

ИННОВАЦИОННЫЕ ТЕХНОЛОГИИ В АНАЛИЗЕ ИНФОРМАЦИИ О КЛИЕНТУРЕ БАНКА. ВОЗМОЖНОСТИ И ПРАКТИКА BIG DATA

*Н.В. КЛИМОВСКИХ, кандидат экономических наук, доцент, доцент кафедры экономической теории, Кубанский государственный аграрный университет имени И.Т. Трубилина
e-mail: Nadin180676@yandex.ru*

*А.В. СОКОЛ, студент учетно-финансового факультета, Кубанский государственный аграрный университет имени И.Т. Трубилина
e-mail: sokol-alex-so-cool@mail.ru*

Аннотация

На смену традиционной банковской аналитики информации о клиентах приходят инновационные технологии, построенные на современной концепции Big Data. В статье приводится сравнение традиционной и инновационной банковской аналитики, выделены существенные признаки и методы построения технологий Big Data, приведены примеры из отечественной практики их применения. Сделан вывод о сглаживании негативного последствия от развития инновационных технологий Big Data.

Ключевые слова: инновация, технология, банк, Big Data, банковский продукт, клиент, данные.

Финансовый сектор современной рыночной экономики является одним из самых значимых потребителей больших массивов информации, использующий для этого самые современные информационные технологии. Высокая конкуренция, особенно в розничном банковском сегменте, постоянно заставляет кредитные организации инвестировать свои финансовые ресурсы в инновационные технологии, в первую очередь для целей обработки и аналитической оценки сложно структурированной банковской информации. В этой связи отметим, что перечень направлений в решении аналитических задач достаточно обширный, но, на наш взгляд, в нем приоритетно выделяются следующие позиции:

– прогнозно-аналитическая оценка, включая прогнозы развития банковского сектора в целом и конкретных кредитных организаций в частности;

– имитационное моделирование (например, влияния изменения процентной ставки на сальдо банковского счета);

– Spatial analysis (пространственный анализ, использующий топологическую, геометрическую и географическую информацию [12, с. 20–25]);

– статистический анализ (в частности, использование A/B-тестирования и методов анализ временных рядов);

– представление информации в виде рисунков с использованием интерактивных и анимационных технологий, в частности, для целей формирования исходных данных в проведении в дальнейшем анализа, т.е. ее визуализация [16, с. 35–40].

Результатом решения коммерческим банком аналитических задач является формирование, заметим, уже систематизированных и структурированных, информационных ресурсов об объекте аналитики – клиенте банка (потенциальном или уже существующем), отражающих определённые характеристики и закономерности, тенденции в них. В последствии на объект банковской аналитики (клиента) может оказываться достаточно сильное информационное и психологическое воздействие, причем сам объект даже может и не догадываться о последнем. В процессе своей жизнедеятельности человек, организация своими действиями оставляют о себе

данные, информацию, которую научились использовать банки в своих целях, приоритетной из которых является извлечение прибыли в текущем и (или) долгосрочном плане. Как говорил Роберт Грин в книге «48 законов власти»: «Можно заставить других зависеть от вас, применяя тактику разведывания секретов. Зная тайны окружающих, владея информацией, распространение которой им хотелось бы предотвратить, вы связываете свою судьбу с их судьбами. Вы неприкосновенны» [8]. Эту истину хорошо усвоили современные коммерческие банки, зачастую откровенно навязывая нам ту или иную банковскую услугу, предварительно уже располагая некоторой информацией о нас.

В своей деятельности отечественные банки (в совокупности) ежедневно совершают миллионы банковских операций и сопутствующих им действий, постоянно накапливают и хранят всю информацию, полученную от непосредственного и (или) косвенного общения с клиентами. Все данные о банковских клиентах либо связанная с ними информация (конвертация валютных средств, оформления кредитов, транзакции по банковским картам, сделки с ценными бумагами, звонки по мобильному телефону, перемещения, финансовое положение, посещаемые Интернет-сайты, потребительские предпочтения, увлечения и т.д.) вполне реально могут быть обработаны с помощью инновационных технологий в банковской деятельности и использоваться в достижении бизнес-целей при корректной обработке. Отметим, что такая идея использования информации (личной, коммерческой) в банковском деле не является чем-то новым и базируется на науке о данных Data Science (иногда даталогия – *datalogy* [1]), появившейся около двух десятков лет назад на стыке, по сути, совершенно разных дисциплин (математики, статистики, информатики, бизнеса, экономики). Объектом исследования в этой науке является Big Data, причем в России под этим термином подразумевают технологии обработки информации. Возможности применения Big Data логично вытекают, на наш взгляд, изначально из принципиальных трактовок этого понятия.

Общепризнанного понятия или какого-то хотя бы относительного единства в трактовке понятия Big Data как в технических, так и в гуманитарных и социальных науках не существует. Вместе с тем с большой долей уверенности на основе компаративистского подхода можно утверждать, что Big Data представляет собой обозначение структурированных и неструктурированных данных огромных объемов и многообразия, которые могут быть эффективно обработаны с помощью горизонтально масштабируемых программным обеспечением инновационных информационных технологий, появившимися как общепринятые механизмы управления базами данных и решения класса Business Intelligence [18; 3; 14]. Для более качественного понимания возможностей и перспектив использования в банковском деле Big Data необходимо выделить и сравнить концептуальные алгоритмические признаки традиционной [10, с. 40–46] и инновационной аналитики, полученной с использованием Big Data.

Перед финансовым менеджментом любого коммерческого банка традиционно в качестве одной из основных ставится задача увеличения доходности от продажи банковских продуктов и совершения финансовых операций. До применения инновационных технологий банковской аналитики на основе Big Data использовались агрессивные, так называемые push-методы продаж банковских продуктов своим клиентам (примечательно, что слово «push» переводится (англ.) как «толкать»). При этом банку были абсолютно не важны потребительские предпочтения его клиентуры, банковские продукты должны были быть проданы любой ценой. Используя данные о финансовом положении, накапливая и анализируя информацию относительно небольших объемов о личности клиентов, банки пытались улучшить свой сервис и одновременно увеличить клиентскую базу. Обобщенно алгоритм работы традиционной банковской аналитики может быть представлен следующим образом (рис. 1)

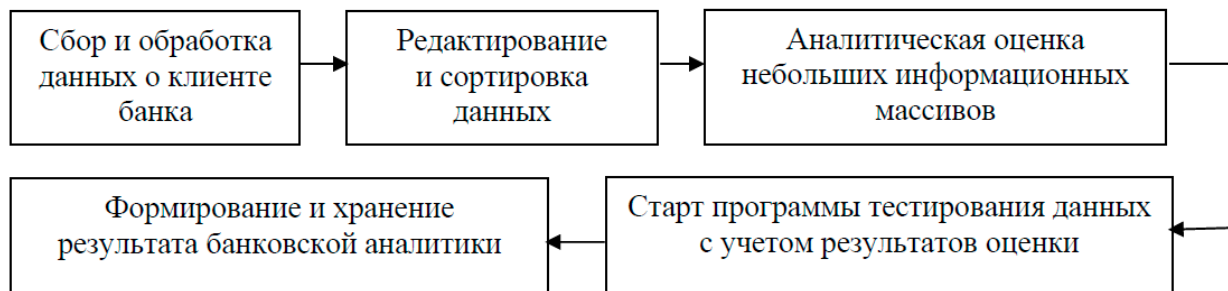


Рис. 1. Обобщенный алгоритм традиционной аналитики клиентуры банка

Алгоритм инновационной аналитики на основе технологий Big Data отличается от представленного выше (традиционного) существенным образом. Так, изначально [11] данная технология обработки данных содержит признаки, обозначаемые как «Три вэ» (на англ. «VVV»), а именно:

- объем (Volume) – информационные данные отражаются в некотором физическом объеме (документов);
- быстрота, скорость (Velocity) – быстрое обновление данных;
- разнообразие (Variety) – обрабатываемые данные имеют неоднородные (разнообразные) форматы, не структурированы (частично структурированы).

В дальнейшем к перечисленным признакам Big Data добавлялись другие и данный метод получал новые названия:

- «Четыре вэ», дополнительный признак Veracity (достоверность, часто использовался в оценке рекламных материалов компании IBM [2]);
- «Пять вэ», дополнительные признаки (к исходному варианту Big Data) Viability (жизнеспособность) и Value (ценность) [5];
- «Семь вэ», дополнительные признаки (к предыдущему варианту Big Data) Variability (переменчивость) и Visualization (визуализация) [13].

В нашем представлении применительно к аналитике банковской клиентуры предлагается комбинация из шести признаков Big Data (рис. 2).

Полагаем, что в данном случае такие признаки, как Viability и Visualization, не являются существенными (принципиальными). В отличие от традиционной обработки информационных данных обобщенный алгоритм на основе технологий Big Data может быть представлен следующим образом (рис. 3)

Принципиально, что обработка данных о банковском клиенте происходит сразу и с созданием в последующем аналитической корреляции (взаимосвязей) по этим данным. Анализ измененных данных осуществляется в реальном режиме времени, что исключено в технологии традиционной аналитики клиентуры коммерческого банка. Подлежат обработке и массивы из неструктурированной информации, чего нет в традиционной аналитике.

В алгоритме инновационной аналитики реализуются особенные методы построения информационной банковской технологии. Среди них:

- методы Data Mining, направленные на обучение ассоциативным правилам (методы категоризации новых данных, применяемые к

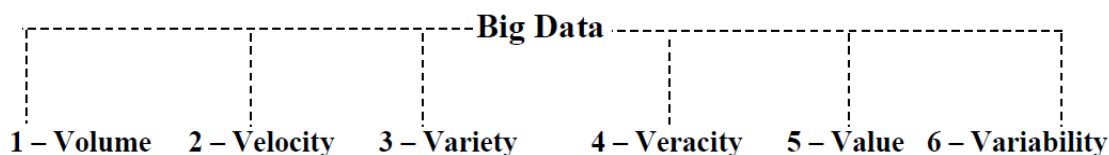


Рис. 2. Авторское представление существенных признаков Big Data

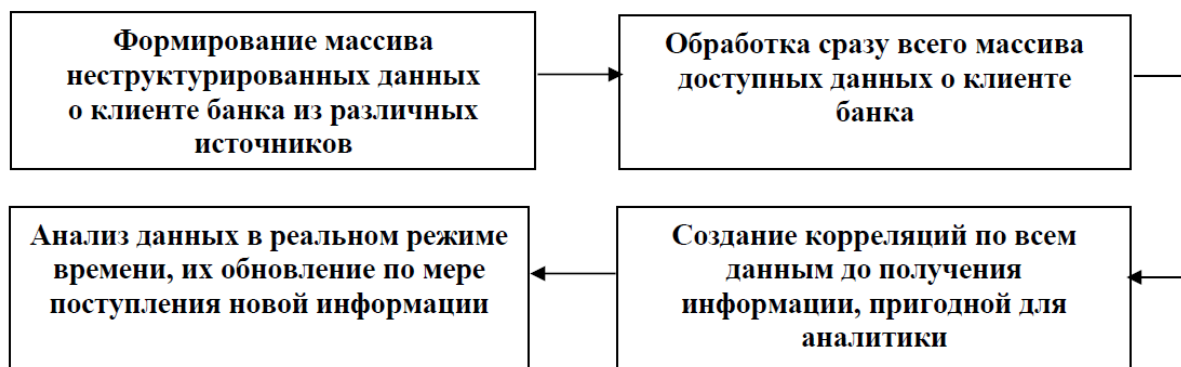


Рис. 3. Обобщенный алгоритм инновационной аналитики клиентуры банка

уже сформированным данным, кластерного и регрессионного анализа [17, с. 55–71];

- краудсорсинг – метод категоризации и обогащения информационных данных путем привлечения неопределённого круга лиц без оформления трудовых отношений как таковых;

- Data fusion and integration – методы смешения и интеграции данных, обеспечивающих интеграцию разнородных данных из разнообразных источников для возможности проведения углубленного анализа информации о клиентуре банка, когда производится цифровая обработка сигналов и естественного языка;

- методы машинного обучения [6, с. 623, 638–646], включая обучение без преподавателя, а также метод Ensemble learning (метод формирования комплексных прогнозов на основе базовых моделей с использованием приемов статистического анализа или машинного обучения [7, с. 168–175]);

- методы искусственных нейронных сетей, сетевого анализа, оптимизации, генетические алгоритмы [15, с. 218–220];

- методы распознавания образов и др.

Применяя в своей деятельности инновационные технологии Big Data в целях анализа информации о клиентуре, коммерческие банки имеют возможность значительно расширить круг решаемых основных аналитических задач, а также решать задачи по уже отработанным алгоритмам, но совершенно на другом качественном уровне в следующих областях [9]:

- сегментация клиентов (Customer Segmentation) по различным признакам [21, с. 141, 194, 264–290];

- обнаружение прецедентов мошенничества (Fraud Detection);

- кредитный скоринг (Credit Scoring) [19, с. 11–12];

- персональные предложения клиенту (Personalized Product Offering);

- прогнозирование ухода клиентов с банковского рынка и возможности привлечения новых клиентов (Customer Loyalty and Marketing);

- формирование отчетности и оценка соответствия нормам банковского регулятора (Compliance and Regulatory Reporting).

Использование ряда решений в области Big Data имеет место в практике российских банков последних лет (см. таблицу).

Примеры из практики внедрения инновационных технологий Big Data в отечественном банковском секторе [20].

Использование банками результатов анализа, полученных с помощью инновационных технологий Big Data, имеет позитивный эффект как в отношении клиента банка, так и для самой кредитной организации. Например, с помощью технологий Data Science и Machine Learning могут создаваться достаточно простые алгоритмы, реализация которых позволяет анализировать и «фильтровать» деятельность клиента-пользователя, чтобы создать наиболее привлекательное для него предложение. Данные технологии, дополненные практическими рекомендациями, де-

Опыт внедрения инновационных технологий Big Data в банковском секторе РФ

Наименование банка	Решаемые аналитические задачи	Программные продукты и технологии Big Data
Альфа-банк	Анализ клиентов в социальных сетях, их поведения как пользователей сайтов Оценка кредитоспособности и платежеспособности клиентов Прогноз объемов оттока клиентов из банка Персонализация содержимого кредитной истории клиента	Oracle Exadata Database Machine Oracle Big Data Appliance Платформа Hadoop
ВТБ24	Сегментация клиентуры банка Формирование финансовой отчетности Анализ отзывов клиентов о банке в социальных сетях и на форумах	Продукты компании Teradata SAS Visual Analytics SAS Marketing Optimizer
Газпромбанк	Доскоринговая проверка репутации потенциальных заемщиков Противодействие мошенничеству клиентов Оперативное получение отчетности клиентов (предприятий) Персонализация предложений клиентам Предоставление результатов анализа информации банковскому регулятору	SAS Visual Analytics SAP BusinessObjects
ОТП Банк	-Оптимизация технологического процесса продажи банковских продуктов	IBM Campaign
Нордеа Банк	-Формирование периодической отчетности банка	HP Vertica
Райффайзен банк	-Формирования текущей (ежедневной) банковской отчетности	Oracle (хранилище данных GDWH, Master Data Management)
Сбербанк РФ	Повышение качества обслуживания банковских клиентов Управление рисками по кредитным продуктам Оптимизация затрат на банковское обслуживание клиентов Борьба с мошенничеством клиентов Сегментация и оценка кредитоспособности клиентов Управление персоналом банка в целях улучшения банковского сервиса Прогнозирование очередей клиентов в отделениях Сбербанка Расчет бонусов для сотрудников банка (в том числе, по критерию лояльного отношения к клиентам)	Продукты компании Teradata Продукты компании Cloudera Сервисы Yandex Data Factory

Окончание таблицы

Наименование банка	Решаемые аналитические задачи	Программные продукты и технологии Big Data
Ситибанк	Оптимизация внутренних процессов коммерческого банка, включая аналитику	Продукты компании Teradata
Тинькофф Банк	Управление рисками в продаже банковских продуктов Анализ потребностей потенциальных и существующих клиентов Скоринг Маркетинговые программы	EMC Greenplum Продукты компании Cloudera SAS Visual Analytics Платформа Hadoop
Тройка Диалог	Прогнозная аналитика клиентуры Брокерские отчеты	Продукты компании EMC Greenplum Продукты компании SAS
УРАЛСИБ	Анализ результатов проведения маркетинговых мероприятий	IBM Contact Optimization
Уральский банк реконструкций и развития	Создание персонализированных кредитных предложений клиентам	SAP Business Communication Management
ХМБ Открытие	Анализ поведения клиентов	Продукты компании SAP HP Vertica

монстрируют возможности в повышении заинтересованности клиентов в приобретении и использовании предлагаемого банковского продукта. При этом, чтобы создать эффективный механизм формирования рекомендаций, технологии Data Scientist анализируют и обрабатывают большие массивы неструктурированной информации, идентифицируют профили клиентов и охватывают данные, которые показывают взаимодействия между ними. Вероятность повторения банком предложений клиенту сводится к минимуму.

При всех очевидных положительных моментах и последствиях от применения инновационных технологий Big Data необходимо сказать о том, что их дальнейшее развитие делает конфиденциальную информацию о клиенте банка все более доступной для злоумышленников. Так, если банки уже сегодня имеют возможность получать весьма подробные данные о своих клиентах и их потребительском поведении, обладают информацией о совершаемых (безналичным способом) ими

покупках, то почему бы банкам не начать диктовать принятие решений клиентами исключительно в свою пользу?

Высказанное нами опасение имеет теоретическое обоснование. Например, результаты исследования М. Козински (Высшая школа бизнеса, Стэнфорд, США) по психометрии говорят о том, что можно отнести человека к тому или иному типу поведения, предсказывать его действия на основе социологических данных, количества «лайков». По мнению ученого анализ всего 68 лайков позволяет определить цвет кожи человека, его политические и сексуальные предпочтения. Анализ от 70 до 150 «лайков» позволяет получить такой же объем информации, который известен только близким знакомым исследуемого индивидуума, а оценка до 300 «лайков» покажет информацию о человеке в объеме даже больше, чем знают об этом его родители. Оценка свыше же 300 «лайков» позволяет узнать индивида так же, как знает он себя сам. Данную технологию применяло информаци-

онное агентство Cambridge Analytica в целях персональной политической рекламы во время выборов президента США, а что мешает сегодня применять ее и коммерческим банкам в отношении своей клиентуры?

В заключение скажем, что данная проблема информационной безопасности в банковской аналитике может быть разрешена, только если банки будут следовать свое кредо «Scientia potentia est» (лат. «знание – сила») в области информационной политики, придерживаясь как норм банковского законодательства, так и норм этики при проведении аналитической работы с личными данными своих клиентов.

Библиографический список

1. Беркана А. Что такое Big data: собрали всё самое важное о больших данных. 16 мая 2017. URL: <http://rb.ru/howto/что-такое-big-data>.
2. Бьен Н. Исчезающее V's в Big Data: Жизнестойкость и Ценность. // URL: <https://www.wired.com/insights/2013/05/the-missing-vs-in-big-data-viability-and-value/>
3. Гасты Т., Тибширани Р., Фридман Дж. Элементы статистического обучения: поиск данных, вторжение и предсказания. 2-е изд., Springer, 2017.
4. Герон А. Практическое машинное обучение при помощи Scikit-Learn и TensorFlow: концепты, инструменты и техники построения организационных систем. O'Reilly, 2017.
5. Грин Р. «48 законов власти». Закон № 11. «Старайся, чтобы люди зависели от тебя». URL: <https://48power.bib.bz/zakon-11-staraysya-chtoby-lyudi-zaviseli-ot-tebua>
6. Данрадджани С. Искусственный интеллект и Аналитика: Прогрессивные бизнес решения. Wiley, 2018.
7. Конотопов П.Ю., Курносков Ю.В. Аналитика: методология, технология и организация информационно-аналитической работы. М., 2004.
8. Лэни Д. 3D-менеджмент данных: контроль объема данных, скорости и вариативности. Технический отчет. META Group. 2001. URL: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>
9. Смит М. Дж. де, Гудчайлд М. Ф., Лонг-ли П.А. Геопространственный анализ: Полный сборник законов, техник и программных механизмов. 2-е издание. Winchelsea, 2007.
10. МакНалти Е. Понимая Big Data: 7 V's. Май 22. 2014. // URL: <https://dataconomy.com/2014/05/seven-vs-big-data/>
11. Маньика Дж. [и др.] Большие данные: Следующий рубеж инноваций, соревнований и продуктивности. McKinsey Global Institute. 2011
12. Митчелл Т. Изучение машин. McGraw-Hill Science/Engineering/Math.
13. Нуссбаумер Кнафлик К. Повествование через данные: Визуальный справочник по данным для бизнес профессионалов. Wiley, 2018.
14. О'Нилл К., Шутт Р. Изучение науки о данных. Разговор лицом к лицу с передовой линии. O'Reilly. 2014.
15. Праймесбергер К. Yahoo, Большие данные. Светлое будущее бизнес организаций. 2011. URL: <https://www.eweek.com/storage/hadoop-yahoo-big-data-brighten-bi-future>
16. Сиддики Н. Кредитный риск протоколов результата: Развитие и применение организационных кредитных протоколов. New Jersey. 2006.
17. Специальное исследование. Большие данные в финансовой отрасли: обзор и оценка перспектив развития мирового и российского рынков. IDC. 2015 URL: <file:///C:/Users/Пользователь/Downloads/idc-26012016.pdf>
18. Тсинтсис К., Корианопулос А. Технологии поиска данных в системе управления взаимоотношений между клиентами: внутренняя сегментация покупателя. 2009.

19. Фельдман С. Минута в интернете 2019 г. URL: <https://www.statista.com/chart/17518/internet-use-one-minute/>

20. *Gil Press*. Data Scientists: The Definition of Sexy. 2012. URL: <https://www.forbes.com/sites/gilpress/2012/09/27/data-scientists-the-definition-of-sexy/#13c277d35f96>

21. The Four V's of Big Data. IBM 2011. URL: https://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg

22. *Alan Morrison etc.* Big Data: how to extract information from them; Technological forecast // Quarterly magazine. Russian edition. 2010. № 3. PricewaterhouseCoopers.