

Определение расстояния между точкой и множеством, представленным бинарной диаграммой решений

Курсовая работа

студента 345 группы Зубаревича Дмитрия

Научный
руководитель:

к.ф.-м.н., доцент
кафедры
информатики
Бугайченко Д.Ю.

Цель работы

Выявление закономерностей в многомерных данных с использованием *бинарных диаграмм решений*.

Задачи:

- Разработка алгоритма
- Реализация алгоритма
- Тестирование алгоритма

Бинарные диаграммы решений

- Кодирование объекта графовой структурой
- Медленный рост графа при наличии закономерностей в объекте
- Операции за полиномиальное время от размера графовых кодировок операндов

Расстояние между точкой и множеством

$S = \{s_1, \dots, s_{|S|}\}$ – множество,
 $\rho : T \times T \rightarrow R$ – метрика, где $T \supset S$.

def Расстояние – функционал: $\delta_{\rho,S}(y) = \begin{cases} 0, & \text{если } y \in S, \\ \frac{\sigma_{\rho,S}(y)}{|S|}, & \text{если } y \notin S, \end{cases}$
где $\sigma_{\rho,S}(y) = \sum_{i \in 1:|S|} \rho(s_i, y)$.

В БДР работа с множеством S осуществляется с помощью характеристической функции $\chi_S \Rightarrow$

$$\sigma_{\rho,S}(y) = \sum_{x \in T} \chi_S(x) * \rho(x, y).$$

Вычисление расстояния

$$\left. \begin{array}{l} \chi_s(x) \rightarrow A - \text{вектор-строка над } \mathbb{R}, \text{ где } A_{1x} = \chi_s(x) \\ \rho(x, y) \rightarrow B - \text{матрица над } \mathbb{R}, \text{ где } B_{xy} = \rho(x, y) \end{array} \right\} \Rightarrow$$

$$(A * B)_{1y} = \sum_{x \in T} A_{1x} * B_{xy} = \sum_{x \in T} \chi_s(x) * \rho(x, y) = \sigma_{\rho, s}(y) \Rightarrow$$

$$A * B \rightarrow \sigma_{\rho, s}(y) \Rightarrow$$

$$\sigma_{\rho, s}(y) = (\chi_s \times \rho)(y), \text{ где } \times - \text{матричное умножение}$$

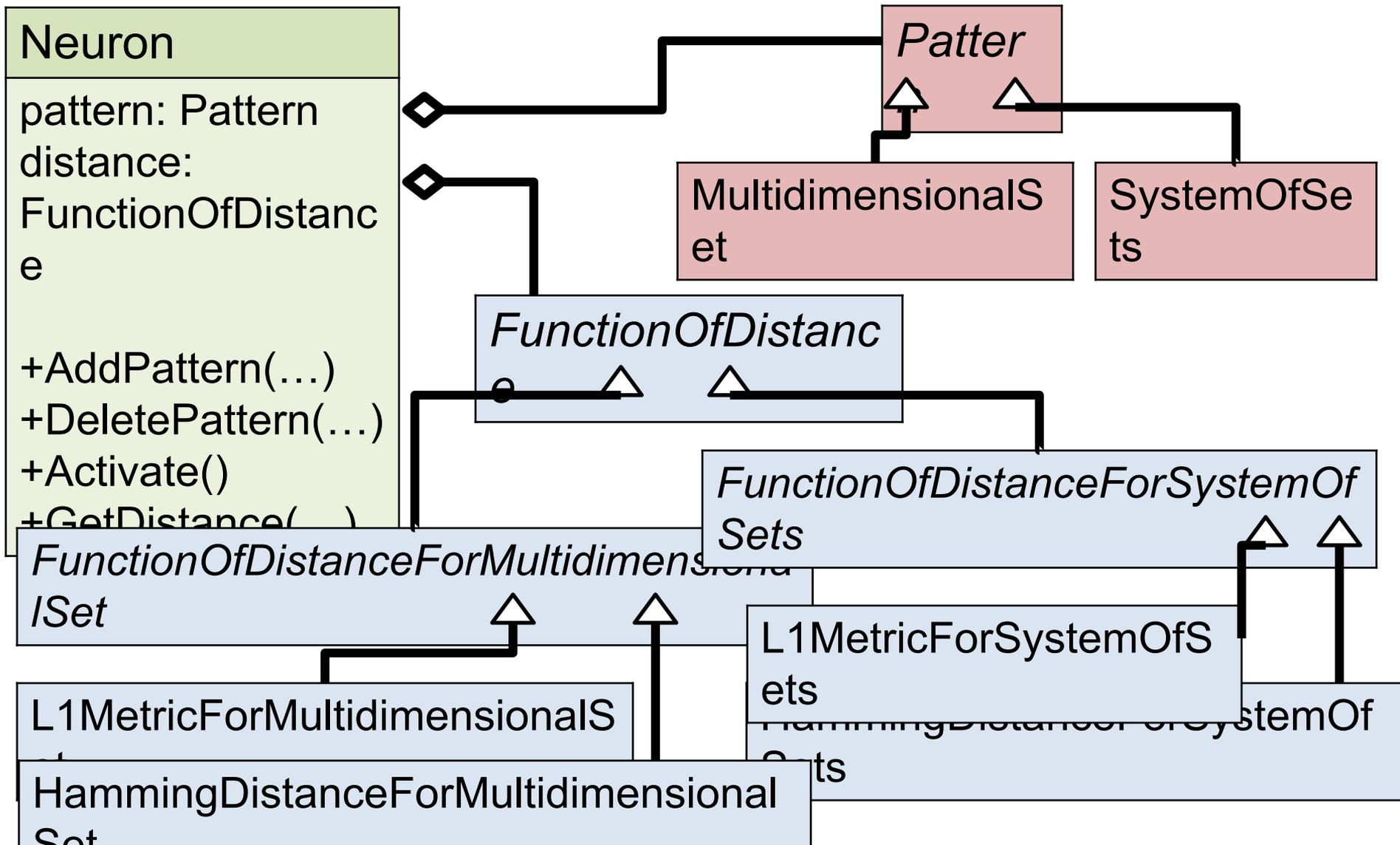
Идея работы

Критерий	БДР	ИНС
Цель	Аппроксимация некоторой функции	
Метод	Кодирование таблицы значений	Итеративный подбор параметров
Ресурсоемкость	Высокая	Средняя
Представление данных	Явное	Неявное

Идея: создать нейрон на основе БДР

- Анализ “знаний” ИНС
- Уменьшение ресурсоемкости БДР

Архитектура



Результаты тестирования

- Создана ИНС из 10 нейронов, по одному на каждую цифру от 0 до 9
- Сеть обучена на 4-ёх шрифтах и умеет распознавать 5-ый.

Итоги

- Разработан алгоритм определения расстояния от точки до множества
- Алгоритм реализован на основе библиотеки BDDFunctions
- Успешно проведены тесты алгоритма
- Исходный код можно увидеть по адресу:
<http://code.google.com/p/bddffunctions/>