**Московский авиационный институт(Национальныйисследовательскийуниверситет)**

Факультет: «Информационные технологии и прикладная математика»Кафедра: 806 «Вычислительная математика ипрограммирование»Дисциплина:«Искусственный интеллект»

Лабораторнаяработа№1

Тема:«Машинное обучение»

Студент: Рыженко Иван Александрович

Группа:М80-408Б-18

Преподаватель: Ахмед Самир Халид Дата:

Оценка:

Москва,2021

1. Постановка задачи

Найти себе набор данных (датасет), для следующей лабораторнойработы,ипроанализироватьего.Выявитьпроблемынабораданных,устранитьих.Визуализироватьзависимости,показатьраспределениянекоторыхпризнаков.РеализоватьалгоритмыКближайшихсоседасиспользованиемвесовиНаивныйБайесовскийклассификаторисравнитьс реализациейбиблиотеки sklearn.

1. Описаниепрограммы

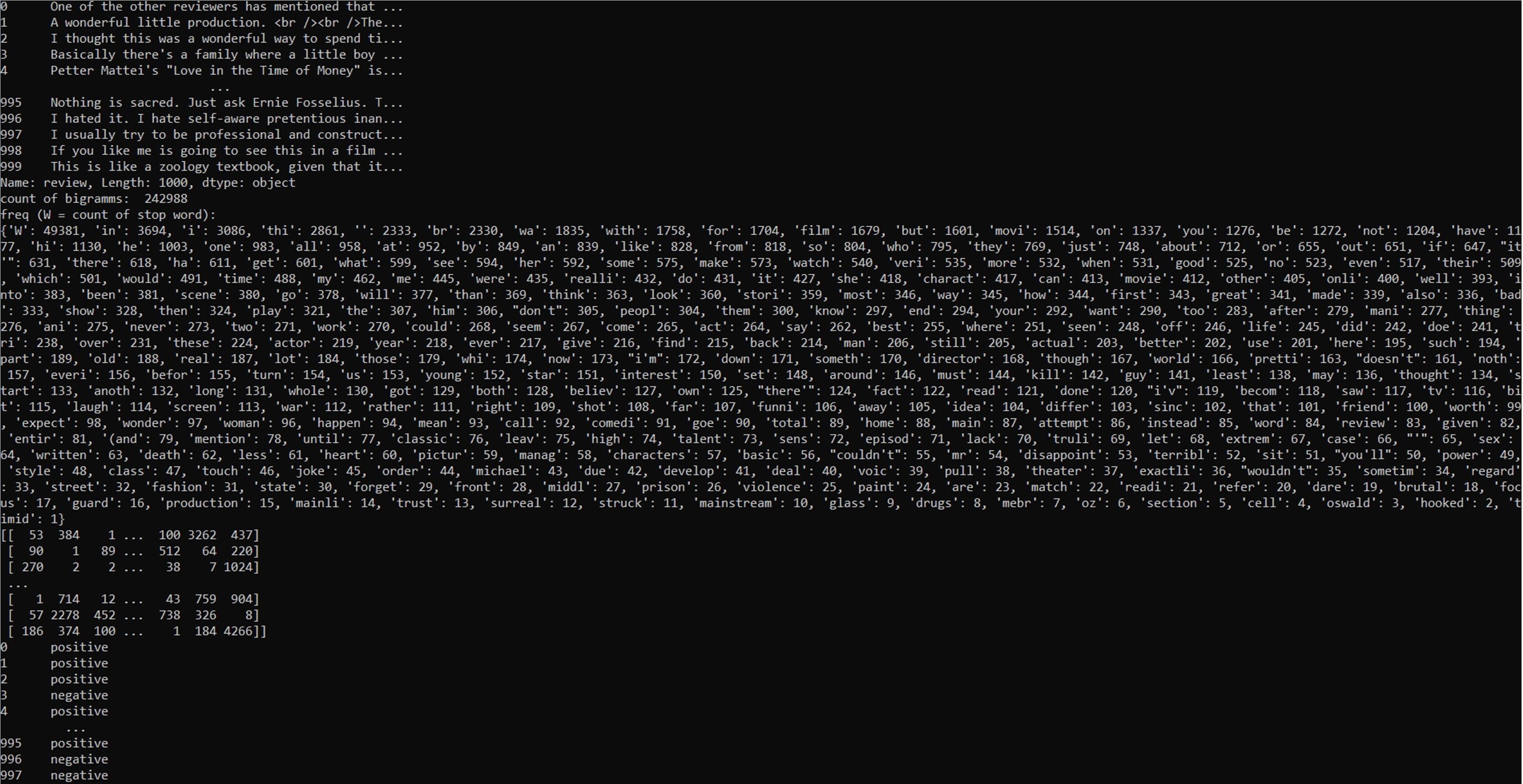
Для данной лабораторной работы я нашел базу данных плохих ихороших отзывов о фильмах на сайте Kaggle. Датасет представляет из себятаблицусдвумяколоннами:отзывиегокачество:позитивныйилинегативный.

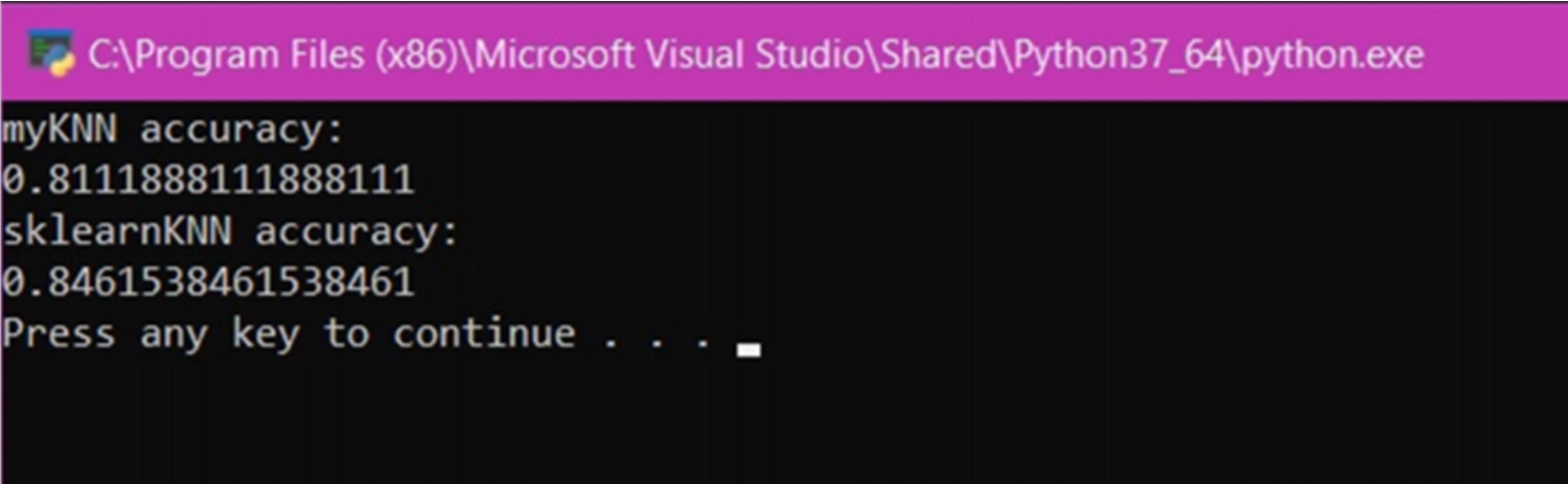
Яобработалэтубазуданныхдляееиспользованиявмашинномобучении. Для начала я проверил, полна ли база данных, проанализировавеенапустыезначения.Ксчастью,онасразуоказалась полной.Послеэтогояотфильтровалвесьтекст,избавившисьотлишнихненужныхсимволов и преобразовав некоторые слова в более удобную форму.Затемятокенизировалвсеслова,представивихнатуральнымичисламивнекоторомуказанноммнойдиапазоне.Данныйчисловойформатужепригодендля использования в машинном обучении.

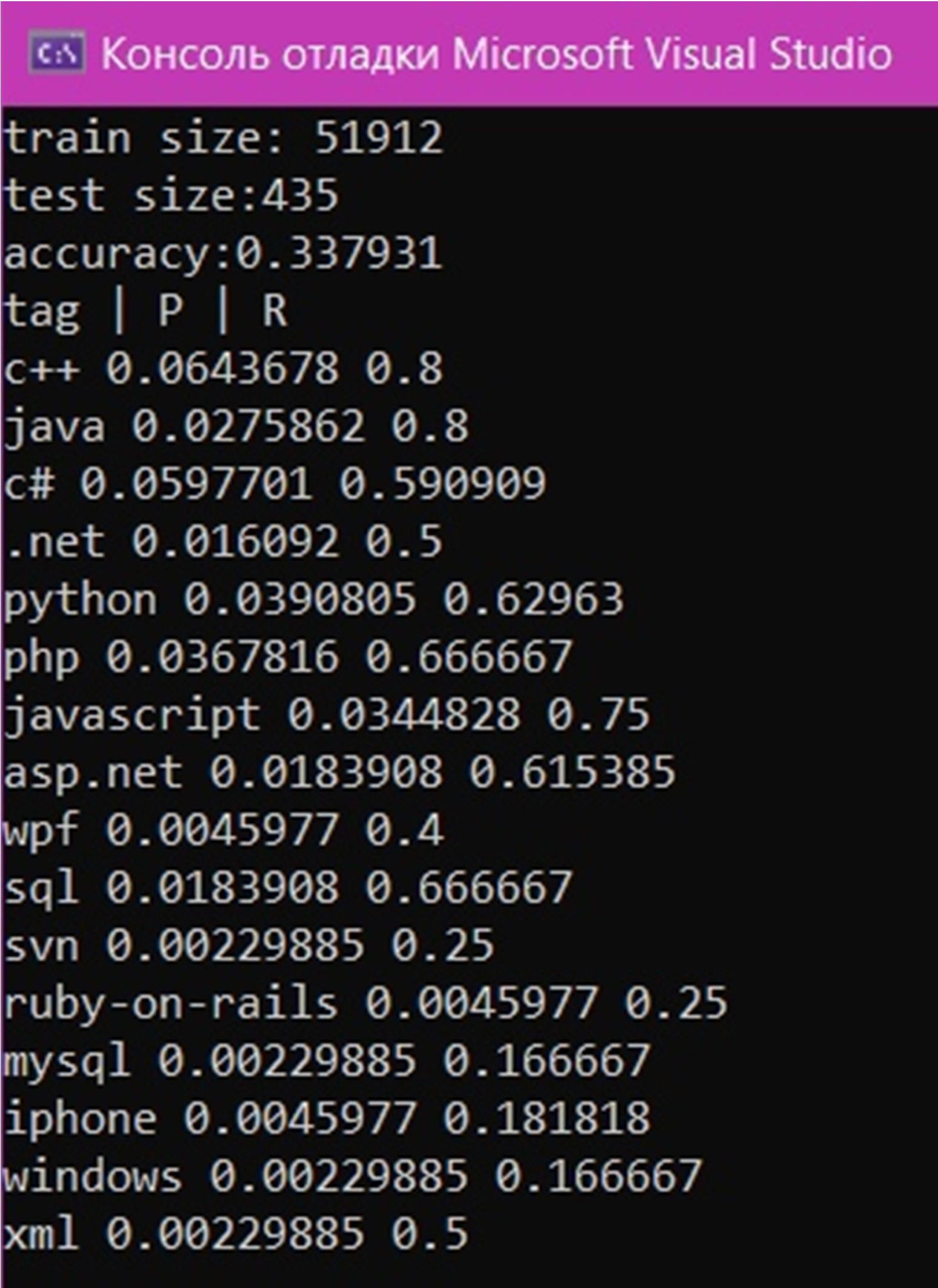
ДалееяреализовалалгоритмКближайшегососеда,использовавстандартнуюдокументациюиязыкPython.Чтобыпроверитьегокорректность,ясравнилсвоюреализациюпрограммысреализациеймодуляsklearn.Полученныерезультаты оказалисьсхожи.

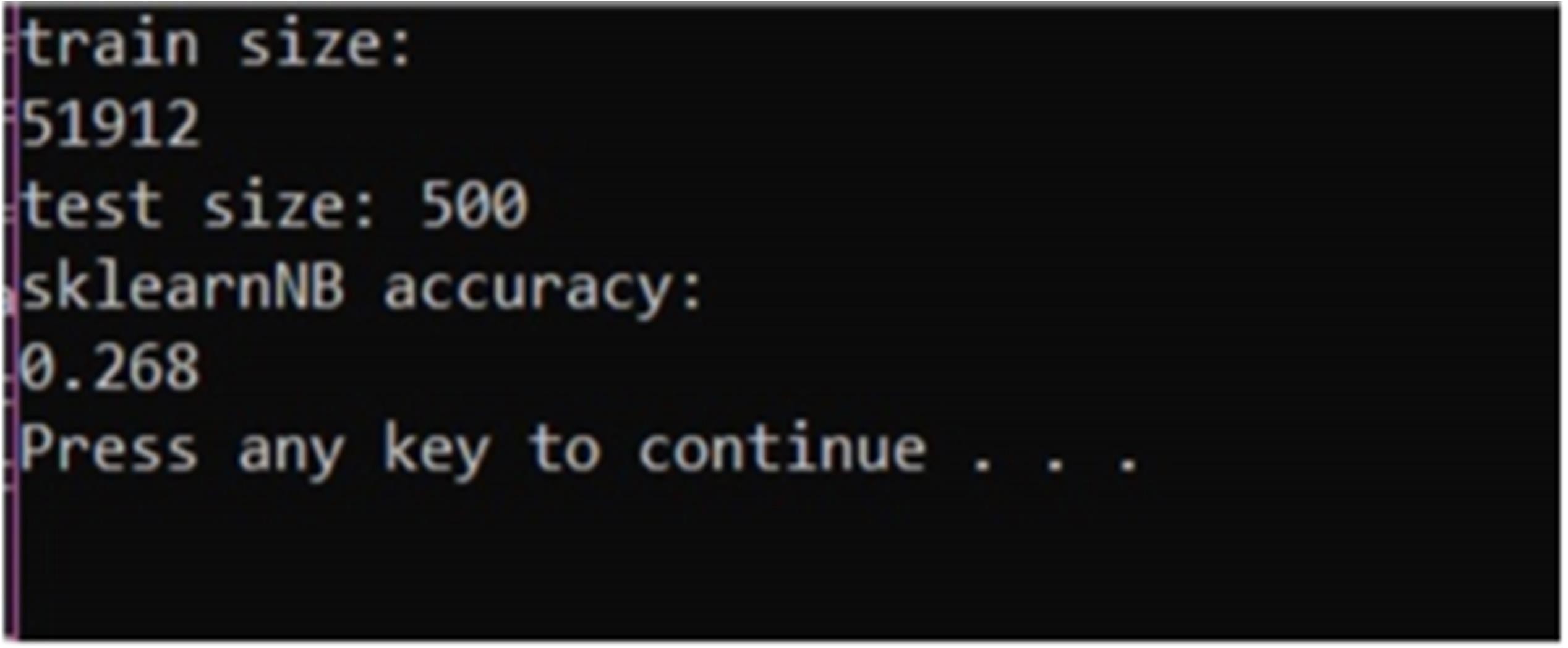
И, в заключение, я создал свой Байесовский классификатор на базеязыка C++, аналогично, сравнив его работу с уже готовым Байесовскимклассификатором в библиотеке sklearn. Полученные результаты оказалисьсхожи.

1. Результаты работы программы1)Программаанализирующаяданные



1. KNN
2. NB

Мояпрограмма:

Программаsklearn:

1. Вывод

В данной лабораторной работе я отыскал некоторую базу данных и,проанализировавее,подготовилкалгоритмаммашинногообучения.Также я реализовал и сами алгоритмы машинного обучения, такие как: КближайшихсоседаиБайесовскийклассификатор.Сравнивмоюреализацию с реализацией sklearn, я убедился в достоверной реализациисвоих программ.

В заключение хочу сказать, что полученные знания и опыт в даннойлабораторнойработемне,несомненно,пригодятся,каквбудущихстуденческихпроектах, так идалеко запределамиинститута.