Факультет информатики и экономики

Дисциплина «Проектирование интеллектуальных информационных систем»

Группа 1243

Дисциплина «Интеллектуальные системы и технологии»

Группа 1247

**Лабораторная работа №2. Классификация. Метод K ближайших соседей**

1. Есть три группы объектов (класс №1, класс №2 и класс №3), которые характеризуются двумя условными параметрами *x*1 и *x*2. Данные об этих объектах представлены в файле **ДатасетМетодБлижайшихСоседей.xlsx**.

Общее количество объектов каждого вида – 100.

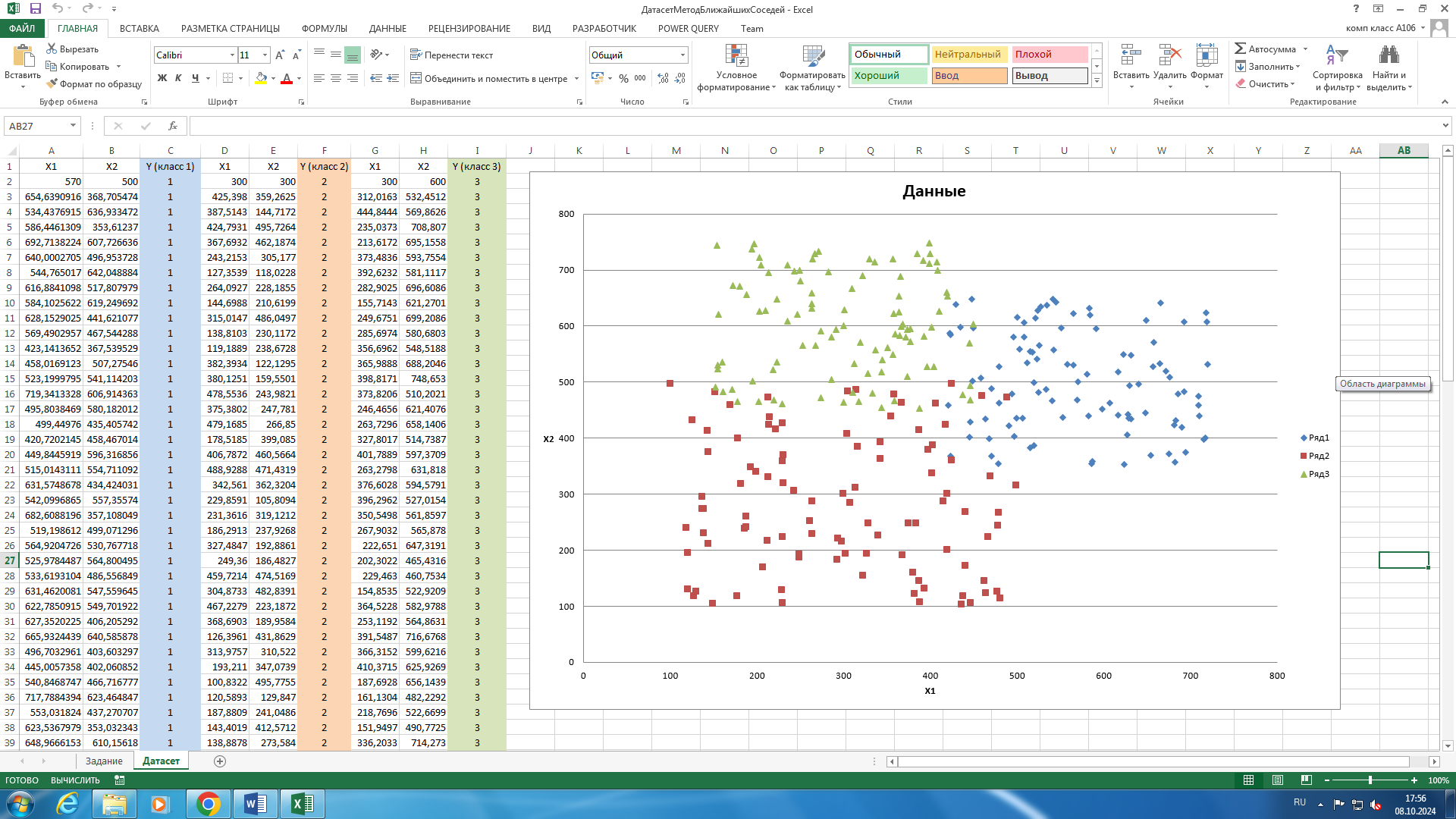


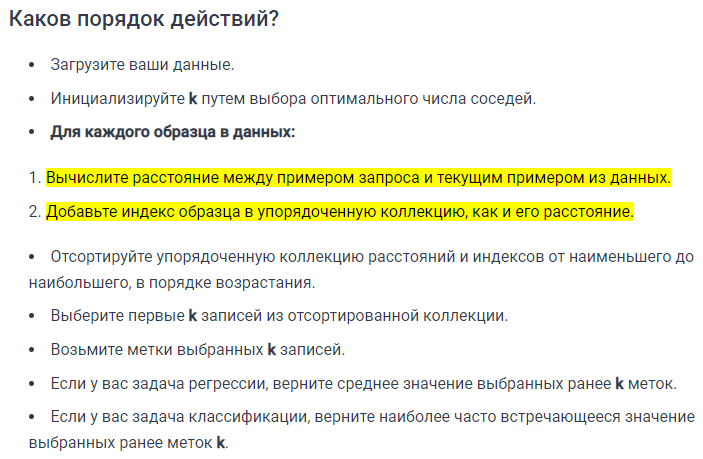
Рис. 1. Данные об объектах

Нужно разбить всю выборку случайным образом на 2 части: основная часть – 80%, тестовая часть – 20%. Объекты из тестовой части выборки нужно классифицировать, используя метод К-ближайших соседей.

1. Программная реализация математического метода.

Язык программирования: Python (с библиотеками numpy, pandas, matplotlib, tkinter).

Описание метода K-ближайших соседей: https://proglib.io/p/metod-k-blizhayshih-sosedey-k-nearest-neighbour-2021-07-19



Необходимо определить суммарную погрешность метода ** для нескольких значений K. Например, для К = 3, 5, 7, 9, 11, 13, 15. Вывести полученные результаты в виде графика зависимости **= *f* (K). Указать оптимальное количество соседей, которые необходимо использовать (локальный минимум на графике).

1. Реализация графического пользовательского интерфейса.

Язык программирования: Python (библиотека tkinter или другие).

Приложение с GUI должно содержать управляющие элементы: для выбора данных, для разбиения данных на обучающую и тестовую в определенной пропорции, для отображения результатов, для графического отображения данных (точек на плоскости) и т.д., для вывода погрешности модели.