МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «КубГУ»)
Экономический факультет
Кафедра экономики и управления инновационными системами

КУРСОВАЯ РАБОТА

ПРИМЕНЕНИЕ BIG DATA В УПРАВЛЕНИИ ЭКОНОМИЧЕСКИМИ СИСТЕМАМИ

Работу выполнил _____ Г.Р. Лесков
(подпись)
Направление подготовки 27.03.03. Системный анализ и управление
Направленность (профиль) Интеллектуальная бизнес-аналитика и управление экономическими процессами
Научный руководитель
канд. экон. наук, доц. _____ Г.Н. Библя
(подпись)
Нормоконтролер
канд. экон. наук, доц. _____ Г.Н. Библя
(подпись)

Краснодар
2025

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

В настоящее время технологии Big Data и искусственный интеллект занимают важное место в информационных технологиях. Спектр их применения, действительно, очень широк и затрагивает многие сферы деятельности людей, к примеру бизнес, медицину и промышленность. Большие данные и их анализ позволяют эффективно использовать ресурсы, находя оптимальные пути решения тех или иных проблем, например проблемы логистики и ее огромной транспортной цепочки с целью оптимизации данного процесса.

Целью курсовой работы является исследование технологий Big Data и машинного обучения для улучшения учёта товарных остатков и разработки методов их интеграции в бизнес-процессы компании ООО «ЛЕ МОНЛИД».

В соответствии с указанной целью необходимо решить следующие задачи:

1) изучить теоретические основы Big Data и их роль в экономике,
2) проанализировать состояние и выявить тенденции современного российского рынка Big Data,
3) провести характеристику и анализ деятельности ООО «ЛЕ МОНЛИД» как объекта исследования,
4) разработать концептуальное предложеие по внедрению технологий Big Data и искусственного интеллекта для оптимизации бизнес-процессов компании.

Объектом исследования данной работы является процесс управления системой продаж ООО «ЛЕ МОНЛИД», а в частности используемых ими технологий больших данных в управлении логистикой.

Предметом исследования являются количественные и качественные методы и инструменты анализа больших данных, а также организационно-управленческие аспекты их внедрения в систему управления продажами.

Теоретическую основу работы составили исследования отечественных и зарубежных ученых и практиков в области системного анализа, информационных технологий и управления: Юрия Ильича Черняка, Кеннета Боулдинга, Юнира Абдулловича Урманцева, Александра Александровича Богданова Виктора Майер-Шёнбергера, Кеннета Кука и других.

Методологическая база исследования включает в себя систематизацию данных, классификацию, сравнительный анализ, а также методы структурирования и обобщения информации.

Структура курсовой работы определяется логикой исследования и включает введение, три главы, заключение и список использованных источников.

Во введении обоснована актуальность работы, поставлена цель и задачи, объект и предмет данной работы.

В первой главе раскрываются теоретические аспекты Big Data и их характеристика.

Во второй главе представлен анализ объекта исследования — компании ООО «ЛЕ МОНЛИД», ее организационной структуры и текущих бизнес-процессов.

В третьей главе предложена концепция внедрения технологий Big Data и искусственного интеллекта для оптимизации деятельности предприятия.

В заключении сформулирован основной вывод работы.

1 Теоретическая основа применения Big Data в экономике

1.1 Понятие Big Data: определение, история, характеристики

Большие данные в нашей современной жизни играют ключевую роль в качестве инструмента, управляющего сотни тысячами процессов в различных системах и позволяющему анализировать те или иные процессы. Исследования, связанные с большими данными, изучают способы выделения знаний из этих данных. Они ведутся в рамках огромного количества сфер и областей, таких как информационные науки, экономические науки, машинное обучение и т.д.

Сам же термин «Big Data» используется для обозначения огромных объемов информации, которые настолько велики, разнообразны, и поступают буквально каждую секунду, что их невозможно хранить и обрабатывать привычными нам методами и способами. Вместе с тем, Big Data – это не только метод хранения и обработки информации, это еще и способ получения новых знаний, методов, практик и исследований на основе тех объемов данных, которые мы обрабатываем в процессе исследования. Это новый подход к работе с данными: их сбор, хранение, переработка, интерпретация и извлечение ценности.

Big Data — это возможность анализировать огромные массивы информации и преобразовывать их в форму, доступную для поиска, обработки и извлечения ценной информации.

Алекс Горелик описывает Big Data как данные, которые не могут быть эффективно обработаны традиционными системами из-за их объема, разнообразия и скорости поступления. (Корпоративное озеро данных, стр 28-35).

Чтобы подробнее разобраться с Big Data, далеко не лишним будет изучить историю взаимодействия людьми с данными, не важно, большими или малыми. С древних времен наши предки сталкивались и решали проблемы с хранением и обработкой информации.

Еще 10-20 тысяч лет назад прародители современного человека использовали наскальную живопись, кости от остатков собранных запасов для записи какой-либо информации на них. Последние же использовались, предположительно, в качестве метода ведения торговой деятельности и возможности иметь прогнозируемый остаток на нужды собственного пропитания. Если смотреть позднее – появление Вавилонских библиотек в 2000-х годах до нашей эры. Эти методы, в свою очередь, использовались для сохранения знаний и мудростей во время набегов врага. Как неудачный пример – захваченная римлянами Александрия утратила большую часть хранящейся в ее библиотеке литературы.

Говорить про анализ данных в те времена, конечно, бессмысленно. Какие-либо отголоски, похожие на современный анализ данных, не наблюдались вплоть до середины 17 века, когда Джон Грант предположил теорию, использующую аналитику смертности для предсказания начала эпидемии бубонной чумы.

Время шло, количество хранимых человечеством данных росло – с ним и росла вероятность столкновения с проблемами, связанными с хранением и обработкой людьми данных. Так, в 1880 году, перед переписью населения, американское бюро столкнулось с трудностями и объявило, что с текущим подходом к работе с данными они смогут провести перепись лишь за 8 лет, а при следующей переписи в 1890 – дать результаты удастся не раньше чем через 10 лет. Нетрудно догадаться, что в то время результаты уже устареют и мало кто назовет подобную статистику актуальной и полезной в рамках того времени.

Во время Второй мировой войны у человечества возникла нужда в быстром анализе данных, поступающих от разведки. Таким образом в 1943 году британцами была создана машина „Colossus”, которая ускоряла расшифровку сообщений в разы: ранее для этого требовалось вплоть до нескольких недель, а после создания машины – несколько часов.

С появлением Интернета возник вопрос поиска по многочисленной информации, размещающейся на центральных хранилищах данных. Для решения этой проблемы IT-компании начали разрабатывать собственные поисковые системы. Одной из таких была „AltaVista”, уникальность которой заключалась в использовании лингвистического алгоритма, разбивающего заданную фразу на слова и проводя поиски по существующим индексам для ранжирования результата. В течение двух лет количество запросов возрасло с 300 тысяч до 80 миллионов.

Дуг Лейни (Doug Laney), работник Meta Group, 6 февраля 2001 года выпустил документ, который описывает основные проблемы, связанные с повышенными требованиями к хранилищам данных в те времена ввиду бурного роста e-commerce (электронная торговля). В нем он выделил три основных направления, на которые стоит обратить внимание для решения вопросов управления данными: **Volume** (объем), **Velocity** (скорость) и **Variety** (разнообразие). Годами позже эти понятия стали основой для описания модели больших данных, закрепив за собой название 3V (VVV). Погрузимся глубже в суть концепции и обсудим каждое из направлений.

Volume – характеристика больших данных, описывающая хранение огромных массивов информации, которые в силу своего размера не помещаются в привычные нам системы хранения данных и требуют для этого специализированные решения, такие как распределенные файловые системы и облачные хранилища. Алекс Горелик в своей книге ”Корпоративное озеро данных: новый подход к использованию Big Data и Data Science в бизнесе” подчеркивает, что:

- объем данных растет экспоненциально,
- традиционные подходы не справляются с такими масштабами.

Velocity – харакретистика больших данных, описывающая скорость, с которой данные создаются, обрабатываются в системе и изменяются (при необходимости). Современные хранилища данных с их системами создают

огромные потоки событий, логов, транзакций и других данных, которые требуют немедленной обработки для принятия своевременных решений.

Variety - характеристика больших данных, описывающая разнообразие источников и типов данных. С появлением Big Data анализ данных перестал ограничиваться только лишь табличными данными, ведь источниками данных теперь становится буквально все существующие типы передачи информации: текстовые документы, фотографии, видеозаписи, аудиозаписи и даже геолокационные данные. Том Фоусет и Фостер Провост в „Data Science для бизнеса. Как собирать, анализировать и использовать данные" подчеркивают, что Variety делает аналитику более мощной, но требует других методов, таких как машинное обучение, работы с графами, потоками и т.д.

Позже к классической 3V начали добавлять другие характеристики, образующие новые наборы признаков. Например:

– 4V: Volume, Velocity, Variety и Veracity,

– 5V: Volume, Velocity, Variety, Viability и Value,

– 7V: Volume, Velocity, Variety, Viability, Value, Variability и Visualization

Хоть базовыми характеристиками Больших Данных до сих пор являются Volume, Velocity и Variety, предложенные годами позже дополнительные характеристики по значимости не уступают первоначальным.

Veracity – дополнительная характеристика больших данных, которая определяет достоверность, качество и надежность данных, используемых для анализа. Не каждый набор больших данных может быть действительно качественным и надежным. При всех преимуществах больших данных присутствует шанс того, что данные могут быть ошибочными, неполными, искаженными или вовсе содержать шум. Подобные данные будут просто непригодными для дальнейшего анализа и получения из них пригодных знаний.

Viability – дополнительная характеристика больших данных, описывающая способность данных быть полезными и применимыми в

решении разнообразных задач. Основной смысл этой характеристики заключается в том, что не все большие данные могут быть полезными и применимыми в решении некоторых задач. Грубо говоря, она отсеивает лишний информационный шум в больших данных, оставляя место только действительно нужным и полезным данным для последующих обработки и анализа.

Variability – дополнительная характеристика больших данных, которая отражает изменчивость и нестабильность данных во времени, контексте или пространстве. Данные могут поступать регулярно и стабильно, но отличаться по содержанию, структуре и смыслу в зависимости от времени или условий.

Visualization – дополнительная характеристика больших данных, описывающая способность описать данные в наглядной, простой и интуитивной форме для последующего анализа и принятия решений. Алекс Горелик в своей книге „Корпоративное озеро данных: новый подход к использованию Big Data и Data Science в бизнесе” комментирует эту характеристику следующим образом: данные часто остаются „сырыми” и только аналитика, визуализация и отчеты превращают их в полезный актив

В этом подразделе были рассмотрены основные определения Big Data, как научные, так и прикладные, их характеристики и историю их зарождения. В следующем параграфе будут описаны инструменты и методы обработки и анализа больших данных.

1.2 Инструменты и методы обработки и анализа больших данных

Большие данные в виде огромных массивов быстроизменяющейся информации хранятся в датацентрах по всему миру, прямо влияя на нашу жизнь. Но ведь одного хранения данных недостаточно. Как говорил великий физик-теоретик Альберт Эйнштейн: „Информация в чистом виде – это не знание. Настоящий источник знания – это опыт”. Как получать опыт из знаний? Как получать знания из огромного количества ежесекундно меняющихся данных, называемых Big Data? Этот параграф будет полностью посвящен инструментам анализа больших данных и их методам обработки.

Обычно работа с большими данными в компаниях проходит в три этапа:

– Интеграция,
– Управление,
– Проведение анализа

На этапе интеграции организация проводит внедрение инструментов информационных технология для накопления, хранения и обновления больших данных. Пожалуй, это самый важный этап, ведь без него невозможно приступить к последующим этапам.

На этапе управления решается вопрос хранения больших данных. Как было рассмотрено в прошлом параграфе, хранить большие данные классическими методами затрудняется ввиду неподготовленности технической составляющей используемого оборудования. Для этого компании зачастую используют услуги «хостингеров» дата-центров для выделения на их мощностях определенного сектора, который компания может использовать для работы с большими данными. Как правило, дата-центры выделяют организациям выделенные сервера, но иногда организации выбирают виртуальные сервера, которые создаются на основе уже выделенного.

На этапе анализа организация привлекает в процесс специально обученных специалистов – Big Data аналитиков, и уже они применяют методы обработки, анализа и извлечения из больших данных действительно полезной информации. Ведь большие данные сами по себе в „сыром” виде не имеют какой-либо ценности до того момента, как с ними не начнут работать Big Data аналитики.

Рассматривая инструменты для работы с Big Data для начала нужно разобраться с тем как данные распределяются. Для распределения больших данных было придумано достаточно много моделей, но прародителями и основными из них по сей день считаются две: „MapReduce” и „Hadoop”. Чтобы разобраться в них более подробнее окунемся в их историю, принцип работы, и виды реализации.

Модель „MapReduce” была предложена компанией Google в 2005 году для распределения больших данных, чтобы решить проблему с индексацией веб-страниц, возникшую в одноименной поисковой системе компании. Принцип работы заключается в разделении процесса распределения данных на три этапа: Map (разбиение данных), Shuffle (перемешка данных) и Reduce (агрегация данных).

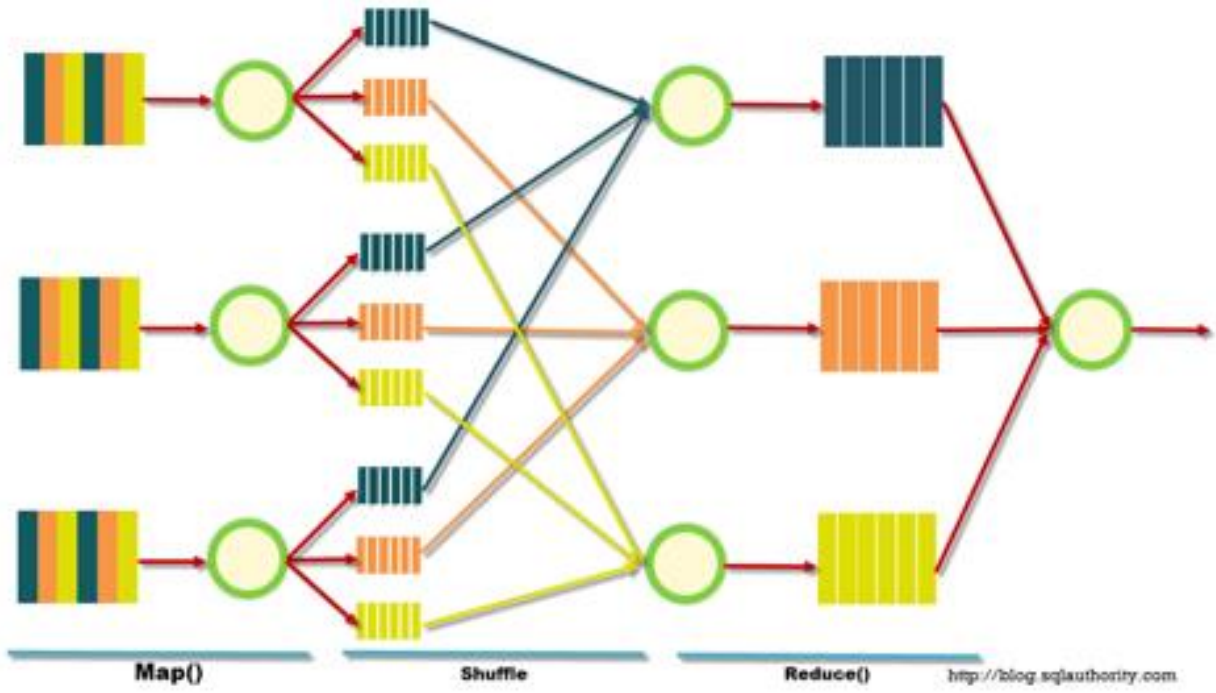

Рисунок 1 – Схема работы модели MapReduce

На первом этапе входные данные преобразуются при помощи функции map(), которую задает сам пользователь, а функция по итогу выдает множество пар ключ-значений. Это значит, что функция может выдать только одну запись, несколько записей, а может не выдать ничего.

На втором этапе вывод функции map() как бы сортируется в соответствии с ключами, присвоенными значениям вывода функции. Все данные, которые были сгенерированы в формате пар ключ-значение, сортируются по ключу и группируются по одинаковым ключам.

На третьем, завершающем этапе, происходит агрегация данных, полученных в результате работы второго этапа. То есть для каждой группы одинаковых ключей пар значений вызывается какая-либо функция (например, суммирование, объединение или подсчет).

Годом позже, в 2006, Дугласом Куттингом и Майклом Дж. Коннором на основе модели MapReduce был разработан фреймворк Hadoop как часть проекта Apache. По сути этот фреймворк является во много раз улучшенной версией модели MapReduce от Google, о которой было расписано ранее. Разработанный на Java, проект решал задачи в рамках парадигмы MapReduce, когда приложение делится на большое количество элементарных заданий, которые решаются на распределенных компьютерах кластера и сводятся к единому результату.

Если говорить о методах обработки больших данных, а не об инструментах, то из них можно подчеркнуть как machine learning (машинное обучение, искусственный интеллект), так и Data Mining (метод, используемый для „добычи” из больших данных действительно полезной информации). Рассмотрим оба метода обработки больших данных подробнее.

Уже далеко не секрет, что искусственный интеллект занял далеко не последнее место во взаимодействии с человечеством и влиянии на нашу современную жизнь. Но как именно он связан с большими данными? На самом деле, очень тесно, ведь именно благодаря большим данным появилась возможность тренировать большие интеллектуальные модели, которые по сути и являются искусственным интеллектом в его прямом виде. Искусственный интеллект требует огромных объемов данных для обучения. Большие данные предоставляют именно те объемы, которые необходимы для создания и обучения моделей машинного обучения и искусственного интеллекта.

Искусственный интеллект, в свою очередь, помогает анализировать большие объемы данных, автоматизируя процессы, которые раньше требовали значительных ресурсов и времени. Например, AI может помочь анализировать данные о поведении клиентов и прогнозировать их предпочтения.

Говоря о методах обработки больших данных невозможно упустить и следующее понятие, тесно связанное с моделированием и прогнозированием данных – Data Mining.

Data Mining (извлечение данных) - это процесс анализа больших объёмов данных с целью выявления скрытых закономерностей, зависимостей и паттернов, которые могут быть полезными для принятия решений в различных областях. (Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking).

В этом параграфе были рассмотрены основные инструменты и методы работы с Big Data. В следующем параграфе будут описаны роль и место Big Data в управлении экономическими системами.

1.3 Роль и место Big Data в управлении экономическими системами

Big Data имеет широкий спектр применения: от машинного обучения до управления экономическими процессами. Как раз на последнее идет особый упор, ведь с помощью больших данных и систем аналитики можно оптимизировать многие экономические процессы, сократив на это время и ресурсы.

По сути своей, основной метод применения Big Data в экономике заключается в сборе информации компании о своих клиентах, сделках и всех основных процессах, происходящих при взаимодействии организации с внешними процессами. Ее использование позволит принимать обоснованные решения. Разберем на примере конкретных ситуаций.

Рисунок 2 – Степень влияния больших данных на экономику

Использование больших данных в маркетинге и рекламе позволяет компаниям анализировать данных о покупках, поисковых запросах, социальных сетях и других источниках, чтобы лучше понимать потребительские предпочтения. Учит их создавать точные маркетинговые кампании.

Иной пример – таргетированная реклама: на основе данных потребителей можно создавать персонализированные рекламные кампании.

Прогнозирование и аналитика тоже не обошли стороной работу с большими данными – с их помощью появляется возможность создавать модели для прогнозирования экономических тенденций и изменений на рынке. В этом случае большие данные позволяют анализировать и прогнозировать те или иные экономические процессы с целью снижения потенциальных рисков.

Большие объемы данных могут быть сгенерированы и собраны с помощью различных устройств IoT (интернет вещей), таких как датчики, умные дома, автомобили и другие, и использоваться для оптимизации бизнес-процессов, производственной деятельности, логистики и других аспектов экономической деятельности.

Также большие данные могут помочь в сфере логистики, например, оптимизировать цепь поставок, транспортную логистику, что может снизить затраты и улучшить эффективность доставки товаров и услуг.

Использование Big Data в экономике открывает большие возможности для улучшения бизнес-процессов, принятия взвешенных решений и повышения продуктивности организаций. Благодаря современным технологиям, таким как Hadoop, компании способны обрабатывать и анализировать значительные объёмы данных, обнаруживать скрытые закономерности и тенденции, а также предсказывать будущие изменения и события.

1.4 Текущее состояние и роль Big Data в России

Большие данные весьма стремительно, семимильными шагами врывается в нашу с вами жизнь. Буквально во все научные сферы и сферы общественной жизни так или иначе забрались большие данные и их анализ: физика, биология, статистика, политика и др.

Но в данной работе нас, конечно, больше всего интересует именно экономика. Мы разберемся, какую роль большие данные занимают в нашей отечественной экономике, экономике всей страны.

Все современные организации в той или иной степени собирают и постоянно используют информацию о состоянии собственной деятельности, состоянии отечественного и мирового рынка в целом, поведение конкурентов и клиентов в целом. Объем данной информации стал расти в геометрической прогрессии с внедрением информационных технологий и в наше время эти данные уже можно называть большими, ведь их настолько много, они достаточно быстро меняется и огромное количество полезной информации. Сами же алгоритмы больших данных помогают превратить гигантский объем информации в весьма ценный ресурс, который, точно ограненный алмаз, с каждым этапом обработки становится все полезнее и привлекательнее для бизнеса.

Согласно исследованию ИСИЭЗ НИУ ВШЭ (Института Статистических Исследований и Экономики Знаний НИУ ВШЭ) в 2021 году технологии добычи, сбора и анализа информации больших данных применяли только 25,8% организаций.

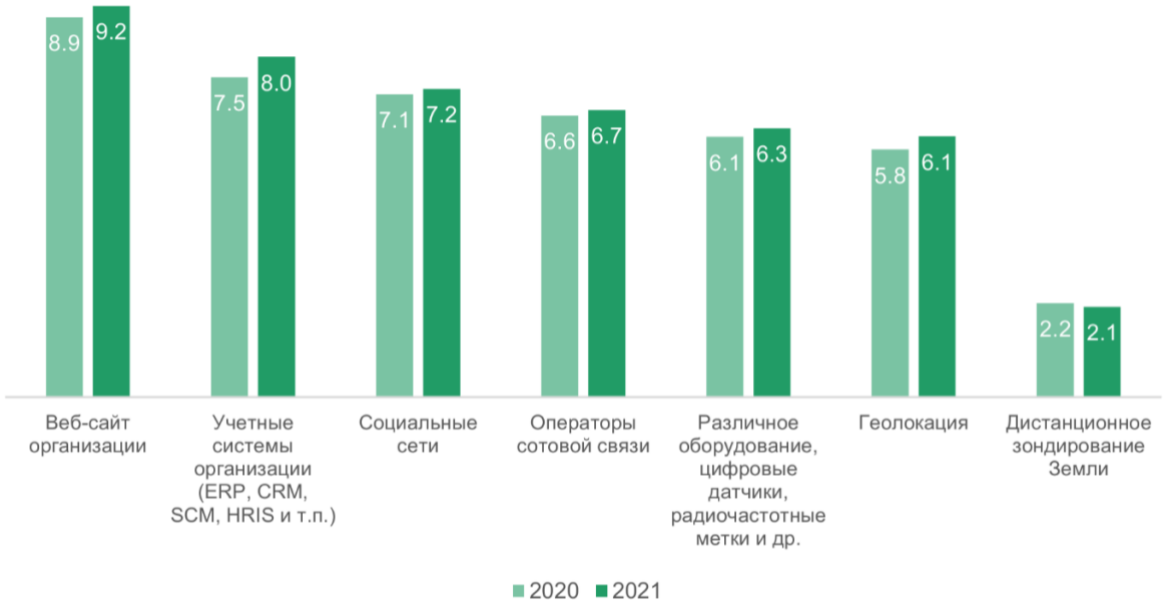

Рисунок 3 – Статистика источников больших данных в организациях, в % к общему числу организаций

По популярности применения и использования больших данных в той или иной сфере, среди организаций почти всегда преимущественно находится маркетинг и продажи: почти 60% компаний задействуют их в организации продаж и проведении маркетинга. Это связано с тем, что большие данные позволяют компаниям глубже понять поведение и предпочтения клиентов, оптимизировать рекламные кампании, персонализировать предложения и эффективно управлять клиентской базой, что в итоге приводит к увеличению выручки и конкурентоспособности. Четверть компаний использует их напрямую в производственном процессе, а каждая пятая организация использует большие данные для обеспечения безопасности как в организации, так и на ее предприятиях, например, посредством дистанционного зондирования Земли.

Рисунок 4 – Направления использования технологий сбора, обработки и анализа больших данных, в % к общему числу организаций

В заключении по текущему состоянию Big Data на российском рынке хотелось бы сделать некие выводы, а именно:

– стремительное внедрение в экономику РФ больших данных, особенно в сферы маркетинга и продаж (60% компаний),

– самым используемым источником данных до сих пор остаются собственные веб-сайты организаций (9,2% организаций)

– социальные сети (7,2% организаций) и учетные системы (8% организаций) на данный момент остаются менее популярными источниками данных

– большие данные также применяются в производстве (25%) и сфере безопасности (20%), включая технологии дистанционного зондирования

2 Анализ и характеристика объекта исследования

2.1 Характеристика предприятия

В данной курсовой работе в качестве объекта исследования была выбрана компания, не имеющая в своем пользовании технологии Big Data. Так, ООО «ЛЕ МОНЛИД» (ранее известная как “Леруа Мерлен”) является крупнейшей сетью розничных магазинов в России в сегменте товаров для строительства, ремонта и обустройства дома. Рассмотрим данную организацию поближе, чтобы разобраться кому мы предлагаем внедрить технологии Big Data для управления экономическими и бизнес процессами.

Общество с ограниченной ответственностью ООО «ЛЕ МОНЛИД» занимается следующими видами деятельности (по ОКВЭД): Торговля розничная мебелью, осветительными приборами и прочими бытовыми изделиями в специализированных магазинах (47.59), Производство готовых текстильных изделий, кроме одежды (13.92), Производство деревянных рам для картин, фотографий, зеркал или аналогичных предметов и прочих изделий из дерева (16.29.14), Сбор и обработка сточных вод (37.00). Компания осуществляет свою деятельность с 2003 года. Среднесписочная численность работников в компании по состоянию на 2024 год превышает 45 тыс. человек.

Уставный капитал ООО «ЛЕ МОНЛИД» составляет 5.98 млрд. рублей. По состоянию на конец 2024 года, выручка компании составила 456.632 млрд рублей, что почти в полтора раза больше в сравнении с отчетностью на конец 2023 года – 347.163 млрд рублей. Что касается чистой прибыли – по состоянию на конец 2024 года компания имеет чистой прибыли в размере 35.15 млрд рублей, что почти в 8 раз больше в сравнении результатов финансовой отчетности по этому же показателю по состоянию на конец 2023 года – 4.395 млрд рублей.

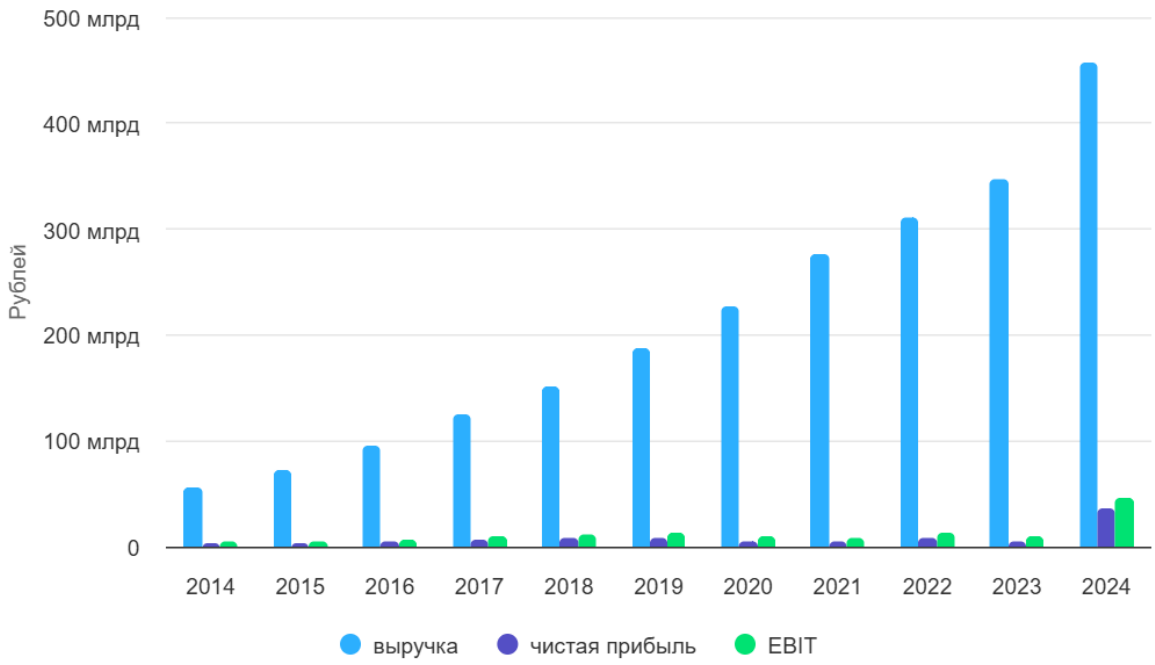

Рисунок 5 – Динамика основных показателей финансовой отчетности ООО «ЛЕ МОНЛИД»

Также хочется обратить внимание на еще один финансовый показатель – EBIT. Его размер по состоянию на конец 2024 года составлял 46.636 млрд рублей, что в сравнении с результатами конца 2023 года (10.348 млрд рублей) почти в 4 раза меньше.

Исследовав материалы сети интернет, удалось выявить нескольких компаний-конкурентов ООО «ЛЕ МОНЛИД»: ООО “ОБИ ФЦ”, ООО "МАКСИДОМ" и ООО "БАУЦЕНТР РУС". Построим табличку:

Таблица 1 – Сравнение ООО «ЛЕ МОНЛИД» с конкурентами

Организация	ООО «ЛЕ МОНЛИД»	ООО “ОБИ”	ООО “МАКСИДОМ”	ООО "БАУЦЕНТР”
Объем выручки	Высокий	Низкий	Средний	Средний
Целевая аудитория	Частные клиенты	Частные, малые бизнесы	Частные клиенты	Частные, малые бизнесы
Цены	Доступные	Доступные	Доступные	Доступные
Узнаваемость бренда	Высокая	Средняя	Средняя	Средняя

Отличительной особенностью ООО «ЛЕ МОНЛИД» является высокая узнаваемость бренда в связи с международной популярностью организации, а также полный упор на аудиторию частных клиентов, что делает компанию “народной” в глазах потребителя и с большей вероятностью при выборе он отдаст предпочтение именно этой организации. Обратив внимание на уровень выручки в сравнении с объемом выручки компаний-конкурентов хочется подметить, что ООО «ЛЕ МОНЛИД» оторвалась от них в значительной степени и именно поэтому занимает лидирующие позиции в рейтингах лучших сетей магазинов строительных материалов и фурнитуры для дома.

Для дальнейшего анализа организации и выявления проблем требуется разобраться и понять организационно-управленческую структуру предприятия. Но прежде чем это сделать нам требуется выяснить что это такое и в чем заключается ее цель.

Организация – это набор стабильных социально-структурированных отношений, созданных преднамеренно, с ясной целью решения конкретных подзадач и задач. Ее основной чертой является структурированность.

Управление – целенаправленная способность некоторых специфических задач.

Составим блок-схему, описывающую организационно-управленческую структуру ООО «ЛЕ МОНЛИД»

Рисунок 7 – Организационно-управленческая структура ООО «ЛЕ МОНЛИД»

Разберемся подробнее в том, чем занимается каждый директор и каждый отдел в организации

Генеральный директор – руководит процессами всей компании, принимает стратегические решения и разрабатывает долгосрочные планы развития.

Директор по закупкам и логистике – отвечает за закупку товаров, доставку товаров в магазины, оптимизацию логической цепочки организации.

Директор по продажам и маркетингу – ответственен за создание и воплощение стратегии продаж, за маркетинговые идеи.

Финансовый директор – руководит всеми финансовыми аспектами бизнеса.

Директор по персоналу – отвечает за подбор персонала в организацию.

Далее рассмотрим по отделам, за которые отвечают директора:

Отдел закупок – отвечает за закупку товаров, работу с поставщиками и сроки поставок.

Отдел логистики – отвечает за управление транспортировкой товара и его хранение во время логистики.

Отдел продаж – отвечает за выполнение планов продаж, оптимизацию работы точек организации.

Маркетинговый отдел – отвечает за разработку рекламных кампаний, исследует рынок и продвигает продукцию.

Финансово-бухгалтерский отдел – ведет бухгалтерию, производит учет финансовых операций в организации и готовит отчетность.

HR-отдел – подбирает персонал и адаптирует его в организации.

Отдел корпоративной культуры – поддерживает корпоративные ценности в организации, организует мероприятия для сотрудников.

2.2 Анализ существующей IT-инфраструктуры

С целью эффективного управления бизнес-процессами и обеспечения высокого качества обслуживания клиентов ООО «ЛЕ МОНЛИД» использует комплексную и распределённую IT-инфраструктуру. Она включает в себя разнообразные системы и технологии, обеспечивающие сбор, хранение, передачу и обработку больших объёмов данных, поступающих из множества торговых точек, складов и онлайн-платформ.

Рассмотрим ключевые компоненты этой инфраструктуры, которые формируют основу обработки данных и поддерживают IT-процессы компании

То, без чего не может существовать буквально любая IT-платформа, занимающаяся в той или иной степени обработкой и хранением больших данных – это корпоративное озеро данных (Data Lake).

В качестве корпоративного озера данных в ООО «ЛЕ МОНЛИД» используется платформа Greenplum – масштабируемое решение на основе PostgreSQL, позволяющее хранить и обрабатывать огромные массивы структурированных данных. Она служит централизованных хранилищем данных из разных источников: магазинов, складов и интернет-платформы. С его помощью выполняются функции агрегирования и подготовки больших массивов данных для дальнейшего анализа и построения отчетности.

Но ведь данные бывают не только чистыми, структурированными и понятные – в „руки” системы также нередко попадают неструктурированные и полуструктурированные данные. В работе с ними в компании используется MongoDB – документо-ориентированная NoSQL база данных, используемая для быстрого доступа и работе с разнообразной информацией о товарах, характеристиках и логистике. Высокая скорость доступа позволяет быстро получать данные для машинного обучения и оперативного принятия решений.

Для высокоскоростной обработки транзакций в ООО «ЛЕ МОНЛИД» используется Tarantool – резидентая СУДБ с поддержкой технологий ACID и SQL. Эта база данных, помимо высокоскоростной обработки, обеспечивает хранение результатов аналитики и машинного обучения, доступных для конечных приложений и мобильных клиентов. Эта система особенно полезна в сценариях, требующих низкой задержки и высокой производительности на уровне распределенных систем.

Бизнес-логика, включая алгоритмы и модели машинного обучения в ООО «ЛЕ МОНЛИД» развертываются в специальных контейнерах Docker дя изоляции и упрощения сопровождения. На помощь в управлении контейнерами приходит Kubernetes – технология, позволяющая управлять масштабированием и отказоустойчивостью отдельных модулей и приложений, что позволяет компании оперативно адаптировать ИТ-систему под изменяющиеся нагрузки и требования.

За интеграцию данных и их обмен в компании отвечает Apache Kafka – распределенная платформа обмена сообщениями. Kafka играет роль централизованного канала обмена данными между разными системами и приложениями в компании, также именуется „корпоративной шиной” (ESB). То есть это система, буквально соединяющая все подсистемы основной системы воедино.

Так как компания расширяется на рынке, она также предоставляет возможности заказа товаров онлайн через собственный интернет-магазин. Рассмотрим как именно работает система интернет-магазина, погрузившись в подробности того, откуда именно компания берет товары, заказанные клиентами через сайт, и отправляет их впоследствии клиенту.

Компания реализует сложный мультиканальный подход к учету и управлению товарными запасами, что включает несколько различных типов торговых точек и каналов. В структуру бизнеса входят несколько торговых точек и каналов продаж: крупные офлайн-магазины, склады и дарксторы. Разберемся в этой структуре подробнее:

Офлайн-магазины – крупные супермаркеты, в которых покупатели могут лично ознакомиться с ассортиментом и сразу же приобрести понравившийся товар. Эти магазины играют ключевую роль, так как именно из них формируется основная часть заказов интернет-магазина.

Склады – логистические центры, которые предназначены для хранения больших объемов товаров, которые используются для пополнения магазинов или формирования заказов с длительным сроком доставки.

Дарскторы (dark stores) – специальные склады, ориентированные исключительно на сбор интернет-заказов, недоступные для обычных покупателей. Это позволяет оптимизировать доставку и скорость обработки заказов.

Особенностью бизнес-процесса является то, что около 98% заказов, сделанных через интернет-магазин, собираются именно из офлайн-магазинов, а не со складов или даркстороров. Это накладывает серьезные требования к точности данных о наличии товара в каждом магазине и синхронизации информации между разными системами. В следующем параграфе займемся анализом проблем с существующей ИТ-инфраструктурой и постановкой целей для их решения.

2.3 Анализ проблем и постановка цели

Проведя характеристику предприятия и детальный анализ ИТ-инфраструктуры ООО «ЛЕ МОНЛИД», ее ключевых компонентов и бизнес-процессов, мы переходим к определению основных задач и направлений развития данной инфраструктуры. Для этого нам необходимо ответить на следующие вопросы: какие проблемы существуют в текущей ИТ-инфраструктуре компании? Каковы приоритеты и вектор дальнейшего развития систем? Какие внутренние ограничения и слабые места требуют устранения? Какие внешние факторы оказывают влияние на эффективность и устойчивость ИТ-инфраструктуры?

Для системного понимания данных аспектов целесообразно провести SWOT-анализ.

SWOT-анализ (англ. SWOT analysis) — это метод стратегического планирования, используемый для оценки текущей ситуации компании. Аббревиатура SWOT означает: Strengths (сильные стороны), Weaknesses (слабые стороны), Opportunities (возможности) и Threats (угрозы).

SWOT-анализ является одним из ключевых инструментов стратегического планирования, который помогает компаниям всесторонне оценить своё текущее положение на рынке и определить оптимальные направления развития. Благодаря систематическому изучению сильных и слабых сторон внутри организации, а также возможностей и угроз со стороны внешней среды, SWOT-анализ позволяет сформировать целостное представление о внутренних ресурсах и внешних факторах, влияющих на деятельность компании.

Рисунок 8 – SWOT-анализ ООО «ЛЕ МОНЛИД»

Исходя из SWOT-анализа можно сделать несколько выводов:

– компания обладает современной и гибкой архитектурой Big Data с эффективной потоковой обработкой данных. Использование контейнеризации и оркестрации обеспечивает масштабируемость и отказоустойчивость системы. Наличие инструментов автоматизации способствует упрощению и ускорению рабочих процессов,

– основной вызов — сложность интеграции разнородных систем, что приводит к задержкам и несогласованности данных. Кроме того, отсутствует полная автоматизация контроля качества данных, что увеличивает риск ошибок и замедляет процесс их исправления,

– компания имеет хорошие перспективы для внедрения методов искусственного интеллекта, что позволит повысить качество и точность данных.

В результате комплексного анализа ИТ-инфраструктуры и бизнес-процессов ООО «ЛЕ МОНЛИД» выявлены ключевые проблемы, связанные с точностью поступаемых данных о продажах и скоростью их обработки.

В частности эти недочеты затрагивают систему учета товарных запасов. Анализ показал, что сложность интеграции множества систем, отсутствие полноценной автоматизации контроля качества данных и крайне высокие требования к синхронизации данных создают существенные риски расхождения в информации. В сочетании с ростом онлайн-продаж и ожиданиями клиентов данные недочеты приводят к распространенной проблеме несобранных заказов и снижению уровня обслуживания, что подтверждается исследованиями в области ритейла и практическими кейсами компании. В связи с этим требуется поставить цели для решения вышеизложенных проблем. В этом нам помогут метод целеполагания „дерево целей” и метод оценки поставленных целей – КОВ анализ

Рисунок 9 – Метод “дерево целей” для ООО «ЛЕ МОНЛИД»

Таблица 2 – Первый уровень анализа квадрата разности

Подцели	Эксперты					Сумма рангов	d	d^2
	1	2	3	4	5			
G.1 Повысить точность данных о наличии товара для клиентов и сборщиков заказов до конца 2025 года	3	5	4	3	3	18	2	4
G.2 Оптимизировать процессы обработки и управления заказами до конца 2025 года	5	3	4	5	5	22	2	4
	8	8	8	8	8	40		8

Таблица 3 – Первый уровень КОВ анализа

Подцели	Эксперты					Сумма	КОВ
	1	2	3	4	5		
G.1 Повысить точность данных о наличии товара для клиентов и сборщиков заказов до конца 2025 года	3	5	4	3	3	18	0,45
G.2 Оптимизировать процессы обработки и управления заказами до конца 2025 года	5	3	4	5	5	22	0,55
	8	8	8	8	8	40	1,00

Далее перейдем ко второму уровню ковариационного анализа. На нем важно не только понимать общие тренды и зависимости между переменными, но и глубже анализировать взаимодействия между различными факторами, а также влияние этих факторов друг на друга.

Таблица 4 – Второй уровень анализа квадрата разности

Подцели	Эксперты					Сумма рангов	d	d^2
	1	2	3	4	5			
G 1.1 Внедрить систему машинного обучения	1	5	5	5	1	17	1,33	1,78
G 1.2 Обеспечить актуализацию данных в режиме реального времени	5	1	2	1	5	14	1,33	1,78
G 1.3 Разработать интерфейс обратной связи	5	5	4	5	5	24	-5,67	32,11
	11	11	11	11	11	55		35,67

Подцели	Эксперты					Сумма	КОВ	Нормированный КОВ
	1	2	3	4	5			
G 1.1 Внедрить систему машинного обучения	1	5	5	5	1	17	0,31	0,14
G 1.2 Обеспечить актуализацию данных в режиме реального времени	5	1	2	1	5	14	0,25	0,11
G 1.3 Разработать интерфейс обратной связи	5	5	4	5	5	24	0,44	0,20
	11	11	11	11	11	55	1,00	

Таблица 5 – Второй уровень КОВ анализа

Подцели	Эксперты					Сумма рангов	d	d^2
	1	2	3	4	5			
G 2.1 Автоматизировать контроль и мониторинг состояния заказов	2	4	4	4	2	16	4,00	16,00
G 2.2 Снизить задержки и ошибки в процессе комплектования заказов	5	3	3	4	5	20	4,00	16,00
G 2.3 Внедрить систему уведомлений и предупреждений	5	5	5	4	5	24	-4,00	16,00
	12	12	12	12	12	60		48,00

Подцели	Эксперты					Сумма	КОВ	Нормированный КОВ
	1	2	3	4	5			
G 2.1 Автоматизировать контроль и мониторинг состояния заказов	2	4	4	4	2	16	0,27	0,15
G 2.2 Снизить задержки и ошибки в процессе комплектования заказов	5	3	3	4	5	20	0,33	0,18
G 2.3 Внедрить систему уведомлений и предупреждений	5	5	5	4	5	24	0,40	0,22
	12	12	12	12	12	60	1,00	

Построив дерево целей и проведя КОВ-анализ для ООО «ЛЕ МОНЛИД», можно сделать следующие выводы:

По основным целям (G1 и G2):

– Оптимизация процессов обработки и управления заказами (G2) получила наивысший коэффициент относительной важности (КОВ – 0.55), что свидетельствует о её ключевой роли в проекте

– Повышение точности данных о наличии товара (G1) занимает значимое место в структуре приоритетов проекта с коэффициентом относительной важности (КОВ) 0.45. Это подчёркивает, что достоверность и своевременность информации о товарных запасах являются фундаментальными условиями для эффективного управления процессом обработки заказов.

По подцелям G1:

– Разработка интерфейса обратной связи для сборщиков заказов и клиентов (G1.3) получила самый высокий КОВ (0.20), что подчёркивает важность прозрачной коммуникации для оперативного реагирования на проблемы с наличием товара.

– Внедрение системы машинного обучения для обнаружения аномалий (G1.1) занимает второе место по приоритету (КОВ – 0.14), что свидетельствует о значительном потенциале ИИ.

– Обеспечение актуализации данных в режиме реального времени (G1.2) имеет несколько меньший вес (нормированный КОВ – 0.11), но также остаётся важной задачей для поддержки синхронизации информации.

По подцелям G2:

– Внедрение системы уведомлений и предупреждений о проблемах с заказами (G2.3) является наиболее приоритетной подцелью (нормированный КОВ – 0.22), так как позволяет быстро реагировать на сбои и минимизировать последствия.

– Снижение задержек и ошибок в комплектации заказов за счёт улучшения логистики и координации (G2.2) занимает второе место по

важности (нормированный КОВ – 0.18), влияя напрямую на скорость и точность выполнения заказов.

При проведении анализов бизнец-процессов нам также удалось поработать совместно с командой ООО «ЛЕ МОНЛИД», вследствие чего благодаря совместным усилиям была выявлена главная проблема, напрямую связанная с бизнес-процессами и ИТ-инфраструктурой предприятия – проблема оперативной инвентаризации. Каждую минуту через интернет-магазин в компанию поступает огромное количество заказов на те или иные позиции. Зачастую проблема связана напрямую с наличием позиций на складах компании. Пройдемся по этой проблеме чуть подробнее.

Проблема инвентаризации в наше время актуальна для любого предприятия. В ООО «ЛЕ МОНЛИД» она особенно усугублялась тем, что помимо крупных оффлайн-точек компании, у них также есть склады и так называемые дарк-сторы. Заказ из интернет-магазина может собираться как на складе или в дарк-сторе, так и в крупном оффлайн-магазине. На практике 98% заказов собираются из товаров в торговых залах оффлайн-магазинов. Немаловажным в этой ситуации является человеческий фактор – быстро найти товар из 50 тысяч позиций на территории 10000 квадратных метров не всегда получается. А уж тем более не всегда получается провести грамотную инвентаризацию всех позиций, находящихся в оффлайн точках. В связи с этим в глобальной системе учета позиций и наличия их на складах часто оказываются неверные данные по наличию того или иного товара в магазине.

Во время анализа существующих данных инвентаризации было выяснено, что расхождение между реальным и фактическим количеством товара возникает по следующим причинам:

– Выставочные образцы (недоступные для продажи) при инвентаризации зачастую указывают как доступные к продаже. Однако на самом деле гарантия на подобные позиции не распространяется вовсе, магазин не может их продать, поэтому клиент не должен иметь возможность заказать такой товар, доступный к продаже на сайте.

– Обратная ситуация в виде огромного количества товара, указанного в виде витринного образца недоступного для продажи. По факту, все эти позиции являются реальными товарами, хранящимися на складах с распространенной на них гарантией, а интернет-магазин в данной ситуации попросту не может дать возможности клиенту заказать товар, указывая его отсутствие на складе в связи с тем, что он является „витринным образцом”.

Рисунок 10 – Пример ошибок в данных о товарных остатках компании

На финальных этапах анализа было выяснено, что для решения данной проблемы компании требуется автоматизировать контроль товарных остатков и разработать систему, связанную с обнаружением подобных аномалий и расхождений между фактическими и реальными значениями в отчетах по наличию товара на складах. Решение этой проблемы значительно снизит количество несобранных заказов, повысит удовлетворенность клиентов, а также поможет оптимизировать внутренние бизнес-процессы организации, такие как инвентаризация, управление запасами и логистика.

3 Решение проблемы и внедрение проекта

3.1 Предложение концептуального решения

Проанализировав текущую обстановку на рынке, проведя характеристику предприятия и выявив его “болевые точки” настало время решать текущую проблему предприятия, предлагая концепт проекта, способного устранить актуальную для компании проблему – неточный учет данных по наличию товара на складах компании.

Опираясь на выводы, сделанные по итогу анализа предприятия, представляется следующий проект: “Внедрение системы автоматизированного учёта товарных остатков с использованием технологий Big Data и ИИ”. Что предполагает проект?

Для решения проблемы неточности данных о товарных остатках, которая негативно сказывается на точности отображения информации на сайте и в системах ООО «ЛЕ МОНЛИД», было принято решение внедрить модель машинного обучения, способную автоматически обнаруживать аномалии в данных о товарных остатках и оперативно их корректировать. Учитывая большое количество магазинов и широкое разнообразие товаров (в общей сложности более 40 000 наименований), традиционные методы обработки данных с использованием линейных алгоритмов не обеспечивают необходимой точности и скорости. Поэтому был выбран подход, основанный на применении алгоритмов машинного обучения, который способен эффективно обрабатывать большие объёмы данных и выявлять даже самые сложные аномалии в учёте товаров.

Подробно рассмотрим наше концептуальное решение, которое будет включать внедрение модели машинного обучения для автоматического обнаружения аномалий в данных о товарных остатках, интеграцию потоковых данных для синхронизации информации и автоматизацию процессов контроля и валидации данных, что обеспечит точность и оперативность обновлений.

Основой решения нашей проблемы будет являться метод градиентного бустинга на деревьях решений, реализованный на библиотеке CatBoost, которая была разработана Яндексом в далеком 2017 году. Этот алгоритм является одним из наиболее эффективных инструментов для работы с табличными данными, особенно когда данные содержат как числовые, так и категориальные признаки. Градиентный бустинг позволяет строить ансамбли деревьев решений, где каждое новое дерево корректирует ошибки предыдущих, что делает алгоритм мощным инструментом для выявления аномалий в сложных и многомерных данных.

Для обучения модели будут использованы данные, полученные из нескольких источников:

– Данные о товарных движениях в магазинах и складах: поступления, продажи, возвраты, перемещения товаров между магазинами и складами.

– Инвентаризационные данные, полученные из ежедневных и ежегодных инвентаризаций, что позволяет точно отслеживать физическое наличие товаров в различных точках компании.

– Данные об отменённых заказах, которые помогают выявить несоответствия между доступностью товаров в системах и фактическим наличием.

Модель использует более 70 признаков (предикторов), которые охватывают широкий спектр данных о товарных движениях, продажах, возвратах, а также характеристиках товаров, таких как категория, бренд, сезонность и другие важные параметры. Помимо этого, в расчет также принимается информация о заказах, инвентаризации, а также о тенденциях в потребительских предпочтениях и изменениях в рыночной ситуации. Такое разнообразие признаков позволяет модели значительно повысить свою точность, эффективно выявляя аномалии даже в условиях многозадачности и большого объёма данных.

Для подготовки модели будет использован метод разделения данных на две выборки: обучающую (80%) и тестовую (20%). Обучающая выборка служит для построения модели, а тестовая — для её проверки на новых данных, которые модель не использовала в процессе обучения. Этот подход позволяет убедиться, что модель не тренируется под какие-либо конкретные и заготовленные данные и может работать с абсолютно новыми.

Для повышения точности и качества работы модели будет использоваться метод кросс-валидации. Этот подход позволяет более надежно оценить производительность алгоритма, минимизируя риск переобучения и обеспечивая стабильность результатов на различных подмножествах данных.

Таким образом, модель сможет лучше обобщать закономерности и демонстрировать устойчивую работу на новых, ранее не встречавшихся данных.

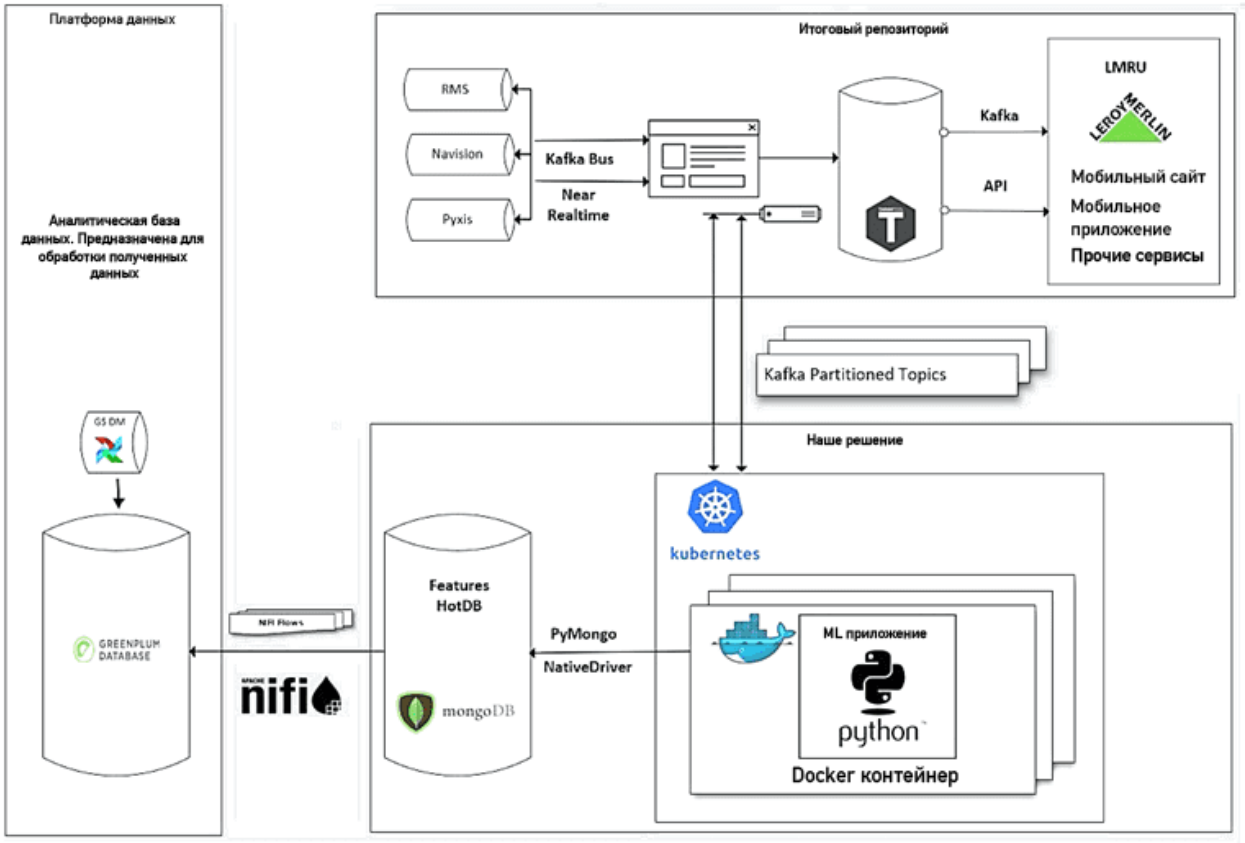

Рисунок 11 – Архитектура Big Data системы, предлагаемой в качестве концептуального решения

Для того чтобы интегрировать модель машинного обучения с текущими бизнес-процессами и системами компании, будет использована платформа Apache Kafka, которая обеспечивает потоковую передачу данных между различными компонентами системы в реальном времени. Kafka используется как корпоративная шина сообщений (Enterprise Service Bus, ESB), что позволяет синхронизировать данные о товарных остатках между различными подразделениями компании, такими как интернет-магазин, склады и магазины.

Для хранения и обработки данных в рамках системы будет использоваться Greenplum — распределённая СУБД, которая служит корпоративным хранилищем данных компании. Все данные о товарных остатках, продажах, возвратах и инвентаризациях будут храниться в этой базе, что обеспечит быстрый и удобный доступ к необходимой информации для анализа и принятия решений.

Результаты работы модели машинного обучения, включая выявленные аномалии и предсказания, будут храниться в Tarantool, который представляет собой высокоскоростную распределённую СУБД с поддержкой SQL-запросов и ACID-транзакций. Это позволяет оперативно хранить и извлекать результаты анализа для использования в дальнейшем, в том числе для отображения на сайте и в мобильных приложениях.

Далее перейдем к плану внедрения нашего концептуального решения с подробным описанием и сроками всех процессов, необходимых для реализации нашего проекта

3.2 План внедрения и использования ресурсов

Столь амбициозный проект мало просто описать, нужно также и продумать план и этапы внедрения, чем мы и займемся далее – планом по внедрению вышеупомянутых технологий и систем. Всем известно, что все новое таит в себе различного рода риски, однако за этими рисками скрывается целое поле возможностей для более эффективной работы бизнеса, поэтому переход на новую систему работы должен быть плавным и размеренным.

В этом нам поможет построение сетевого графика. Для начала потребуется план работ, которые будут произведены в ходе процесса внедрения планируемой нами архитектуры, а именно – PaaS (Platform-as-a-Service). После анализа подобных случаев на примерах других компаний получился следующий перечень задач:

– Провести анализ текущих данных о товарных остатках и выявить основные проблемы с точностью и актуальностью информации, которая поступает из магазинов и складов

– Определить, какие данные и процессы можно интегрировать с моделью машинного обучения, а какие останутся под контролем существующих систем.

– Подготовка плана обучения модели машинного обучения и создания набора данных для обучения, включая данные о продажах, возвратах и инвентаризации

– Определить роли и ответственности участников проекта, включая разработчиков, специалистов по машинному обучению, ИТ-инженеров, аналитиков данных и бизнес-аналитиков

– Провести тестирование модели машинного обучения с использованием исторических и новых данных

– Интегрировать модель машинного обучения с платформами для управления запасами и торговыми системами

Составим технологическую таблицу комплекса предстоящих работ:

Таблица 3 - Технологическая таблица работ

№ работы	Работа	Предшествующие работы	Длительность (дни)	Ресурс	События
A1	Подготовка анализа существующих данных	-	10	Технические ресурсы	0, 1
A2	Разработка модели ИИ	1	20	Технические ресурсы	1, 2
A3	Интеграция модели ИИ	1	15	Человеческие ресурсы	1, 3
A4	Настройка потоковой обработки данных с использованием Apache Kafka	2	12	Технические ресурсы	2, 4
A5	Разработка системы уведомлений и мониторинга данных	4	6	Технические ресурсы	3, 5
A6	Тестирование модели на пилотных точках	2, 3	7	Человеческие ресурсы	4, 6
A7	Разработка интерфейса для взаимодействия с клиентами	1, 2, 3, 4, 5	9	Материальные ресурсы	5, 6
A8	Масштабирование системы на все магазины и склады	6	14	Человеческие ресурсы	6, 7
A9	Оптимизация алгоритмов на основе полученных данных о товарных остатках	6	10	Технические ресурсы	7, 8
A10	Полный запуск системы	7	5	Финансовые ресурсы	8, 9

Следующим этапом станет создание сетевого графа, который отображает последовательность и зависимость работ, позволяет осуществлять базовый контроль над ходом работ проекта, их сроками и исполнением бюджета.

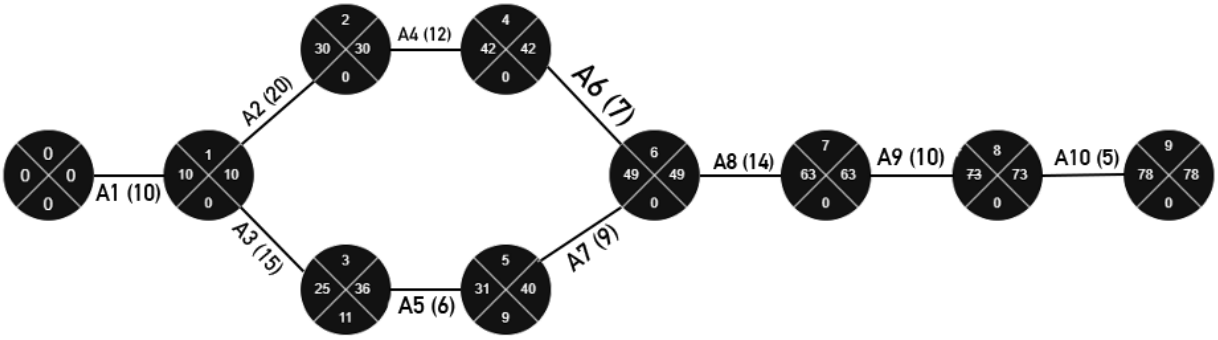

Рисунок 6 – Сетевой график работ

Критический путь - полный путь от исходного до завершающего события, имеющий наибольшую длину (продолжительность) из всех полных путей. Его временная длина определяет срок выполнения всех работ в сетевом графике. Критический путь всего мероприятия в данном случает составляет 78 дней.

В этом разделе мы предложили решение для проблемы с неточностью данных о товарных остатках в компании ООО «ЛЕ МОНЛИД». Мы решили использовать машинное обучение и Big Data, чтобы модель могла автоматически выявлять ошибки в данных и оперативно их исправлять.

Мы описали, как будет работать модель, какие данные она будет использовать и какие технологии будут задействованы для синхронизации данных. Важно, что интеграция с текущими системами компании через Apache Kafka и использование облачных решений должны помочь сделать систему более гибкой и эффективной.

Составление подробного плана внедрения и анализ всех этапов помогает нам чётко понимать, какие задачи нужно решить и какие ресурсы для этого понадобятся.

ЗАКЛЮЧЕНИЕ

В ходе написания курсовой работы мы изучили основные термины и характеристики технологий Big Data и управлением с их помощью экономическими процессами. На основе информации полученной из различных источников нам удалось провести анализ российского рынка в данной области, узнать его текущее состояние и тенденции развития.

Следующей нашей целью стала попытка внедрения технологий больших данных в одну из компаний, которая их не имеет. Так, в качестве объекта исследования выступило компания ООО «ЛЕ МОНЛИД», которая подходило под необходимые условия и имело острую актуальную проблему. Перед тем, как приступить к составлению плана работ по внедрению готовых решений в ООО «ЛЕ МОНЛИД» нами было принято решение провести исследование организационной и функциональной структуры компании, чтобы более детально ее изучить.

Заключительным этапом нашего исследования стало составление примерного план работ для реализации проекта, решающего проблему с аномальными и неверными данными, указанными в отчетах по наличию товаров на складах. Для этого мы представили концепт идеи, рассмотрели сроки реализации, составили технологическую таблицу для определения сроков реализации каждого этапа проекта, а также составили сетевой график для определения общих временных затрат. Все это послужит фундаментом для наших планируемых нами в будущем мероприятий.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Майер-Шёнбергер В., Кук К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / В. Майер-Шёнбергер, К. Кукьер. — М.: МИФ, 2014. — 256 с.

2. Горелик А. Корпоративное озеро данных: новый подход к использованию Big Data и Data Science в бизнесе / А. Горелик. — Москва: Эксмо, 2023. — 272 с.

3. Провост Ф., Фоусет Т. Data Science для бизнеса. Как собирать, анализировать и использовать данные / Ф. Провост, Т. Фоусет. — Себастопол: O'Reilly Media, Inc., 2013. — 414 с.

4. Журавлев А. Е. Большие данные. Big Data. Учебник для вузов : учебное пособие / Издательство ЛАНЬ. — Санкт-Петербург : БГУИР, 2022. — 188 с.

5. История больших данных (Big Data). Часть 1 // Компьютерра. — URL: https://www.computerra.ru/234239/istoriya-bolshih-dannyh-big-data-chast-1/ (дата обращения: 18.05.2025).

6. Большие данные (Big Data) // Википедия. — URL: https://ru.wikipedia.org/wiki/Большие_данные#VVV (дата обращения: 18.05.2025).

7. Большие данные: свойства, методы обработки и описание // Otus Journal. — URL: https://otus.ru/journal/bolshie-dannye-svojstva-metody-obrabotki-opisanie/ (дата обращения: 18.05.2025).

8. Методы Data Mining: обзор и классификация // НИУ ВШЭ. — URL: https://hsbi.hse.ru/articles/metody-data-mining-obzor-i-klassifikatsiya/ (дата обращения: 18.05.2025).

9. Big Data vs AI: взаимосвязь и различия // VC.ru. — URL: https://vc.ru/ai/1534003-big-data-vs-ai-vzaimosvyaz-i-razlichiya (дата обращения: 18.05.2025).

10. Большие данные в экономике // Decosystems.ru. — URL: https://www.decosystems.ru/big-data-v-ekonomike/ (дата обращения: 18.05.2025).

11. 5 сценариев применения Big Data в бизнесе // e-MBA.ru. — URL: https://e-mba.ru/knowledge-base/5-stsenariev-primeneniya-big-data-v-biznese (дата обращения: 18.05.2025).

12. Большие данные. Хабр // Habr.com. — URL: https://habr.com/ru/articles/267361/ (дата обращения: 18.05.2025).

13. Новости и исследования НИУ ВШЭ // Иссек НИУ ВШЭ. — URL: https://issek.hse.ru/news/776383019.html (дата обращения: 18.05.2025).