УНИВЕРСИТЕТ ИТМО

Факультет программной инженерии и компьютерной техники

Направление подготовки 09.03.04 Программная инженерия

Дисциплина «Системы искусственного интеллекта»

**Лабораторная работа №4**

*Вариант 11*

Студент

*Кузнецов М. А.*

*P33131*

Преподаватель

*Авдюшина А. Е.*

Санкт-Петербург, 2022 г.

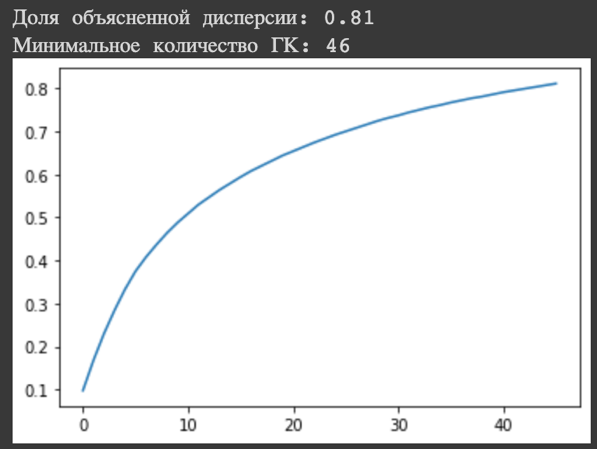
Описание задания

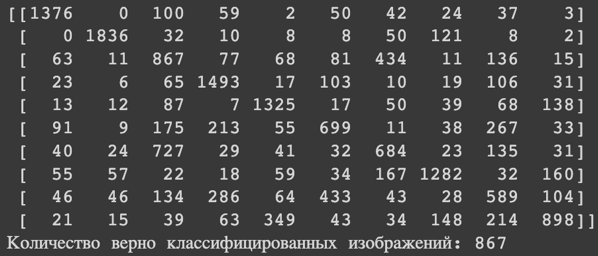
Цель: решить задачу многоклассовой классификации, используя в качестве  
тренировочного набора данных - набор данных MNIST, содержащий образы рукописных  
цифр.

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+номер\_в\_списке%10.
2. Построить график зависимости доли объясненной дисперсии от количества используемых  
   ГК.
3. Введите количество верно классифицированных объектов класса номер\_в\_списке%9  
   для тестовых данных.
4. Введите вероятность отнесения 5 любых изображений из тестового набора к  
   назначенному классу.
5. Определите Accuracy, Precision, Recall or F1 для обученной модели.
6. Сделайте вывод про обученную модель.

Выполнение

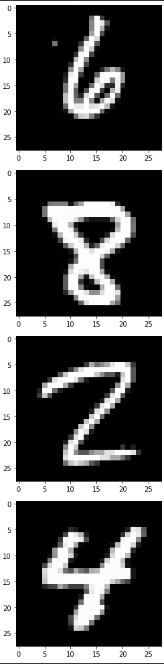
Весь код доступен по ссылке: <https://shorturl.at/hBO29>

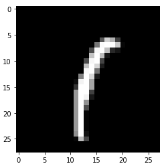
* 1. Необходимо минимум 46 главных компонент, чтобы доля объясненной дисперсии не превышала 0.81
  2. 
  3. Confusion\_matrix (Класс 2)



* 1. Изображение выглядит как текст

     Автоматически созданное описание





* 1. Изображение выглядит как текст, электроника, калькулятор

     Автоматически созданное описание

Выводы

В ходе выполнения данной лабораторной работы я:

* построил график зависимости для объясненной дисперсии
* реализовал многоклассовую классификацию образов рукописных цифром на алгоритме RandomForest.
* нашел параметры accuracy, precision и recall полученной модели