|  |  |
| --- | --- |
| **Gerb-BMSTU_01** | **Министерство науки и высшего образования Российской Федерации**  Калужский филиал  федерального государственного бюджетного  образовательного учреждения высшего образования  ***«Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)»***  ***(КФ МГТУ им. Н.Э. Баумана)*** |

**ФАКУЛЬТЕТ** ***ИУК «Информатика и управление»***

**КАФЕДРА** \_\_***ИУК4 «Программное обеспечение ЭВМ, информационные технологии»***

**ЛАБОРАТОРНАЯ РАБОТА №7**

**«Наивный байесовский классификатор. Виды классификаторов»**

**ДИСЦИПЛИНА: «Технологии анализа данных»**

|  |  |  |
| --- | --- | --- |
| Выполнил: студент гр. ИУК4-82Б | | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ( Карельский М.К. )  (Подпись) |
| Проверил: | | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ( Ерохин И.И. )  (Подпись) |
| Дата сдачи (защиты):  Результаты сдачи (защиты): | | |
|  | - Балльная оценка:  - Оценка: | |

Калуга, 2024

**Цель:** формирование практических навыков проведения классификации с использованием наивного Байесовского классификатора.

**Задачи:**

1. Составить модель на основе наивного Байесовского классификатора.
2. Изучить виды наивного Байесовского классификатора

**Задание:**

Для всех вариантов необходимо провести классификацию с помощью наивного Байесовского классификатора. Составить модель для классификации. Сравнить данные, полученные моделью, с реальными данными. Данные для вариантов необходимо брать из набора sklearn.datasets

**Вариант 5**

Используя дополненный (complement) наивный Байес, построить модель (на обучающей выборке) и произвести классификацию (на тестовой выборке) цветков ириса. Визуализировать полученную классификацию. Построить матрицу несоответствий и интерпретировать результаты. Оценить точность модели (для оценки точности необходимо взять среднее арифметическое оценки 10 испытаний).

**Листинг:**

import matplotlib.pyplot as plt

from sklearn.datasets import load\_iris

from sklearn.model\_selection import train\_test\_split

from sklearn.naive\_bayes import ComplementNB

from sklearn.metrics import confusion\_matrix, accuracy\_score

iris = load\_iris()

X = iris.data

y = iris.target

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

model = ComplementNB()

model.fit(X\_train, y\_train)

y\_pred = model.predict(X\_test)

accuracy = accuracy\_score(y\_test, y\_pred)

print("Accuracy:", accuracy)

conf\_matrix = confusion\_matrix(y\_test, y\_pred)

print("Confusion Matrix:")

print(conf\_matrix)

plt.figure(figsize=(8, 6))

plt.scatter(X\_test[:, 0], X\_test[:, 1], c=y\_pred, cmap='viridis', s=50, alpha=0.8)

plt.xlabel('Sepal Length (cm)')

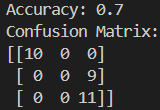
plt.ylabel('Sepal Width (cm)')

plt.title('Classification of Iris Flowers')

plt.colorbar(label='Predicted Class')

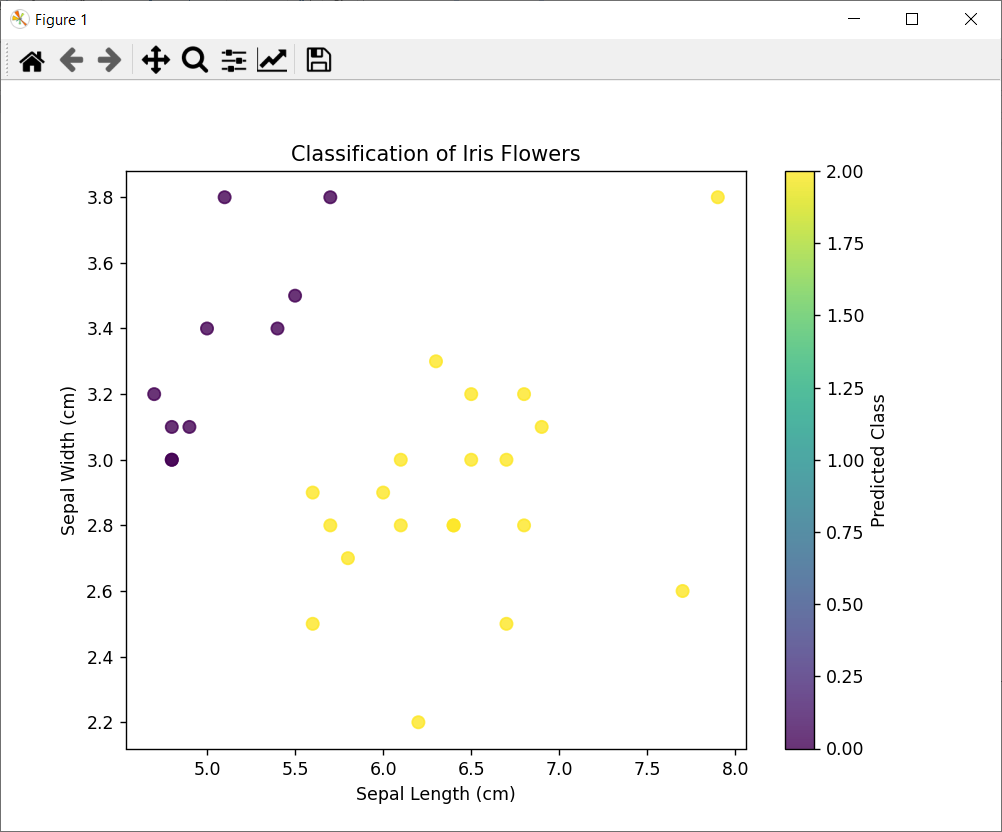
plt.show()

**Результат:**



**Рис. 1.** Оценка точности и матрица несоответствий

Из матрицы видно, что модель имеет трудности с классификацией образцов второго класса и чаще ошибочно относит их к третьему классу. Модель хорошо справляется с классификацией образцов первого и третьего классов.



**Рис. 2.** Классификация

**Вывод:** в ходе выполнения лабораторной работы были получены практические навыки проведения классификации с использованием наивного Байесовского классификатора.