Metric based classification

ограничение по времени на тест: 20 секунд

ограничение по памяти на тест: 512 мегабайт

ввод: стандартный ввод

вывод: стандартный вывод

# Задача

Требуется построить метрический классификатор на обучающем наборе данных с известными классами и классифицировать с его помощью тестовый набор данных с неизвестными классами. Ваш классификатор должен содержать несколько метрик, сглаживающих ядер и стратегий выбора ширины окна ядра (числа ближайших соседей). Требуется выбрать оптимальную комбинацию гипер-параметров для каждого отдельного набора данных.

# Входные данные

Первая строка содержит целое число ***M*** (5 ≤ ***M*** ≤ 200) — число признаков у объектов исключая класс.

Вторая строка содержит целое число ***K*** (2 ≤ ***K*** ≤ 25) — число классов.

Третья строка содержит целое число ***N*** (50 ≤ ***N*** ≤ 400) — число объектов в обучающем множестве.

Следующие ***N*** строк содержат описание объектов. ***i***-тая из этих строк содержит описание ***i***-того объекта ***M***+1 целых чисел: первые ***M*** чисел ***Ai,j*** (|***Ai,j***| ≤ 109) — признаки объекта, последнее число ***Ci*** (1 ≤ ***Ci*** ≤ ***K***) — его класс.

Следующая строка содержит целое число ***Q*** (50 ≤ ***Q*** ≤ 400) — число объектов в тестовом множестве.

Следующие ***Q*** строк содержат описание объектов. ***t***-тая из этих строк содержит описание ***t***-того объекта: ***M*** целых чисел ***At,j*** (|***At,j***| ≤ 109) — признаки объекта.

# Выходные данные

Выведите ***Q*** строк. Каждая ***t***-тая строка из них должна содержать результат классификации ***t***-того объекта из тестового множества: целое число ***St*** (1 ≤ ***St*** ≤ 20) — число соседей классифицируемого объекта, затем следует ***St*** пар чисел ***i*** и ***w*** (1 ≤ ***i*** ≤ ***N***, 0 ≤ **w** ≤ 106), где ***i*** — целое число, индекс объекта из тренировочного множества, а ***w*** — вещественное число с плавающей точкой, вес с которым учитывается этот объект. Числа ***St*** не обязательно должны быть одинаковыми для всех объектов.

# Система оценки

Для каждого объекта ***t*** будет рассчитан его предсказанный класс ***pt*** = argmax(∑***w***∙[***Ci*** = *c*]) исходя из соответствующего множества пар (***i***,***w***), в случае неопределённости класс выбирается псевдослучайно. Далее на основании предсказанных и реальных классов вычисляется усреднённая по классам микро F мера. Тест считается пройденным, если эта F мера будет выше определённого порога, рассчитанного с 5% запасом с помощью базового метрического классификатора.

# Пример

|  |  |
| --- | --- |
| **Входные данные** | **Выходные данные** |
| 2  2  4  2 2 1  6 2 1  4 1 2  4 3 2  2  1 2  5 2 | 3 1 0.75 3 0.21 4 0.21  3 2 0.75 3 0.65 4 0.65 |

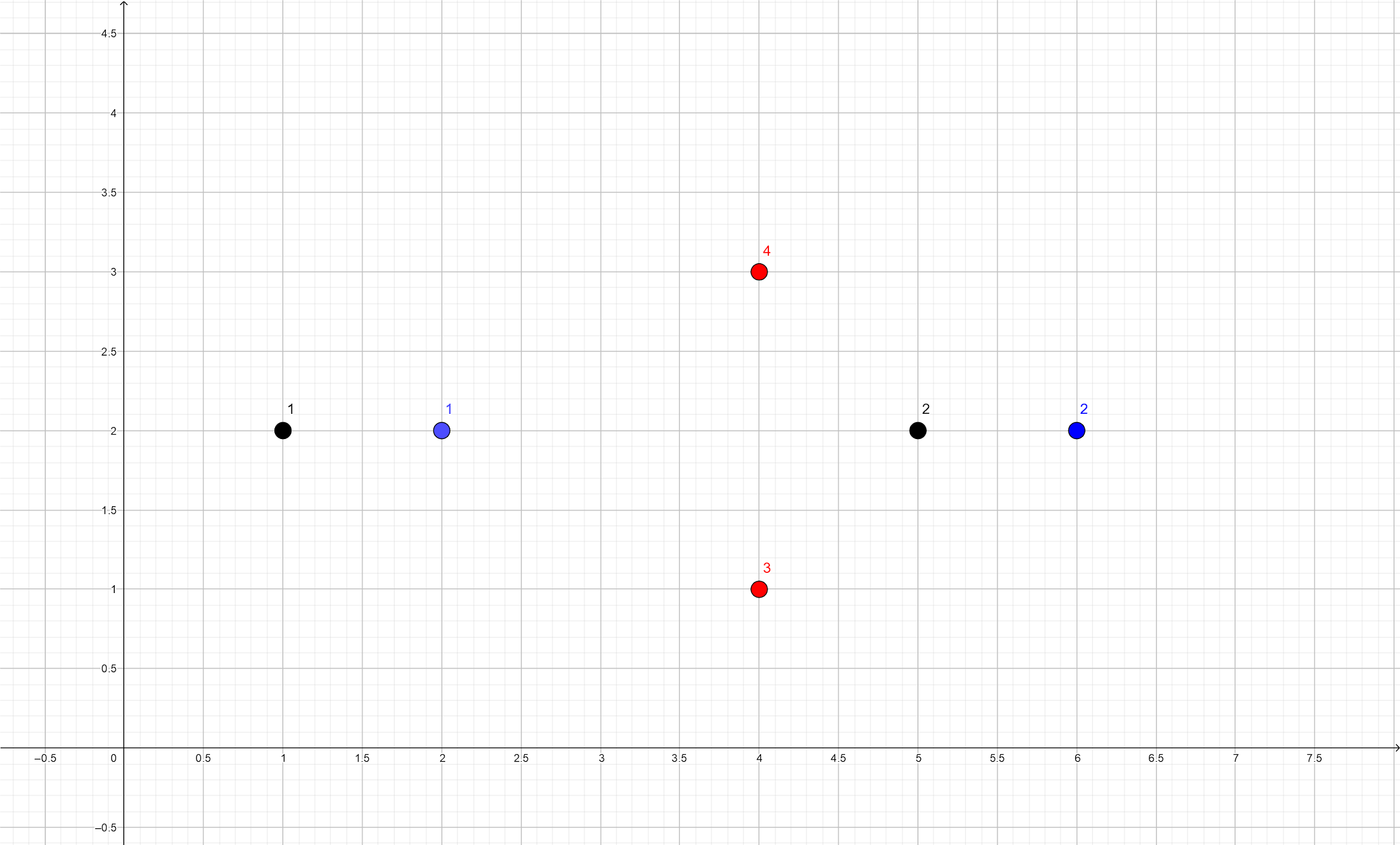


Иллюстрация примера

## Пояснение

Обратите внимание, что экземпляр задачи из данного примера не подходит под нижние ограничения на число объектов и признаков!

В данном примере в качестве метрики используется Евклидово расстояние, а в качестве взвешивающей функции Треугольное ядро. Используется три ближайших соседа с шириной окна 4.

Для первого запроса:

|  |  |  |  |
| --- | --- | --- | --- |
| Номер объекта: | 1 | 3 | 4 |
| Расстояние: | 1 | √10 | √10 |
| Полученный вес: | 0.75 | 0.21 | 0.21 |
| Класс: | 1 | 2 | 2 |

Суммарный вес первого класса **0.75**, второго 0.42.

Для второго запроса:

|  |  |  |  |
| --- | --- | --- | --- |
| Номер объекта: | 2 | 3 | 4 |
| Расстояние: | 1 | √2 | √2 |
| Полученный вес: | 0.75 | 0.65 | 0.65 |
| Класс: | 1 | 2 | 2 |

Суммарный вес первого класса 0.75, второго **1.3**.