

УДК 330.46

СОВЕРШЕНСТВОВАНИЕ ПРОЦЕССА ПРИНЯТИЯ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ В ЭКОНОМИКЕ И БИЗНЕСЕ НА ОСНОВЕ ПРИМЕНЕНИЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

¹Васильев Е.П., ²Орешков В.И.

¹ФГБОУ ВПО «Рязанский государственный агротехнологический университет им. П.А. Костычева», Рязань, e-mail: university@rgatu.ru;

²ФГБОУ ВПО «Рязанский государственный радиотехнический университет», Рязань, e-mail rgrtu@rsreu.ru.

Эффективная организация поддержки принятия решений является одной из составляющих конкурентоспособности компаний. Обоснованное управленческое решение невозможно принять без достоверных и нетривиальных знаний о предметной области. Источником таких знаний являются данные, описывающие свойственные ей экономические и бизнес процессы и явления. Основным инструментом поиска знаний в базах данных предприятий и организаций является интеллектуальный анализ данных (Data Mining), в основе которого лежит использование интеллектуальных моделей, основанных на обучении, которые позволяют в автоматическом режиме обнаруживать в данных скрытые зависимости, закономерности и структуры. Их интерпретация и осмысление, позволяют специалисту предметной области ему генерировать выводы и суждения, на основе которых могут быть приняты качественные управленческие решения.

Ключевые слова: управленческое решение, знание, интеллектуальный капитал, интеллектуальный анализ, онтология, машинное обучение, аналитическая система

IMPROVING MANAGEMENT DECISION MAKING IN ECONOMICS AND BUSINESS THROUGH THE USE OF INTELLIGENT DATA ANALYSIS

¹Vasiliev E.P., ²Oreshkov V.I.

¹Ryazan State Agrotechnological University n.a. P.A. Costychev, Ryazan, e-mail: university@rgatu.ru;

²Ryazan State Radio Engineering University, Ryazan, e-mail rgrtu@rsreu.ru.

Effective organization of decision support is one of the components-term competitiveness of companies. Informed management decisions can not be accepted without a valid and nontrivial knowledge about the domain. The source of such knowledge are the data describing characteristic economic and business processes and phenomena of the domain. The main tool for knowledge discovery in databases of businesses and organizations is a data mining, which is based on the use of intelligent models, based on the machine learning, that allow automatically discovered, hidden dependencies, patterns and textures. Their interpretation and understanding of specialist subject area, allow it to generate conclusions and judgments on the basis of which may be adopted quality management solutions.

Keywords: management decision, knowledge, intellectual capital, intellectual analysis, ontology, machine learning, analytical system

Процесс принятия управленческих решений (УР) является важнейшим элементом системы управления в любой компании. От обоснованности решений, от того, насколько полно при их выработке учитывается предыдущий опыт, текущее и прогнозируемое состояние дел на предприятии и во внешнем окружении, зависит, в конечном итоге, эффективность его функционирования и перспективы развития.

УР могут приниматься с целью обеспечения устойчивого функционирования экономической или бизнес-системы в рамках заданных параметров, либо с целью поиска новых возможностей роста и перспектив развития. Независимо от цели, УР должны приниматься только на основе глубоких, достоверных, нетривиальных знаний о предметной области. Важную роль в данном процессе играет интеллектуальный капитал (ИК) организации – совокупность знаний, умений и опыта ее работников. Он является общепризнанным фактором конкурентно-

го преимущества, особенно в условиях современной постиндустриальной экономики – экономикой знаний. Поэтому современные, динамично развивающиеся компании уделяют огромное внимание развитию ИК и вкладывают в его формирование значительные средства, постоянно ищут формы и методы его использования для совершенствования процессов управления.

Однако процесс создания и накопления ИК в современных условиях – весьма непростое дело. Одной из причин этого является высокая динамика экономической и бизнес-среды: стремительно развиваются технологии, появляются новые товары и услуги, формируются и распадаются рынки, разрабатываются новые формы ведения бизнеса. Поэтому текущие знания, умения и опыт, еще вчера позволяющие добиться успеха, просто устаревают и становятся бесполезными. Второй фактор – высокая сложность современных экономических и бизнес-процессов, которые описываются

десятками и сотнями параметров, находящимися в очень сложной взаимосвязи. Поэтому, чтобы разобраться в текущих тенденциях, выявить важные закономерности и принять на их основе своевременное и обоснованное УР, требуется не только использование человеческого ИК, но и новейших достижений IT-технологий.

Роль и место интеллектуальных информационных технологий в процессе поиска знаний. Важной особенностью современной бизнес-среды является смещение центров принятия УР от высших эшелонов менеджмента компаний на уровень специалистов, непосредственно интегрированных в бизнес-процессы. Это связано с наличием у последних более точной и актуальной информации о текущих проблемах, а также требованиями к оперативности принимаемых решений. В результате в процесс «добычи» знаний, необходимых для принятия УР, вовлекается огромное количество специалистов с самым разнообразным уровнем и профилем образования, знанием компьютера и программного обеспечения. Это создает для компаний, разработчиков компьютерных систем, ориентированных на поддержку принятия решений, огромного рынка клиентов в самых разнообразных областях экономики и бизнеса. Анализ данного рынка показывает, что все программные продукты, реализующие те или иные аспекты поддержки принятия УР, можно отнести к двум направлениям:

– классические системы поддержки принятия решений, строящиеся на основе инженерии знаний, экспертные системы;

– системы Knowledge Discovery (открытие знаний), ориентированные на поиск знаний в данных, накапливаемых в БД компаний в процессе электронной регистрации ими фактов хозяйственной деятельности.

Системы 1-го типа реализуют онтологический подход – формализация области знаний с помощью некоторой концептуальной схемы. Формализованные знания организуются в базу знаний (БЗ), откуда с помощью подсистемы логического вывода и интеллектуального интерфейса предоставляются пользователю. Системы 2-го типа реализуют аналитический подход, в основе которого лежит построение компьютерных моделей, отражающих зависимости, закономерности и структуры в данных, интерпретация и осмысление которых человеком, позволяют ему генерировать новые знания о предметной области, описываемой этими данными.

Таким образом, главным отличием онтологического подхода к поиску знаний от аналитического, заключается в самом

представлении знания. Первый предполагает, что знания существуют отдельно от сознания человека и представляют собой совокупность специально организованной информации и правил вывода. В рамках аналитического подхода знание рассматривается как субъективный образ реальности, отражаемый в сознании человека в виде понятий и представлений. Тогда задача заключается в обнаружении в данных, описывающих предметную область, зависимостей, закономерностей и структур, интерпретация которых специалистом позволяет сформировать понятия, выводы и суждения, необходимые для принятия решений.

В условиях высокой оснащенности компаний электронными средствами регистрации, сбора и хранения данных, отражающих их деятельность, все большую актуальность приобретает аналитический подход к организации поддержки принятия решений, т.е. стремление получить знания «от данных». Кроме этого, внедрение систем 1-го типа на уровне пользователей, непосредственно интегрированных в бизнес-процессы, сталкивается с большими проблемами, основными из которых являются:

– высокая трудоемкость формирования БЗ, в процессе которого эксперт и инженер по знаниям вручную описывают и вводят в нее факты предметной области, разрабатывают средства логического вывода. В условиях высокой сложности и динамики современной бизнес-среды формирование такой БЗ может продлиться неопределенно долго;

– БЗ обычно являются предметно-ориентированными и трудно актуализируются;

– работая с экспертной системой, пользователь, фактически, применяет формализованные знания экспертов, которые уже не являются новыми, что во многом обесценивает их роль в получении конкурентных преимуществ.

Аналитический подход к проблеме поиска знаний «от данных» начал развиваться в рамках математической и прикладной статистики в части методов восстановления зависимостей из выборочных данных. Поэтому одним из главных инструментов исследования стали специализированные статистические пакеты (SAS, SPSS, Statistical, SYSTAT, Minitab, STADIA и др.). Работа с ними не вызывает проблем до тех пор, пока использующие их специалисты имеют соответствующую подготовку, решаемые задачи хорошо формализованы, а анализируемые данные содержатся в относительно небольших локальных БД. Однако, начиная примерно с середины 90-х г. XX века, на предприятиях и в организациях стали

накапливаться огромные массивы информации, аналитическая обработка которой потенциально позволяла получить новые практически полезные знания. Использование методов прикладной статистики и статистических пакетов столкнулось с рядом проблем, основными из которых являлись:

- очень большие объемы данных, разнообразие их представлений, форматов и происхождения, требовали включения в процедуру аналитической обработки мощного модуля управления данными, в котором бы осуществлялся доступ к источникам данных, их консолидация и предобработка;

- вовлечение в процессы принятия УР большого числа лиц, не имеющих достаточного уровня знаний в области математики и статистики;

- некоторые статистические подходы оказались плохо приспособлены для работы с «живыми» данными: мало реалистичные предположения о независимости и нормальности признаков, стремление навязать вероятностную сущность вполне детерминированным явлениям, оперирование фиктивными величинами типа средних, плохое отражение причинно-следственных связей, – все это ограничивало возможности использования методов прикладной статистики для извлечения знаний из реальных данных.

Между тем руководителей компаний, вкладывающих значительные средства в развитие IT-служб, в большей степени интересовала не математическая корректность и достоверность результатов, а практическая значимость сделанных выводов и суждений. В тех случаях, когда возможности статистических подходов не позволяли решить поставленную задачу, исследователи стали привлекать методы из других научных областей, таких, как базы данных, теория информации, искусственный интеллект, распознавание образов и машинное обучение, комплексировать их. В результате сформировалось направление, получившее название Data Mining (раскопка, разработка данных) или интеллектуальный анализ данных (ИАД), в котором статистические методы в основном играют вспомогательную роль, и используются только на некоторых этапах процесса анализа. Ядром ИАД являются методы машинного обучения (МО) – нейронные сети, деревья решений, самоорганизующиеся карты признаков, ассоциативные правила и др., которые позволяют автоматически, с минимальным вмешательством пользователя извлекать из данных зависимости и закономерности и визуализировать в наиболее удобном для восприятия и интерпретации виде.

Прикладные области ИАД. В настоящее время ИАД можно рассматривать как своего рода мейнстрим в области информационных технологий поиска знаний. В различных предметных областях формируются концепции бизнес-анализа, использующие интеллектуальные модели для поддержки принятия решений. Типичными примерами таких концепций являются точное земледелие в агробизнесе и анализ рыночной корзины в розничной торговле.

Концепция точного земледелия основана на том, что земельные площади, используемые под те или иные культуры, не являются однородными по своим агрохимическим и агрофизическим свойствам, степени увлажненности и т.д. Следовательно, норма внесения удобрений, средств защиты растений и орошения, должна подбираться индивидуально для каждого участка. Это делается на основе анализа данных, полученных с помощью навигационных систем (GPS, GLONASS), ведомостей агрохимического и агрофизического обследования почв, истории урожайности. Результатами являются экономия ресурсов, повышение урожайности за счет целевого применения удобрений, пестицидов и орошения; защита окружающей среды за счет ограничения применения нитратов, ядохимикатов, забора воды из водоемов и т.д.

Анализ рыночной корзины, использующий ассоциативные правила, позволяет обнаруживать группы товаров, которые чаще всего покупаются совместно. Такие исследования позволяют выявлять поведенческие шаблоны клиентов, которые можно использовать для более полного удовлетворения спроса, стимулирования продаж, оптимизации ассортимента товаров.

Аналогичные решения на основе методов ИАД активно разрабатываются и внедряются в следующих областях экономики и бизнеса:

- в потребительском кредитовании – для оценки кредитоспособности клиентов (кредитный скоринг), разработки новых кредитных продуктов, выявления мошеннических действий с кредитными картами и др.;

- в страховании – для анализа клиентской базы с целью разработки новых продуктов и привлечения клиентов, выявления мошенничеств со страховыми выплатами;

- в телекоммуникациях – анализ клиентской базы с целью определения лояльности клиентов, выработке мер по удержанию имеющихся и привлечению новых клиентов, разработки новых конкурентоспособных продуктов, обнаружения мошеннических действий;

– в медицине и здравоохранении – диагностика заболеваний, анализ эпидемиологической обстановки;

– в науке – определение наиболее достойных претендентов на получение грантов научных фондов;

– в поиске залежей полезных ископаемых, на основе данных геологоразведки, аэрофотосъемки и космических снимков.

Краткий обзор инструментов ИАД.

Десятки IT-компаний во всем мире активно разрабатывают и продвигают собственные решения в области ИАД и бизнес-аналитики. Среди них как признанные лидеры: Microsoft, IBM, Oracle, SAS Institute, SPSS, Silicon Graphics, StatSoft, так и менее известные Angross Software, Neuro Solution, PolyAnalyst, BaseGroup Labs и др. Предлагаются как мощные платформы бизнес-аналитики, позволяющие создавать завершённые аналитические проекты (SAS Interprise Miner, IBM Intelligent Miner, MineSet, Knowledge Studio, Microsoft SQL Server Analysis Services, SPSS Predictive Analytics Software Modeler, Deductor Enterprise, PolyAnalyst), так и программы, ориентированные на конкретную предметную область (FinMetrics, Estard Data Miner).

Стоимость ИАД-приложений варьируется от нескольких тысяч до нескольких десятков

тысяч USD. Существуют и бесплатные продукты, разрабатываемые и поддерживаемые крупными университетами и исследовательскими центрами, например Weka (университет Уайкато, Новая Зеландия), Orange (университет Любляны, Словения), RapidMiner (технический университет Дортмунда) и др. Российская компания BaseGroup Labs (www.basegroup.ru), разработчик платформы Deductor, выпускает бесплатную учебную версию Deductor Academic, которая используется десятками вузов России и ближнего зарубежья в учебном процессе.

Обобщенная структура аналитического бизнес-приложения. Как современное направление информационных технологий ИАД не предписывает использования каких-либо конкретных алгоритмов и методов обработки данных при решении тех или иных задач анализа. Оно лишь определяет методологию поиска знаний в массивах структурированных данных, порядок действий аналитика, который позволит сделать поиск знаний наиболее эффективным. Поэтому, несмотря на то, что аналитические системы могут быть предметно-ориентированными и содержать различные средства обработки данных, их структура содержит примерно одни и те же элементы, представленные на рис. 1.

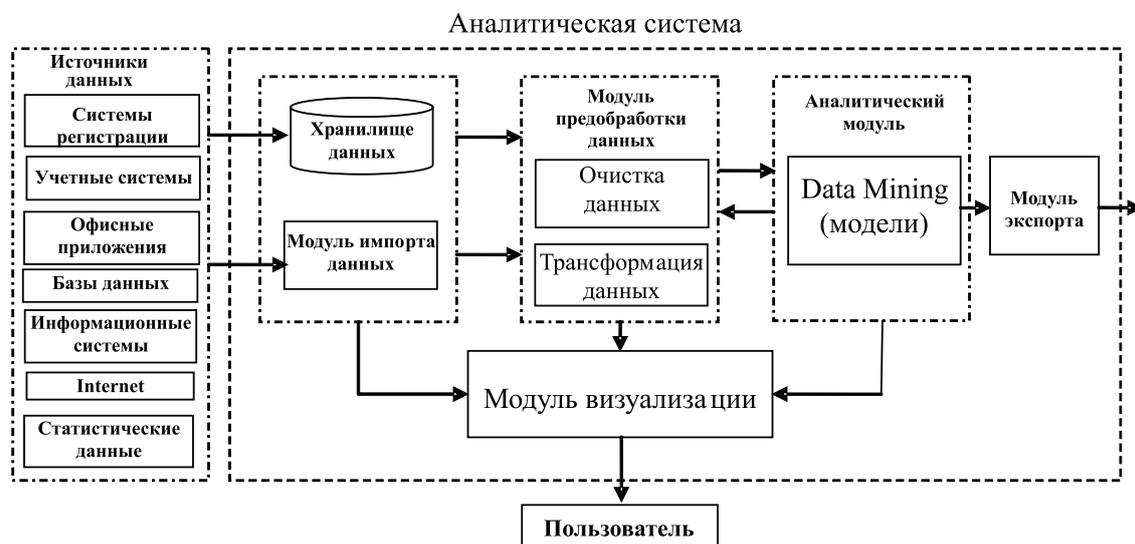


Рис. 1. Обобщенная структура типичного аналитического приложения

Анализируемые данные расположены в различных источниках, таких, как системы оперативной регистрации данных – кассовые терминалы супермаркетов, вокзалов и аэропортов; учетные системы (1С:Предприятие), офисные приложения, базы данных, файлы из Интернета и других внешних источников. Поэтому анали-

тическая система (АС) должна содержать средства импорта, конвертирования и интегрирования данных из файлов самых различных типов и форматов. Данная задача и решается с помощью модуля импорта.

Наиболее продвинутые АС используют специализированные хранилища данных (ХД), которые отличаются от обычных баз

данных и других источников тем, что адаптированы именно для целей и задач ИАД. Они автоматически обеспечивают интегрирование данных из различных источников, поддерживают их целостность, неизменчивость и непротиворечивость, а также хронологию и высокую скорость доступа к данным.

После загрузки данных в АС, к ним необходимо применить процедуру преобразования с целью подготовки к анализу. Такую преобработку можно разделить на два этапа – очистку и трансформацию. Очистка позволяет исключить из данных факторы, мешающие их корректной обработке аналитическими алгоритмами: аномальные и фиктивные значения, пропуски, противоречия, дубликаты, шумы и т.д. Поскольку качество данных очень важно с точки зрения достоверности результатов анализа, данному вопросу уделяется очень большое внимание при практической реализации проектов ИАД. Трансформация данных выполняется для их согласования (типов, форматов, диапазонов и т.д.) с применяемыми аналитическими алгоритмами. Она включает такие операции, как квантование, нормирование, преобразование форматов и типов данных, кодирование, слияние, группировку и т.д. Если исходные данные полностью соответствуют заданным критериям качества, то этап преобработки может быть пропущен и данные могут анализироваться в том виде, в каком они были загружены. Однако на практике такая ситуация встречается крайне редко.

После того, как данные подготовлены, к ним применяются различные алгоритмы анализа. Часть из них относится к математической статистике (метод главных компонент, байесовская классификация, линейная и логистическая регрессия и др.), а часть – к технологиям машинного обучения (нейронные сети, деревья решений, карты Кохонена, ассоциативные правила и др.). Именно последние традиционно считаются аналитическим ядром ИАД, поскольку позволяют в автоматическом режиме извлекать из данных зависимости, закономерности и структуры, интерпретация которых и позволяет генерировать знания. Результатом работы аналитического алгоритма является модель, отражающая зависимость, закономерность или структуру, содержащуюся в исходном множестве данных. После этого модель верифицируется, сохраняется и может быть использована для обработки новых, ранее не известных наблюдений.

В процессе анализа важно не только получить значимые и достоверные результаты, но и представить обнаруженные зависимости и закономерности в наиболее

удобной и интерпретируемой форме. Для этого в АС предусмотрен модуль визуализации, с помощью которого пользователь формирует наиболее информативное представление для определенного вида данных (таблицы, графики, диаграммы, многомерные визуализаторы, визуализаторы связей и т.д.). Результаты, полученные в процессе анализа, могут быть не только представлены пользователю с помощью средств визуализации, но и экспортированы в файлы других приложений – электронные таблицы, базы данных, текстовые документы и другие форматы.

Аналитическая платформа Deductor – современный инструмент интеллектуального анализа данных. Программные средства, комплексно обеспечивающие выполнение всех этапов аналитического процесса ИАД, называют аналитическими платформами (АП). Стоимость АП зарубежных компаний (SAS, IBM, SPSS, Microsoft) достигает десятков тысяч USD. Кроме этого, не всегда доступны их русифицированные версии, отсутствует желаемый уровень сопровождения продуктов и поддержки, реализуемых на их основе бизнес-проектов внутри России. Поэтому данные инструменты недоступны большинству российских компаний, как в плане цены, так и в плане ресурсов, необходимых для развертывания и поддержки масштабных аналитических бизнес-проектов.

В этой связи особый интерес представляют перспективы внедрения отечественных разработок в области интеллектуальных методов бизнес-аналитики. К сожалению, данное направление в России практически не представлено. Единственной полноценной платформой, позволяющей создавать завершенные аналитические проекты на основе технологий Data Mining, является АП Deductor компании BaseGroup Labs (г. Рязань). Функциональность платформы соответствует уровню наиболее современных зарубежных аналогов (SAS Enterprise Miner, IBM Intelligent Miner, SPSS Clementine), но его стоимость почти на порядок ниже. Платформа снабжена русскоязычной справочной системой, исчерпывающим руководством пользователя, руководством по используемым алгоритмам и методам анализа. На сайте разработчика (www.basegroup.ru) представлено большое количество методических материалов и тематических статей как по отдельным методам интеллектуального анализа, так и их практической реализации, дано описание более 100 завершенных аналитических бизнес-проектов.

В АП Deductor реализована технология клиент-сервер, позволяющая производить

обработку удаленных источников данных в режиме реального времени. Система содержит интегрированное хранилище данных Deductor Warehouse, которое позволяет консолидировать данные из различных источников и обеспечивает их целостность и непротиворечивость. Модуль импорта поддерживает загрузку данных из файлов более 10 типов, в том числе, баз данных учетной системы 1С:Предприятие. Подсистема предобработки данных включает очистку данных с возможностью обработки дубликатов и противоречий, корреляци-

онного и факторного анализа, подавления шумов и аномальных значений, восстановления пропусков. «Интеллектуальное» ядро платформы позволяет решать задачи классификации, численного предсказания, кластеризации, ассоциации и прогнозирования с помощью современных технологий машинного обучения – нейронных сетей, деревьев решений, карт Кохонена, ассоциативных правил.

Аналитическая обработка данных в АП Deductor производится по схеме, представленной на рис. 2.

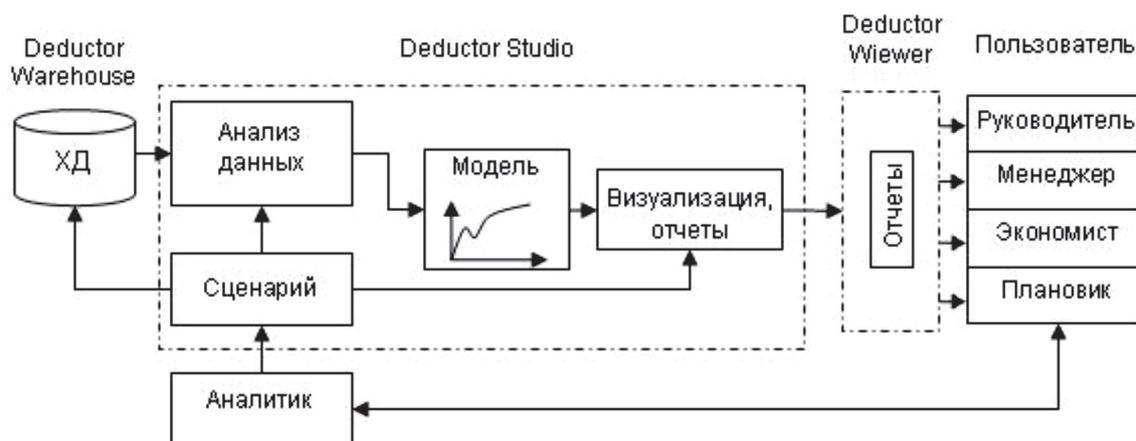


Рис. 2. Схема аналитического процесса в АП Deductor

Данные из различных источников консолидируются в ХД Deductor Warehouse. Проходя необходимую предобработку, они загружаются в аналитическое приложение Deductor Studio, где в соответствии с разработанным сценарием производится их анализ, результатом которого являются модели, реализующие зависимости и закономерности в данных. Сценарий разрабатывается аналитиком – специалистом в области алгоритмов и методов анализа данных, и пользователем – специалистом в предметной области. Построение моделей производится на основе методов машинного обучения с минимальным вмешательством человека. Обнаруженные зависимости и закономерности визуализируются и формируются соответствующие отчеты.

В настоящее время на основе АП Deductor реализовано более 100 законченных проектов в различных областях бизнеса, социальной сферы, науки и государственного управления.

Выводы

Таким образом, наиболее перспективным инструментом поиска знаний, необходимых для поддержки принятия решений в различных сферах экономики и бизнеса,

является интеллектуальный анализ данных – современное направление информационных технологий, объединяющее совокупность методов извлечения знаний из структурированных массивов данных, описывающих экономические и бизнес-процессы предметной области. «Интеллектуальным» ядром направления является моделирование процессов, объектов и явлений с помощью моделей, основанных на машинном обучении, которые восстанавливают зависимости и закономерности в данных и визуализируют их в наиболее удобной для интерпретации форме.

Внедрение и использование интеллектуальных систем бизнес-аналитики в компаниях позволяет им приобретать конкурентные преимущества за счет повышения эффективности процесса принятия решений и накопления интеллектуального капитала. Наиболее перспективным инструментальным средством бизнес-аналитики для внедрения в условиях РФ является аналитическая платформа Deductor российской компании BaseGroup Labs. Не уступая зарубежным аналогам в функциональности, она на порядок дешевле, и адаптирована для совместной работы с отечественными бизнес-приложениями.

Список литературы

1. Васильев Е.П., Орешков В.И. Современные аналитические платформы для задач АПК // Вестник Рязанского гос. агротехнологического университета имени П.А. Костычева. Рязань: РГАТУ, – 2011. – № 1. – С. 68–75.
2. Васильев Е.П., Орешков В.И. Интеллектуальные системы бизнес-аналитики // Интеграция науки с сельскохозяйственным производством: материалы науч. конф. – Рязань: изд. РГАТУ, 2011 – С. 67–71.
3. Орешков В.И. Интеллектуальный анализ данных как важнейший инструмент формирования интеллектуального капитала организаций // Креативная экономика. – 2011. – №12. – С. 84–89.
4. Орешков В.И. Интеллектуальный анализ данных как современный инструмент поддержки управленческих решений // Вестник Рязанского гос. агротехнологического университета имени П.А. Костычева. – Рязань: РГАТУ, 2011. – №4. – С. 55–59.
5. Паклин Н.Б. Бизнес-аналитика: от данных к знаниям (+ CD) / Н.Б. Паклин, В.И. Орешков. – 2-е изд., перераб. и доп. – СПб.: Питер, 2010. – 700 с.

References

1. Vasiliev E.P., Oreshkov V.I. Modern analytical platform for the problems of agriculture. Journal of the Ryazan State Agrotechnological University. Ryazan. 2011. no. 1. pp. 68–75.

2. Vasiliev E.P., Oreshkov V.I. Intelligent systems of business analysis. materialy konferencii «Integraciya nauki s sel'skoxozyajstvenny'm proizvodstvom (Proc. conference Integration of science to agricultural production. Ryazan, 2011, pp. 67–71.

3. Oreshkov V.I. Data mining as an essential tool for the formation of the intellectual capital of organizations. Creative economy. 2011, no. 12. pp. 84–89.

4. Oreshkov V.I. Data mining as a modern tool to support management decisions. Journal of the Ryazan State Agrotechnological University. Ryazan, 2011. no. 4. pp. 55–59.

5. Paklin N.B., Oreshkov V.I. Business Intelligence: From Data to Knowledge. Ed. 2nd, St. Peter Publ, 2010. p. 700.

Рецензенты:

Солдак Ю.М., д.э.н., профессор кафедры экономики, менеджмента и организации производства ФГБОУ «Рязанский государственный радиотехнический университет», г. Рязань;

Текучев В.В., д.э.н., профессор, заведующий кафедрой экономической кибернетики ФГБОУ «Рязанский государственный агротехнологический университет им. П.А. Костычева», г. Рязань.

Работа поступила в редакцию 06.09.2012.