

Звіт  
По лабораторній роботі №3  
з дисципліни  
«Методи обробки та візуалізації багатовимірних даних»

Виконав: студент гр. ПМ-06-2  
Євсюков Ігор  
Перевірила: доцент Мацуга О. М.

Дніпропетровськ  
2009

## **Зміст**

Постановка задачі.....	3
Опис програми.....	4
Реалізація.....	5
Висновки.....	7
Література.....	8
Додаток 1.....	9

## Постановка задачі

Реалізувати ЕМ-алгоритм або SEM-алгоритм для класифікації.

Програма має забезпечувати:

1. Завантаження із файлу вибірки у вигляді

$$\{X_i, y_i; i = \overline{1, n}\} = \{x_{i1}, x_{i2}, K, x_{ip}, y_i; i = \overline{1, n}\}.$$

2. Розбиття початкової вибірки на 2 частини: навчальну та екзаменаційну.
3. Проведення навчання за навчальною вибіркою.
4. Проведення класифікації об'єктів із екзаменаційної вибірки.
5. Перевірку якості класифікації шляхом обчислення відсотка вірно класифікованих об'єктів.
6. Виведення на графік результатів класифікації, де різними кольорами показати об'єкти різних класів та невірно класифіковані об'єкти (на графіку показати результати, одержані за допомогою методу, а не початкову класифікацію!).

## Опис програми

Програма реалізована на мові високого рівня Java з використанням бібліотеки JFreeChart для виведення графіків. Мова є об'єктно-орієнтованою, тому складається з декількох модулів. Модулі в програмі бувають двох типів - для роботи з даними та для зв'язування графічного інтерфейсу програми з обробкою даних.

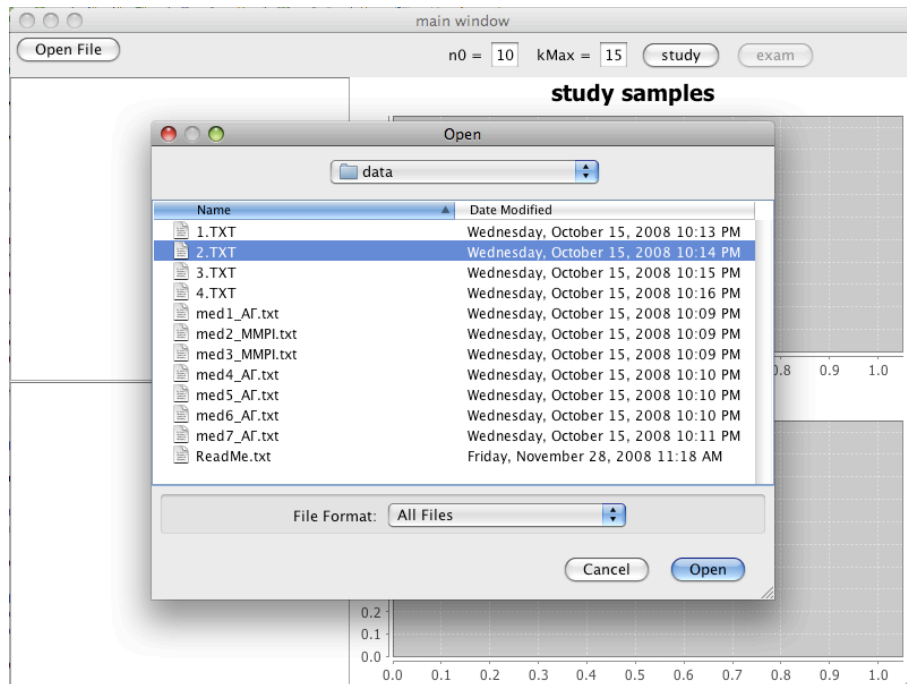
Для зручності були реалізовані наступні класи

- *MDBObject* - клас, що описує простий багатовимірний об'єкт
- *MDSample* - клас, що описує вибірку з об'єктів типу MDBObject заданної розмірності
- *MDComposition* - клас, що представляє суміш нормальних розподілів. В цьому класі, за допомогою SEM алогритму вираховуються параметри нормального розподілу та повертається вірогідність попадання об'єкту в суміш.
- *GaussianCommonWeb* - клас, що описує Гаусівську загальну мережу. Вхідні дані етапу навчання групуються за номером класу і для кожної групи створюється суміш нормальних розподілів.
- *SEM* - в цьому класі реалізовані процедури SEM алгоритму
- *MatrixOperations* - процедури для обробки векторів та масивів

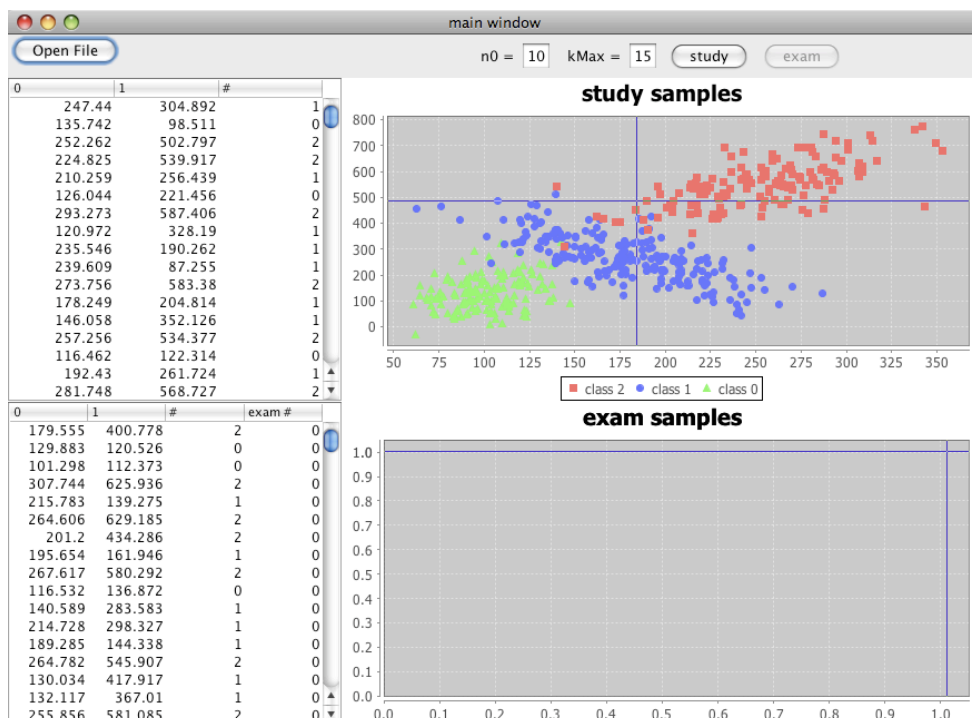
Детальна інформація про класи з описом їх полів та функцій знаходиться у **Додатку 1**

## Реалізація

Після запуску програми необхідно завантажити файл. Діалог вибору файлу з'явиться після натискання на кнопку “Open File”

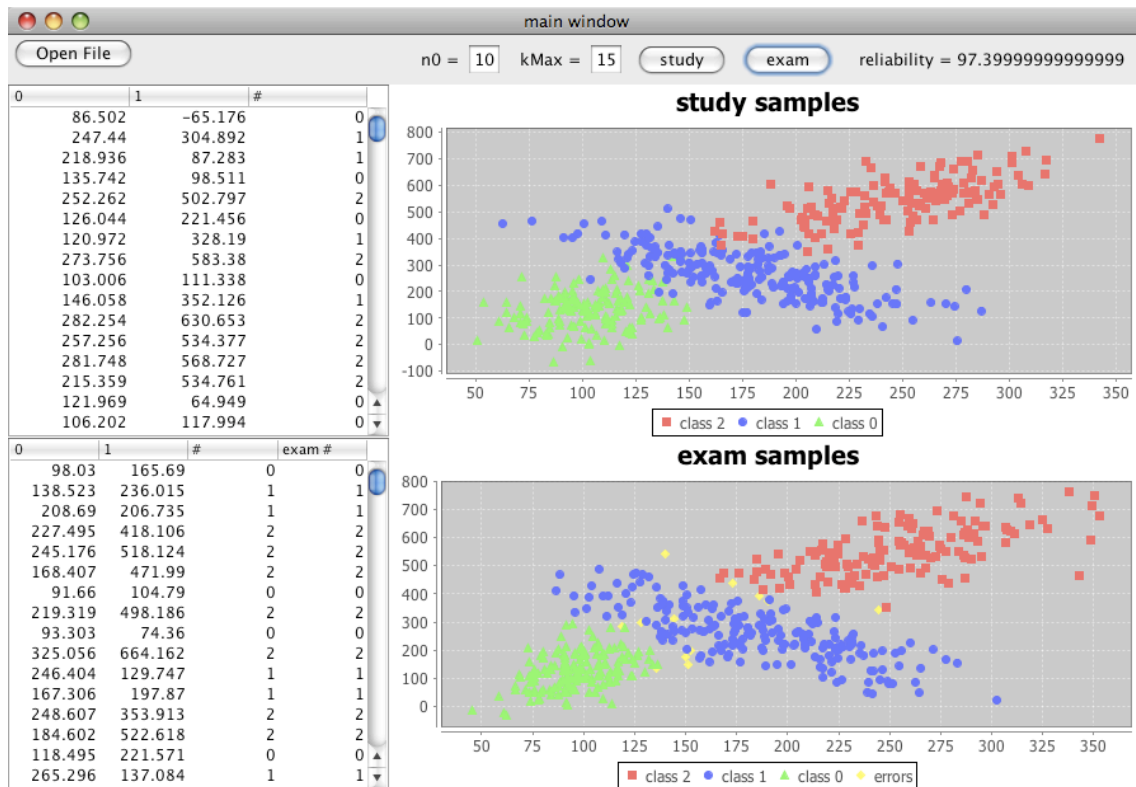


Після завантаження файлу вибірка ділиться на 2 частини порівню, заповняться таблиці з навчальними та екзаменаційними вибірками(номер класу, отриманого на екзамені в таблиці екзаменаційної вибірки буде рівним 0) та графік навчальної вибірки. На графіках по вісі Ох відображаються перша



компонента об'єкта, по вісі Oy - друга. Для проведення навчання необхідно ввести параметри алгоритму SEM - мінімальну кількість елементів в класі та початкову кількість класів у поля напроти "n0" та "kMax" відповідно та натиснути кнопку "Study".

Після проведення навчання, для проведення екзамену необхідно натиснути на кнопку "Exam". Значення класів, отриманих об'єктами на екзамені будуть внесені в таблицю, графік екзаменційної вибірки буде заповнений(помилки екзамену будуть відображені окремою групою) та біля кнопки "Exam" буде



вказана якість визначення у відсотках.

## Висновки

Дана програма повністю відповідає постановці задачі:

- Завантажує із файлу вибірки заданого вигляду
- Розбиває завантажену вибірку на 2 частини
- Проводить навчання за навчаючою вибіркою
- Виводить якість класифікації шляхом обчислення відсотка вірно класифікованих об'єктів
- Виводить на графік результати класифікації, причому невірно класифіковані об'єкти відображаються окремо

## Література

1. Андерсон Т. Введение в многомерный статистический анализ. – М.: Наука, 1963. – 500 с.
2. Дуда Р., Харт П. Распознавание образов и анализ сцен. – М.: Мир, 1976. – 512 с.
3. Миленький А.В. Классификация сигналов в условиях неопределенности. – М.: Сов. радио, 1975. – 328 с.



## **Додаток 1**