

УНИВЕРСИТЕТ ИТМО

Факультет программной инженерии и компьютерной техники

Направление подготовки 09.03.04 Программная инженерия

Дисциплина «Системы искусственного интеллекта»

Лабораторная работа №5

Вариант 12

Студент *Голиков Д.И.*

P33102

Преподаватель

Кугаевских А.В.

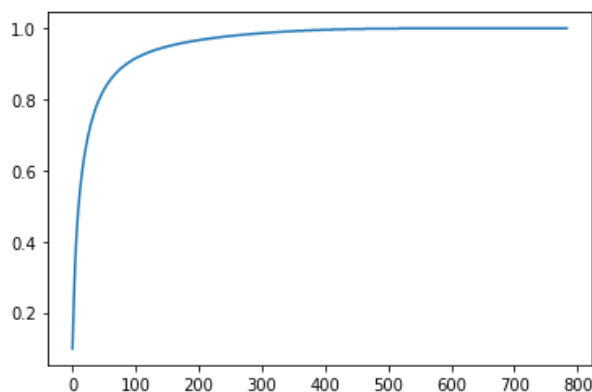
Санкт-Петербург, 2023 г.

Задание:

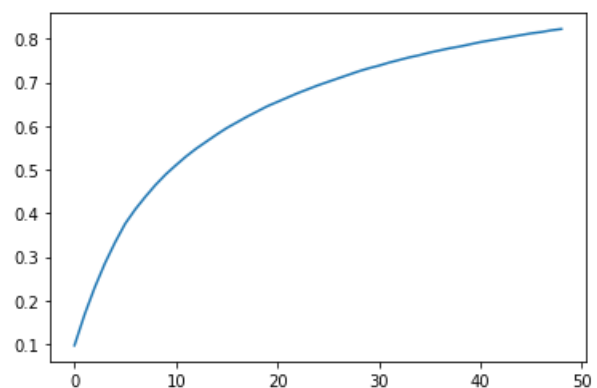
Цель: решить задачу многоклассовой классификации, используя в качестве тренировочного набора данных - набор данных MNIST, содержащий образы рукописных цифр.

1. Используйте метод главных компонент для набора данных MNIST (train dataset объема 60000). Определите, какое минимальное количество главных компонент необходимо использовать, чтобы доля объясненной дисперсии превышала $0.80 + \text{номер_в_списке} \% 10$. Построить график зависимости доли объясненной дисперсии от количества используемых ГК
2. Выведите количество верно классифицированных объектов класса номер_в_списке % 9 для тестовых данных
3. Введите вероятность отнесения 5 любых изображений из тестового набора к назначенному классу
4. Определите Accuracy, Precision, Recall и F1 для обученной модели
5. Сделайте вывод про обученную модель

Задание 1:



Зависимость доли объясненной дисперсии от всех компонент



Зависимость доли объясненной дисперсии от необходимого количества компонент

Чтобы доля объясненной дисперсии была больше 82%, количество компонент должно быть более 49.

Задание 2:

Количество верно классифицированных объектов класса 3: 1521

Задание 3:

Вероятность, что 435 относится к классу 4: 0.5854953233601031

Вероятность, что 521 относится к классу 0: 0.23127981956301175

Вероятность, что 682 относится к классу 7: 0.4775189355974805

Вероятность, что 711 относится к классу 4: 0.2950755350430737

Вероятность, что 888 относится к классу 3: 0.4802208357445512

Задание 4:

Accuracy: 0.6301111111111111

	precision	recall	f1-score	support
Class 0	0.80	0.81	0.80	1693
Class 1	0.93	0.90	0.91	2075
Class 2	0.42	0.56	0.48	1763
Class 3	0.67	0.81	0.73	1873
Class 4	0.67	0.77	0.71	1756
Class 5	0.39	0.35	0.37	1591
Class 6	0.53	0.40	0.46	1766
Class 7	0.76	0.79	0.78	1886
Class 8	0.37	0.32	0.35	1773
Class 9	0.66	0.51	0.58	1824
accuracy			0.63	18000
macro avg	0.62	0.62	0.62	18000
weighted avg	0.63	0.63	0.62	18000

Вывод:

В ходе выполнения лабораторной работы была обучена модель для интерпретирования рукописного ввода из набора MNIST. Было сформировано 49 главных компонент для получения доли объясненной дисперсии более 0,82. Также обучена модель на основе этих компонент и рассчитаны ее метрики.