**ЛАБОРАТОРНА РОБОТА № 1**

**ПОПЕРЕДНЯ ОБРОБКА ТА КОНТРОЛЬОВАНА КЛАСИФІКАЦІЯ ДАНИХ**

**Мета роботи:** використовуючи спеціалізовані бібліотеки та мову програмування Python дослідити попередню обробку та класифікацію даних.

**Хід роботи**

**Завдання 1. Попередня обробка даних**

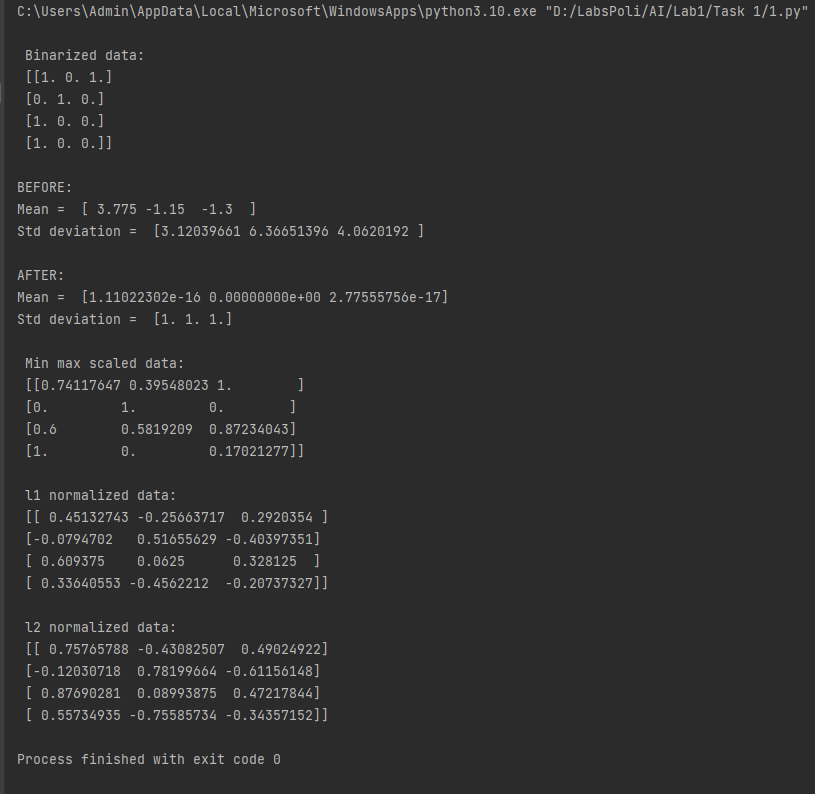


Рис. 1

L1-нормалізація використовує метод найменших абсолютних відхилень, що забезпечує рівність 1 суми абсолютних значень в кожному ряду. L2-нормалізація використовує метод найменших квадратів, що забезпечує рівність 1 суми квадратів 4 значень.



Рис. 2

**Завдання 2. Попередня обробка нових даних**

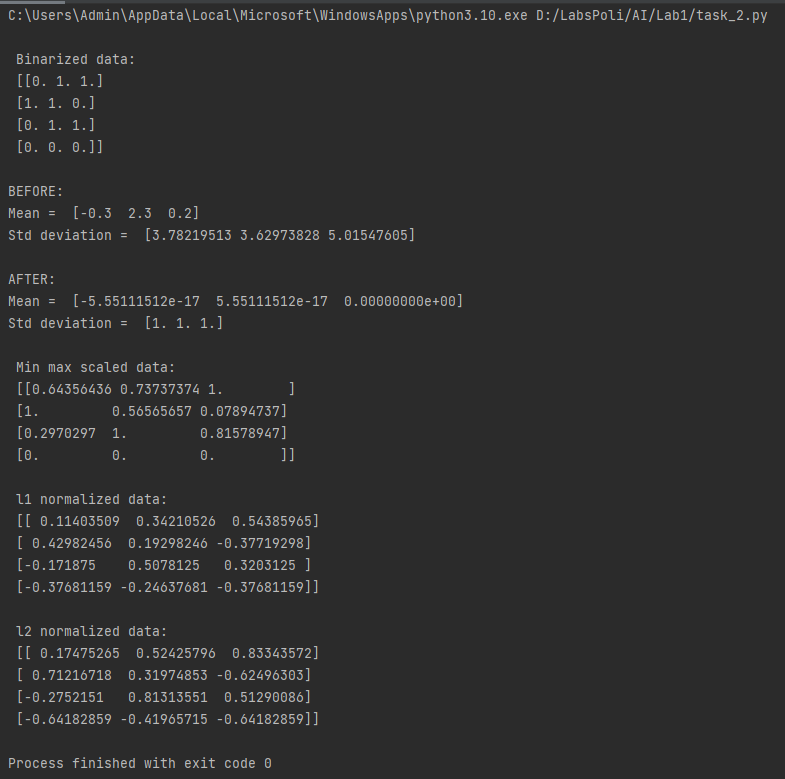


Рис. 3

**Завдання 3. Класифікація логістичною регресією або логістичний класифікатор**

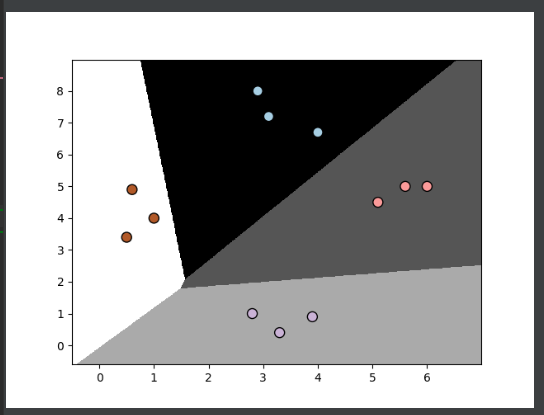


Рис. 4

**Завдання 4. Класифікація наївним байєсовським класифікатором**

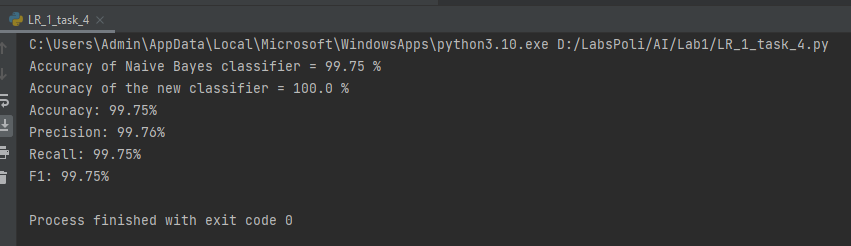
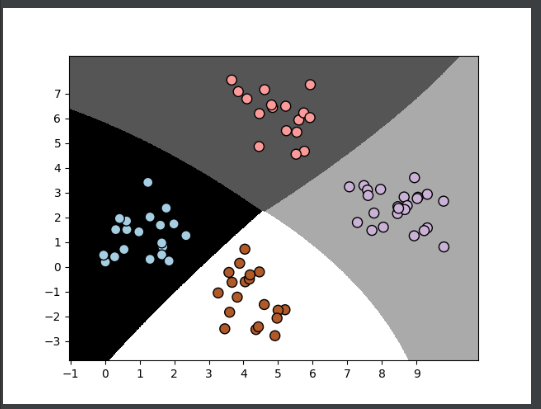
 

Рис. 5

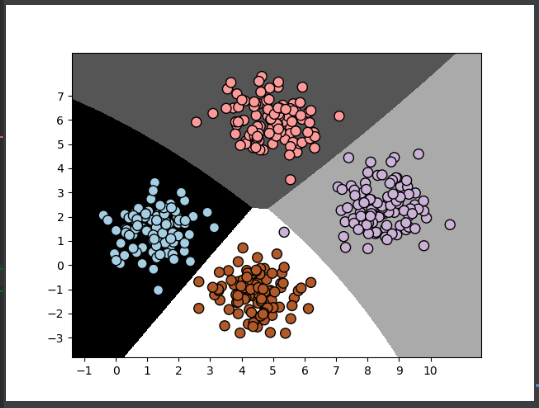
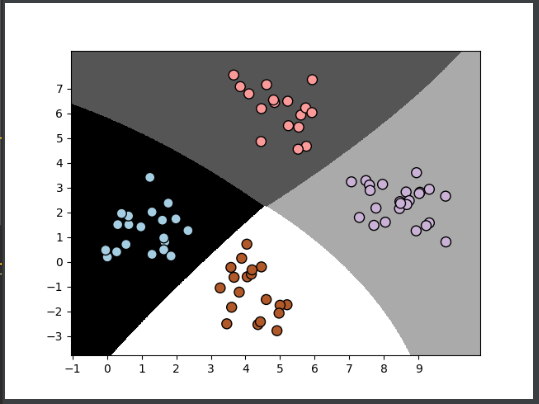


Рис. 6



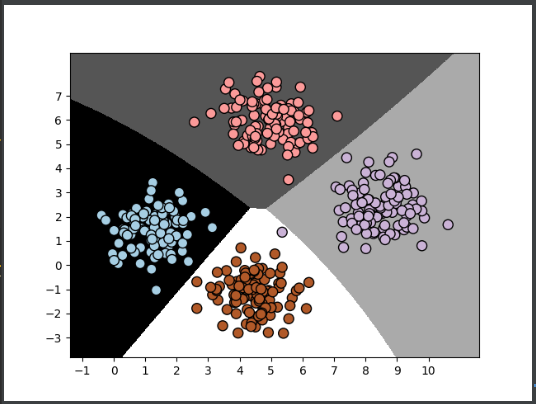


Рис. 7

**Завдання 5. Вивчити метрики якості класифікації**

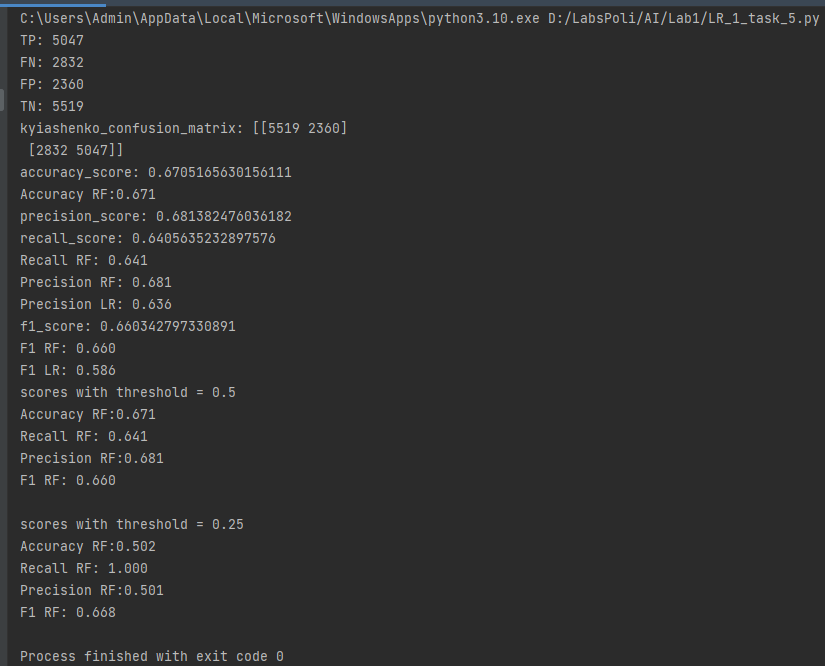
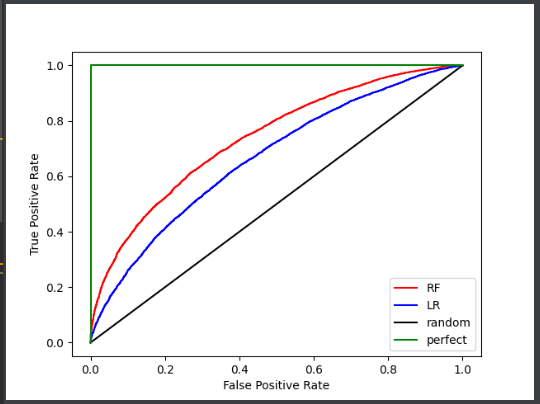
 

Рис. 8

RF модель показала себе більш точною.

**Завдання 6**

Розробіть програму класифікації даних в файлі data\_multivar\_nb.txt за допомогою машини опорних векторів (Support Vector Machine - SVМ). Розрахуйте показники якості класифікації. Порівняйте їх з показниками наївного байєсівського класифікатора. Зробіть висновки яку модель класифікації краще обрати і чому.

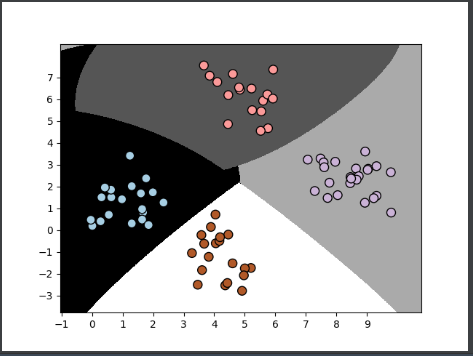
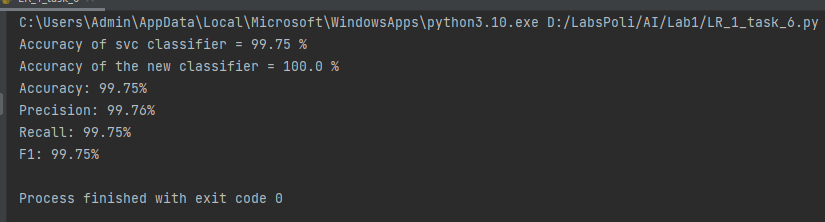
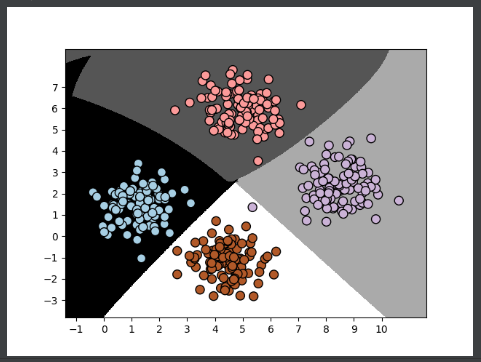
 

Рис. 9

Метод наївного байєсівського класифікатора спрацював точніше і швидше.

Посилання на Git: <https://github.com/Grum74/AI>

**Висновок**

Я, використовуючи спеціалізовані бібліотеки та мову програмування Python дослідив попередню обробку та класифікацію даних.