|  |
| --- |
| МИНОБРНАУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение высшего образования **«МИРЭА − Российский технологический университет»**  **РТУ МИРЭА** |

**Институт информационных технологий (ИИТ)**

**Кафедра практической и прикладной информатики (ППИ)**

**ОТЧЕТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ**

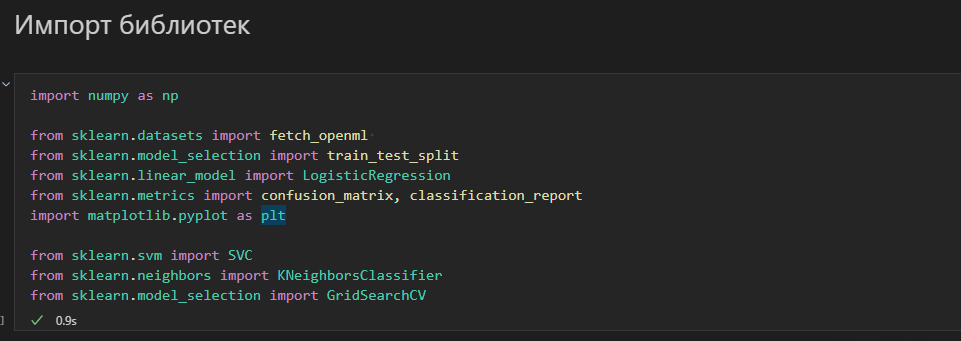
по дисциплине «Технологии и инструментарий анализа больших данных»

**Практическое занятие № 5**

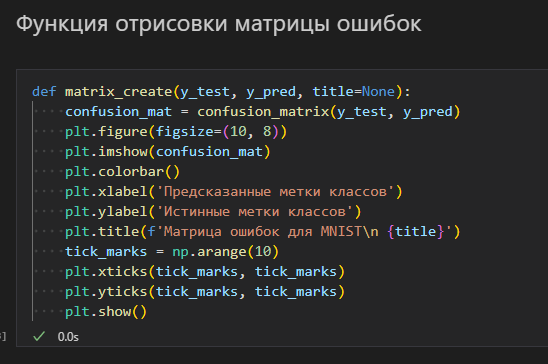
|  |  |  |  |
| --- | --- | --- | --- |
| Студент группы ИНБО-01-17 | *ИНБО-03-20, Першутов Н.С.* | (подпись) | |
| Преподаватель | *Парамонов В.В* | (подпись) | |
| Отчет представлен | «\_\_\_»\_\_\_\_\_\_\_\_2023г. | |  | |

Москва 2023г.

**Результат работы:**

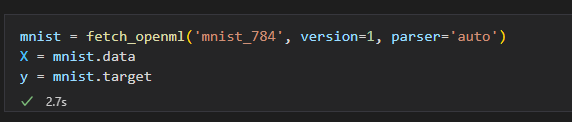
****

**Рисунок – Импорт необходимых библиотек**

****

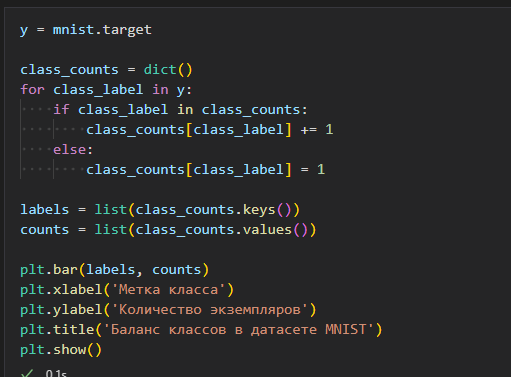
**Рисунок – Функция отрисовки матрицы ошибок**

**Задание 1**

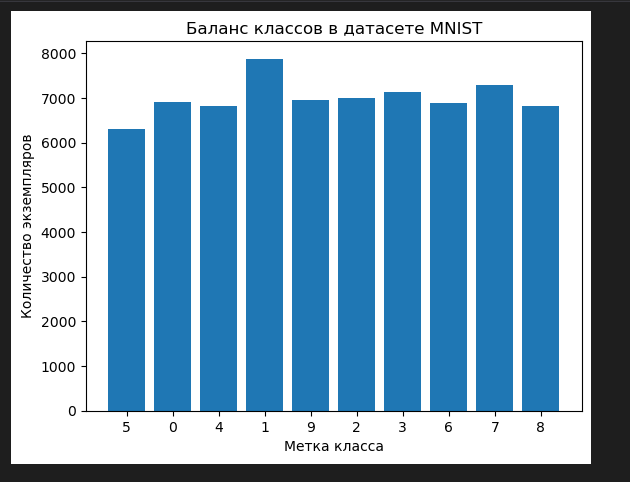
****

**Рисунок 1 – Импортирование данных для классификации**

**Задание 2**

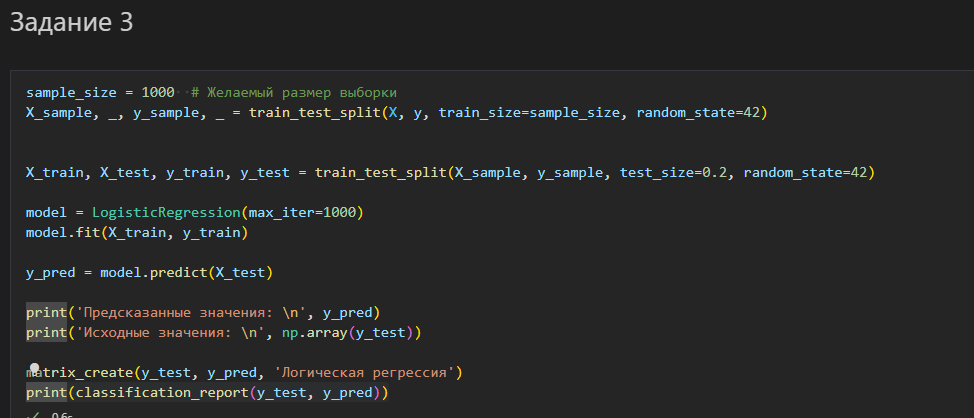
****

**Рисунок 2.1 – Сбор классов в словарь и отрисовка гистограммы баланса классов**

****

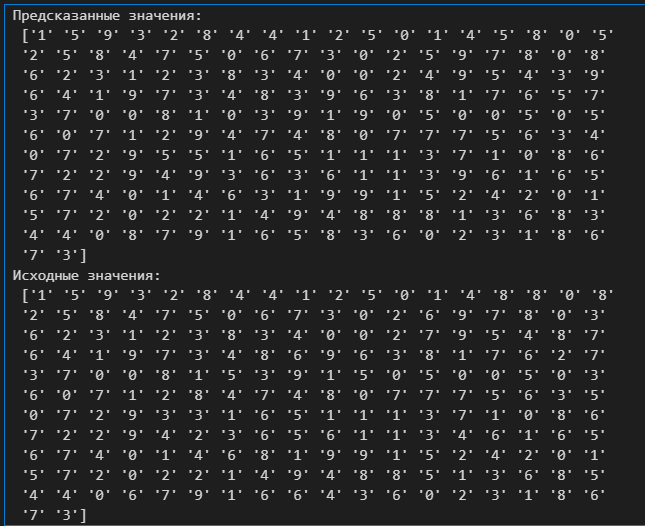
**Рисунок 2.2 – Выполнение кода**

**Задание 3**

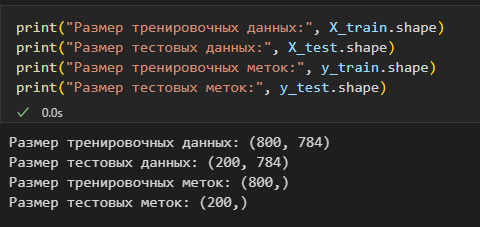
****

**Рисунок 3.1 – Разбиение выборки на тренировочную и тестовую**

Из за огромного размера данных было принято решение уменьшить размер до 1000, соотношение данных 20% к 80% (20% - для тестирования)

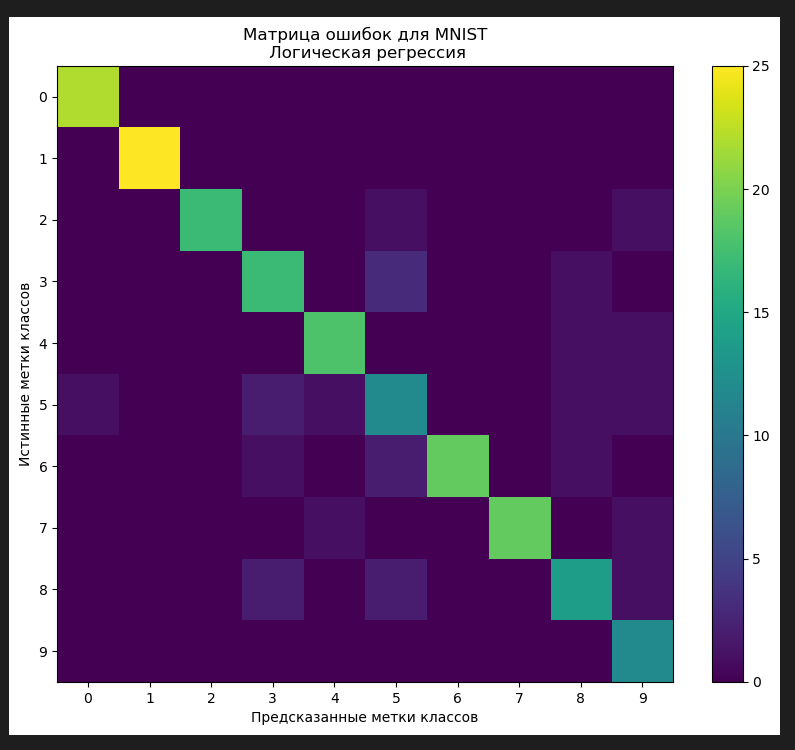
****

**Рисунок 3.2 – Сравнение предсказанных и исходных значений**

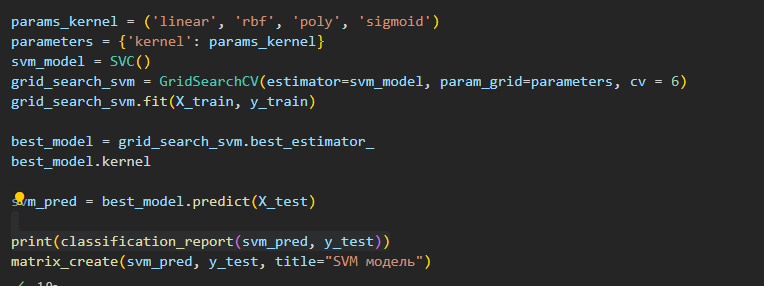
****

**Рисунок 3.3 – Размеры данных**

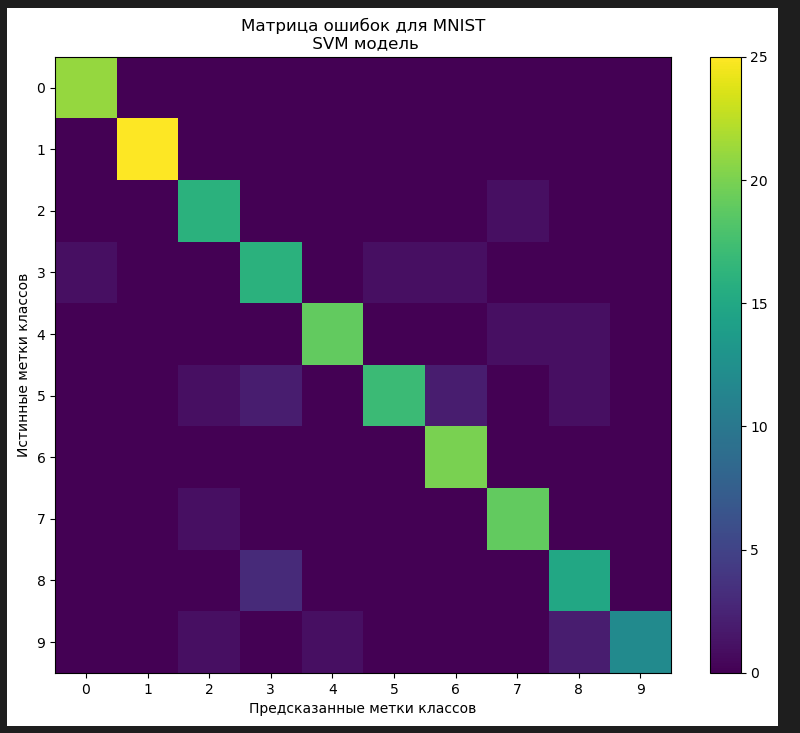
**Задание 4**

****

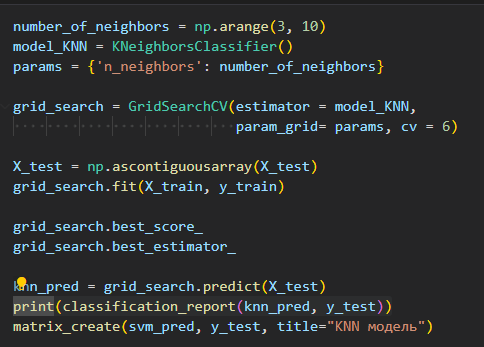
**Рисунок 4.1 – Матрица ошибок для Логической регрессии**

****

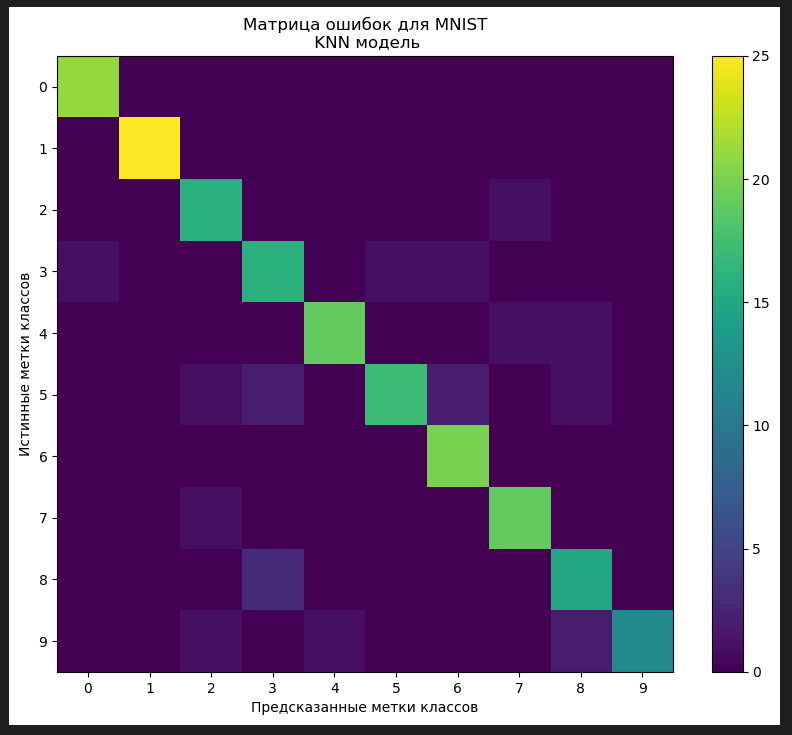
**Рисунок 4.2 – Расчет SVM и отрисовка матрицы ошибок**

****

**Рисунок 4.3 – Матрица ошибок для SVM**

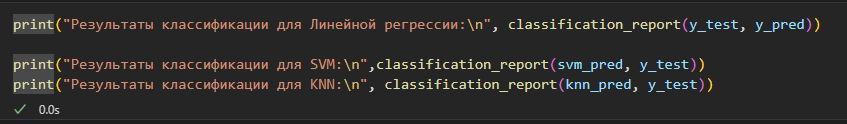
****

**Рисунок 4.4 – Расчет KNN и отрисовка матрицы ошибок**

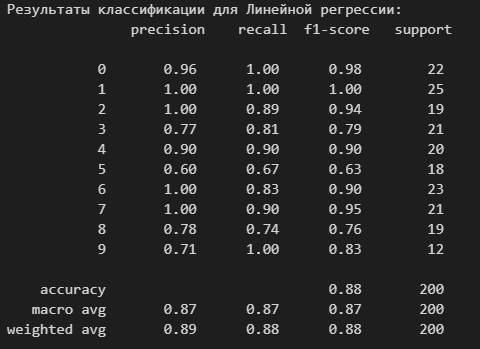
****

**Рисунок 4.5 – Матрица ошибок для KNN**

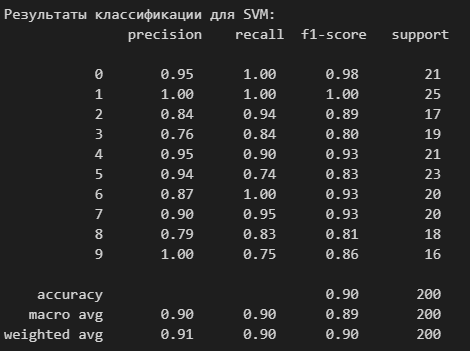
**Задание 5**

****

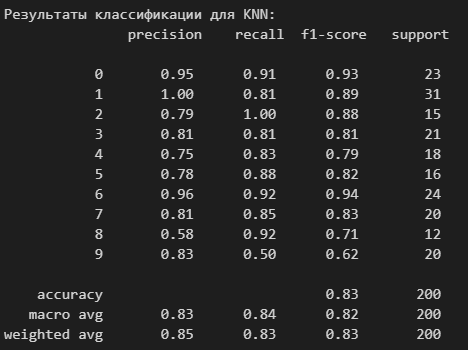
**Рисунок 5.1 – Сравнение результатов классификации используя classification\_report**

****

**Рисунок 5.2 – Результат выполнения**

****

**Рисунок 5.3 – Результат выполнения**

****

**Рисунок 5.4 – Результат выполнения**