

УДК 336.71

М.П. ГАЛИМОВА

ФГБОУ ВО «Уфимский государственный авиационный технический университет», г. Уфа, Россия

E-mail: polli66@mail.ru

В.У. КАРПОВА

ФГБОУ ВО «Уфимский государственный авиационный технический университет», г. Уфа, Россия

E-mail: V.Almukhametova@yandex.ru

**РАЗВИТИЕ ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ BIG DATA В
СОВРЕМЕННОЙ БАНКОВСКОЙ ДЕЯТЕЛЬНОСТИ
THE DEVELOPMENT OF THE USE OF BIG DATA TECHNOLOGY
IN MODERN BANKING**

Аннотация: современное развитие банков предполагает рост ассортимента кредитных продуктов и объемов их реализации как для физических лиц, так и для юридических. Важной проблемой является обеспечение устойчивости банков. А она напрямую зависит как от устойчивости привлекаемых ресурсов, так и от платежеспособности кредитуемых клиентов. Построить эффективные прогнозы и оценить риски позволяет технология big data, которая предполагает формирование структурированных баз данных на основе обработки структурированной, неструктурированной и слабоструктурированной информации. Ее всё чаще применяют крупные банки, имеющие ресурсы на развитие таких систем.

Abstract: the modern development of banks implies an increase in the range of loan products and their sales volumes, both for individuals and legal entities. An important problem is ensuring the stability of banks. And it directly depends on both the stability of the attracted resources and the solvency of the credited customers. The construction of effective forecasts and risk assessment allows the big data technology, which involves the formation of structured databases based on the processing of structured, unstructured and poorly structured information. It is increasingly being used by large banks that have the resources to develop such systems.

Ключевые слова: big data, оценка рисков, прогнозирование, скоринг, базы данных, обработка информации.

Keywords: big data, risk assessment, forecasting, scoring, databases, information processing.

Технология big data («большие данные») представляет собой технологическое решение, которое позволяет обрабатывать накопленную и вновь поступающую информацию для использования этой информации в принятии решений, оценке ситуации и т.п. [1, с.7]. Big data позволяет формировать большие массивы данных в текущем времени, проводить их обработку, чтобы получать актуальные оценки. Но применение этой

технологии требует значительных вложений в разработку самой технологии обработки, сбора информации, а также принятия решения на основе обработки. Требуются и значительные инвестиции, связанные с развитием IT-инфраструктуры для получения, хранения и обработки такой информации (серверы, процессинговые центры и т.п.). Смысл в создании таких систем особенно высок, когда организация осуществляет работу с большим количеством клиентов, имеющих разные характеристики, выполняющими в процессе работы с ними разные действия, при том, что эти действия могут быть зафиксированы и обработаны.

Именно поэтому банки в нашей стране наиболее крупные пользователи технологии big data. Применение такой технологии в этой сфере позволяет качественно оценивать риски, в том числе связанные с кредитованием, поскольку накапливаются и обрабатываются действия клиента с разными банковскими продуктами.

Технология работы big data выглядит следующим образом (рис.1).



Рис. 1. Технология работы Big data

То есть, поступающая информация по структурированным данным (имеющим четкое цифровое выражение) попадает сразу в BI-платформу. BI (Business Intelligence) – это набор IT-технологий для сбора, хранения и анализа данных, позволяющих предоставлять пользователям достоверную аналитику в удобном формате, на основе которой можно принимать эффективные решения для управления бизнес-процессами компании. Неструктурированная информация (на практике ее гораздо больше, она не имеет четкой структуры и состава), а также слабоструктурированная информация проходит обработку с помощью технологии big data, структурируется на основе разработанного алгоритма, из нее выделяются важные характеристики, она трансформируется в интеллектуальный

интегрированный архив и далее также может использоваться VI-платформой для принятия решений.

В статье рассматриваются возможности и проблемы big data в банковской деятельности.

Одним из достаточно распространенных направлений применения big data является кредитный скоринг, т.е. оценка кредитных рисков физического или юридического лица, основанная на статистических методах.

Актуальность кредитного скоринга связана с высокой долей просроченных кредитов (в т.ч. сомнительных, проблемных и безнадежных ссуд) в общих активах российских банков, а также сопутствующий этому рост резервов банков (рис. 2).

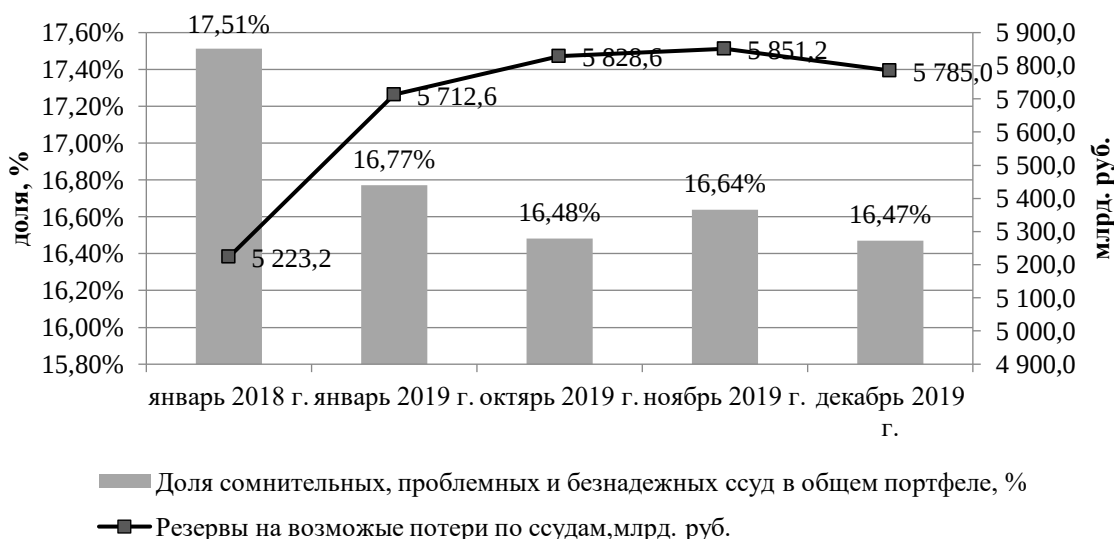


Рис. 2. Доля сомнительных, проблемных и безнадежных ссуд и величина резервов в российском банковском секторе в 2018-2019 гг. [2]

При реализации big data собираются данные не только по конкретному клиенту, которых может и не быть, но они собираются по подобным клиентам, исходя из отрасли, величины бизнеса, численности персонала, длительности ведения бизнеса, учитывается информация и по собственникам бизнеса и их экономической деятельности, включая их предыдущие бизнесы (если речь о кредитовании юридических лиц), предыдущие кредиты в том же банке. До внедрения big data собрать эти данные было практически невозможно, потому использовались только справки с места работы потенциального заемщика (для юридических лиц – отчетность), проверка службой безопасности, операционная история (поступления и расходования со счета), если счет открыт в этом банке, а также данные из бюро кредитных историй. Не исследовались такие данные как отзывы о бизнесе и тенденции развития отрасли, поведение заемщика-физического лица или собственника бизнеса в социальных сетях, информация в СМИ и иных источниках, негативные упоминания о предприятии в СМИ и прочее. С применением big data можно собрать полный комплекс такой информации, обработать, сопоставить данные. На основе этого система формирует вывод о вероятности

несовременного погашения кредита, банкротства клиента и т.п., откуда можно сделать вывод о возможности кредитования [3, 4].

Поэтому и создаются модели, имеющие множество переменных, анализирующие возможные параметры, которые могут повлиять на потенциальный риск невозврата кредитов, выход клиентов на просроченную задолженность [5, с.20]. Научной и практической задачей является обоснование перечня показателей для скоринга. В частности, целесообразным является включение таких показателей для анализа физического лица как информация о его деятельности в Интернет (поиск объявлений по номерам телефонов, например), поиск возможной негативной информации о лице, анализ поведения лица в социальных сетях. Для юридических лиц: анализ аналогичных анализу физических лиц данных, а также отзывы о бизнесе, иная информация о бизнесе, которая есть в Интернет.

Примером применения big data в банковской деятельности может быть технология Сбербанка, внедряемая с конца 2016 г., по беззалоговому кредитованию малого бизнеса на основе анализа больших данных по потенциальному клиенту [1].

Применение big data может быть связана и со сбором информации о просроченной задолженности, а также для взыскания такой задолженности. С учетом сложности поиска должников, которые не погашают просроченный долг, не выходя на связь с банком, технология может помогать искать информацию о них другими способами, в т.ч. обрабатывая данные социальных сетей, форумов и иных ресурсов [6, с. 19].

Технология big data достаточно широко используется и в банковском маркетинге, в том числе, на основе данных о клиентах и их транзакциях.

Сбербанк, как достаточно крупный банк, имеющий большие массивы информации, для развития применения big data в банковском маркетинге приобрел контрольный пакет в RuTarget, которая разработала рекламную платформу Segmento, в результате чего банк получил возможность развития «умной рекламы» [3], а также развития возможностей накопления информации для принятия маркетинговых решений. С начала 2017 г. Сбербанк начал реализацию программы «Открытые данные». Суть данной программы также связана с применением технологии big data, формировании на ее основе структурированной аналитической информации – агрегированных дата-сетов – наборов данных, которые основаны на транзакционных данных наших клиентов. В результате банк может представить статистику по средним суммам и количествам заявок по разным видам кредитов, о средних размерах и количествах новых депозитов, о динамике зарплат и пенсий (поскольку именно Сбербанк является крупнейшим банком, через который такие перечисления производятся). Эта система позволяет также делать финансовые прогнозы на основе формирования моделей в результате обработки данных [7, с.157]. В целом, обработка данных и формирование подобных показателей позволяет развивать новые банковские продукты, которые могут быть востребованы теми или иными группами потребителей.

В настоящее время коммерческие банки все чаще применяют технологии big data для обработки информации в целях развития кредитного скоринга, оценки рисков и маркетингового анализа. Подобные технологии реализованы не только в Сбербанке, но и во втором по размеру активов банке – ВТБ, а также в банке «Тинькофф», который стал одним из самых быстроразвивающихся в стране, увеличив свой рейтинг по величине активов с 25 до 17 места только за 2019 г. [8].

В таблице 1 представлены данные о вложении и об эффекте внедрения технологии big data в деятельности Сбербанка, ВТБ и банка «Тинькофф».

Таблица 1

Результаты внедрения технологий big data в деятельность коммерческих банков Сбербанк, ВТБ и «Тинькофф»

Банки	Область применения big data	Эффект
ПАО «Сбербанк»	Внедрены технологии Teradata, Cloudera Hadoop, Impala, Zettaset, стек продуктов Apache, Segmento RuTarget. Использование технологии для управления рисков и выявления мошенничества, оценки платежеспособности (кредитный скоринг), управления персоналом, анализа и прогнозирования очередей клиентов в офисах обслуживания банка, формирования и продажи статистических данных («Открытые данные» Сбербанка)	В 2017 г. получено прибыли в размере \$2-3 млрд за счет анализа потенциальных заемщиков на основе big data. Прибыль банка за 2019 г. увеличилась на 11,2% и составила 870,1 млрд руб.
ПАО «ВТБ»	Внедрены технологии Teradata, SAS Visual Analytics и SAS Marketing Optimizer. Обработка отзывов из социальных сетей, сегментирование клиентов, прогнозирование доходов и финансовой отчетности. В 2020 г. совместно с «Ростелеком» создано ООО «Платформа больших данных для разработки и монетизации продуктов на основе big data, искусственного интеллекта и машинного обучения	Инвестиции в \$33 млн окупались за 3 года, принесли около 19 млрд руб. чистого дохода за год. В 2019 г. чистая прибыль банка увеличилась на 12,9% – до 208,3 млрд руб.
АО «Тинькофф Банк»	Внедрены технологии EMC Greenplum, SAS Visual Analytics и Hadoop. Формирование управленческой отчетности, предиктивный анализ, кредитный скоринг. С 2019 года в соглашение об обслуживании карт внесен пункт по получению информации о покупках по карте для развития маркетинга на основе big data	Сокращение в 10-100 раз времени на обслуживание внутренних процессов. За 2019 г. чистая прибыль банка составила 27,5 млрд руб. (рост на 61,9% по сравнению с 2017 г.), привлечено более 4,3 млн новых клиентов по кредитным продуктам

Источник: составлено авторами на основе [8, 9, 10, 11]

При этом, проблемами развития технологии big data в банковской сфере является высокая стоимость самой технологии (требуется большие вложения), необходимость внедрения новых программ и недостаточность специалистов в этой сфере.

В таблице 2 представлен SWOT-анализ развития применения технологий big data в банке с позиции кредитора.

Таблица 2

**SWOT-анализ развития технологий big data в банках
(с позиции кредитора)**

Внутренняя среда (применения технологии в банках)	
Сильные стороны (S)	Слабые стороны (W)
Значительные финансовые возможности Значительные технологические возможности развития за счет кооперации с разработчиками Переход от продуктоцентричной к более эффективной клиентоцентричной модели бизнеса Возможность создания больших массивов данных по клиентам и их поведению и за счет этого большая дифференциация банковских услуг Сокращение сроков и трудоемкости анализа кредитоспособности Минимизация рисков банка Гибкость реагирования на изменения рынка	Высокая стоимость внедрения big data Необходимость вложений в инфраструктуру для хранения информации и процессинга Необходимость обеспечения защиты собираемой информации, что увеличивает стоимость системы Дефицит специалистов в области big data в банковском секторе не только в России, но и в мире (неразвитость цифровых компетенций) Финансовые и технологические ограничения применения для средних банков, в т.ч. региональных банков Низкая цифровая зрелость банков Ограничение доступа к информации корпоративных клиентов
Внешняя среда	
Возможности (O)	Угрозы (T)
Развитие технологий Интернет для передачи данных Развитие технологий обработки данных big data крупными разработчиками Развитие безналичных транзакций клиентов банков	Монопольное положение крупных производителей систем big data на рынке Ужесточение государственного регулирования применения big data и ограничения применения использования технологии

Источник: составлено авторами

То есть, имеются определенные возможности и угрозы развития технологии big data в банках, есть и сильные стороны, но они в большей мере характерны для крупных банковских структур. Слабые стороны внедрения big data связаны с высокой стоимостью внедрения как технологии, так и развития инфраструктуры, а также защиты информации. Также проблемой является и дефицит специалистов в этой сфере.

Для заемщиков использование технологии big data также обеспечивает экономические выгоды и риски.

Таблица 3

Экономические выгоды и риски технологий big data с позиций заемщика

Выгоды	Риски
Сокращение сроков предоставления кредитов (экономия времени) Персонализация предложений банка и оптимизация условий и стоимости услуг Возможность получения информации о лучших кредитных и банковских услугах за счет применения цифровых агрегаторов	Риск получения банком ошибочной негативной информации, ухудшающей рейтинг потенциального заемщика Возможная утечка данных клиента в банке и попадание информации в общий доступ или мошенническим структурам

Источник: составлено авторами

Внедрение цифровых технологий, в том числе и big data, требует оценки цифровой зрелости банков. Поэтому актуальной стоит и задача идентификации цифрового профиля [11]. Низким уровнем цифровой зрелости объясняется проблема внедрения big data для региональных банков, не имеющих значительные масштабы и инвестиции для цифровизации. Такое положение порождает цифровое неравенство, что снижает конкурентоспособность региональной банковской системы. Здесь можно предложить использовать механизм совместного участия в разработке и внедрения систем big data для нескольких банков одновременно на основе двустороннего или многостороннего сотрудничества. Актуальным может стать кооперация с частными компаниями, специализирующимися на агрегации различной информации по клиентам, построении поведенческих моделей, моделей платежеспособности тех или иных сегментов и т.д. с предоставлением всем заинтересованным организациям соответствующих платных сервисов.

Технология big data, помимо оценки кредитоспособности заемщика, может использоваться, и уже широко используется внедрившими ее банками, в сфере банковского маркетинга, а также в операциях с ценными бумагами, в управлении обслуживанием клиента, в управлении персоналом банка.

Таким образом, технологии big data представляют возможности на основе широкого и разнообразного массива данных, с учетом обработки по специальному алгоритму, формировать структурированную базу данных, которая используется для оценки рисков кредитования, для оценки потребительского поведения клиентов банка, для организации управления банком. Благодаря внедрению таких систем в оценку кредитоспособности можно улучшить качество ссуд и сократить резервы на возможные потери и сами потери для банка. Несмотря на рост возможностей внедрения big data в банковской сфере в России, имеются проблемы, связанные с внедрением и развитием таких систем, особенно, в тех банках, которые не имеют значительных ресурсов для вложений в эти технологии или разработку собственных продуктов.

Список использованной литературы:

1. Локтионова Е.А., Рагозина А.В. Особенности применения систем анализа больших данных в деятельности коммерческого банка // *Baikal Research Journal*. – 2017. – №2. – С.7-15.
2. Статистика банковского сектора // Банк России: официальный сайт. [Электронный ресурс]. – Режим доступа: https://cbr.ru/banking_sector/statistics/ (дата обращения: 11.04.2020).
3. Носов Н. Big Data в российских банках. Начало большого пути // *IT Week*. – 2015. [Электронный ресурс]. – Режим доступа: <https://www.itweek.ru/idea/article/detail.php?ID=176526> (дата обращения 14.04.2020).
4. Альмухаметова В.У., Галимова М.П. Сравнительный анализ рынка ипотечного кредитования в России и за рубежом // В сборнике: Актуальные проблемы развития социально-экономических систем: теория и практика. Сборник научных статей 9-й Международной научно-практической конференции / под редакцией: Болычевой Е.А. – Т. 1. – Курск: ЮЗГУ, Юго-Зап. гос. ун-т. – 2019. – С. 42-48.
5. Казаков Р.И. Технологии Big Data в управлении крупными банками // *Бизнес-образование в экономике знаний*. – 2015. – №2 (2). – С.19-23.
6. Бердышев А.В. Открытая платформа как технологическая основа развития ПАО «Сбербанк» // *Вестник ГУУ*. – 2018. – №11. – С.154-159.
7. Рейтинг коммерческих банков России // *Банки.ру*. – 2020. [Электронный ресурс]. – Режим доступа: <https://www.banki.ru/banks/ratings> (дата обращения: 12.04.2020).
8. Мейксин С.М. Новые технологии банков // *Вестник науки и образования*. – 2020. – №2-2 (80). – С.29-33.
9. Соколова А. Как устроен рынок big data в России // *RB*. – 2015. [Электронный ресурс]. – Режим доступа: <https://rb.ru/howto/big-data-in-russia/> (дата обращения 10.04.2020).
10. Шилова Е.В., Мороз А.К. Развитие банковской сферы в условиях промышленных революций // *Вестник Прикамского социального института*. – 2019. – №1 (82). – С.128-134.
11. Галимова М.П. Готовность российских предприятий к цифровой трансформации: организационные драйверы и барьеры // *Вестник УГНТУ. Наука, образование, экономика. Серия: Экономика*. – 2019. – №1 (27). – С. 27-37.