

Вывод: Исследование таких основных понятий инноваций, как «инновационный процесс» и «инновационная деятельность», имеет существенное теоретическое значение. Инновации позволяют сделать вывод о том, что специфическое содержание инновации представляют изменения, а основной функцией инновационной деятельности является функция превращений. Поэтому можно сказать, инновация – непрерывный процесс развития и совершенствования научно-технического прогресса.

Таким образом, можно сказать, что точное и правильное понимание сущности основных экономических сроков значительно облегчает дальнейшие исследования в области инноваций.

УДК 519.25

Давид Латышевич
(Республика Беларусь)

Научный руководитель И.И. Станкевич, ст. преподаватель
Белорусский государственный аграрный технический университет

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ

Интеллектуальный анализ данных – это направление информационных технологий, охватывающее всю область проблем, связанных с извлечением знаний из массивов данных.

Часто интеллектуальный анализ данных рассматривают как трансформацию термина Data Mining (раскопка, разработка данных) и считают два этих понятия синонимами, и в широком смысле это вполне допустимо. Однако считать их абсолютно идентичными было бы неверно, поскольку методы Data Mining сосредоточены главным образом на процессе анализа данных и интеллектуальном моделировании. В то же время считается, что интеллектуальный анализ данных рассматривает весь спектр проблем, связанный с процессом извлечения знаний из баз данных.

В основе интеллектуального анализа данных, как и в Data Mining, лежит идея активного применения математических методов, таких как оптимизация, генетические алгоритмы, распознавание образов, статистика, Data Mining и т. д., а также использующих визуальное представление информации. Несколько основных методов, которые используются для интеллектуального анализа данных, описывают

тип анализа и операцию по восстановлению данных. К сожалению, разные компании и решения не всегда используют одни и те же термины, что может усугубить путаницу и кажущуюся сложность.

Рассмотрим некоторые ключевые методы и примеры того, как использовать те или иные инструменты для интеллектуального анализа данных:

Ассоциация (или отношение), вероятно, наиболее известный, знакомый и простой метод интеллектуального анализа данных. Для выявления моделей делается простое сопоставление двух или более элементов, часто одного и того же типа.

Классификацию можно использовать для получения представления о типе покупателей, товаров или объектов, описывая несколько атрибутов для идентификации определенного класса.

Кластеризация. Исследуя один или более атрибутов или классов, можно сгруппировать отдельные элементы данных вместе, получая структурированное заключение. На простом уровне при кластеризации используется один или несколько атрибутов в качестве основы для определения кластера сходных результатов. Кластеризация полезна при определении различной информации, потому что она коррелируется с другими примерами, так что можно увидеть, где подобию и диапазоны согласуются между собой.

Прогнозирование – это широкая тема, которая простирается от предсказания отказов компонентов оборудования до выявления мошенничества и даже прогнозирования прибыли компании. В сочетании с другими методами интеллектуального анализа данных прогнозирование предполагает анализ тенденций, классификацию, сопоставление с моделью и отношения. Анализируя прошлые события или экземпляры, можно предсказывать будущее.

Последовательные модели, которые часто используются для анализа долгосрочных данных. Является полезным методом выявления тенденций или регулярных повторений подобных событий.

Дерево решений, связанное с большинством других методов (главным образом, классификации и прогнозирования), можно использовать либо в рамках критериев отбора, либо для поддержки выбора определенных данных в рамках общей структуры. Дерево решений начинают с простого вопроса, который имеет два ответа (иногда больше). Каждый ответ приводит к следующему вопросу, помогая классифицировать и идентифицировать данные или делать прогнозы.

Комбинации. На практике очень редко используется только один из этих методов. Классификация и кластеризация – подобные методы. Используя кластеризацию для определения ближайших соседей, можно дополнительно уточнить классификацию. Деревья решений часто используются для построения и выявления классификаций, которые можно проследивать на исторических периодах для определения последовательностей и моделей.

Обработка с запоминанием. При всех основных методах часто имеет смысл записывать и впоследствии изучать полученную информацию. Деревья решений редко строятся один раз и никогда не забываются. При выявлении новой информации, событий и точек данных может понадобиться построение дополнительных ветвей или даже совершенно новых деревьев.

Получение и подготовка данных. Сам интеллектуальный анализ данных опирается на построение подходящей модели и структуры, которые можно использовать для обработки, выявления и создания необходимой информации. Независимо от формы и структуры источника данных, информация структурируется и организуется в соответствии с форматом, который позволяет выполнять интеллектуальный анализ данных с максимально эффективной моделью.

Опора на SQL. Наиболее простым из всех подходов часто служит опора на базы данных SQL. SQL (и соответствующая структура таблицы) хорошо понятен, но структуру и формат информации нельзя игнорировать полностью. Например, при изучении поведения пользователей по данным о продажах в модели данных SQL (и интеллектуального анализа данных в целом) существуют два основных формата, которые можно использовать: транзакционный и поведенческо-демографический.

Нетривиальность разыскиваемых шаблонов означает, что найденные шаблоны должны отражать неочевидные, неожиданные (unexpected) регулярности в данных, составляющие так называемые скрытые знания (hidden knowledge). К обществу пришло понимание, что сырые данные (raw data) содержат глубинный пласт знаний.

Управление при должном анализе данных в то же время помогает сократить издержки, устраняет кассовые издержки, а также сокращает время на учет, позволяет улучшить работу предприятия и улучшает атмосферу предприятия.