

# ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ КАК ОСНОВА УПРАВЛЕНИЯ

В связи с совершенствованием технологий записи и хранения данных существует огромное количество неструктурированной информации в самых различных областях человеческой деятельности. В связи с этим используется Data Mining - совокупность методов обнаружения в данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

## ВВЕДЕНИЕ

Современные технологии Data Mining перерабатывают информацию с целью автоматического поиска шаблонов (паттернов), характерных для каких-либо фрагментов неоднородных многомерных данных. Data Mining — это совокупность различных методов обнаружения знаний. Выбор метода часто зависит от типа имеющихся данных и от того, какую информацию вы пытаетесь получить. Некоторые из методов: ассоциация (объединение), классификация, кластеризация, анализ временных рядов и прогнозирование, нейронные сети и другие. В Data Mining для представления полученных знаний служат модели. Виды моделей зависят от методов их создания. Наиболее распространенными являются: правила, деревья решений, кластеры и математические функции.

### I. МЕТОДЫ DATA MINING

Сфера применения Data Mining ничем не ограничена. Data Mining востребован везде, где имеются какие-либо данные. Методы Data Mining позволяют решить многие задачи, с которыми сталкивается аналитик. Из них основными являются: классификация, регрессия, поиск ассоциативных правил и кластеризация. Ниже приведено краткое описание основных задач анализа данных.

Задача классификации сводится к определению класса объекта по его характеристикам. Необходимо заметить, что в этой задаче множество классов, к которым может быть отнесен объект, заранее известно.

Задача регрессии, подобно задаче классификации, позволяет определить по известным характеристикам объекта значение некоторого его параметра. В отличие от задачи классификации значением параметра является не конечное множество классов, а множество действительных чисел.

Задача классификации и регрессии имеет геометрическую интерпретацию. Рассмотрим ее на примере с двумя независимыми переменными,

*Волк Александр Викторович*, магистрант кафедры информационных технологий автоматизированных систем БГУИР, av.volk89@gmail.com.

*Научный руководитель: Гуринович Алевтина Борисовна*, доцент кафедры вычислительных методов и программирования БГУИР, кандидат технических наук, gurinovich@bsuir.by.

что позволит представить ее в двумерном пространстве, представленном на рис.1. Каждому объекту ставится в соответствие точка на плоскости. Символы + и - обозначают принадлежность объекта к одному из двух классов. Данные имеют четко выраженную структуру: все точки класса + сосредоточены в центральной области. Построение классификационной функции сводится к построению поверхности, которая обводит центральную область. Она определяется как функция, имеющая значения + внутри обведенной области и - вне области.

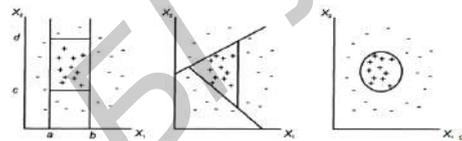


Рис. 1 – Классификация в двухмерном пространстве

Задача ассоциации. При поиске ассоциативных правил целью является нахождение частых зависимостей между объектами или событиями. Найденные зависимости представляются в виде правил и могут быть использованы для лучшего понимания природы анализируемых данных и для предсказания появления событий.

### II. ВЫВОДЫ

Таким образом, использование методов Data Mining на предприятии позволяет получать не просто набор бессвязной информации, а даёт возможность получить необходимую информацию для принятий управленческих решений, что ведет к повышению конкурентоспособности и

1. Барсеян А. Б., Куприянов М. В., Методы и модели анализа данных: OLAP и Data Mining // А. Б. Барсеян, Куприянов М. В. // БХВ-Петербург, С-Петербург 2004. - 336 с.
2. Паклин, Н. Ю. Бизнес-аналитика: от данных к знаниям // Н. Ю. Паклин // Питер, С-Петербург, 2009. - 271 с.