

СРАВНЕНИЕ СИСТЕМЫ «DISCOVERY» с MICROSOFT ASSOCIATION RULES¹

Николай Иванович Фирсов

Институт Систем Информатики им. А.П. Ершова

e-mail: <nickfirsov@gmail.com>

Abstract

We have developed a relational approach (Relational Data Mining) to Data Mining, and software system «Discovery», that realize this approach. The goal of this paper is to provide both theoretical and experimental comparisons of «Discovery» system versus Microsoft Association Rules. We show that «Discovery» system performs better for rules discovery and prediction tasks than the Association Rules, and, can analyze high-noisy data, for example, financial time series.

Аннотация

Нами был разработан реляционный подход (Relational Data Mining) к методам извлечения знаний и программная система «Discovery» [1, 2, 4], снимающие практически все ограничения с методов KDD&DM (Knowledge Discovery in Data Bases and Data Mining).

Система Discovery, обладает следующими важными теоретическими свойствами: может обнаруживать теорию предметной области, может обнаруживать все правила, имеющие максимальные условные вероятности, может обнаруживать непротиворечивую вероятностную аппроксимацию теории предметной области [2], обнаруживает все максимально специфические правила, позволяющие предсказывать без противоречий [3].

В данной работе проводится теоретическое и экспериментальное сравнение системы «Discovery» с алгоритмом Microsoft Association Rules. Мы показали, что система «Discovery» больше подходит для обнаружения закономерностей и прогнозирования, чем Association Rules, а также, в отличие от алгоритма Association Rules, метод Discovery позволяет обнаруживать знания в сильно зашумленных данных, какими, например, являются финансовые временные ряды [4].

ЛИТЕРАТУРА

- [1] Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. – НГУ, Новосибирск, 2006. – 293 с.
- [2] Evgenii Vityaev, Boris Kovalerchuk, Empirical Theories Discovery based on the Measurement Theory. Mind and Machine, v.14, #4, 551-573, 2004)
- [3] Evgenii Vityaev. The logic of prediction. In: Mathematical Logic in Asia. Proceedings of the 9th Asian Logic Conference (August 16-19, 2005, Novosibirsk, Russia), World Scientific, Singapore, 2006, pp.263-276).
- [4] Kovalerchuk, B.Y., Vityaev E.E. Data mining in finance: Relational and hybrid methods, Kluwer, 2000, 308 p.

¹ Работа поддержана грантом РФФИ РФФИ №11-07-00560а; интеграционными проектами СО РАН № 47, 115, 119, а также работа выполнена при финансовой поддержке Совета по грантам Президента РФ и государственной поддержке ведущих научных школ (проект НШ-3606.2010.1.)