

Научная статья

УДК 004.6; 339.13

EDN [DWZNTH](#)

DOI 10.17150/2411-6262.2022.13(2).27

**М.Ю. Иванов¹** , **М.В. Сыгодина¹** , **В.В. Надршин²** , **А.В. Дербенёва¹** ¹ Братский государственный университет, г. Братск, Российская Федерация² Иркутский национальный исследовательский технический университет, г. Иркутск, Российская ФедерацияАвтор, ответственный за переписку: М.Ю. Иванов, nis@brstu.ru

ТЕХНОЛОГИИ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ В РЕШЕНИИ ЭКОНОМИЧЕСКИХ ЗАДАЧ

АННОТАЦИЯ. Повышение эффективности управления бизнес-процессами является сложной и крайне ответственной задачей, решение которой немислимо без использования современных информационных систем и технологий поддержки принятия управленческих решений. Целью данной работы является исследование технологий интеллектуального анализа данных в сравнении с такими популярными методами как онлайн обработка транзакций и аналитическая онлайн обработка. В статье предложено обобщенное иерархическое представление методологий обработки данных от интеграции разнородных источников до принятия управленческих решений. Систематизированы базовые методы интеллектуального анализа данных: классификация, регрессия, кластеризация, прогнозирование, взаимозависимость, визуализация, определение отклонений, оценка, отбор значимых признаков. С иллюстрацией на примерах описаны возможности применения интеллектуального анализа данных в сфере информационных технологий, маркетинге, торговле, финансовой и страховой деятельности.

Подробно рассмотрено практическое использование технологий интеллектуального анализа данных в экономике с помощью приложения Microsoft® Excel® и специальной надстройки системы управления реляционными базами данных Microsoft® SQL Server®, способных на выявление неявных (скрытых) факторов, влияющих или, что также не менее важно, не влияющих на объемы продаж спортивной продукции.

КЛЮЧЕВЫЕ СЛОВА. Интеллектуальный анализ данных, классификация, кластеризация, прогнозирование, взаимозависимость, определение отклонений, отбор значимых признаков.

ИНФОРМАЦИЯ О СТАТЬЕ. Дата поступления 23 марта 2022 г.; дата принятия к печати 25 мая 2022 г.; дата онлайн-размещения 10 июня 2022 г.

Original article

М.Yu. Ivanov¹ , **M.V. Sygotina¹** , **V.V. Nadrshin²** , **A.V. Derbeneva¹** ¹ Bratsk State University, Bratsk, Russian Federation² Irkutsk National Research Technical University, Irkutsk, Russian FederationCorresponding author: M.Yu. Ivanov, nis@brstu.ru

DATA MINING TECHNOLOGIES IN SOLVING ECONOMIC PROBLEMS

ABSTRACT. Improving the efficiency of business process management is a complex and extremely important task, the solution of which is unthinkable without the use of advanced information systems and management decision technologies. The paper presents the results of research on data mining technologies using On-Line Transaction Processing and On-Line Analytical Processing methods. The study proposed a generalized hierarchical representation of data processing methodologies with integration of heterogeneous sources to management decision making. Data mining

© Иванов М.Ю., Сыгодина М.В., Надршин В.В., Дербенёва А.Д., 2022

Baikal Research Journal

электронный научный журнал Байкальского государственного университета

base methods are systematized: classification, regression, prediction, clustering, interdependence, visualization, deviation (anomaly) detection, estimation, and feature selection (engineering). We described the possibilities of using data mining in the field of information technology, marketing, trade, financial and insurance activities. The study examined data mining technologies in sport products business using the Microsoft® Excel® application and a special add-in of the Microsoft® SQL Server® relational database management system, capable of identifying implicit (hidden) factors that affect or, equally important, do not affect sales volumes.

KEYWORDS. Data mining, classification, clustering, prediction, interdependence, deviation (anomaly) detection, feature selection (engineering).

ARTICLE INFO. Received March 23, 2022; accepted May 25, 2022; available online June 10, 2022.

Введение

В настоящее время на мировом рынке сложилась достаточно специфическая, точнее сказать, патовая ситуация: доходность предприятий снижается из-за наличия сходных товаров и услуг с идентичными потребительскими качествами, сложностей при организации экспорта, но, владельцы бизнеса, как и раньше, требуют от менеджеров увеличения объемов продаж и прибыли.

Несмотря на повсеместное использование самых передовых инструментов маркетинга эффективность методов привлечения клиентов невысока, а процесс сбыта при этом становится только дороже. Рынки замедляют свой рост, миграция клиентов ускоряется и, что самое главное, наблюдается пресыщенность данными при одновременной нехватке информации для принятия решений.

Известно, что применение информационных технологий и специализированного программного обеспечения способствует повышению эффективности любой сферы деятельности [1–4]. Так, в конце прошлого века вышеописанные проблемы решали с помощью универсальных инструментов для извлечения, хранения и управления большими объемами структурированных данных, в частности, реляционных систем управления базами данных (далее — БД).

С ростом активности информационно-телекоммуникационной сети «Интернет» получили развитие и технологии построения распределенных БД. В связи с этим многократно возросла интенсивность формирования и архивирования разнородных данных в больших объемах. При этом обнаружилась потребность анализа имеющихся данных в разновременном аспекте с возможностью построения произвольных запросов и при условии обработки данных из различных источников. Применение для решения этих задач корпоративных информационных систем и БД крайне затруднительно.

Например, в традиционной регистрирующей системе информация актуальна только на момент обращения к БД, а в следующий момент времени по такому же запросу вполне возможно получение совсем иных данных. Интерфейс таких систем рассчитан на проведение определенных типизированных операций и вероятность поиска ответа на нерегламентированный произвольный запрос невысока. Возможности обработки больших массивов также могут быть ограничены из-за ориентации систем управления БД на нормализованные данные, характерные для стандартных реляционных регистрирующих БД.

Больших успехов удалось добиться благодаря методам онлайн обработки транзакций (англ. On-Line Transaction Processing) и он-лайн аналитической обработки (англ. On-Line Analytical Processing), предполагающих некоторую предварительную обработку данных и их интеграцию [5].

Такие инструменты ориентированы только на хорошо нормализованные табличные данные и не предполагают использование целого ряда дополнительных анали-

тических механизмов типа кластеризации, классификации, моделирования, регрессионного анализа, интерпретации многомерных массивов, прогнозирования и т.п.

Таким образом, на сегодняшний день наблюдается высокий уровень развития масштабируемой аппаратно-программной информационно-телекоммуникационной инфраструктуры, позволяющей увеличивать и без того значительные архивы данных. Имеется достаточно существенный задел в области информационных технологий, проработаны прикладные аспекты теории вероятности и математической статистики. Однако при этом снова следует признать, что присутствует заметный избыток данных при дефиците информации и знаний.

Быстро растущие объемы накопленных и автоматически пополняемых архивов данных пока существенно превышают способности человека в их практически полезной обработке. Для лучшего понимания сложившейся ситуации применимо выражение, что большие БД стали «могилами», которые редко посещаются [6]. Как следствие, критически важные решения зачастую принимаются не с учетом полученных из имеющихся огромных объемов данных полезных знаний, а, как и много лет назад, на основе интуиции человека.

Именно поэтому относительно недавно стала привлекать к себе особое внимание область поддержки эффективной управленческой деятельности, связанная с высокопроизводительным интеллектуальным анализом данных (далее — ИАД) и направленная на оперативное извлечение из значительных массивов, накопленных и постоянно поступающих данных ценных экспертных знаний [7].

На рис. 1 приведен пример обобщенного иерархического представления методологий обработки данных, начиная от интеграции разнородных источников и завершая использованием методов ИАД для принятия управленческих решений.

Основное назначение методов ИАД заключается в выявлении скрытых связей и закономерностей, имеющих самое непосредственное влияние на стратегии развития бизнеса, маркетинговые мероприятия, рекламу и прочие сферы экономики, успех которых основан на анализе и работе с имеющимися данными [8].



Рис. 1. Обобщенное иерархическое представление методологий обработки данных

Теоретические основы ИАД

Наибольший интерес к технологиям ИАД, в первую очередь, проявляют предприятия, работающие в условиях высокой конкуренции и имеющие четко определенную группу потребителей (розничная торговля, связь, маркетинг). Компании пытаются найти взаимосвязь между «внутренними» (цена, востребованность товара, компетентность персонала) и «внешними» (экономические показатели, конкуренция, демография клиентов) факторами. Такой подход позволяет оценить или спрогнозировать объем продаж и степень удовлетворенности клиентов, размер доходов, а также сформулировать на основе совокупности всей имеющейся информации практически полезные выводы и рекомендации. Выгода от инструментов ИАД может быть значительной при сравнительно невысокой стоимости их применения.

Рассмотрим основные методы ИАД.

Классификация. Задача классификации заключается в определении для каждого объекта какого-либо класса (типа, категории), которому он принадлежит в соответствии с некоторым известным набором атрибутов и массивом соответствующих этим атрибутам данных.

Решение задач классификации востребовано при оценке кредитоспособности потенциального заемщика (назначаемые классы «кредитоспособен» и «некредитоспособен»). Характеристиками (признаками) исследуемого (классифицируемого) объекта (заемщика) являются возраст, место работы, уровень дохода, семейное положение и т.д. Классификация важна при формировании объемов продаж на основе текущих показателей, выявлении лояльных или нелояльных клиентов операторов сотовой связи [9], держателей банковских карт и т.п.

Регрессия. Задача регрессии похожа на классификацию данных, но в ходе ее решения производится поиск определенных шаблонов, например, расчет допустимого лимита кредита.

Кластеризация (сегментация). Кластеризация заключается в разделении множества объектов на группы (кластеры или сегменты) со схожими параметрами. При этом, в отличие от классификации, число кластеров и их характеристики могут быть заранее неизвестны и определяться в ходе построения кластеров исходя из степени схожести объединяемых объектов по совокупности параметров.

Решение задач кластеризации необходимо при поиске новых рынков сбыта, формировании специальных предложений для выделенных групп клиентов с учетом их потребностей, интересов, возраста, выявлении случаев подлога в страховании и т.п.

Прогнозирование. Задача прогнозирования основана на ИАД, между значениями которых наблюдается корреляция за определенный отрезок времени, и является одной из наиболее востребованных в бизнесе. Этот метод применяется для оценки спроса на товары и услуги, характеризующиеся сезонными колебаниями, позволяет оценить ожидаемую потребность в кадрах, выбрать оптимальный момент купли-продажи на фондовых рынках и биржах [10].

Прогнозирование востребовано в розничной торговле продовольственными товарами. В скоротечных условиях заказа, поставки, хранения и продажи продуктов крайне важно выдержать баланс между удовлетворенностью клиента, имеющего достаточный выбор и возможность найти желаемое, и запасами товара, занимающего место на складе и в торговом зале, но в итоге нерезализованного и утилизированного по истечению срока годности [11].

Взаимозависимость. Задача определения взаимосвязей или поиска ассоциативных правил заключается в определении часто встречающихся наборов объектов. Классическим примером взаимозависимости является поиск интересных

ассоциаций и (или) корреляционных связей, имеющих в существующем наборе данных. Так, при анализе потребительской корзины выявляются сопутствующие товары, чаще всего встречающиеся в одном заказе или в одном чеке. Полученная информация может быть использована при размещении продукции на витринах и прилавках или при формировании специальных предложений для группы связанных товаров. А анализ последовательности переходов по Интернет-страницам позволяет разработчикам web-сайтов формировать наиболее востребованный для посетителей контент [12].

Визуализация. С использованием визуальных инструментов ИАД создаются графические образы анализируемых данных, что, несомненно, способствует их лучшему восприятию лицом, принимающим решения.

Определение отклонений. ИАД позволяет обнаруживать среди множества данных или фрагментов данных те, которые существенно отличаются от нормативных или усредненных значений, выявлять нехарактерные паттерны [13].

Отклонение может свидетельствовать о каком-то необычном событии: неожиданный результат расчета, мошенническая операция по банковской карте, ошибка ввода данных и т.п.

Оценка. Данный метод ИАД имеет своей целью предсказание непрерывных значений данных по какому-либо признаку: производительность центрального процессора при решении определенных задач по ряду параметров ЭВМ, оценка числа детей в семье по уровню образования родителей, оценка дохода семьи по количеству в ней автомобилей, оценка стоимости недвижимости в зависимости от ее удаленности от центра города и т.д.

Отбор значимых признаков. Как правило, применяется как вспомогательный метод на этапе предварительной обработки данных, а также для повышения эффективности методов визуализации в многомерных признаковых пространствах, например, при отборе кандидатов на замещение вакантной должности, допущенных до собеседования с представителем работодателя.

Практическое использование ИАД в экономике

С учетом разнообразия форм представления данных, используемых алгоритмов и сфер применения, ИАД может проводиться с помощью следующих разновидностей программного обеспечения:

- специализированные («коробочные») программные продукты;
- математические пакеты;
- электронные таблицы и различного рода надстройки над ними;
- аналитические средства, интегрированные в системы управления БД;
- другие программные продукты [14].

В настоящей работе показана возможность ИАД с помощью приложения Microsoft® Excel® и специальной надстройки от разработчиков системы управления реляционными базами данных Microsoft® SQL Server®, позволяющей выявлять неявные закономерности в наборах данных.

Традиционные БД состоят из совокупности записей-строк, содержащих различные поля (в частности, сведения о клиентах магазина). Именно эти поля и представляют собой анализируемые признаки, которые менеджеры и маркетологи используют для выявления зависимостей, оценки потребительского спроса и формирования соответствующего предложения. В ходе ИАД исследуется множество объектов или вариантов и устанавливается или не устанавливается зависимость между результирующим показателем и воздействующими на него факторами. Такой подход может быть использован и для решения нетривиальных задач. Например, мотивы, которыми руководству-

ется человек, оставляя за солнечным козырьком своего автомобиля запасной комплект ключей?

Как отмечалось выше, в экономике наиболее популярной ситуацией является оценка кредитоспособности заемщика. Любой банк нередко сталкивается с необходимостью анализа больших массивов данных и далеко не всегда эти выборки являются абсолютно репрезентативными [15]. Если в отношении возраста, размера заработной платы, семейного положения вопросов не возникает, то стоимость движимого и недвижимого имущества заемщика нередко может измеряться не конкретными цифрами, а их диапазоном, поскольку рыночная оценка автомобиля или квартиры не является величиной постоянной и подвержена значительным колебаниям.

В данном исследовании рассматривается база данных физических лиц, включающая 1000 клиентов магазина спортивных товаров, с помощью технологий ИАД выявляются факторы, влияющие на продажи, определяются категории покупателей и выделяются исключения. Вся эта информация необходима для формирования сценариев будущих объемов продаж велосипедов.

Характеристиками покупателя являются семейное положение, пол, размер дохода, наличие детей, образование, занимаемая должность, владение частным жильем, количество автомобилей, расстояние от дома до работы, страна проживания, возраст (рис. 2–3). Именно по этим полям осуществляется поиск закономерностей и предпринимается попытка определить факторы явно или неявно способствующие приобретению человеком велосипеда.

ID	Семейное положение	Пол	Доход	Дети	Образование	Тип работы	Дом
12496	Женатый, замужняя	Женский	40000	1	Бакалавр	Квалифицированный ручной труд	Да
24107	Женатый, замужняя	Мужской	30000	3	Неоконченное высшее	Офисный работник	Да
14177	Женатый, замужняя	Мужской	80000	5	Неоконченное высшее	Профессионал	Нет
24381	Одинокий(ая)	Мужской	70000	0	Бакалавр	Профессионал	Да
25597	Одинокий(ая)	Мужской	30000	0	Бакалавр	Офисный работник	Нет
13507	Женатый, замужняя	Женский	10000	2	Неоконченное высшее	Ручной труд	Да
27974	Одинокий(ая)	Мужской	160000	2	Среднее	Управление	Да

Рис. 2. Начальный фрагмент БД покупателей

Для решения задачи необходимо загрузить надстройку Microsoft® SQL Server® Master Data Services для Microsoft® Excel®. Надстройка SQLServerDMXLAAddIn распространяется свободно и доступна на сайте компании Microsoft®. Установка надстройки SQLServerDMXLAAddIn никаких сложностей не вызывает и не нуждается в описании.

Для запуска надстройки следует выполнить действия в пунктах меню приложения Microsoft® Excel® в следующем порядке: *Параметры \ Надстройки \ Управление (Надстройки COM) \ Надстройки для модели компонентных объектов COM \ SQLServerDMXLAAddIn.*

	Н	И	Ж	К	Л	М	Н	С
1	деления исключений и анализа сценариев							
2								
3	Домовладелец	Кол-во авто	Расстояние до работы	Регион	Возраст	Приобрел велосипед		
4	Да		0-1 км	Европа	42	Нет		
5	Да		1-1 км	Европа	43	Нет		
6	Нет		2-5 км	Европа	60	Нет		
7	Да		5-10 км	Россия	41	Да		
8	Нет		0-1 км	Европа	36	Да		
9	Да		0-1 км	Европа	50	Нет		
10	Да		4-1 км	Россия	33	Да		

Рис. 3. Заключительный фрагмент БД покупателей

Надстройки служб Master Data Services позволяет нескольким пользователям обновлять основные данные с помощью привычных средств, не нарушая в то же время целостность данных.

Надстройка SQLServerDMXLAAddIn не является активной по умолчанию и запускается применительно к конкретной БД (таблице). После активации надстройки на странице приложения Microsoft® Excel® появляется соответствующая закладка для анализа данных. Закладка для анализа данных включает в себя восемь инструментов, реализующих различные алгоритмы обработки значений и поиска скрытых закономерностей (рис. 4). В рассматриваемом примере необходим только первый инструмент «Анализ ключевых факторов влияния».

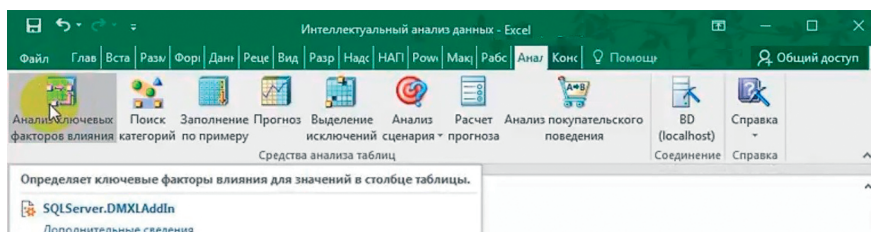


Рис. 4. Средства анализа таблиц в приложении Microsoft® Excel®

Также для использования надстройки SQLServerDMXLAAddIn необходимо запустить систему управления реляционными БД Microsoft® SQL Server®. Система управления реляционными БД Microsoft® SQL Server® включает в себя несколько основных служб, в том числе и Analysis Services (рис. 5). После запуска данной службы необходимо создать БД и обязательно предоставить к ней доступ.

Путем запроса свойств сервера (закладка «Общие», имя скрипта DataMining \ AllowSessionMiningModels) устанавливаем значение «true» (рис. 6).

После выполнения этих операций появляется возможность использовать данный сервер для импорта данных из приложения Microsoft® Excel®.

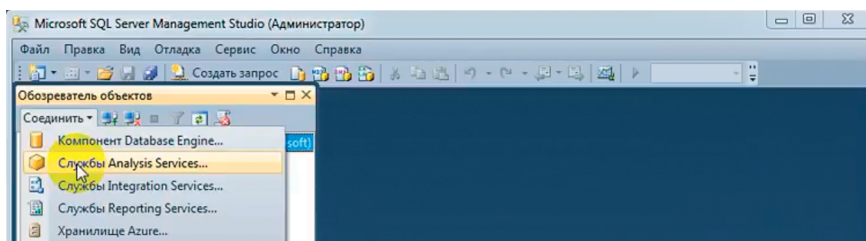


Рис. 5. Выбор службы анализа данных в системе управления реляционными БД Microsoft® SQL Server®

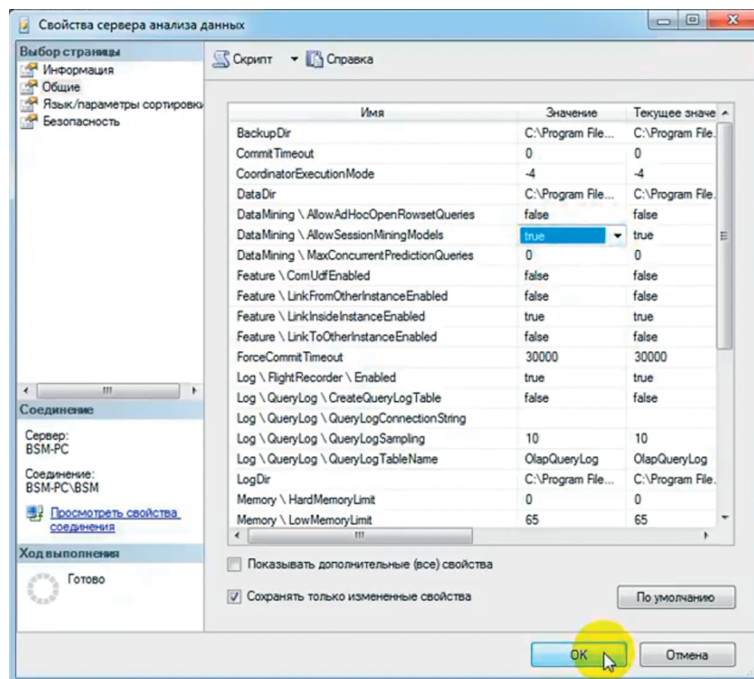


Рис. 6. Запрос свойств сервера

Далее в приложении Microsoft® Excel® осуществляется соединение с текущим сервером (вкладка BD (localhost) \ Создать). Обязательно заполняются поля «Имя сервера» (можно внести имя локального компьютера или указать стандартное значение «localhost»). В поле «Имя каталога» указывается наименование созданной в Microsoft® SQL Server® базы данных (рис. 7).

Затем на панели инструментов приложения Microsoft® Excel® активируется закладка «Анализ ключевых факторов влияния» и выбираются анализируемые столбцы для ключевых факторов влияния на результирующий показатель «Приобрел велосипед» (рис. 8).

Первый отчет создается автоматически, но при необходимости можно сгенерировать и другой отчет (рис. 9).

Так, в результате ИАД выявлена закономерность приобретения велосипеда при отсутствии у клиента автомобиля (выделено красным цветом). Абсолютно противоположная ситуация при наличии у клиента магазина нескольких (не ме-

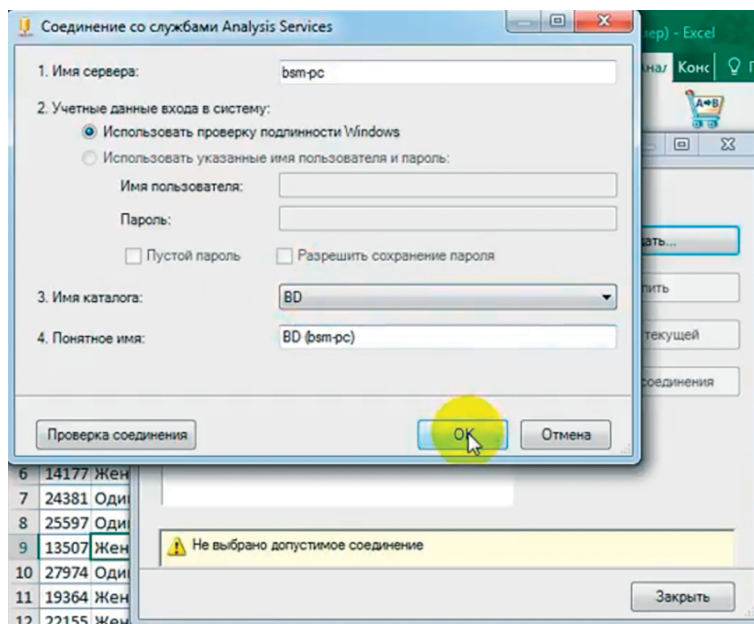


Рис. 7. Установка соединения сервера со службами анализа

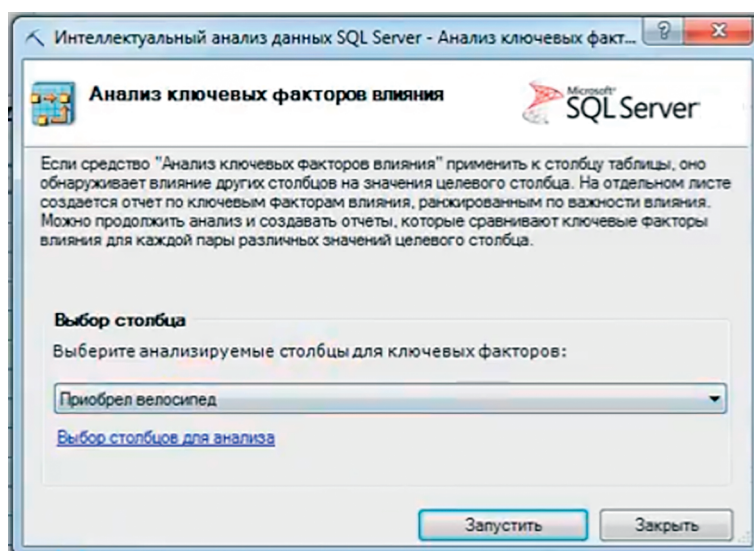


Рис. 8. Подготовка к запуску ИАД при поиске ключевых факторов влияния на результирующий показатель

нее двух авто). Такие люди абсолютно интактны к спортивной продукции (выделены синим цветом). Граждане России покупают велосипеды чаще граждан США, незначительное влияние оказывает семейное положение, но одинокие люди покупают велосипеды, все же, охотнее (заметна корреляция) (рис. 10).

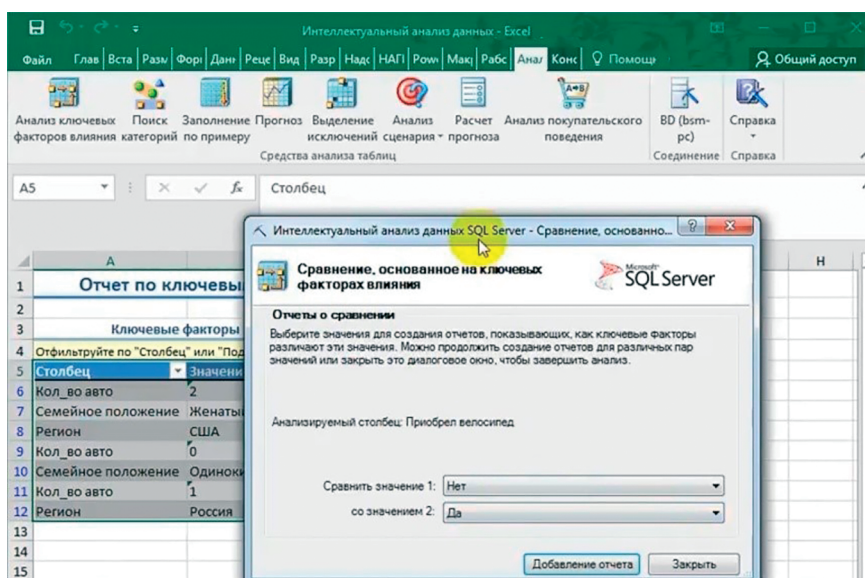


Рис. 9. Создание отчета по результатам ИАД

D	A	B	C	E	F	G	H
1	Отчет по ключевым факторам влияния для "Приобрел велосипед"						
2							
3	Ключевые факторы влияния и их воздействие на значения "Приобрел велосипед"						
4	Отфильтруйте по "Столбец" или "Подходит", чтобы увидеть, как разные столбцы влияют на "Приобрел велосипед"						
5	Столбец	Значение	Подходит	Относительное влияние			
6	Кол_во авто	2	Нет				
7	Семейное положение	Женатый, замужняя	Нет				
8	Регион	США	Нет				
9	Кол_во авто	0	Да				
10	Семейное положение	Одинокий(ая)	Да				
11	Кол_во авто	1	Да				
12	Регион	Россия	Да				
13							
14							
15							
16	Сравнение факторов, ведущих к значениям "Нет" и "Да"						
17	Отфильтруйте по "Столбец", чтобы увидеть, как разные значения подходят "Нет" или "Да"						
18	Столбец	Значение	Подходит Нет	Подходит Да			
19	Кол_во авто	2					
20	Кол_во авто	0					
21	Семейное положение	Женатый, замужняя					
22	Семейное положение	Одинокий(ая)					
23	Кол_во авто	1					
24	Регион	Россия					
25	Регион	США					
26							

Рис. 10. Отчет по поиску ключевых факторов влияния на результирующий показатель

Абсолютно не выявлена корреляция с гендерным признаком, размером дохода, наличием детей, занимаемой должностью, владением частным жильем, расстоянием от дома до работы, возрастом и образованием. На решение человека о приобретении велосипеда перечисленные факторы не влияют, соответственно, в сформированный отчет эта информация не попадает вовсе.

Заключение

Как показывает практика, сбор и хранение данных о клиентах уже недостаточны для повышения доходов предприятия. При этом нередки ситуации, в которых необходимо действовать обязательно. Вопрос лишь в том, принимаются ли управленческие решения на основе объективной информации или интуитивно.

В работе показана возможность использования ИАД в экономике без избыточного «погружения» лица, принимающего решения, в детали базовых данных или промежуточной аналитики.

ИАД представляет собой процесс обнаружения пригодных к использованию сведений, содержащихся в крупных наборах данных. Математический ИАД позволяет выявлять тенденции, существующие в данных. Обычно такие закономерности нельзя обнаружить при простом просмотре данных, поскольку связи слишком сложны, а объем выборки чрезмерно велик.

Таким образом, технологии ИАД являются эффективным инструментом повышения эффективности деятельности предприятий, предоставляя перспективные возможности оценки большой совокупности факторов, моделирования и прогнозирования доходов.

Список использованной литературы

1. Иванов М.Ю. Экспертные системы для оценки деятельности хозяйствующего субъекта / М.Ю. Иванов. — EDN [RTDVJP](#) // Проблемы социально-экономического развития Сибири. — 2012. — № 3 (9). — С. 23–27.
2. Modern Approach to Enterprise Information Systems / A.I. Alchinov, Z.K. Tavbulatova, O.V. Dudareva, M.Yu. Ivanov. — DOI 10.1088/1742-6596/1661/1/012164 // Journal of Physics: Conference Series. — 2020. — Vol. 1661, no. 1. — Art. 012164.
3. Vakhrusheva M.Yu. Barclays' Application of Information System in Manufacturing Process / M.Yu. Vakhrusheva, M.S.-U. Khaliev, E.O. Pokhomchikova. — DOI 10.1088/1742-6596/2032/1/012129 // Journal of Physics: Conference Series. — 2021. — Vol. 2032, no. 1. — Art. no. 012129.
4. Malsagov B.S. Structural Features of Accounting Automation Application / B.S. Malsagov, M.Yu. Ivanov, L.F. Natalevich. — DOI 10.1088/1742-6596/2032/1/012128 // Journal of Physics: Conference Series. — 2021. — Vol. 2032, no. 1. — Art. 012128.
5. Управление инновационным процессом развития малого бизнеса в регионе на основе интеллектуального анализа данных (технология Big Data) / Д.В. Горбунов, С.И. Нестерова, В.М. Рамзаев [и др.]. — EDN [VYVJQN](#) // Фундаментальные исследования. — 2016. — № 4-2. — С. 381–386.
6. Пискарев Д.М. Интеллектуальный анализ данных и машинное обучение как методы управления рисками соответствия / Д.М. Пискарев. — EDN [KMJJGH](#) // Инновации и инвестиции. — 2019. — № 10. — С. 130–134.
7. Lynch C.A. Big Data: How do your Data Grow? / C.A. Lynch. — DOI 10.1038/455028a // Nature. — 2008. — Vol. 455, no. 7209. — P. 9–28.
8. Problem Solution of Optimal Pathfinding for the Movement of Vehicles over Rough Mountainous Areas / K. Zhigalov, D.K.-S. Bataev, E. Klochkova [et al.]. — DOI 10.1088/1757-899X/1111/1/012033 // IOP Conference Series: Materials Science and Engineering. — 2021. — Vol. 1111, no. 1. — Art. 012033.
9. Daudov I.M. Liable Bluetooth Tracking Technology for Enhancement of Location-Based Services / I.M. Daudov, Zh.L. Gavrilova, V.A. Kudashkin. — DOI 10.1088/1757-899X/1111/1/012043 // IOP Conference Series: Materials Science and Engineering. — 2021. — Vol. 1111, no. 1. — Art. 012043.
10. Набережная А.Т. Моделирование эффективности инвестирования в акционные пакеты (на основе интеллектуального анализа данных Data Mining) / А.Т. Набережная, В.Э. Потапова. — EDN [YAQIJP](#) // Modern Economy Success. — 2019. — № 2. — С. 43–47.
11. Ребенок И.И. Методы интеллектуального анализа и прогнозирования данных стационарной розничной торговой сети / И.И. Ребенок, М.П. Малыхина. — EDN [SYZLUB](#) //

Современные проблемы науки и образования. — 2014. — № 3. — URL: <https://www.elibrary.ru/item.asp?id=22527955>.

12. Ильина А.А. Интеллектуальный анализ данных из социальной сети «ВКонтакте» средствами машинного обучения / А.А. Ильина. — EDN [ZYTGTS](#) // Научно-практические исследования. — 2020. — № 12-3 (35). — С. 20–22.

13. Фомичева Т.Л. Применение методов интеллектуального анализа данных и Machine Learning в борьбе с мошенничеством в банках / Т.Л. Фомичева, Р.М. Магомедов, Е.А. Викулина. — EDN [TZFPEI](#) // Самоуправление. — 2019. — Т. 2, № 3 (116). — С. 337–339.

14. Наумов Р.К. Способы интеллектуального анализа данных средствами СУБД / Р.К. Наумов, М.С. Самылкин, М.В. Копейкин. — DOI 10.18413/2518-1092-2021-6-2-0-5. — EDN [MKTNQU](#) // Научный результат. Информационные технологии. — 2021. — Т. 6, № 2. — С. 32–40.

15. Иванченко О.В. Интеллектуальный анализ больших данных в развитии маркетинга отношений в банковской сфере / О.В. Иванченко. — EDN [DUGMCQ](#) // Региональные проблемы преобразования экономики. — 2019. — № 10 (108). — С. 283–288.

References

1. Ivanov M.Yu. Expert Systems for Economic Entity Activity Assessment. *Problemy sotsial'no-ekonomicheskogo razvitiya Sibiri = Issues of Social-Economic Development of Siberia*, 2012, no. 3, pp. 23–27. (In Russian). EDN: [RTDVJP](#).

2. Alchinov A.I., Tavbulatova Z.K., Dudareva O.V., Ivanov M.Yu. Modern Approach to Enterprise Information Systems. *Journal of Physics: Conference Series*, 2020, vol. 1661, no. 1, art. 012164. DOI:10.1088/1742-6596/1661/1/012164.

3. Vakhrusheva M.Yu., Khaliev M.S.-U., Pokhomchikova E.O. Barclays' Application of Information System in Manufacturing Process. *Journal of Physics: Conference Series*, 2021, vol. 2032, no. 1, art. 012129. DOI: 10.1088/1742-6596/2032/1/012129.

4. Malsagov B.S., Ivanov M.Yu., Natalevich L.F. Structural Features of Accounting Automation Application. *Journal of Physics: Conference Series*, 2021, vol. 2032, no. 1, art. 012128. DOI: 10.1088/1742-6596/2032/1/012128.

5. Gorbunov D.V., Nesterova S.I., Ramzaev V.M., Khaymovich I.N., Chumak V.G. Management of Innovative Processes Development of Small Business in the Region Based on Intelligent Data Analysis (Big Data). *Fundamental'nye issledovaniya = Fundamental Research*, 2016, no. 4-2, pp. 381–386. (In Russian). EDN: [VVYJQN](#).

6. Piskarev D.M. Data Mining and Machine Learning as Compliance Risk Management Techniques. *Innovatsii i investitsii = Innovation and Investment*, 2019, no. 10, pp. 130–134. EDN: [KMJJGH](#).

7. Lynch C.A. Big Data: How do your Data Grow? *Nature*, 2008, vol. 455, no. 7209, pp. 9–28. DOI: 10.1038/455028a.

8. Zhigalov K., Bataev D.K.-S., Klochkova E., Svirbutovich O.A., Ivashchenko G.A. Problem Solution of Optimal Pathfinding for the Movement of Vehicles over Rough Mountainous Areas. *IOP Conference Series: Materials Science and Engineering*, 2021, vol. 1111, no. 1, art. 012033. DOI: 10.1088/1757-899X/1111/1/012033.

9. Daudov I.M., Gavrilova Zh.L., Kudashkin V.A. Liable Bluetooth Tracking Technology for Enhancement of Location-Based Services. *IOP Conference Series: Materials Science and Engineering*, 2021, vol. 1111, no. 1, art. 012043. DOI: 10.1088/1757-899X/1111/1/012043.

10. Naberezhnaya A.T., Potapova V.E. Modeling of Efficiency of Investing in Promotional Packages (Based on Intellectual Data Mining Analysis). *Modern Economy Success*, 2019, no. 2, pp. 43–47. (In Russian). EDN: [YAQIJP](#).

11. Rebenok I.I., Malykhina M.P. Data Mining Methods and Forecasting Data in Fixed Retail Chain. *Sovremennye problemy nauki i obrazovaniya = Modern Problems of Science and Education*, 2014, no. 3. Available at: <https://www.elibrary.ru/item.asp?id=22527955>. (In Russian). EDN: [SYZLUB](#).


12. Iljina A.A. Intelligent Analysis of Data from the Social Network "VKontakte" using Machine Learning. *Nauchno-prakticheskie issledovaniya = Scientific and Practical Research*, 2020, no. 12-3, pp. 20–22. (In Russian). EDN: [ZYTGTS](#).


13. Fomicheva T.L., Magomedov R.M., Vikulina E.A. Application of Data Mining and Machine Learning Methods in Fight Against Fraud in Banks. *Samoupravlenie = Self-Government*, 2019, vol. 2, no. 3, pp. 337–339. (In Russian). EDN: [TZFPEI](#).


14. Naumov R.K., Samylkin M.S., Kopeikin M.V. Data Mining Methods Using DBMS Tools. *Nauchnyi rezul'tat. Informatsionnye tekhnologii = Research Result. Information Technologies*, 2021, vol. 6, no. 2, pp. 32–40. (In Russian). EDN: [MKTNQU](#). DOI: 10.18413/2518-1092-2021-6-2-0-5.


15. Ivanchenko O.V. Big Data Mining in Development Marketing Relations in the Banking Sector. *Regional'nye problemy preobrazovaniya ekonomiki = Regional Problems of Economic Transformation*, 2019, no. 10, pp. 283–288. (In Russian). EDN: [DUGMCQ](#).

Информация об авторах


Иванов Михаил Юрьевич — кандидат технических наук, доцент, базовая кафедра менеджмента и информационных технологий, Братский государственный университет, г. Братск, Российская Федерация, nis@brstu.ru,  <https://orcid.org/0000-0003-0538-7083>, SPIN-код: 4579-7257, Scopus Author ID: 57220022497, ResearcherID: A-3970-2014.


Сыготи́на Марина Владимировна — кандидат технических наук, доцент, базовая кафедра менеджмента и информационных технологий, Братский государственный университет, г. Братск, Российская Федерация, msygotina@bk.ru,  <https://orcid.org/0000-0001-8617-9766>, SPIN-код: 6658-5548, Scopus Author ID: 57220036656.

Надршин Владимир Вагизович — кандидат технических наук, доцент, кафедра экономики и цифровых бизнес-технологий, Иркутский национальный исследовательский технический университет, г. Иркутск, Российская Федерация, nadrshin@istu.edu,  <https://orcid.org/0000-0002-2375-9672>, SPIN-код: 2246-7392, Scopus Author ID: 57218676294.


Дербенёва Анжелика Викторовна — студент, факультет экономики и строительства, профиль подготовки «Прикладная информатика в экономике», Братский государственный университет, г. Братск, Российская Федерация, anjelika.derbenyova@mail.ru,  <https://orcid.org/0000-0002-7880-5285>.

Authors

Mikhail Yu. Ivanov — PhD in Technical Sciences, Associate Professor, Department of Management and Information Technologies, Bratsk State University, Bratsk, Russian Federation, is@brstu.ru,  <https://orcid.org/0000-0003-0538-7083>, SPIN-Code: 4579-7257, Scopus Author ID: 57220022497, ResearcherID: A-3970-2014.

Marina V. Sygotina — PhD in Technical Sciences, Associate Professor, Department of Management and Information Technologies, Bratsk State University, Bratsk, Russian Federation, msygotina@bk.ru,  <https://orcid.org/0000-0001-8617-9766>, SPIN-Code: 6658-5548, Scopus Author ID: 57220036656.

Vladimir V. Nadrshin — PhD in Technical Sciences, Associate Professor, Department of Economics and Digital Business Technologies, Irkutsk National Research Technical University, Irkutsk, Russian Federation, nadrshin@istu.edu,  <https://orcid.org/0000-0002-2375-9672>, SPIN-Code: 2246-7392, Scopus Author ID: 57218676294.

Anzhelika V. Derbeneva — Student, Faculty of Economics and Construction, Training Profile «Applied Informatics in Economics», Bratsk State University, Bratsk, Russian Federation, anjelika.derbenyova@mail.ru,  <https://orcid.org/0000-0002-7880-5285>.

Вклад авторов

Все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Contribution of the Authors

The authors contributed equally to this article. The authors declare no conflicts of interests.

Для цитирования

Иванов М.Ю. Технологии интеллектуального анализа данных в решении экономических задач / М.Ю. Иванов, М.В. Сыготина, В.В. Надршин, А.В. Дербенёва. — DOI 10.17150/2411-6262.2022.13(2).27. — EDN [DWZNTH](#) // Baikal Research Journal. — 2022. — Т. 13, № 2.

For Citation

Ivanov M.Yu., Sygotina M.V., Nadrshin V.V., Derbeneva A.V. Data Mining Technologies in Solving Economic Problems. *Baikal Research Journal*, 2022, vol. 13, no. 2. (In Russian). EDN: [DWZNTH](#). DOI: 10.17150/2411-6262.2022.13(2).27.