Университет ИТМО

Факультет программной инженерии и компьютерной техники

Курс «Искусственный интеллект»

Лабораторная работа № 5

Вариант: 1

|  |  |
| --- | --- |
|  | **Работу выполнил**  Хуан Сыюань  P33101  **Преподаватель**  Игорь Александрович Бессмертный |

Санкт-Петербург

2021

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 + 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

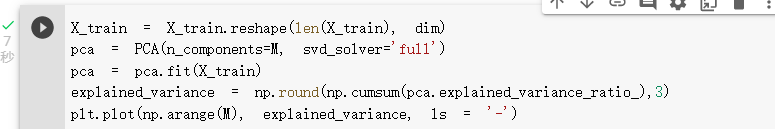
# Выполнение лабораторной работы

Question 1

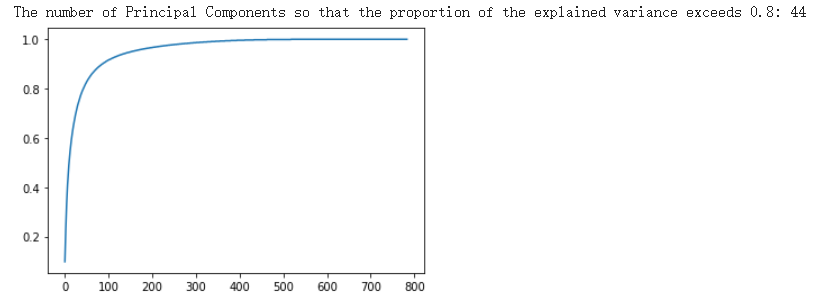
Codes:

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成



Answers:

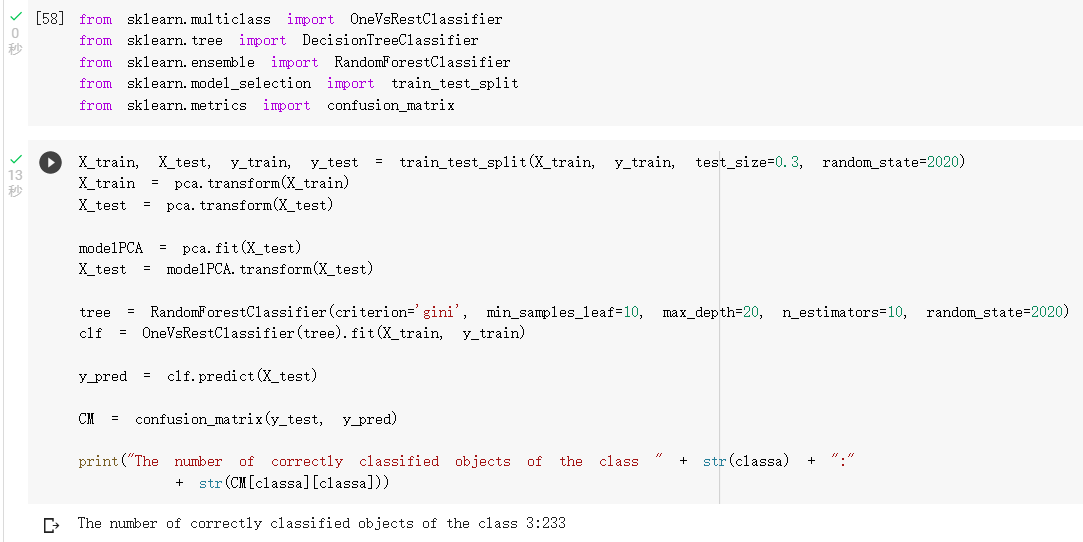


Dependence of the proportion of the explained variance on all:图表

描述已自动生成

Question 2

Codes and answers:



Question 3

Codes and answers:

图片包含 散点图

描述已自动生成

Question 4

Codes and answers:

表格

描述已自动生成

# Вывод

As a result of the laboratory work, a model was trained to predict the drawn numbers on the MNIST set using 44 principal components out of 784 available, resulting in a share of explained variance of 0.8 on the test sample. The resulting model has a precision of 0.684 and does a good job of determining the digits 0, 1, 3, 4, 7, and 9 in comparison to the rest of the digits, for which the more informative measures Precision, Recall, and F1 are much lower and less than 0.5.