



**А. Р. ХАСАНОВ**  
Аспирант первого  
года обучения ФГОБУ  
ВО «Финансовый  
университет  
при Правительстве  
Российской Федерации».  
Область научных  
интересов: стратегический  
менеджмент, инновации,  
Индустрия 4.0,  
промышленный Интернет  
вещей.

E-mail: ARHasanov@fa.ru

# ВЛИЯНИЕ ПРЕДИКТИВНОЙ АНАЛИТИКИ НА ДЕЯТЕЛЬНОСТЬ КОМПАНИЙ<sup>1</sup>

## АННОТАЦИЯ

Для анализа влияния предиктивной аналитики на деятельность компаний проведен обзор литературы. Предметно рассмотрены существующие виды аналитики на основе больших данных (Big Data): описательная, диагностическая, предписывающая и предиктивная аналитика. Рассмотрены основные инструменты предиктивной аналитики и представленные на рынке технические решения. Благодаря инструментам предиктивной аналитики компании могут анализировать и прогнозировать протекающие во времени процессы, выявлять тенденции, предвидеть изменения и, следовательно, более эффективно планировать будущее.

## КЛЮЧЕВЫЕ СЛОВА

ИНДУСТРИЯ 4.0, УМНОЕ ПРОИЗВОДСТВО, БОЛЬШИЕ ДАННЫЕ, ПРЕДИКТИВНАЯ АНАЛИТИКА, ЦИФРОВИЗАЦИЯ.

## ВВЕДЕНИЕ

Интенсивная глобальная конкуренция, неопределенность в отношении стоимости энергии и экспоненциальный рост информационных технологий подталкивают отрасли к освоению гибкого, высокопроизводительного и устойчивого (энергоэффективного) производства.

Для того чтобы справиться с многочисленными производственными задачами (например, с вопросами обеспечения гибкости, использования ресурсов и пр.), компании внедряют «умное производство». Для него характерно интенсивное применение продвинутой интеллектуальных систем, динамическое реагирование и оптимизация выпуска в режиме реального времени. Ключевым фак-

тором применения «умного производства» является анализ больших данных.

## АНАЛИЗ БОЛЬШИХ ДАННЫХ

Термином «большие данные» обозначают информацию, которая хранится в цифровых хранилищах компаний. На сегодняшний день – это ресурс, который организации используют малоэффективно, в то время как эти данные можно использовать для выявления тенденций, шаблонов, прогнозирования, поиска корреляций и т.д. Изучение массивов данных с целью нахождения невидной, но полезной с точки зрения развития компании информации лежит в основе бизнес-аналитики, к которой современные компании обращаются все чаще, так как это может помочь снизить издержки, увеличить выручку, повысить эффективность процессов и даже добиться роста конкурентоспособности.

Традиционно интеллектуальная аналитика достаточно трудоемка. На рис. 1 показана типичная последовательность действий при предиктивной аналитике.

Требуемые данные идентифицируются и собираются из различных источников (например, из системы управления ресурсами компании (Enterprise Resource Planning, ERP) или набора инструментов для управления клиентами, продажами, контроля работы менеджеров, автоматизации бизнес-процессов (Customer Relationship Management, CRM-системы) или берутся из хранилища данных. У разных аналитических инструментов разные требования к тому, как лучше обрабатывать данные. Обычно необходимо некоторое преобразование данных в формат, поддерживаемый конкретной аналитической системой, чтобы информация могла быть корректно обработана. После анализа данных делаются выводы, на основе которых в дальнейшем проводят изменения, например сегментацию клиентов или кластеризацию продуктов.

Поскольку в эпоху глобальной цифровизации компании собирают большое количество самой разной информации на регулярной основе, для интеллектуального анализа больших данных используется специальное программное обеспечение. За многие годы оно значительно улучшилось,

что привело к тому, что современные компьютеры могут анализировать все большие объемы данных, быстрее обрабатывать запросы и выполнять более сложные алгоритмы.

На рис. 2 представлена схема с описанием видов бизнес-аналитики с указанием вопросов, на которые добытая и проанализированная компанией на каждом этапе информация позволяет ответить.

Описательная аналитика призвана сообщать, что произошло. Простые отчеты и визуализации, которые показывают, что произошло в определенный момент или в течение определенного периода.

Диагностическая аналитика должна объяснить коренные причины произошедшего. Для этого используются более продвинутые инструменты, чем для описательной аналитики.

Предписывающая аналитика показывает, что компания должна делать, чтобы достичь желаемого результата. Сегодня на рынке сравнительно немного решений такого уровня, поскольку для них нужны серьезные ресурсы машинного обучения.

Предиктивная аналитика сегодня наиболее популярна. Интеллектуальные аналитические инструменты используют высокоразвитые алгоритмы для прогнозирования того, что может случиться в будущем. Часто эти инструменты используют искусственный интеллект и технологию машинного обучения, предполагающие самостоятельное (без описания человеком подробного алгоритма действий) выполнение компьютером поставленных перед ним задач по поиску закономерностей и решений на основе предложенных данных. Интерес к предиктивной аналитике обусловлен тем, что исследователи и компании озабочены вопросами прогнозирования будущего.

Предиктивная аналитика использует целый ряд инструментов статистики, интеллектуального анализа данных и теории игр. С предиктивной аналитикой связано распространенное заблуждение, что предсказания связаны только с будущим. Однако существует концептуальная классификация, которая решает эту проблему. Согласно этой классификации, существует два типа предсказательной аналитики: прогнозирование настоящего и формирование будущего.

Анализируемые инструментами предиктивной аналитики изменения развиваются подобно S-образным кривым. Однажды возникшие события начинают повторяться все чаще и чаще, образуя через некоторое время новый тренд или новую парадигму, которая становится лучшей практикой. В какой-то момент происходит что-то неожиданное, например, появляется новая технология, новый сильный игрок на рынке, экономический кризис и т.д. Происходит структурный сдвиг, возникает новая S-образная кривая, характеризующая новую парадигму.

Разные типы аналитики подходят для разных прогнозов (рис. 3). Прогнозирование настоящего необходимо для определения шаблонов поведения, выявления закономерностей в настоящем времени, то есть в рамках текущей парадигмы. Формирование будущего, напротив, призвано аккумулировать новые нетипичные для текущего состояния систем данные, с тем чтобы не только предсказать структурный сдвиг, но и определить содержание новой парадигмы.

Таким образом, возможности технологий предиктивной аналитики шире, чем может показаться на первый взгляд,

Рис. 1. Последовательность работы с данными в предиктивной аналитике



Рис. 2. Виды аналитики больших данных



<sup>1</sup> Статья подготовлена на основе результатов исследования «Индустрия 4.0: исследование влияния развития передовых производственных технологий на производительность российских промышленных компаний», проведенного за счет средств бюджетного финансирования в рамках госзадания Финансового университета при Правительстве Российской Федерации, 2018.



и позволяют не только делать прогнозы на основе получаемой в режиме реального времени информации, но и собирать новые данные, способные оказать влияние на текущую ситуацию в будущем. В табл. 1 приведен обзор типичных инструментов, основанных на функциональности анализа данных.

Таблица 1  
Функциональность Oracle Data Mining [Buytendijk F., Trepanier L., 2010]

Функциональность	Алгоритм	Применимость
Классификация	<ul style="list-style-type: none"><li>• Логистическая регрессия;</li><li>• деревья принятия решений;</li><li>• наивный байесовский классификатор;</li><li>• метод опорных векторов</li></ul>	<ul style="list-style-type: none"><li>• Моделирование ответов;</li><li>• рекомендация «следующий вероятный продукт»;</li><li>• создание эффективной стратегии удержания работников;</li><li>• моделирование кредитного дефолта</li></ul>
Регрессия	<ul style="list-style-type: none"><li>• Множественная регрессия;</li><li>• метод опорных векторов</li></ul>	<ul style="list-style-type: none"><li>• Оценка репутации/долгов;</li><li>• моделирование рентабельности клиентов</li></ul>
Обнаружение аномалий	<ul style="list-style-type: none"><li>• One-Class Support Vector Model – модуль в системе Azure Machine Learning для создания моделей обнаружения аномалий</li></ul>	<ul style="list-style-type: none"><li>• Предотвращение мошенничества и сетевого вторжения</li></ul>
Важность атрибута	<ul style="list-style-type: none"><li>• Минимальная длина описания</li></ul>	<ul style="list-style-type: none"><li>• Хирургическая подготовка;</li><li>• индекс потребительской лояльности</li></ul>
Правила ассоциации	<ul style="list-style-type: none"><li>• Алгоритм поиска ассоциативных правил</li></ul>	<ul style="list-style-type: none"><li>• Анализ потребительской корзины;</li><li>• анализ закономерностей потребительского поведения</li></ul>
Кластеризация	<ul style="list-style-type: none"><li>• Иерархический алгоритм K-Means – алгоритм кластеризации, предполагающий заранее указанное количество кластеров и случайным образом выбранные начальные центроиды;</li><li>• иерархический алгоритм O-Cluser – модель кластеризации на основе сетки</li></ul>	<ul style="list-style-type: none"><li>• Сегментация клиентов;</li><li>• Анализ генов и белков</li></ul>
Выделение функций	<ul style="list-style-type: none"><li>• Факторизация неотрицательных матриц</li></ul>	<ul style="list-style-type: none"><li>• Анализ текста, поиск</li></ul>

Как отмечалось выше, прогнозирование будущего помогает организациям повышать собственную конкурентоспособность за счет своевременного реагирования на изменения внешней и внутренней среды. Однако такое реагирование всегда предполагает принятие определенных управленческих решений, совокупность которых формирует процессы управления, направленные на создание и внедрение новых стратегий на основе мониторинга. Таким образом, получение благодаря предиктивной аналитике более достоверных и своевременных данных позволяет повысить качество процессов управления.

Одним из главных прогнозных инструментов является стратегическая карта, являющаяся частью сбалансированной системы показателей. Основная задача последней состоит в том, чтобы показать, как решения, принятые в настоящем, могут повлиять на будущие результаты. Это делается путем связывания индикаторов опережения и отставания. Первый прогнозирует будущую производительность, второй сообщает о прошлых результатах.

Тем не менее данный подход имеет недостаток: в случае некорректных значений индикаторов отставания построенная стратегическая карта потеряет свою актуальность. Кроме того, современные реалии меняются очень быстро, поэтому опыт прошлого далеко не всегда применим к настоящему.

Очевидно, что необходим иной подход к сбору данных в настоящем для прогнозирования будущего с точки зрения новых предположений, вопросов «что если?» и сценариев. Именно в этих предположениях можно зафиксировать риск и неопределенность, которые являются частью любого

Рис. 4. Результаты опроса компаний относительно направлений использования в их деятельности предиктивной аналитики [Halper F., 2014]



стратегического планирования. Чтобы получить актуальный прогноз, организация должна выполнить три важных шага:

- составить реалистичные предположения;
- выделить наиболее важные предположения;
- использовать драйверы, которые можно контролировать, и мониторить те, которые нельзя контролировать.

## ИСПОЛЬЗОВАНИЕ ПРЕДИКТИВНОЙ АНАЛИТИКИ

Компании используют предиктивную аналитику для решения сложных проблем и поиска новых возможностей – от прогнозирования поведения потребителей до поддержки обслуживания оборудования (рис. 4). Пока предиктивная

аналитика используется преимущественно в маркетинге и продажах. Компании хотят прогнозировать поведение потребителя при использовании той или иной маркетинговой кампании, оценивать возможность применения up sell продаж (продажа более дорогого товара: мотивация покупателя потратить больше денег в вашем магазине, например купить более дорогую модель того же продукта, добавить опции или услуги к приобретаемому продукту, перекрестные продажи: мотивация покупателя потратить больше денег но уже через продажу товаров из других категорий, нежели изначально выбранная пользователем), улучшать построение взаимоотношений с клиентами и удерживать их. Постепенно предиктивная аналитика находит применение при анализе портфеля продуктов, рисков и пр. Показательно, что около 80% респондентов планируют использовать предиктивную

Рис. 5. Преимущества от использования предиктивной аналитики (по данным Intel Corporation)





аналитику для оптимизации как минимум в течение ближайших трех лет. Компании начинают использовать предиктивную аналитику в операционном менеджменте, производстве, сервисном обслуживании и пр.

Предиктивная аналитика активно используется для поддержки основных стратегических решений и положитель-

ным образом влияет на основные показатели компании (рис. 5). Ее допустимо задействовать и для принятия краткосрочных тактических решений в рамках операционной деятельности.

Консалтинговое агентство Forrester прогнозирует рост рынка предиктивной аналитики и машинного обучения

Таблица 2  
Обзор предложений на рынке

Продукт	Сильные стороны	Слабые стороны
SASAnalyticsSuite (SAS)	<ul style="list-style-type: none"><li>Решения для компаний любого масштаба из любых отраслей;</li><li>инфраструктура и спектр возможностей;</li><li>встроенные решения;</li><li>углубленная аналитика;</li><li>простое развертывание</li></ul>	<ul style="list-style-type: none"><li>Сложность управления;</li><li>низкие оценки наборов методов интеллектуального анализа данных SAS Enterprise Miner и SAS Visual Analytics;</li><li>необходимость приобретения нескольких продуктов для обеспечения полной функциональности;</li><li>высокая стоимость</li></ul>
IBM SPSS Modeler (IBM)	<ul style="list-style-type: none"><li>Широкая клиентская база и постоянное внедрение инноваций;</li><li>приверженность технологиям с открытым исходным кодом;</li><li>поддержка широкого спектра типов данных (анализ текста, анализ сущностей, управление решениями и их оптимизация и др.);</li><li>управление моделью (точность и прозрачность рабочих процессов, развертывание модели, мониторинг деградации и пр.)</li></ul>	<ul style="list-style-type: none"><li>Проблемы совместимости с другими приложениями;</li><li>высокая степень бюрократизации;</li><li>слабое сервисное обслуживание</li></ul>
SAP BusinessObjects Predictive Analytics, SAP HANA SPS (SAP)	<ul style="list-style-type: none"><li>Интеграция с другими предложениями SAP, что обеспечивает значительную функциональность;</li><li>возможность масштабирования системы;</li><li>возможность добавить новые компоненты: систему бизнес-анализа SAP Business Objects, инструмент управления интеллектуальными системами SAP Predictive Factory, каталог расширений SAP Analytics</li></ul>	<ul style="list-style-type: none"><li>Непопулярность;</li><li>низкая удовлетворенность клиентов;</li><li>зависимость от высокопроизводительной NewSQL платформы для хранения и обработки данных SAP HANA (High-Performance Analytic Appliance – высокопроизводительное аналитическое устройство)</li></ul>
KnowledgeSTUDIO (Angoss)	<ul style="list-style-type: none"><li>Выполнение широкого спектра аналитических задач в единой среде;</li><li>интуитивно понятное в использовании программное обеспечение;</li><li>готовые решения для конкретных отраслей</li></ul>	<ul style="list-style-type: none"><li>Медленное развитие продукта на рынке;</li><li>сложность обработки больших объемов данных</li></ul>
Rapid Miner Platform (Rapid Miner)	<ul style="list-style-type: none"><li>Широкий спектр применения;</li><li>простота использования;</li></ul>	<ul style="list-style-type: none"><li>Ограничение использования данных (в рамках конкретных алгоритмов);</li><li>отсутствие глобальной развитой сети обслуживания</li></ul>
KNIME Analytics Platform (KNIME)	<ul style="list-style-type: none"><li>Относительно низкая стоимость решения;</li><li>гибкость, открытость и расширяемость за счет открытого исходного кода;</li><li>доступ пользователей через систему к данным и их трансформации;</li><li>развитые партнерские отношения</li></ul>	<ul style="list-style-type: none"><li>Сложность в управлении;</li><li>проблемы с масштабированием;</li><li>ограниченные возможности визуализации данных</li></ul>
FICO Decision Management Suite (FICO)	<ul style="list-style-type: none"><li>Широкий спектр применения в финансовой сфере;</li><li>функциональность системы (продукта) в ключевых областях управления принятием решений;</li><li>интуитивно понятное управление моделью, платформой и проектами</li></ul>	<ul style="list-style-type: none"><li>Низкий уровень производительности;</li><li>низкий уровень поддержки инструментов с открытым исходным кодом;</li><li>ограниченный выбор алгоритмов</li></ul>

в среднем на уровне 15% в год к 2021 году. Это обусловлено совместным использованием предиктивной аналитики и машинного обучения с инструментами Индустрии 4.0 (искусственный интеллект, глубинное обучение, Интернет вещей и пр.). В табл. 2 представлена сравнительная характеристика основных решений, предлагаемых на рынке.

## ВЫВОДЫ И ДАЛЬНЕЙШИЕ НАПРАВЛЕНИЯ ИССЛЕДОВАНИЯ

Предиктивная аналитика является одним из направлений по обработке больших данных, позволяет компаниям принимать более взвешенные и корректные решения сегодня для достижения лучших результатов завтра. Путем анализа данных компании получают ценную информацию и могут выстраивать крепкие отношения с потребителями, находить новые возможности, предвидеть угрозы, предотвращать мошенничество, защищая доходы и репутацию. Остается открытым вопрос сохранения данных, обеспечения безопасности информационных систем, организованных внутри компаний, а также адекватной интерпретации данных, полученных из разных источников. Кроме того, детального изучения требует вопрос оценки экономических последствий внедрения предиктивной аналитики.

## ЛИТЕРАТУРА

1. *Abbott D.* (2014) Applied predictive analytics: Principles and techniques for the professional data analyst. Indianapolis, IN: John Wiley & Sons. 456 p.
2. *Adomavicius G., Tuzhilin A.* (2005) Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions// IEEE Transactions on Knowledge and Data Engineering. Vol. 17, № 6. P. 734–749.
3. *Alstete J. W., Cannarozzi E. G. M.* (2014) Big data in managerial decision-making: concerns and concepts to reduce risk. International Journal of Business Continuity and Risk Management. Vol. 5, № 1. P. 57–71.
4. *Bag S., Anand N.* (2015) Modelling barriers of sustainable supply chain network design using interpretive structural modelling: an insight from food processing sector in India. International Journal of Automation and Logistics. Vol. 1, № 3. P. 234–255.
5. *Batra S.* (2014) Big Data Analytics and its Reflections on DIKW Hierarchy // Review of Management. Vol. 4, № 1/2. P. 5.
6. *Buytendijk F., Trepanier L.* (2010) Predictive Analytics: Bringing the tools to the data/Oracle Corporation. Redwood Shores, CA.
7. *Halper F.* (2014) Predictive analytics for business advantage // TDWI Research. URL: <http://tdwi.org/research/2013/12/best-practices-report-predictive-analytics-for-business-advantage.aspx?tc=page0>
8. *Harford T.* (2014) Big data: A big mistake? //Significance. Vol. 11, № 5. P. 14–19.
9. *Harris D.* (2013) Why Apple, eBay, and Walmart have some of the biggest data warehouses you’ve ever seen

// Gigaom.URL:<https://gigaom.com/2013/03/27/why-apple-ebay-and-walmart-have-some-of-the-biggest-data-warehouses-youve-ever-seen/>.

10. *Hashem I. A. T., Yaqoob I., Anuar N. B. et al.* (2015) The rise of «big data» on cloud computing: Review and open research issues // Information Systems. Vol. 47. P. 98–115.
11. *Jones-Farmer L. A., Ezell J. D., Hazen B. T.* (2014) Applying control chart methods to enhance data quality // Technometrics. Vol. 56, № 1. P. 29–41.
12. *Kanamoria Y., Matsuokab Y.* (2008) Development of a Model for Estimation of Household Consumption and Environmental Load Generation // Refereed Sessions I–II. Monday 10 March. P. 347.
13. *Marshall A., Mueck S., Shockley R.* (2015). How leading organizations use big data and analytics to innovate // Strategy Leadership. Vol. 43, № 5. P. 32–39.
14. *Myers S. A., Sharma A., Gupta P. et al.* (2014) Information network or social network?: the structure of the twitter follow graph // Proceedings of the 23rd International Conference on World Wide Web. P. 493–498.
15. *Nair P. R.* (2012) Supply Chain Analytics // CSI Communications. Vol. 33, № 9. P. 11.
16. *Tan K. H., Zhan Y., Ji G. et al.* (2015) Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph // International Journal of Production Economics. Vol. 165. P. 223–233.
17. *The Four V’s of Big Data* (2016) // IBM Big Data & Analytics Hub. URL: <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>.
18. *Weber G. M., Mandl K. D., Kohane I. S.* (2014) Finding the missing link for big biomedical data // Jama. Vol. 311, № 24. P. 2479–2480.
19. *Wu X., Zhu X., Wu G.-Q. et al.* Data mining with big data // IEEE transactions on knowledge and data engineering. 2014. Vol. 26, № 1. P. 97–107.