

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 336.025

Алиш
Мария Станиславовна

**ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ BIG DATA В ФИНАНСОВОМ
СЕКТОРЕ РЕСПУБЛИКИ БЕЛАРУСЬ: ПРОБЛЕМЫ И
ПЕРСПЕКТИВЫ**

АВТОРЕФЕРАТ

на соискание степени магистра экономических наук

По специальности 1-25 80 04 «Экономика и управление народным
хозяйством»

Научный руководитель

Марахина Инна Викторовна

Кандидат экономических наук, доцент

Минск, 2020

ВВЕДЕНИЕ

Термин «Big data» (Big data) активно используется в литературе, но единой модели использования, терминологии понимания пока не сформировалось. В данной работе структурирована разрозненная информация о Big data и определено, какую роль играют эти данные в диджитализации (англ. Digitalization - перевод всех видов информации в цифровую форму) финансового сектора Республики Беларусь. С развитием информационных технологий эволюция отразилась и в области аналитики данных. Поэтому необходимость учета Big data при развитии финансового сектора неизбежна. Однако, в настоящее время при внедрении Big data возникают некоторые вопросы в части информационной безопасности, опасения обесценивания человеческого капитала при внедрении новых технологий, требования к аппаратным системам по мощности и высокой скорости обработки данных, и иные задачи проблематики Big data. В данной работе определены проблемы во внедрении анализа Big data. При рассмотрении вопроса проблемы визуализации результатов анализа данных обозначены методы выделения и извлечения систематизации данных, а также рассмотрены критерии оценки видов отображения данных. В ходе работы кратко рассмотрены самые популярные аналитические платформы Business Intelligence, позволяющие визуализировать и анализировать Big data, что говорит о высоком потенциале использования Big data. Данная работа позволит определить, можно ли создавать новые сервисы и продукты, оптимизировать бизнес, повышать точность прогнозирования, следовательно, достигать нужного экономического эффекта с помощью технологии Big data, сможет ли данная технология усовершенствовать процессы принятия решений и повысить конкурентоспособность на мировом рынке, и дает ли это реальные стратегические преимущества от накопленной информации.

Выбранное направление исследования весьма актуально в настоящих условиях в связи с тем, что темпы роста объемов информации с каждым годом возрастают и для финансового рынка очень важно правильно использовать имеющиеся данные в целях развития бизнеса. Важнейшим фактором является то, что финансовая система приобретает новые качества, становится более сложной и динамичной.

Тема диссертации соответствует приоритетным направлениям цифровизации экономики в Республике Беларусь. Результаты исследования были использованы в банковской сфере в части построения модели проведения качественных маркетинговых кампаний, модели выявления клиентов с максимальным риском отказа от продукта, модели сегментации клиентов по предпочтениям.

Целью диссертационного исследования является определение проблем и перспектив использования технологии Big data в финансовом секторе Республики Беларусь, а также разработка рекомендаций по практическому применению данной технологии в выбранной области. Объектом исследования является финансовый сектор Республики Беларусь. Предметом исследования является технология Big data и особенности ее применения в финансовом секторе.

Цель работы обусловила постановку и решение следующих задач:

- 1) структурировать знания в сфере развития технологии Big data и возможности ее применения в финансовом секторе Республики Беларусь.
- 2) выработать авторский подход к понятию «Big data», формирующий методологическую основу для его анализа и оценки;
- 3) проанализировать сферы финансового сектора, в которых возможно применение технологии Big data;
- 4) выявить основные проблемы и задачи Big data;
- 5) проанализировать использование Data mining в финансовом секторе;

- б) проанализировать развитие технологии Big data в финансовом секторе Республики Беларусь;
- 7) изучить правовые аспекты применения Big data в финансовом секторе;
- 8) выработать рекомендации по практическому применению технологии Big data в финансовом секторе.

Основой при написании диссертации послужили публикации в различных источниках финансового сектора, статистические данные компаний IDC Enerprice, Accenture, Capgemini, Insights, Евразийского экономического союза, исследований мирового банка, а также внутренние данные одного из банков Республики Беларусь.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования Целью диссертационного исследования является определение проблем и перспектив использования технологии Big data в финансовом секторе Республики Беларусь, а также разработка рекомендаций по практическому применению данной технологии в выбранной области. Объектом исследования является финансовый сектор Республики Беларусь. Предметом исследования является технология Big data и особенности ее применения в финансовом секторе.

Цель работы обусловила постановку и решение следующих задач:

- 1) структурировать знания в сфере развития технологии Big data и возможности ее применения в финансовом секторе Республики Беларусь.
- 2) выработать авторский подход к понятию «Big data», формирующий методологическую основу для его анализа и оценки;
- 3) проанализировать сферы финансового сектора, в которых возможно применение технологии Big data;
- 4) выявить основные проблемы и задачи Big data;
- 5) проанализировать использование Data mining в финансовом секторе;

- б) проанализировать развитие технологии Big data в финансовом секторе Республики Беларусь;
- 7) изучить правовые аспекты применения Big data в финансовом секторе;
- 8) выработать рекомендации по практическому применению технологии Big data в финансовом секторе.

Новизна полученных результатов Внедрение результатов исследования было осуществлено при разработке рекомендаций по использованию технологии Big data в финансовом секторе Республики Беларусь. Результаты исследования были использованы в банковской сфере в части использования Big data клиентской базы и построения модели проведения качественных маркетинговых кампаний, модели выявления клиентов с максимальным риском отказа от продукта, модели сегментации клиентов по предпочтениям.

Положения, выносимые на защиту.

1. Теоретические основы технологии Big data и эффективности ее использования в финансовом секторе Республики Беларусь, включая:

а) авторская позиция в части определения и понятий Big data и Data Mining и их взаимосвязь. Big data — это различные инструменты, подходы и методы для обработки структурированных и неструктурированных данных огромных объемов и их значительного многообразия, и взаимосвязей с целью их использования для решения конкретных задач бизнеса.

Data Mining - это компьютерный процесс для преобразования больших объемов данных в информацию и знания путем нахождения шаблонов в данных с использованием различных методов визуализации, уменьшения размерности, классификации и построения моделей. Data Mining - это междисциплинарная дисциплина, охватывающая сочетание дисциплин в области статистики, искусственного интеллекта и науки, и информационных систем управления для распознавания образов, математического моделирования и работы с базами данных.

Иными словами, разобрав понятия Big data и Data Mining можно сказать, что термин Big data определяет что обрабатываем (наборы данных, их связи и т.д.), а термин Data Mining - это процесс сортировки данных, чтобы найти что-то стоящее, определяет как мы обрабатываем (методы, инструменты и т.д.)

б) синергетический эффект использования технологии Big data и машинного обучения. Машинное обучение (Machine learning) тесно связано с наукой о данных. Это техника искусственного обучения, которую применяют для сбора Big data. По-простому - это возможность обучить систему или алгоритм получать различные представления из массива, аналитика данных. При машинном обучении для построения модели прогнозирования целевых переменных используется некий первичный набор знаний. Машинное обучение применимо к различным типам сложных систем: от регрессионных моделей и метода опорных векторов до нейронных сетей. Здесь центром является компьютер, который обучается распознавать и прогнозировать.

Примеры алгоритмов:

- Модели, которые могут прогнозировать поведение пользователя.
- Классификационные модели, которые могут узнать и отфильтровать спам.
- Рекомендательные системы — изучают предпочтения пользователя и пытаются угадать, что ему может понадобиться.
- Нейронные сети — не только распознают образы, но и сами могут их создавать.

С увеличением объема анализируемых данных, модифицировались и системы аналитики. С помощью машинного обучения можно строить сложные модели, с учётом всех возможных влияний, что повышает точность предсказаний и дает конкурентное преимущество.

в) выявлены перспективные направления финансового сектора, которые необходимо развивать с помощью технологии Big data:

- Мобильный банкинг

Для поддержания конкурентоспособности в век цифровых технологий банки обязаны развивать онлайн—сервисы. Культура пользователей изменяется. Если еще лет 5-7 назад смартфоны были роскошью, то сегодня они есть у каждого подростка. Люди ориентируются на технологии, т.к. они во многом упрощают жизнь.

Одновременно развитие онлайн—сервисов способно оптимизировать и процессы самого банка. Все это огромный массив обрабатываемых цифровых данных клиентов, которые необходимо анализировать и обрабатывать в доли секунды, что непосредственно и предполагает использование технологий Big data.

- Искусственный интеллект

Big data и искусственный интеллект помогают лучше узнать, понять клиента и предложить то, что ему действительно нужно.

Примеры технологий финансовой структуры, построенные на Big data:

- Биометрическая идентификация.
- Идентификация по голосу.
- Роботы в отделениях банков.
- Отделения без сотрудников.

Стоит отметить, что в Республике Беларусь внедрение искусственного интеллекта находит широкое применение.

- Кредитный скоринг

Робот с помощью машинного обучения способен на следующие действия:

- Умеет учиться; обучается на данных клиента.
- Самостоятельно создает модели в течение минут.
- Показывает результаты лучше, чем логистическая регрессия.
- Работает с сырыми данными.
- Не требует привлечения специалиста по данным и дорогого ПО.

г) выявлены основные проблемы при внедрении технологии Big data:

- Вопросы безопасности (проблемы защиты данных и обеспечения их целостности; риск постороннего вмешательства и утечки конфиденциальной информации; ненадлежащее хранение конфиденциальной информации; риск нецелевого использования)
- Ограниченный бюджет
- Нехватка персонала
- Сложности интеграции
- Ограниченное число поставщиков
- Предприятие не готово к внедрению новой технологии

д) определены задачи технологии Big data (поступление, хранение и обработка данных, организация неструктурированных данных, анализ и визуализация данных)

2. Анализ развития технологии Big data в финансовом секторе Республики Беларусь (уровень цифровизации финансового сектора; выявление факторов, стимулирующих цифровизацию (осознание ее важности высшим руководством, создание нормативно-правовой базы, деятельность ПВТ); определение ограничивающих факторов (недостаточная разработанность законодательства и отсутствие системной программы цифровизации страны, высокая стоимость технологий, консерватизм и непонимание роли цифровых технологий в трансформационных процессах чиновниками, бизнесменами и населением, отсутствие достаточного количества квалифицированных кадров и недооценка роли развития цифровых компетенций среди населения).

3. Правовые аспекты использования технологии Big data в финансовом секторе Республики Беларусь (GDPR, проект закона о персональных данных в Республике Беларусь, использование Cookies)

4. Использование технологии Big data в маркетинговых кампаниях

5. Модель для проведения маркетинговых кампаний, основанная на проблеме качественного проведения рекламной кампании, и адресности предложения. Для решения данной задачи необходимо проанализировать данные по уже проведенным рекламным кампаниям и их результатах, что позволит обучить модель по выявлению пула клиентов с наибольшей долей вероятности отклика на предложение.

6. Модель для выявления клиентов с максимальным риском отказа от продукта, основанная на потребности предприятий финансового сектора удержать существующего клиента, поскольку привлечение нового клиента обходится от семи до десяти раз дороже, чем удержание уже существующего. Поэтому компаниям необходимо выявить в своей клиентской базе ценных клиентов с наибольшим риском отказа от продукта, для дальнейшего проведения целевой кампании по их удержанию.

7. Методика кластеризации клиентской базы. Основная идея кластеризации — определить группы клиентов без использования предварительных гипотез о структуре клиентской базы, найти натуральные кластеры среди свойств клиентов исходя из имеющихся данных. Преимущество этого подхода следующее- кластеризация покажет разрез поведения и свойств клиентов, на который ранее не обращали внимания. Так же лучше будут понятны взаимосвязи и взаимоотношения между разными клиентами, что в дальнейшем поможет повысить транзакционную активность, повысить долю безналичных платежей, свести партнеров компании с различными сферами деятельности и клиентов.

Личный вклад соискателя ученой степени. Диссертационная работа является целостным и законченным научным исследованием, выполненным автором самостоятельно на основе изучения и обобщения обширного теоретического и эмпирического материала, апробации собственных разработок в практике. Основные положения, содержащиеся в

диссертационном исследовании и выносимые на защиту, разработаны соискателем лично.

Апробация результатов диссертации Результаты исследований, включенные в магистерскую диссертацию были представлены на V Международной научно-практической конференции «BIG DATA and Advanced Analytics.BIG DATA и анализ высокого уровня», проводимой Министерством образования Республики Беларусь, учреждением образования «Белорусский государственный университет информатики и радиоэлектроники», Национальной академией наук Беларуси, Объединенным институтом проблем информатики, DHTechnologies & Data Nubes (Austin, USA), Univercity of Technology (Kielce, Poland), BEZNext (Chicago, USA), Invisi BV (Netherlands), IBM (NY, USA), ИООО «ЭПАМ Системз» в г. Минске , Республике Беларусь, 13-14 марта 2019 года, а также на Международном научном форуме «Образование. Наука. Культура», проводимом 20 ноября 2019 года Министерством науки и высшего образования Российской Федерации и Гжельским государственным университетом.

Магистерская диссертация выполнена самостоятельно, проверена в системе «Антиплагиат». Процент оригинальности соответствует норме не менее 70%, установленной кафедрой экономики. Цитирования обозначены ссылками на публикации, указанные в списке использованных источников.

Опубликованность результатов исследования Результаты работы опубликованы в сборнике материалов V Международной научно-практической конференции «BIG DATA and Advanced Analytics.BIG DATA и анализ высокого уровня», а также в сборнике материалов Международного научного форума «Образование. Наука. Культура».

Структура и объем диссертации Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка литературы, включающего 51 источник, приложения. Работа изложена на 95 страницах.

Объем, занимаемый 35 рисунками, 2 таблицами, приложением, составляет 18 страниц.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе дается авторская позиция в части определения и понятий Big data и Data Mining и их взаимосвязь (рисунок 1).



Рисунок 1- Взаимосвязь понятий Big data и Data Mining

В данной главе систематизирована теоретическая информация о технологии Big data, рассмотрен синергетический эффект использования технологии Big data и машинного обучения, рассмотрена модификация систем аналитики (рисунок 2), а также определены перспективные направления финансового сектора- мобильный банкинг, кредитный скоринг, идентификация клиентов- которые необходимо развивать с помощью технологии Big data и проблемы, которые могут возникнуть при внедрении (рисунок 3):

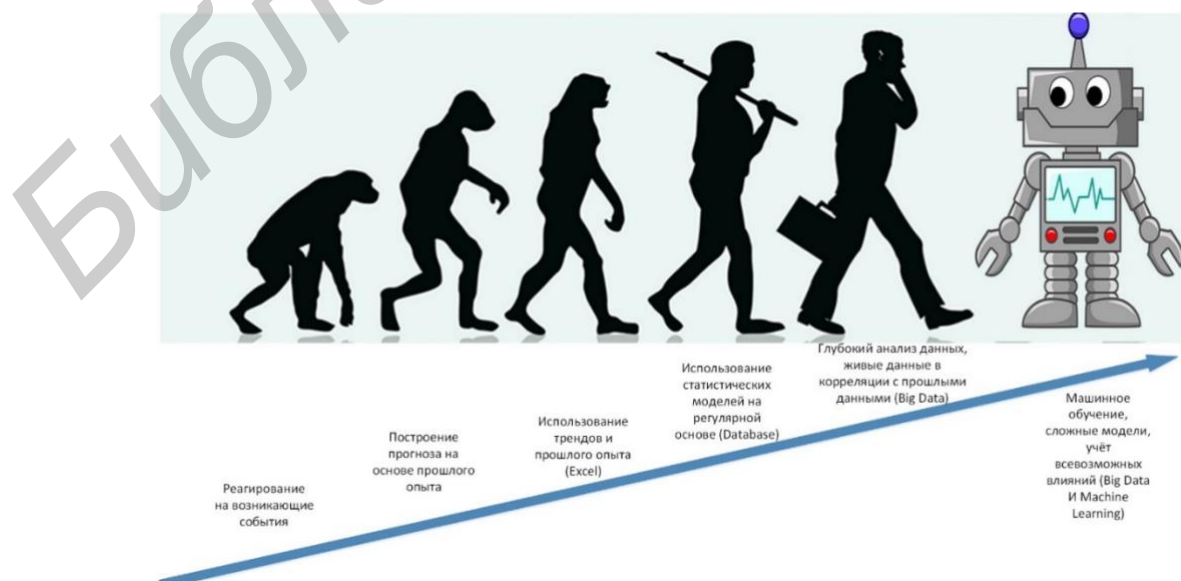


Рисунок 2- Эволюция систем аналитики



Рисунок 3- Основные проблемы при внедрении проектов Big data

В главе 1 также определены основные задачи, решаемые с использованием технологии Big data, определен принцип работы технологии Big data (рисунок 4)

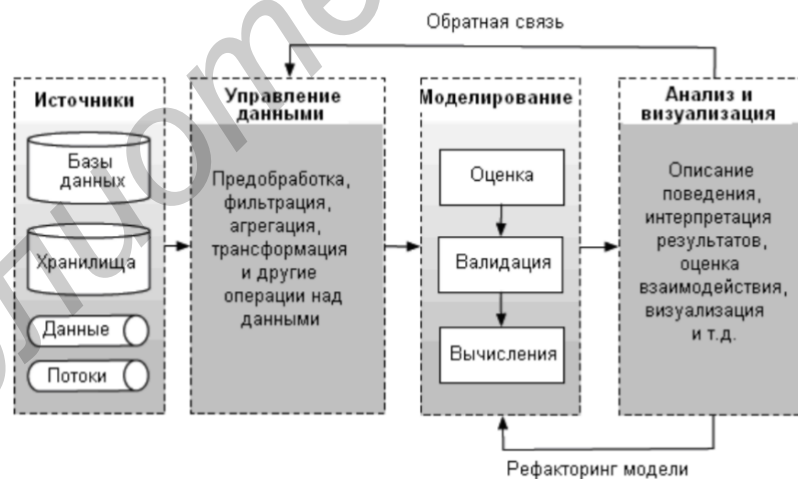


Рисунок 4- Принцип работы с Big data

Рассмотрено Data mining как средство визуализации данных, проведен сравнительный анализ основных Business Intelligence систем.

А также отмечается высокая важность возможности внедрения технологии Big data в финансовый сектор Республики Беларусь, поскольку использование данной технологии дает огромные возможности.

В главе 2 проведен анализ развития технологий Big data в финансовом секторе Республики Беларусь, определен уровень цифровизации экономики страны, и финансового сектора в частности (рисунок 5).

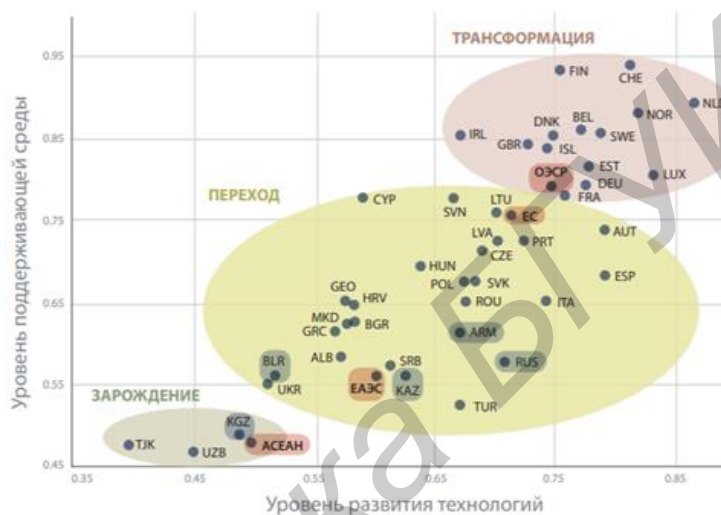


Рисунок 5- Развитие цифровой экономики в странах Европы и Центральной Азии

Отмеченные факторы дают четкое понимание того, что уровень использования технологии Big data в Республике Беларусь достаточно низкий.

В главе 2 также рассмотрены правовые аспекты применения технологии Big data в финансовом секторе Республики Беларусь: рассмотрены основные принципы GDPR, проекта закона о персональных данных в Республике Беларусь, особенности работы с файлами cookies.

В главе 3 приведены рекомендации по практическому применению технологии Big data в маркетинговых кампаниях. Финансовые структуры обладают уникальными маркетинговыми данными- где люди проводят время, что покупают, какой средний чек их покупок, частота покупок и прочее - это основной ресурс информации для банков – клиентские данные, и здесь важность несет любая информация о клиенте (рисунок 6)



Рисунок 6- Клиентские данные

С использованием технологии Big data предлагается 3 модели:

1. Модель для проведения маркетинговых кампаний
2. Модель для выявления клиентов с максимальным риском отказа от продукта
3. Методика кластеризации клиентской базы

С помощью модели 1, построенной на основе Big data компании, предоставляется возможным решить следующие бизнес- задачи:

- Повышение эффективности проведения рекламных кампаний, придание рекламной кампании адресного характера;
- Составление рейтинга клиентов по вероятности отклика на предлагаемое рекламное предложение продукта;
- Отбор целевых групп для проведения кампаний.

Бизнес- задачи, которые можно решить построением модели 2:

- Выявление групп клиентов с максимальным риском отказа от продукта
- Определение факторов, влияющих на отказ
- Выбор клиентов для проведения кампаний по удержанию

С целью повышения транзакционной активности, а также для понимания с какими партнерами заключать договора, предлагается методика кластеризации клиентской базы, с помощью которой все клиенты будут разделены на сегменты для принятия целевых действий. Основная цель

сегментации клиентской базы – предложение клиентам точечных адресных предложений.

Таким образом, использование Big data и машинного обучения, дают возможность решения следующего минимума задач в финансовом секторе (рисунок 7):



Рисунок 7- Задачи, решаемые применением технологии Big data

Оценка эффективности предложенных мероприятий доказывает, что в банковской сфере выиграют игроки, которые смогут построить устойчивую и лояльную клиентскую базу, и в этом несомненную роль играет использование технологии Big data.

ЗАКЛЮЧЕНИЕ

К наиболее значимым результатам диссертационного исследования можно отнести следующие:

1. определены наиболее перспективные направления финансового сектора, в которых внедрение технологии Big Data будет наиболее эффективно;
2. определены основные проблемы, возникающие при использовании технологии Big data в финансовом секторе Республики Беларусь;
3. проведен анализ развития технологии Big data в финансовом секторе;

4. определен уровень цифровизации финансового сектора Республики Беларусь;
5. рассмотрены правовые аспекты применения технологии Big data в части исполнения закона GDPR и Проекта закона о персональных данных в Республике Беларусь;
6. разработаны практические рекомендации по применению технологии Big data в финансовом секторе, позволяющие проводить эффективные маркетинговые кампании предприятий финансового сектора, которые обеспечат стабильность финансово-экономического развития предприятий.

Практическая значимость результатов исследования заключается в том, что сформулированные в работе рекомендации по использованию технологии Big data, предназначенные для возможности построения модели проведения качественных маркетинговых кампаний, модели выявления клиентов с максимальным риском отказа от продукта, модели сегментации клиентов по предпочтениям, могут быть использованы для эффективной оценки и анализа клиентской базы предприятиями финансового сектора Республики Беларусь.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Алиш, М. С. Анализ использования технологии Big data в финансовом секторе Республики Беларусь: проблемы и перспективы / М. С. Алиш // Сб. материалов V Международной научно-практической конференции «BIG DATA and Advanced Analytics. BIG DATA и анализ высокого уровня»/ БГУИР; редкол.: В. А. Богуш [и др.]. — Минск, 2019. — С. 148–166.
2. Алиш, М. С. Анализ использования технологии Big data в финансовом секторе Республики Беларусь: проблемы и перспективы / М. С. Алиш // Сб. материалов Международном научном форуме «Образование. Наука. Культура» / ГГУ; редкол.: Н. Н. Уварова [и др.]. — Гжель, 2019.