|  |
| --- |
|  |
| МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ |
| 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-20-19 | Анваржонов Ж Т |
| Принял семинарист кафедры ИиППО | Десятников А.А. |

|  |  |  |
| --- | --- | --- |
| Практическая работы выполнены | « » 2022г. |  |
| «Зачтено» | « » 2022 г. |  |

Москва 2022

**Постановка задачи**

Выполнить визуализацию многомерных данных, используя t-SNE. Необходимо использовать набор данных MNIST или fashion MNIST (можно использовать и другие готовые наборы данных, где можно наблюдать разделение объектов по кластерам). Рассмотреть результаты визуализации для разных значений перплексии.

Выполнить визуализацию многомерных данных, используя UMAP с различными параметрами n\_neighbors и min\_dist. Рассчитать время работы алгоритма с помощью библиотеки time и сравнить его с временем работы t-SNE.

Оформить отчет о проделанной работе. Отчет должен содержать результаты визуализации для разных значений параметров и выводы.

**Программный код с комментариями и результаты его выполнения**

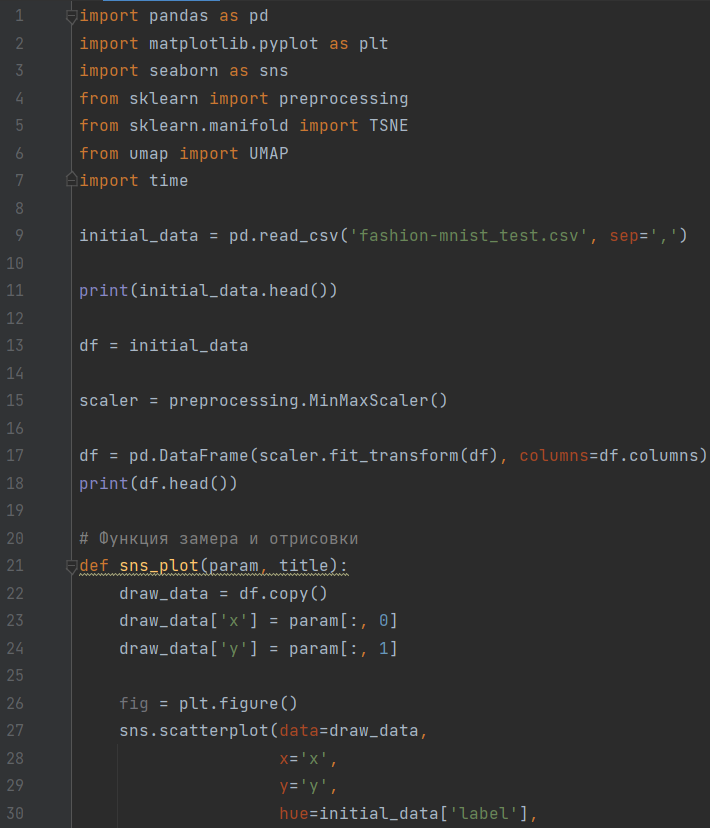


Рисунок 1 – Скриншот программного кода (Часть 1)

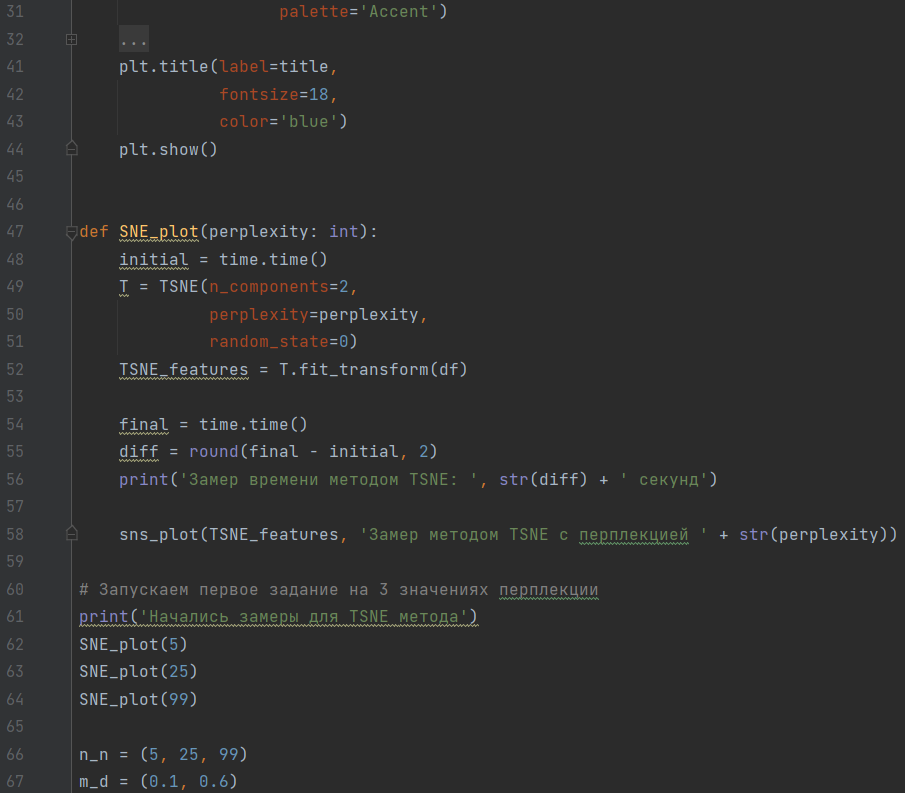


Рисунок 2 – Скриншот программного кода (Часть 2)

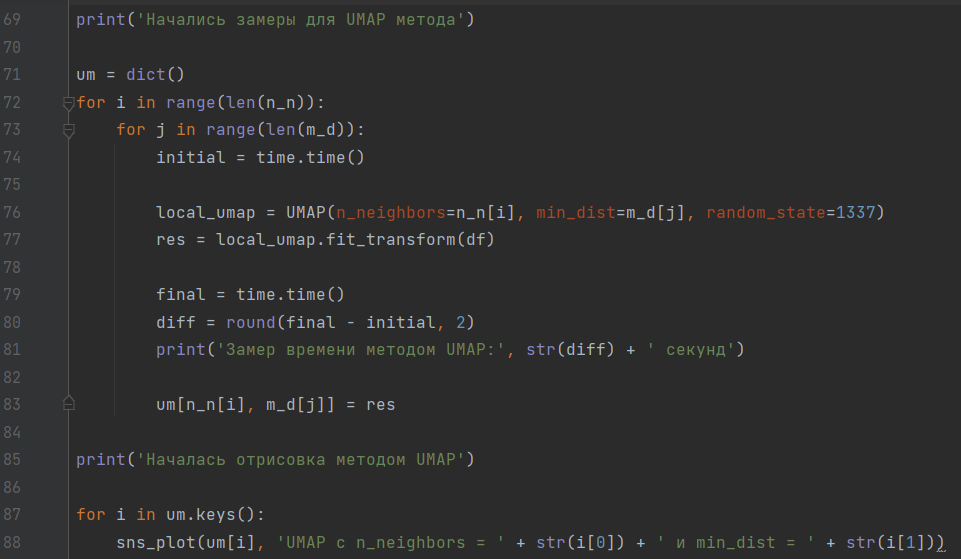


Рисунок 3 – Скриншот программного кода (Часть 3)

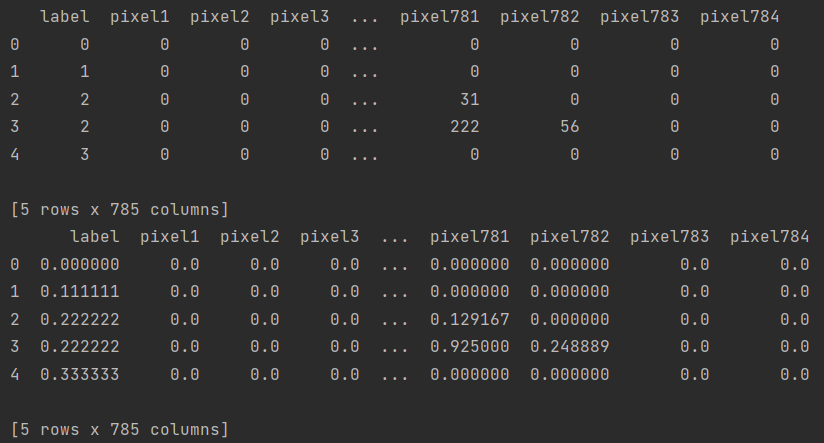


Рисунок 5 – Скриншот результата выполнения заданий (вывод первых строк таблиц данных)

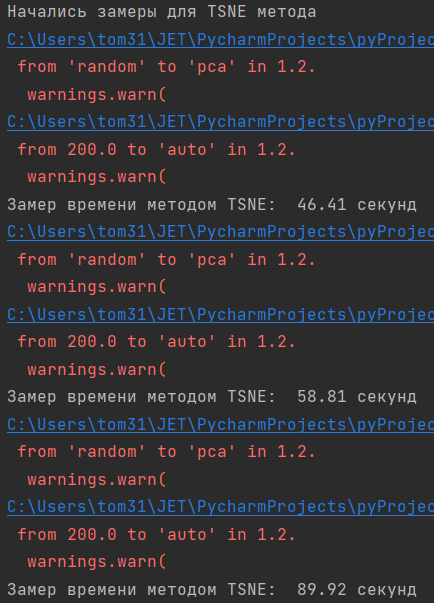


Рисунок 6 – Скриншот результата выполнения заданий (результаты временных замеров методом TSNE)

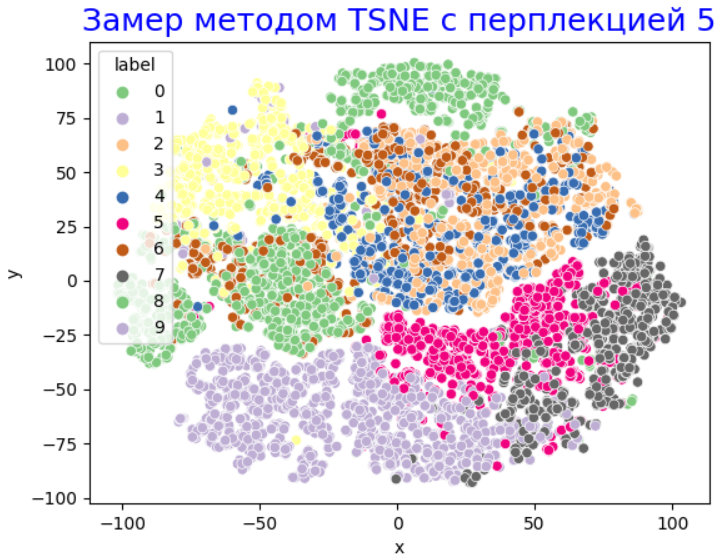


Рисунок 7 – Скриншот результата выполнения заданий (графическое отображение методом TSNE с перплекцией 5)

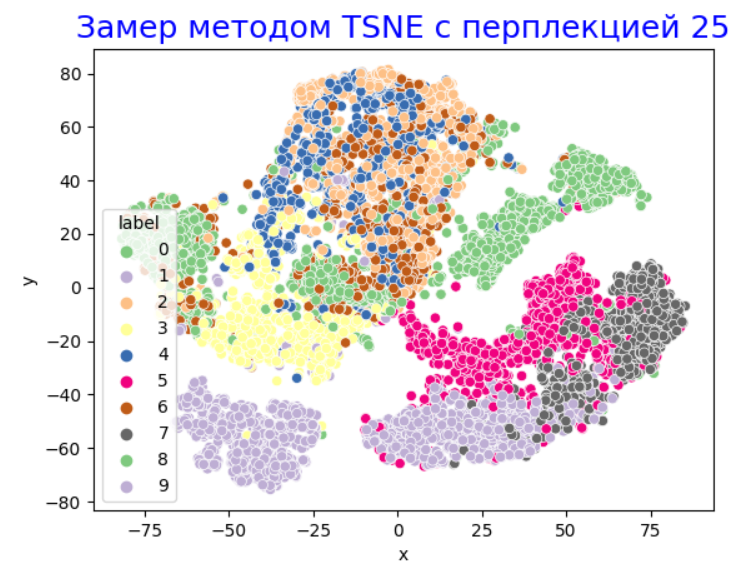


Рисунок 8 – Скриншот результата выполнения заданий (графическое отображение методом TSNE с перплекцией 25)

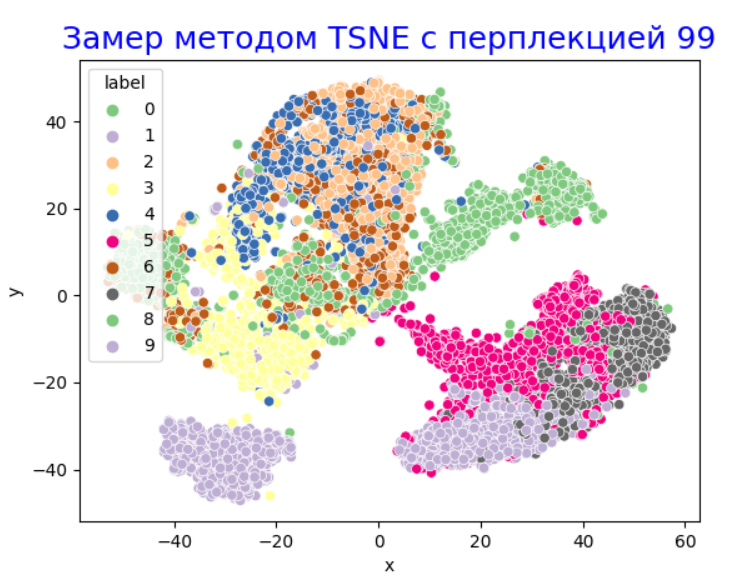


Рисунок 10 – Скриншот результата выполнения заданий (графическое отображение методом TSNE с перплекцией 99)

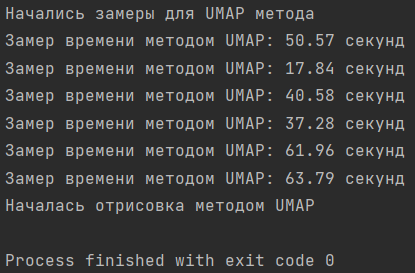


Рисунок 10 – Скриншот результата выполнения заданий (результаты временных замеров методом UMAP)

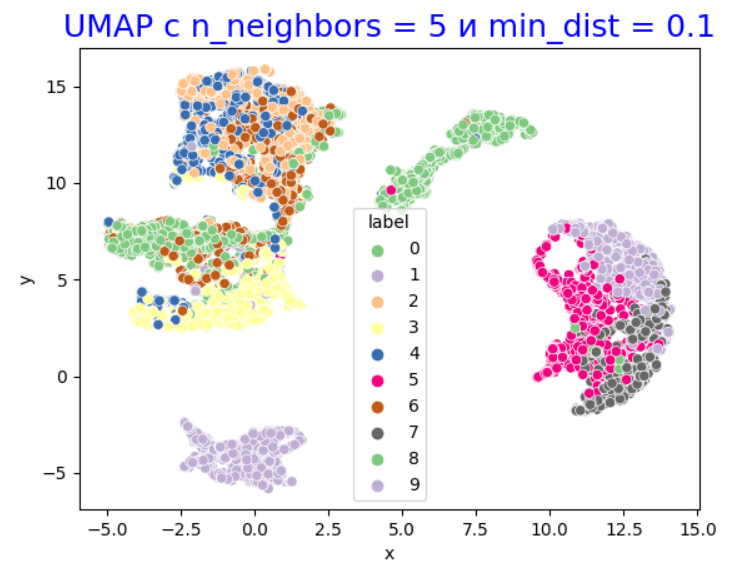


Рисунок 11 – Скриншот результата выполнения заданий (графическое отображение методом TSNE с кол-вом соседей 5 и мин. дистанцией 0.1)

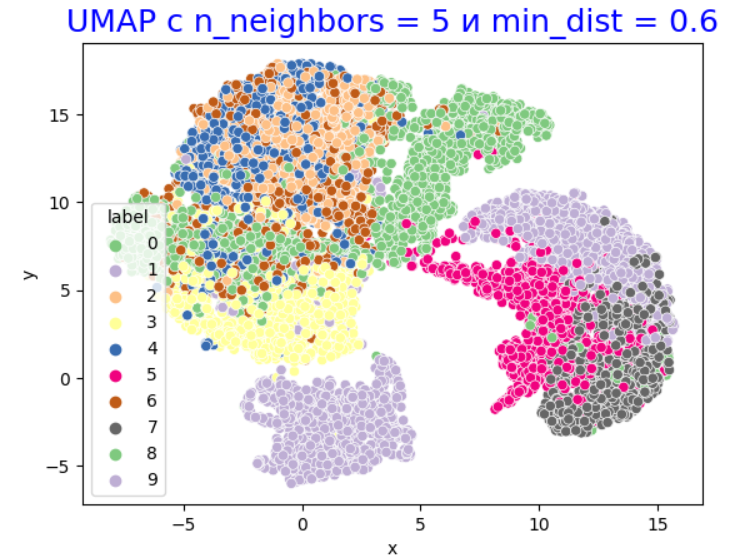


Рисунок 12 – Скриншот результата выполнения заданий (графическое отображение методом TSNE с кол-вом соседей 5 и мин. дистанцией 0.6)

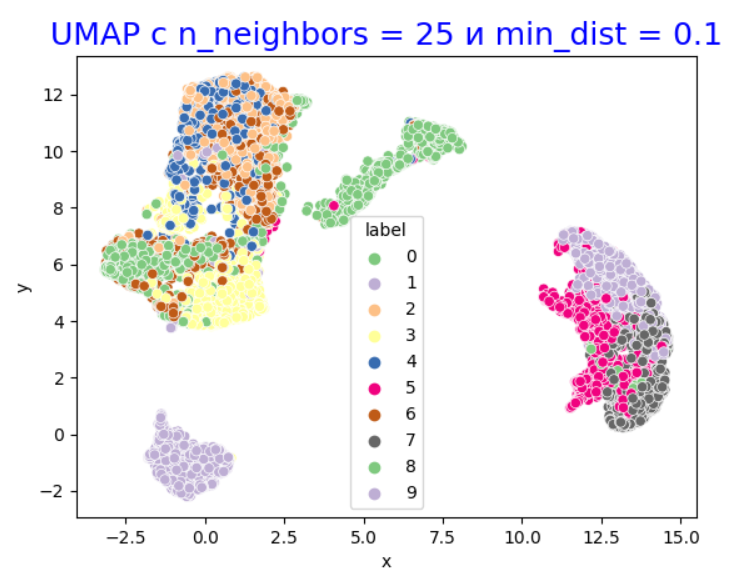


Рисунок 13 – Скриншот результата выполнения заданий (графическое отображение методом TSNE с кол-вом соседей 25 и мин. дистанцией 0.1)

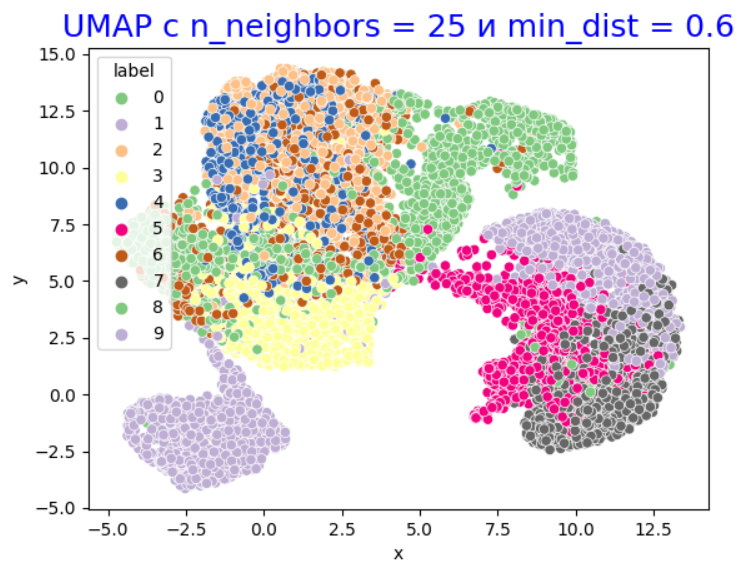


Рисунок 14 – Скриншот результата выполнения заданий (графическое отображение методом TSNE с кол-вом соседей 25 и мин. дистанцией 0.6)

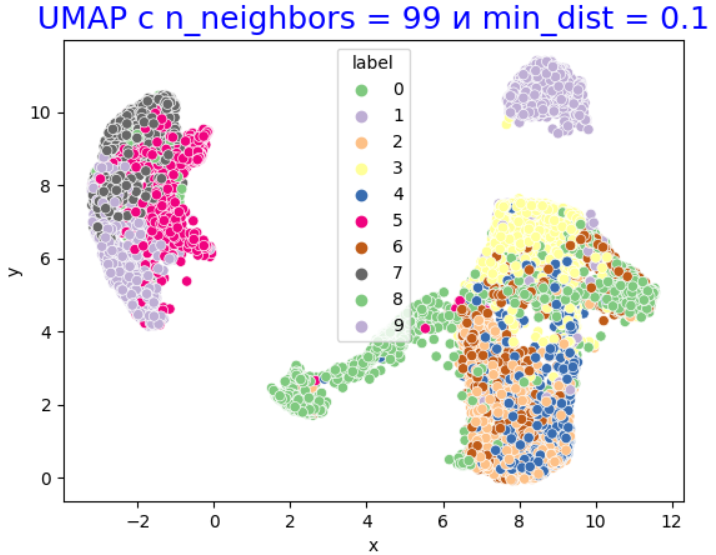


Рисунок 15 – Скриншот результата выполнения заданий (графическое отображение методом TSNE с кол-вом соседей 99 и мин. дистанцией 0.1)

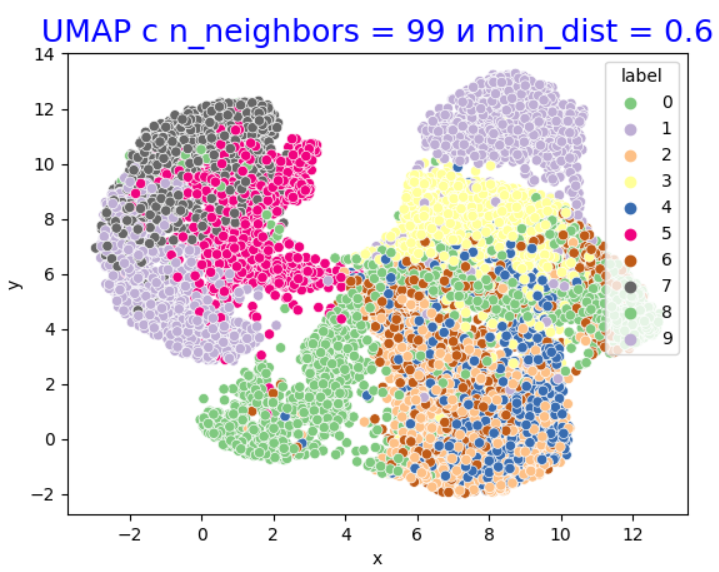


Рисунок 16 – Скриншот результата выполнения заданий (графическое отображение методом TSNE с кол-вом соседей 99 и мин. дистанцией 0.6)

**Вывод**

В результате выполнения данной практической работы были выполнены визуализации многомерных данных, используя t-SNE, а также использованы наборы данных MNIST, были выполнены визуализации данных с использованием UMAP с различными параметрами, а также получен ценный практический и теоретический опыт использования вышеупомянутых инструментов обработки и анализа больших данных.