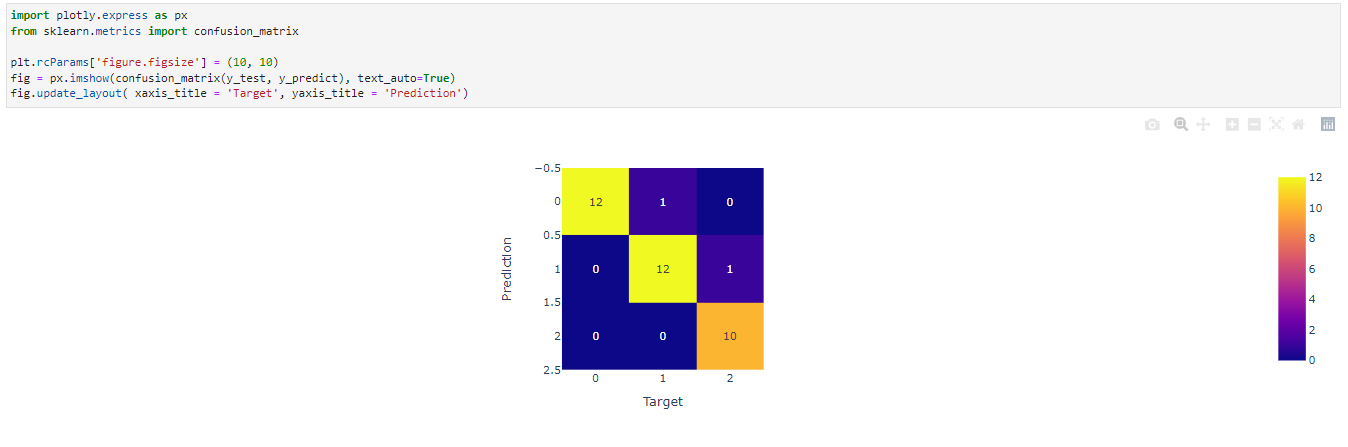




Рисунок 2.6 – Задания 4-5

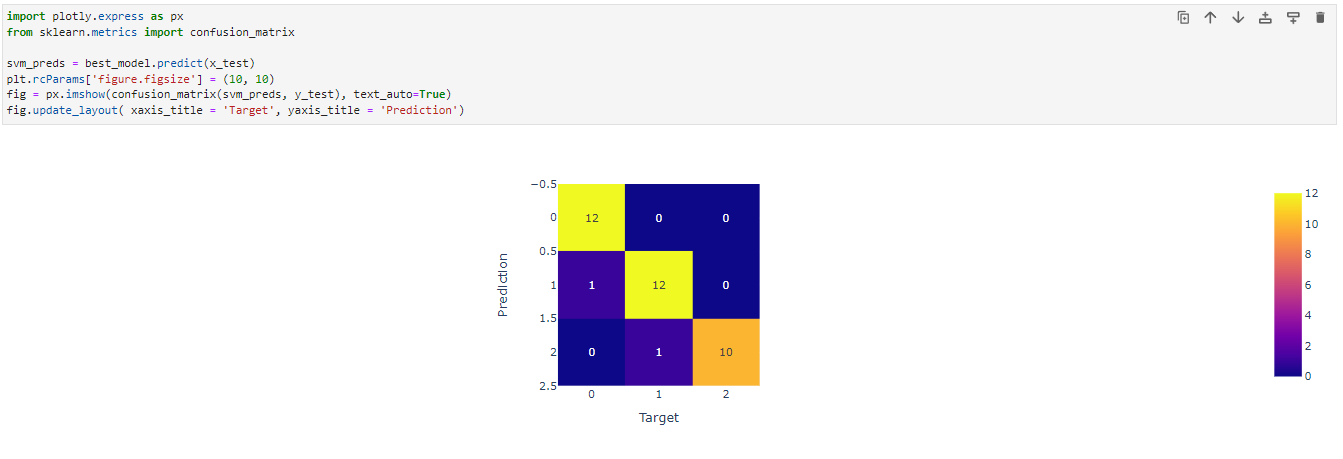
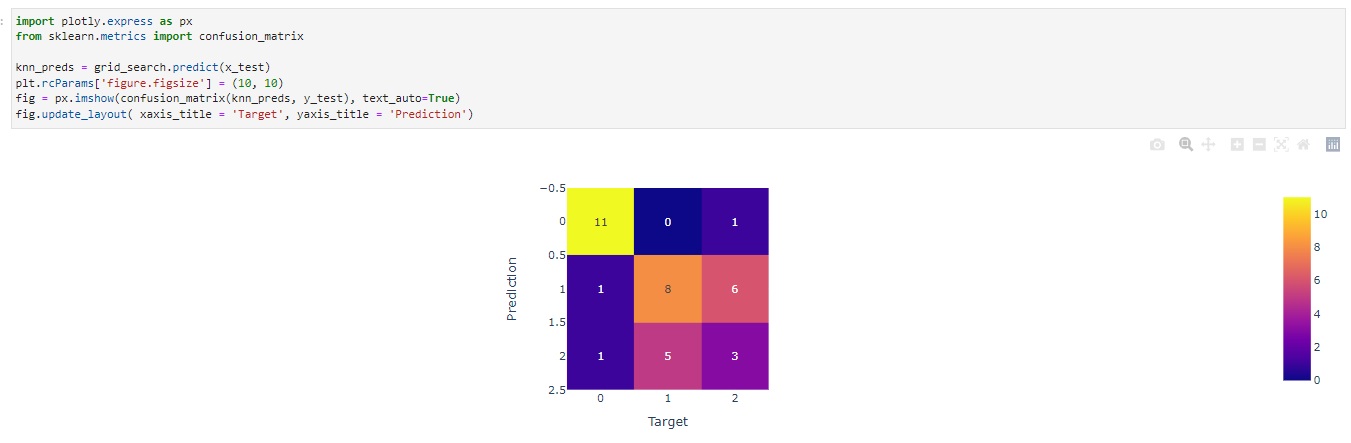




Рисунок 2.7 – Задания 4-5



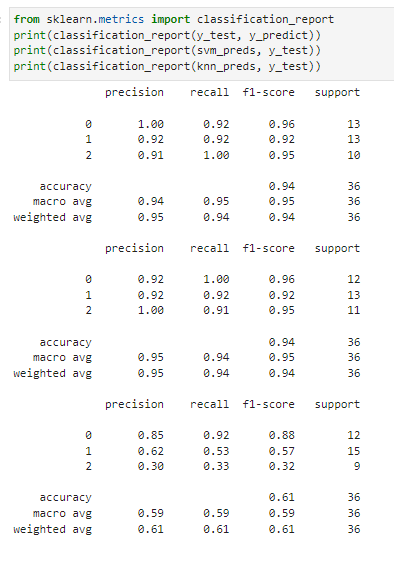


Рисунок 2.8 – Задания 4-5