

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина»
Институт радиоэлектроники и информационных технологий – РТФ
Школа профессионального и академического образования

ДОПУСТИТЬ К ЗАЩИТЕ ПЕРЕД ГЭК

Директор ШПиАО
 Д.В. Денисов
(подпись) (Ф.И.О.)
« 03 » июня 2024 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

РАЗРАБОТКА ПРОТОТИПА СИСТЕМЫ УПРАВЛЕНИЯ В СФЕРЕ
ЛОГИСТИКИ И ТРАНСПОРТА, ОСНОВАННОЙ НА ТЕХНОЛОГИИ
ОБРАБОТКИ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ML

Научный руководитель: Коломыцева Анна Олеговна
к.э.н., доцент


подпись

Нормоконтролер: Огуренко Егор Владимирович


подпись

Студент группы: РИМ-220963 Иванов Иван Александрович


подпись

Екатеринбург
2024

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение высшего образования
**«Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина»**

Институт радиоэлектроники и информационных технологий – РТФ
Школа профессионального и академического образования
Направление подготовки 09.04.01 Информатика и вычислительная техника
Образовательная программа 09.04.01/33.03 Инженерия машинного обучения

ЗАДАНИЕ

на выполнение выпускной квалификационной работы

студента Иванова Ивана Александровича группы РИМ-220963
(фамилия, имя, отчество)

1. Тема выпускной квалификационной работы

Разработка прототипа системы управления в сфере логистики и транспорта, основанной на технологии обработки данных с использованием ML

Утверждена распоряжением по институту от «4» декабря 2023 г. № 33.02-05/298

2. Научный руководитель

Коломыцева Анна Олеговна, доцент, кандидат экономических наук, доцент
(Ф.И.О., должность, ученая степень, ученое звание)

3. Исходные данные к работе

Техническое задание. датасет

4. Перечень демонстрационных материалов

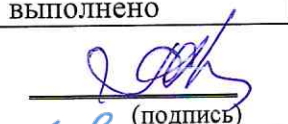
Презентация, архитектура прототипа системы, архитектура модели машинного обучения

5. Календарный план

№ п/п	Наименование этапов выполнения работы	Срок выполнения этапов работы	Отметка о выполнении
1.	<i>1 раздел (глава)</i>	до 23.03.2024 г.	выполнено
2.	<i>2 раздел (глава)</i>	до 29.04.2024 г.	выполнено
3.	<i>3–4 раздел (глава)</i>	до 20.05.2024 г.	выполнено
4.	<i>ВКР в целом</i>	до 24.05.2024 г.	выполнено

Научный руководитель Коломыцева Анна Олеговна
Ф.И.О.

Студент задание принял к исполнению 4.12.2023
дата


(подпись)

(подпись)

6. Допустить Иванова Ивана Александровича к защите выпускной квалификационной работы в экзаменационной комиссии

Директор ШПиАО


(подпись)

Д.В. Денисов
Ф.И.О.

РЕФЕРАТ

Выпускная квалификационная работа магистра 60 стр., 4 рис., 47 источников, 2 прил.

Ключевые слова: ETA, LTL, ML, ETA, АВТОМАТИЗАЦИЯ, ГРАФОВЫЕ НЕЙРОННЫЕ СЕТИ, МАШИННОЕ ОБУЧЕНИЕ, НЕЙРОННЫЕ СЕТИ, ТРАНСФОРМЕРЫ, ОНЛАЙН-БРОНИРОВАНИЕ, УПРАВЛЕНИЕ ЦЕПЯМИ ПОСТАВОК, ЦИФРОВАЯ ПЛАТФОРМА.

Цель работы – повышение уровня обоснованности принимаемых управленческих решений за счет разработки системы управления в сфере логистики и транспорта, базирующейся на комплексе моделей анализа данных и машинного обучения. Объектом исследования являются логистические процессы и процессы разработки прототипа системы управления в сфере логистики и транспорта, основанной на технологии обработки данных с использованием ML. Рассматриваются основные модели машинного обучения, позволяющие определять расчетное время прибытия груза. Для целей разработки системы разработана и имплементирована сложносоставная модель на основе графовых трансформеров. Оценена точность предсказания применяемой модели ML.

СОДЕРЖАНИЕ

РЕФЕРАТ	3
ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ.....	6
ВВЕДЕНИЕ	7
1 ТЕОРЕТИЧЕСКИЕ ОСНОВЫ РАЗРАБОТКИ ПРОТОТИПА СИСТЕМЫ УПРАВЛЕНИЯ В СФЕРЕ ЛОГИСТИКИ И ТРАНСПОРТА, ОСНОВАННОЙ НА ТЕХНОЛОГИИ ОБРАБОТКИ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ML.	11
1.1 Теоретический анализ методов и моделей для разработки приложений на основе технологии управления данными в сфере логистики и транспорта.....	11
1.2 Анализ предметной области и классификации алгоритмов ML для решения задач в сфере логистики и транспорта.....	13
1.2.1 Алгоритмы и методы, связанные с исследованием вопросов оценки времени прибытия (ETA).....	14
1.2.2 Графовые трансформеры	16
1.3 Преимущества и специфика внедрения перспективной модели машинного обучения в информационное приложение в сфере распределительной логистики	17
2 РАЗРАБОТКА МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ СИСТЕМЫ УПРАВЛЕНИЯ В СФЕРЕ ЛОГИСТИКИ И ТРАНСПОРТА.....	19
2.1 Практика и опыт создания ИТ-решений в сфере ИИ-систем управления цепочками поставок (AI Supply Chain Management)	19
2.1.1 Описание рынка ИТ-решений в сфере ИИ-систем управления цепочками поставок (AI Supply Chain Management)	19
2.1.2 Объем и емкость рынка ИТ-решений в сфере ИИ-систем управления цепочками поставок (AI Supply Chain Management).....	21
2.1.3 Анализ и систематизация ИТ-решений в сфере ИИ-систем управления цепочками поставок (AI Supply Chain Management).....	22

2.2 Формализация основных математических и аналитических зависимостей для ML-модели системы управления в сфере логистики и транспорта.....	25
2.2.1 Формальная постановка задачи по созданию модели.....	25
2.2.2. Инжиниринг модуля прогнозирования сроков доставки ETAKnow	26
3 ТЕСТИРОВАНИЕ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ ETAKNOW	38
3.1 Экспериментальные наборы данных.....	38
3.2 Описание данных и статистики	38
3.3 Методы и метрики для сравнения.....	38
3.4 Сравнение производительности	39
4 РАЗРАБОТКА ПРОТОТИПА СИСТЕМЫ УПРАВЛЕНИЯ В СФЕРЕ ЛОГИСТИКИ И ТРАНСПОРТА, ОСНОВАННОЙ НА ТЕХНОЛОГИИ ОБРАБОТКИ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ML И ЭКОНОМИЧЕСКАЯ ОЦЕНКА РЕАЛИЗАЦИИ ПРОЕКТА	42
4.1 Краткое описание создаваемого прототипа системы	42
4.2 Формализация архитектуры создаваемого цифрового решения	42
4.3 Основные планируемые функции и характеристики прототипа	46
4.4 Стек разработки создаваемой системы.....	47
4.5 Проблемы, которые решает создаваемая система управления.....	50
4.6 Экономическая оценка реализации проекта.....	51
ЗАКЛЮЧЕНИЕ.....	54
БИБЛИОГРАФИЧЕСКИЙ СПИСОК	56
ПРИЛОЖЕНИЕ А	61
ПРИЛОЖЕНИЕ Б	75

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

AI (Artificial Intelligence) – искусственный интеллект (ИИ)

ML (Machine Learning) – машинное обучение

OD (Origin-Destination) – исследование методов оценки времени доставки между начальной и конечной точкой

GNN (Graph Neural Network) – графовые нейронные сети

DNN (Deep Neural Network) – глубокие нейронные сети

WDR (Wide-Deep-Recurrent) – широкие-глубокие-рекуррентные сети

ConSTGAT (Contextual Spatio-Temporal Graph Attention Network) – графовые адаптивные трансформеры контекстуальной пространственно–временной внимательности

GAT (Graph Attention Networks) – сети внимания на графах

Learning Representations – обучение представлениям

GCN (Graph Convolutional Network) – графовая сверточная сеть

CNN (Convolutional Neural Network) – сверточная нейронная сеть

GRU (Gated Recurrent Unit) – управляемые рекуррентные блоки

RNN (Recurrent Neural Network) – рекуррентная нейронная сеть

LSTM (Long Short-Term Memory) – нейронная сеть долгой краткосрочной памяти

ВВЕДЕНИЕ

Компании принимают решения о пополнении товарных остатков на объектах хранения на основе хорошо известных методик определения безопасных запасов. В этих методиках ключевым предположением является нормальное распределение времени доставки грузов. Несмотря на то, что многие исследования показали ненормальность распределений времени доставки грузов на многих видах транспорта, все еще неясно, как распределено такое время при трансграничных и мультимодальных перевозках и насколько снизится объем запасов, если оценки времени доставки будут более точными.

Более того, хотя методики определения безопасных запасов являются принятыми практиками, их входные данные иногда получаются из статических и недостаточно сложных графиков времени доставки.

Для решения этих ограничений в данном исследовании будет проведен анализ методов и моделей, возможных к использованию для разработки приложений с применением алгоритмов, направленных на решение таких логистических оптимизационных задач (прежде всего в части определения расчетного времени прибытия груза (ETA, от англ. Estimated Time of Arrival)).

Целью исследования является повышение уровня обоснованности принимаемых управленческих решений за счет разработки системы управления в сфере логистики и транспорта, базирующейся на комплексе инструментов автоматизации, моделей анализа данных и машинного обучения (ML, от англ. Machine Learning). В связи с этим для достижения указанной цели необходимо решить ряд задач:

1. Выполнение теоретического анализа методов и моделей для разработки приложений на основе технологии управления данными в сфере логистики и транспорта.
2. Проведение анализа предметной области, осуществление классификации алгоритмов ML для решения задачи определения ETA.

3. Описание рыночной специфики и условий организации проекта разработки информационного приложения на основе ML в сфере распределительной логистики.

4. Выполнение формализации основных математических и аналитических зависимостей применительно к модели машинного обучения, способной прогнозировать ETA.

5. Построение процессных моделей и реализация этапов разработки формализованной на предыдущем этапе модели машинного обучения, выполнение тестирования модели как принципиально новой технологии обработки данных о распределительных процессах с использованием ML.

6. Описание программно-аппаратного базиса для реализации поставленной задачи по разработке прототипа системы управления в сфере логистики и транспорта с выделением архитектуры, стека и ключевых преимуществ выбранных программных средств, проведение экономической оценки реализации проекта.

Методы исследования основываются на фундаментальных положениях анализа данных, системного анализа, теории временных рядов, теории графов, машинного обучения.

Объектом исследования являются логистические процессы и процессы разработки прототипов систем управления в сфере логистики и транспорта с использованием элементов ML.

Предметом исследования являются методы и алгоритмы прогнозирования ETA.

Научная новизна исследования состоит в решении актуальной задачи предсказания времени доставки грузов на основе исторической информации о сроке доставке заказов. При этом актуальность работы также обусловлена необходимостью предложения нового подхода при предсказании сроков доставки, разработанного и адаптированного к современным условиям управления в сфере логистики и транспорта.

Структура и содержание работы. Выпускная квалификационная работа состоит из введения, четырех глав, заключения, списка литературы и приложений.

Общий объем работы составляет 60 страниц.

Во введении отражена актуальность рассматриваемых проблем, приводятся формулировки цели и задач исследования, представлено описание структуры работы.

Первая глава посвящена актуальности и описанию современных достижений в области исследований определения расчетного времени прибытия груза. Представлен теоретический анализ и обзор методов, применяемых для решения рассматриваемой задачи в данной предметной области (в том числе, методы на основе маршрута, OD-ориентированные методы, графовые и глубокие нейронные сети, графовые трансформеры), предложена классификация таких методов, алгоритмов и моделей, рассмотрена их специфика, достоинства и недостатки. Кроме того, описана специфика разработки прототипа системы управления в сфере логистики и транспорта, основанной на технологии обработки данных с использованием ML, с учетом выявленных требований к таким проектам и рыночных условий.

Вторая глава отражает описание процесса разработки прототипа модели машинного обучения, предлагаемой для имплементации в создаваемую систему управления в сфере логистики и транспорта: формализована постановка задачи создания целевой модели, математические и аналитические зависимости, осуществлен инжиниринг сложносоставного модуля прогнозирования сроков доставки, математически описаны его компоненты, проведено обучение модели машинного обучения ETAskow.

Третья глава содержит детализацию процесса тестирования предложенного алгоритма машинного обучения, в том числе описание экспериментальных наборов данных и статистики, методы и метрики оценки эффективности, сравнение производительности созданной модели с базовыми моделями на выбранном наборе данных.

Четвертая глава описывает разработку прототипа непосредственно системы управления в сфере логистики и транспорта, которая благодаря совокупности выполненных в третьей главе итераций будет основана на технологии обработки данных с использованием ML. В частности, приводится краткое описание и формализованный вид архитектуры создаваемого цифрового решения, основные планируемые функции и характеристики прототипа, стек разработки, а также решаемые системой проблемы. Помимо этого, в рамках указанной главы и соответствующих приложений осуществляется всесторонняя экономическая оценка реализации проекта.

В заключении приведены основные результаты, выносимые на защиту выпускной квалификационной работы, проведен анализ и сравнение с другими методами и алгоритмами, также намечены направления для дальнейшего развития данного исследования.

1 ТЕОРЕТИЧЕСКИЕ ОСНОВЫ РАЗРАБОТКИ ПРОТОТИПА СИСТЕМЫ УПРАВЛЕНИЯ В СФЕРЕ ЛОГИСТИКИ И ТРАНСПОРТА, ОСНОВАННОЙ НА ТЕХНОЛОГИИ ОБРАБОТКИ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ML

1.1 Теоретический анализ методов и моделей для разработки приложений на основе технологии управления данными в сфере логистики и транспорта

В целом вопрос разработки прототипа системы управления в сфере логистики и транспорта, основанной на технологии обработки данных с использованием моделей машинного обучения является междисциплинарным и требует использования и учета в работе над проектом результатов множества исследований как в сфере наук о данных, так и в сферах экономики и логистики.

Изобретение всемирной паутины позволяет пользователям делиться информацией и продуктами в Интернете, что в итоге способствует развитию трансграничной торговли и платформ электронной коммерции. Быстрое развитие маркетплейсов и других способов электронной коммерции приводит к миллионам посылок, доставляемых ежедневно по всему миру. Розничные торговцы на платформах электронной коммерции предоставляют потребителям различную информацию о своих продуктах и услугах, такую как цены, описания, а в последнее время – еще и ожидаемое время доставки. Зачастую время доставки оценивается важнее, чем цена, когда у потребителей есть срок для получения своих посылок. Недавние исследования показали, что этот показатель влияет на решения 87% онлайн-потребителей.

Таким образом, точная оценка времени доставки влияет не только на доходы от электронной коммерции, но и формирует ожидания потребителей. Ложные обещания или недостаточные обещания могут привести к ущербу для

всех звеньев цепочек поставок – производителей, посредников, онлайн-розничных торговцев и потребителей.

Стоит отметить, что прогнозирование времени прибытия является важной задачей во многих областях, включая транспорт, логистику и сервисы по заказу такси. Точные прогнозы такого показателя, как ETA могут повысить эффективность и надежность этих сервисов, а также улучшить общее впечатление от обслуживания клиентов. Поэтому в этой области продолжается много исследований, с регулярным разработкой новых методов и моделей.

С учетом изложенного, ввиду ограниченности масштаба исследования и ресурсов для его проведения в рамках настоящего исследования сосредоточимся на анализе основных достижений в области исследований определения расчетного времени прибытия ETA и разработаем собственную модель машинного обучения для его расчета для имплементации в создаваемый цифровой продукт.

Существующие научные работы по исследованию методов оценки времени доставки между начальной и конечной точкой (OD, от англ. Origin-Destination) либо предполагают, что все начальные и конечные точки могут быть наблюдаемы для построения пространственной сети, либо прямо используют исходные признаки.

Методы на основе графов используют мощь графовых нейронных сетей (GNN, от англ. Graph Neural Network) для изучения представлений поездок. Однако эти подходы требуют, чтобы все адреса в графе OD присутствовали во время обучения модели, что делает их сложными для использования применительно к новым маршрутам.

Еще одно направление научных исследований использует глубокие нейронные сети (DNN, от англ. Deep Neural Network) для прогнозирования ETA на основе исходных признаков. Однако они не могут уловить высокоуровневую семантическую информацию, возникающую во взаимодействии между атрибутами маршрута (например, время работы

таможенного терминала, загруженность тех или иных маршрутов, погодные условия и другие важные факторы), что также может повлиять на точность прогнозирования ETA.

В результате выяснилось, что ближайшее к нашей проблематике исследование – это оценка времени доставки еды, проведенная группой Alibaba на платформе Ele.me. Невзирая на то, что оно фокусируется на доставке еды на уровне города, оно подчеркивает свой вклад в инженерии признаков существующих моделей.

В целом, несмотря на то что существует много различных подходов к прогнозированию времени прибытия, применительно к разработке приложений на основе технологии управления данными в сфере логистики и транспорта, конечная цель заключается в разработке моделей, которые могут точно оценить время прибытия заданного маршрута или доставки. Используя пространственную и временную информацию, а также другие соответствующие функции, эти модели могут помочь повысить эффективность транспортных и логистических сервисов, оказывая пользу как поставщикам услуг, так и клиентам.

С учетом изложенного далее перейдем непосредственно к анализу методов и алгоритмов ML, которые используются в таких моделях, для решения задач в сфере логистики и транспорта.

1.2 Анализ предметной области и классификации алгоритмов ML для решения задач в сфере логистики и транспорта

Современные проблемы логистики и транспорта требуют эффективных алгоритмов прогнозирования ETA, что играет ключевую роль в оптимизации процессов доставки и удовлетворении потребностей клиентов. Как указано выше, в данном контексте широко применяются как классические методы и алгоритмы, связанные с прогнозированием времени прибытия (ETA), так

методы машинного обучения (ML), включая графовые нейронные сети (GNN) и трансформеры, для решения задач оценки времени доставки.

По результатам проведенного анализа предметной области мы классифицировали такие алгоритмы, разделив на две основные группы: алгоритмы, связанные с исследованием вопросов оценки времени прибытия (ETA) и графовые трансформеры. Ниже приведено их более подробное описание.

1.2.1 Алгоритмы и методы, связанные с исследованием вопросов оценки времени прибытия (ETA)

В ходе анализа работ по данной тематике установлено, что методы прогнозирования ETA можно разделить на две основные группы:

- 1) методы на основе маршрута;
- 2) OD-ориентированные методы.

Рассмотрим данные группы методов более подробно.

1.2.1.1 Методы на основе маршрута

Как следует из названия, методы на основе маршрута непосредственно учитывают информацию о маршруте. Если информация о маршруте доступна, то для решения задачи по определению исследуемого параметра ETA исследователи предлагают использовать (графовые) нейронные сети для кодирования пространственной и временной зависимостей.

Например, архитектура Wide-Deep-Recurrent (WDR) разработана на основе последовательных признаков, чтобы решить задачу прогнозирования ETA как задачу регрессии [34, С. 862].

В свою очередь, CompactETA [13, С. 3340] и ConSTGAT [12, С. 2698] используют сети внимания на графах (GAT) [30, С. 22] на

спациотемпоральном графе для обучения представлениям (Learning Representations),

NetETA [16, С. 2451] моделирует разные повороты в гетерогенном информационном графе, а затем изучает представление графа для прогнозирования ETA. Обратим внимание, что структура дорог в NetETA считается статичной, что отличается от доставки мелких грузов, где графы эволюционируют динамически.

В логистических системах DeepETA используется пространственно-временная последовательная нейронная сеть, чтобы зафиксировать регулярность паттерна доставки и последовательность посылок в маршруте доставки [36, С. 776].

Все вышеупомянутые сети предполагают, что узлы присутствуют при обучении эмбедингов и не применимы в индуктивных настройках.

1.2.1.2 OD-ориентированные алгоритмы и методы

Во многих сценариях приложений в сфере логистики и транспорта из соображений конфиденциальности часто бывает сложно получить информацию о маршруте. В таком случае предлагаются OD-ориентированные методы, которые используют признаки места отправления и места назначения для прогнозирования времени прибытия.

В исследовании, результаты которого опубликованы в статье «A Simple Baseline for Travel Time Estimation Using Large-Scale Trip Data» [32, С. 12] авторы предлагают простую модель для прогнозирования времени поездки, взвешивая время поездки соседних поездок. Соседи отбираются в соответствии как с пространственной информацией о месте отправления и месте назначения, так и с временной информацией о времени начала поездки. Цзэнхуй Ли и соавторы разработали новую рамочную модель обучения для прогнозирования времени поездки, которая использует дорожную сеть, пространственно-временное сглаживание и дополнительные атрибуты

поездки для представления атрибутов поездки в реальном мире [23, С. 1699]. Модель может усваивать информацию о дорожной сети в исторических поездках для точного прогнозирования времени прибытия. Однако она предполагает, что информация о маршруте может быть наблюдаема в исторических поездках, в связи с чем такая модель, по нашему мнению, не обладает необходимой обобщающей способностью.

Самой релевантной к нашему исследованию существующей работой в данной группе методов, по нашему мнению, является прогнозирование времени прибытия еды [43, С. 2574], которое также сложно, как и наша проблема, поскольку оно зависит от множества параметров, связанных с курьерами, ресторанами и транспортом. Однако данная работа в основном фокусируется на использовании инженерии признаков для обучения существующих моделей.

1.2.2 Графовые трансформеры

Использование трансформеров [29, С. 6002], графовых нейронных сетей (GNN) и их вариантов достигли большого успеха во многих областях.

Несколько лет назад была предложена идея для объединения двух подходов в графовый трансформер. Например, Юн и соавторы разработали графовые трансформерные сети на гетерогенных графах с целью преобразования заданного гетерогенного графа в граф на основе метапути и последующего выполнения свертки [38, С. 11990]. В фокусе данного исследования находится использование механизма внимания для интерпретации сгенерированных метапутей. В другой работе [18, С. 2705] был разработан графовый трансформер на гетерогенных информационных сетях, который использует трансформер для замены механизма внимания в графовую сеть внимания (GAT, от англ. Graph Attention Network). Более того, Чжоу и соавторы предложили трансформерную генеративную модель, которая генерирует временные графы, изучая динамическую информацию в сетях [39,

С. 406]. В еще одной работе [11, С. 110] авторы предлагают обобщение архитектуры трансформерных нейронных сетей для однородных графов произвольной структуры. Вышеуказанные существующие методы в основном используют механизм внимания в трансформере для агрегации информации. Однако, их применение к индуктивным графам затруднительно.

Таким образом, в рамках настоящего пункта были рассмотрены различные алгоритмы и методы для прогнозирования времени прибытия (ETA) в сфере логистики и транспорта. Была проведена классификация методов, включая методы на основе маршрута и OD-ориентированные методы. Мы также рассмотрели графовые трансформеры, в том числе различные подходы к их применению в индуктивных графах. Как видим, все указанные подходы не лишены недостатков.

1.3 Преимущества и специфика внедрения перспективной модели машинного обучения в информационное приложение в сфере распределительной логистики

В части внедрения возможных алгоритмов в информационное приложение на основе ML в сфере распределительной логистики с учетом результатов анализа существующих работ в данной сфере мы предлагаем новый подход для решения поставленной задачи – это использование индуктивного графового трансформера для оценки времени доставки.

В отличие от текущих совмещенных графовых трансформеров, которые используют трансформер как замену информационного агрегатора в GNN, предлагается осуществить разделение конвейера GNN и трансформера в IGT. Конкретно, предлагается использование трансформера как регрессионной функции и структурирование маршрутов/заказов в предложения NLP. Аналогично словам в предложении для обработки языка, мы предлагаем обозначать параметры маршрутов как «элемент» в заказе. Таким образом, мы сможем представить заказы как гетерогенный граф с различными типами

узлов (т.е. элементами) и захватить высокоуровневые семантические отношения с помощью распространения информации GNN. Кроме того, предлагается адаптировать данную модель к масштабным сценариям, упростив GNN за счет удаления нелинейной активации и матрицы линейного преобразования.

По нашему мнению, с вышеуказанным упрощенной GNN встраивания невидимых элементов в заказ можно аппроксимировать из наблюдаемых соседних элементов. На практике поставки товаров имеют периодичность и повторяющиеся характеристики. Мы захватываем временную эволюцию элементов с помощью управляемого рекуррентного блока (GRU). Наконец, предлагается объединить встраивание узлов и исходные характеристики маршрутов/заказов в архитектуру, аналогичную трансформеру, для прогнозирования ETA.

Представляется, что использование данного подхода при прототипировании приложения обеспечит существенное превосходство результатов прогнозирования в сравнении с другими методами и алгоритмами.

Указанный подход отличается тем, что использование декомпозиции GNN и трансформеров, по нашему мнению, позволит эффективно учитывать как извлеченную гетерогенную информацию из GNN, так и исходные признаки заказов и маршрутов для прогнозирования ETA. Представляется, что предложенная модель может быть успешно применена не только в задачах доставки, но и в других аналогичных сценариях в сфере логистики и транспорта.

2 РАЗРАБОТКА МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ СИСТЕМЫ УПРАВЛЕНИЯ В СФЕРЕ ЛОГИСТИКИ И ТРАНСПОРТА

2.1 Практика и опыт создания ИТ-решений в сфере ИИ-систем управления цепочками поставок (AI Supply Chain Management)

2.1.1 Описание рынка ИТ-решений в сфере ИИ-систем управления цепочками поставок (AI Supply Chain Management)

Искусственный интеллект (ИИ) и аналитика данных становятся все более важными для компаний, стремящихся оптимизировать свои цепи поставок, снизить затраты и улучшить операционную эффективность.

Индустрии логистики и электронной коммерции ориентируется на развитие гибких и настраиваемых решений, которые могут быть адаптированы к уникальным потребностям каждой компании. Компании ищут решения, которые интегрируются легко и эффективно, а также могут масштабироваться по мере роста и изменения требований.

Большое количество компаний уже начали применять ИИ в управлении цепями поставок для автоматизации процессов прогнозирования, оптимизации запасов, управления рисками и повышения общей видимости и эффективности цепи поставок.

Пандемия COVID-19, введение санкций и экспортно-импортных ограничений также оказали значительное влияние на рынок ИТ-решений в сфере ИИ-систем управления цепочками поставок. Неожиданные сбои в поставках и изменение спроса подчеркнули важность гибкости, резилиентности и прозрачности в управлении цепями поставок, и компании все больше обращаются к ИИ и аналитике данных, чтобы справиться с такими вызовами.

Также одной из тенденций на рынке является интеграция ИИ с другими технологиями, такими как интернет вещей (IoT), блокчейн и облачные

вычисления, для создания более полных и инновационных решений в области управления цепями поставок.

Также наблюдается рост интереса к развитию устойчивых и экологически ответственных цепей поставок. Компании стремятся оптимизировать маршруты доставки, уменьшить выбросы и энергопотребление, а также сотрудничать с экологически ответственными поставщиками. В этой области ИИ может быть использован для оптимизации процессов и принятия решений, например, путем моделирования и анализа данных, связанных с устойчивостью.

Большое внимание также уделяется улучшению прозрачности и трассируемости в цепях поставок. Потребители все больше интересуются сроками поставок, происхождением продуктов, условиями производства и использования ресурсов. В этой сфере ИИ может быть применен для создания систем, позволяющих отслеживать и документировать каждый этап поставки продукта, что способствует установлению доверия и снижению рисков.

Интеграция различных акторов в цепи поставок также остается важным вопросом. В современных цепях поставок часто принимают участие множество поставщиков, производителей, дистрибьюторов, розничных продавцов, экспедиторов, перевозчиков, таможенных служб, страховых компаний и других организаций. ИИ может помочь в управлении и координации этих разнообразных акторов, например, путем предоставления информации в режиме реального времени, прогнозирования спроса и оптимизации запасов.

Однако, несмотря на многообещающие возможности, на рынке ИТ-решений в сфере ИИ-систем управления цепочками поставок остаются некоторые вызовы и препятствия. К ним относятся сложность интеграции ИИ-решений с существующими системами и базами данных, а также необходимость обучения и поддержки персонала для работы с новыми технологиями.

Компании, опережающие конкурентов в применении ИИ в управлении цепями поставок, могут получить значительное конкурентное преимущество.

С каждым днем рынок становится все более конкурентным, поскольку появляются новые поставщики решений и инновации.

На текущий момент на рынке уже представлено несколько поставщиков ИТ-решений в области ИИ-систем управления цепочками поставок, включая крупные технологические компании и специализированные стартапы. Эти решения предлагают широкий спектр функциональности, включая оптимизацию запасов, прогнозирование спроса, управление рисками, планирование производства и дистрибуции, аналитику данных и многое другое.

Тем не менее, в целом, рынок ИТ-решений в данной сфере представляет значительный потенциал для роста и развития. С развитием технологий и увеличением осведомленности о преимуществах применения ИИ в управлении цепями поставок, компании все больше признают его значимость и активно ищут инновационные решения.

2.1.2 Объем и емкость рынка ИТ-решений в сфере ИИ-систем управления цепочками поставок (AI Supply Chain Management)

По результатам 2022 года объем мирового рынка ИИ-систем управления цепочками поставок оценен в размере 1 458 669,56 млн долларов США. В период с 2022 года по 2028 год ожидается значительный рост объема данного рынка. Поскольку компании всё больше осознают потенциал и преимущества, которые предоставляют ИИ и аналитика данных в управлении цепями поставок, прогнозируется, что этот рынок будет расти со среднегодовым темпом роста 21,67%, достигнув 4 731 719,69 млн долларов США к 2028 году [1].

2.1.3 Анализ и систематизация ИТ-решений в сфере ИИ-систем управления цепочками поставок (AI Supply Chain Management)

В современном мире, где глобальная экономика становится все более сложной и динамичной, управление цепочками поставок становится ключевым элементом успеха для предприятий всех отраслей. С развитием технологий искусственного интеллекта (ИИ) и его интеграции в бизнес-процессы, возникает новая парадигма управления цепочками поставок – ИИ-системы управления цепочками поставок (AI Supply Chain Management).

В данной главе мы проведем анализ и систематизацию информационно-технологических решений в области ИИ-систем управления цепочками поставок. Будут рассмотрены как мировые, так и российские цифровые продукты, предоставляющие инструменты для оптимизации логистических процессов, выстраивания эффективных маршрутов доставки и предсказания сроков поставок.

Цель данного исследования – выявить особенности различных ИТ-решений, их технические характеристики, а также преимущества и недостатки по сравнению с разрабатываемым нами цифровым продуктом. Путем систематизации и анализа данных мы сможем оценить, какие технические инновации и функциональные возможности вносят вклад в современные системы управления цепочками поставок и какие из них могут применяться в российских условиях для повышения эффективности бизнес-процессов.

Теперь перейдем к подробному анализу основных аспектов ИТ-решений в области управления цепочками поставок, начиная с представления мировых лидеров в данной сфере и их технических характеристик.

Так к основным аналогичным иностранным сервисам можно отнести: Aftership, Searates, Freightos, Shiphub, Kontainers, Flexport, Magaya Supply Chain, GoFreight, Tailwind TMS.

Основной их недостаток – не оказывают сервиса в России (уход с российского рынка). Тем не менее, для целей выявления технологической

конкурентоспособности, рассмотрим основные технические характеристики продуктов иностранных конкурентов:

Указанные сервисы предоставляет определенные возможности по оптимизации маршрутов, но их системы не такие точные и эффективные, как у разрабатываемого нами продукта. Их интеграция с различными сервисами также ограничена по сравнению с нашим решением.

Далее перейдем к анализу российских ИТ-решений в области управления цепочками поставок. В частности, похожие на создаваемый цифровой продукт сервисы в России: Agora freight, Metaship, Apiship, Catapulto. Вместе с тем, ни один из них не использует технологии искусственного интеллекта (ИИ) и машинного обучения (ML) для решения задач оптимизации логистических процессов клиентов.

Среди остальных недостатки российских аналогов также:

- небольшое количество предлагаемых вариантов доставки (в некоторые направления и вовсе отсутствуют какие-либо варианты перевозки);
- отсутствие прозрачного ценообразования стоимости доставки;
- высокие цены среди доступных вариантов перевозки;
- отсутствие возможности подключения своего собственного договора (выбора тарифа) с соответствующей логистической компанией (если у клиента уже есть такой договор);
- отсутствие возможности подключения каналов продаж (маркетплейсов, интернет-магазинов) для автоматизации торговых операций клиентов.

Детализированный сравнительный анализ аналогичных российских ИТ-решений приведен в таблице 1.

Таблица 1 – Сравнительный анализ аналогичных российских ИТ-решений

Метрики	Создаваемый продукт	Agora Freight (Россия)	Магистрали (Россия)
Трансграничные маршруты	весь мир	весь мир	только Россия, Белоруссия, Казахстан
Мультимодальные маршруты	авто, авиа, ж/д, море	авто, ж/д, море	только авто
Оптимизация маршрутов по времени	мономодальные и мультимодальные маршруты	мономодальные и мультимодальные маршруты	только мономодальные маршруты
Интеграция с перевозчиками	российские, зарубежные	только российские	только российские
Интеграция с маркетплейсами	российские, зарубежные	нет	нет
Интеграция с фулфилмент-сервисами	российские, зарубежные	нет	нет
Интеграция с системами заказчика	ERP, CRM, WMS	ERP, CRM	ERP, CRM
Сервис по онлайн-бронированию отправок сборных грузов	да	нет	нет
Сервис по прогнозированию сроков доставки	да	нет	нет
Скорость подключения клиента к платформе	5 минут	5 минут	более 24 часов

Сравнив существующие решения на рынке России и в мире, хотелось бы выделить три основных момента, по которым создаваемый нами цифровой продукт отличается от конкурентов:

- оптимизация маршрутов по времени и стоимости на основе предобученной нейронной сети;
- интеграции с перевозчиками, фулфилмент-сервисами, маркетплейсами, а также возможность интеграции с системами заказчика.

Тем самым, позволим предположить, что наша система будет на порядок лучше систем конкурентов и позволит клиентам сократить расходы на логистику и увеличить обороты и прибыль за счет автоматизации логистических операций и оптимизации доставки.

2.2 Формализация основных математических и аналитических зависимостей для ML-модели системы управления в сфере логистики и транспорта

2.2.1 Формальная постановка задачи по созданию модели

С учетом того, что в упрощенном виде цель настоящего исследования – решить проблему предсказания времени доставки грузов на основе исторической информации о сроке доставке заказов, формализуем определение заказа и задачи оценки времени доставки следующим образом:

Определение 3.1: заказ.

Отправление представим кортежем: $x_i = (r_i, o_i, d_i)$, где r_i обозначает перевозчика, o_i, d_i – местоположение отправления и место назначения заказа соответственно;

Каждый элемент в кортеже такого отправления сопровождается вектором динамических характеристик, описывающих его статистическую информацию.

Таким образом, мы можем обозначить набор данных с N исторических заказов $X = \{x_i | i = 1, 2, \dots, N\}$, а связанное время доставки $y = \{y_i | i = 1, 2, \dots, N\}$.

Определение 3.2: проблема определения времени доставки.

Учитывая набор исторических заказов X , время доставки y и входные характеристики Z , мы стремимся оценить время доставки y_q для запроса на заказ x_q .

2.2.2. Инжиниринг модуля прогнозирования сроков доставки ETaKnow

С момента первоначального упоминания в научных исследованиях описания графовой сверточной сети (GCN) [21, С. 48], предназначенной для обобщения сверточных нейронных сетей (CNN), основанных на структурированных графических данных, GCN и его вариации были успешно применены в различных областях [3, 33, 37, 40-42].

Поскольку граф заказов можно легко построить на основе исторических данных о заказах, каждый элемент в кортеже заказа представлен в виде узла в графе заказа, и два узла связаны, если они встречаются в одном и том же заказе. Учитывая вычислительную сложность обработки графов заказов большого масштаба, мы дополнительно ограничиваем связи смежными элементами заказа.

Определение 4.1: Построение гетерогенного графа.

Конкретно, для учитывая заказ (r_i, o_i, d_i) :

- узел перевозчика r_i может соединиться только с узлом пункта отправления o_i ,
- узел пункта отправления o_i может быть связан как с узлом перевозчика r_i , так и с узлом пункта назначения d_i и т.д.

То есть каждый заказ будет генерировать только 6 ребер вместо 12. Таким образом, плотность графа может быть уменьшена вдвое. В то же время,

информация с одного узла может передаваться к любым другим узлам путем добавления большего числа сверточных слоев GCN.

С помощью построенного таким образом гетерогенного графа мы можем декомпонировать граф на набор двудольных графов и выполнять графовую свертку на этих двудольных графах.

Оператор преобразования и слой активации в GCN требуют непомерно высоких вычислительных затрат для графов большого масштаба, в связи с чем мы упрощаем слой графовой свертки с помощью линейных вычислений.

Однако, помимо структурной информации, узлы в графе могут демонстрировать разные временные паттерны, в связи с чем для анализа временных корреляций на временной оси на уровне узлов мы предлагаем использовать управляемые рекуррентные блоки (Gated Recurrent Unit, GRU). GRU решает проблему затухающих градиентов стандартной рекуррентной нейронной сети (Recurrent Neural Network, RNN), с помощью обновляющего и сбрасывающего гейтов.

По сравнению с нейронной сетью долгой краткосрочной памяти (Long Short-Term Memory, LSTM), GRU быстрее обучается за счет меньшего количества весов и параметров, которые необходимо обновлять во время обучения. Выходные данные GRU, а также исходные входные признаки передаются в вариант трансформера для финального предсказания.

На рисунке 1 показана общая структура предлагаемой модели ETAKnow. Она состоит из двух основных модулей:

- 1) Темпорально-гетерогенный графовый блок (THGblock, от англ. Temporal heterogeneous graph block), т.е. временной и гетерогенный GCN и
- 2) Блок на основе трансформера Tblock (от англ. Transformer-based block), т.е. модуль предсказания значения ETA с помощью трансформеров).

Далее рассмотрим каждый из составляющих данную модель модулей более подробно.

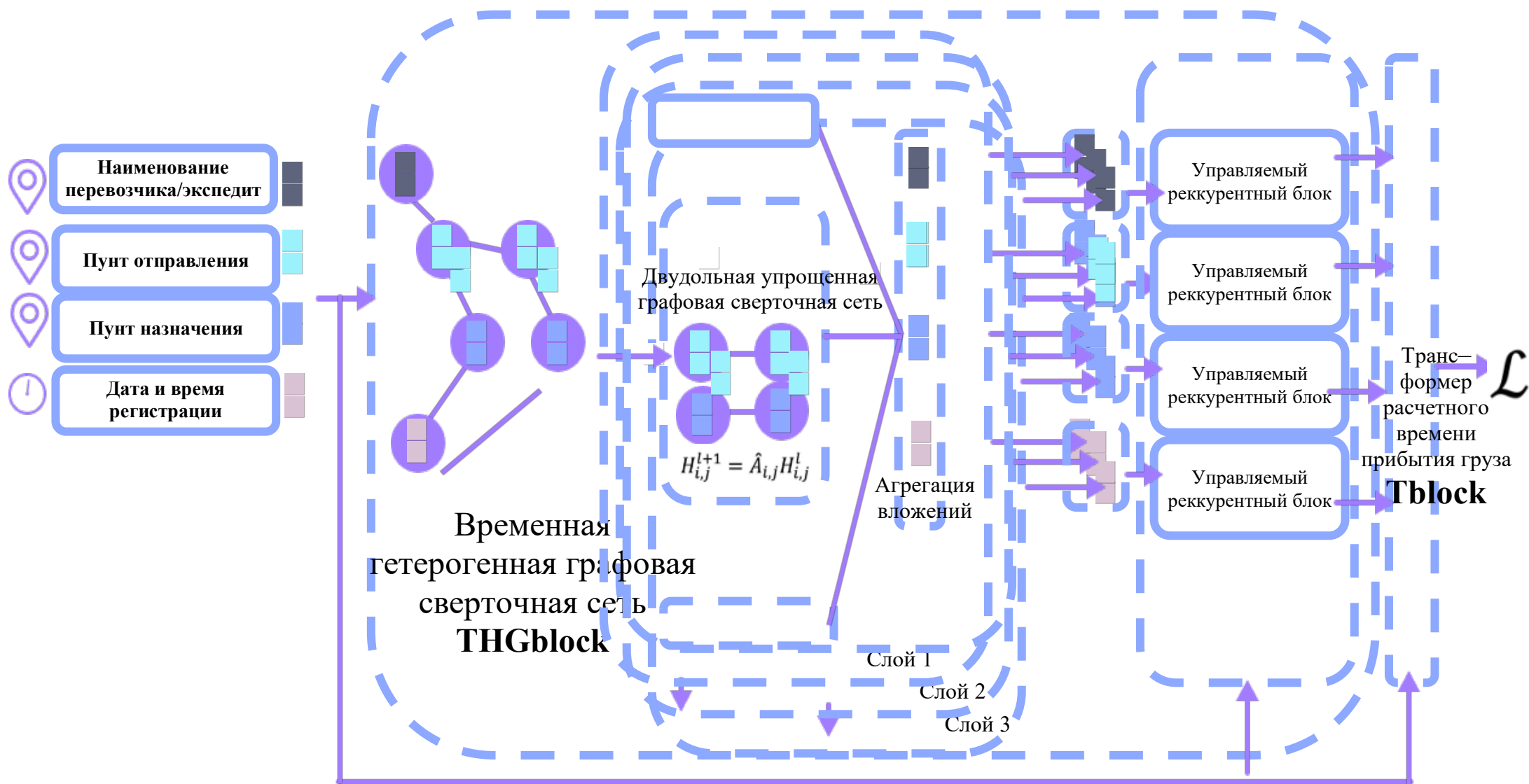


Рисунок 1 – Общая структура предлагаемой модели машинного обучения ETAnow

2.2.2.1. Структурный компонент создаваемой модели машинного обучения THGblock

Предположим, что $G = (V, E)$ – это граф заказов, полученный на основе исторических данных о заказах, где V и E обозначают множество узлов и множество ребер соответственно. Каждому узлу $v \in V$ и каждому ребру $e \in E$ сопоставлены функции отображения типов. Наш сценарий включает четыре типа узлов и один тип ребер. Следовательно, построенный граф будет гетерогенным графом. В связи с этим, сначала мы рассмотрим, как работает GCN на однородных графах, а затем покажем, как по нашему мнению, можно применить свертку к гетерогенным графам.

Типичная GCN на однородных графах, которая рекурсивно обновляет скрытое вложение H^l на l -м уровне, определяется следующим образом:

$$H^l = \sigma(\hat{A}H^{l-1}W^l), \quad (1)$$

где W – обучаемая матрица преобразования, $\sigma(\cdot)$ – нелинейная функция, например, функция ReLU. $\hat{A} = \hat{D}^{-1/2}(A + I)\hat{D}^{-1/2}$ – ренормализация матрицы смежности A . $\hat{D} = \text{diag}(d_1, \dots, d_{|V|})$ – диагональная матрица степеней $A+I$, где каждый элемент на диагонали равен сумме строки матрицы смежности $d_i = 1 + \sum_j a_{ij}$. Каждый узел v_i в графе имеет соответствующий вектор скрытых признаков размерности $h_i \in R^D$. Начальная матрица признаков $H^0 = [h_0, \dots, h_{|V|}]^T$ объединяет $|V|$ векторов признаков один над другим.

Чтобы применить оператор графовой свертки к гетерогенным графам, в рамках настоящего исследования были рассмотрены несколько вариантов GCN, ранее использованных в других областях применения [27, 44]. Однако они оба являются вычислительно затратными и склонными к переобучению.

Вдохновившись исследованием простой графовой свертки (Simple Graph Convolution, SGC) [35, С. 6865], был разработан гетерогенный GCN, который ориентирован на вычислительную эффективность без ущерба для производительности. По сравнению с SGC [35, С. 6866], была дополнительно

убрана матрица преобразования, чтобы сократить пространство поиска параметров. Это позволит данной модели обучаться на крупных наборах данных.

Имея гетерогенный граф G , мы сначала строим набор двудольных подграфов на основе комбинаций типов узлов. Предположим, что количество типов узлов равно n , количество двудольных подграфов не превысит $\binom{n}{2} = \frac{(n)!}{2!(n-2)!}$. В частности, мы строим подграф $G_{i,j}$ из оригинального графа G на основе двух типов узлов i и j . В $G_{i,j}$ тип узла принадлежит либо i , либо j , и любые два узла в подграфе $G_{i,j}$ будут связаны, если они связаны в оригинальном графе G . Поскольку $G_{i,j}$ неориентирован, мы имеем $G_{i,j} = G_{j,i}$. Обозначим матрицу смежности $G_{i,j}$ как $A_{i,j}$:

$$A_{i,j} = \begin{pmatrix} 0 & R & R^T & 0 \end{pmatrix}, \quad (2)$$

где каждый элемент $R_{uv} \in \mathbb{R}$ равен 1, если узел $u \in G_{i,j}$ связан с узлом $v \in G_{i,j}$ в G ; в противном случае R_{uv} устанавливается равным 0.

Следуя [21, С. 48], мы также выполняем трюк ре-нормализации на $A_{i,j}$, результатом которого является $\hat{A}_{i,j}$. Чтобы выполнить операцию свертки на двудольных графах, обновим стандартную GCN в уравнении (1) следующим образом:

$$H_{i,j}^l = \hat{A}_{i,j} H_{i,j}^{l-1}, \quad (3)$$

где $H_{i,j}^l$ обозначает эмбединги узлов в $G_{i,j}$ на l -ом слое GCN. Предположим, что количество узлов типа i равно N_i , тогда $H_i \in \mathbb{R}^{N_i \times D}$ является матрицей признаков узла типа i , получаемая путем сложения этого типа узлов друг с другом. Аналогично, $H_{i,j}^l \in \mathbb{R}^{(N_i+N_j) \times D}$.

Обратим внимание, что уравнение (3) упрощает стандартную GCN, удаляя как весовую матрицу свертки, так и нелинейный слой, что дополнительно упрощает SGC [34, С. 860] за счет удаления нелинейной активации. В то же время, сокращение количества обучаемых параметров предотвращает переобучение модели. Линейное распространение

информации позволяет модели быть эффективной в вычислительном отношении. Назовем эту упрощенную свертку двудольных графов «Двудольная упрощенная графовая сверточная сеть», как указано на рисунке 1.

Поскольку один тип узлов может быть задействован в нескольких двудольных графах и одновременно участвовать в передаче информации, предлагаем агрегировать передаваемую информацию для каждого типа узлов в каждом слое следующим образом:

$$H_i^l = AGGREGATE(\{H_{i,j}^l[:N_i] \mid j = 1, \dots, n, \text{ и } G_{i,j} \neq \emptyset\}), \quad (4)$$

где AGGREGATE – это функция агрегирования, n – количество типов узлов. Здесь мы используем агрегирование сумм. Другие операторы агрегации [4, С. 1730] также могут применяться для агрегирования эмбедингов узлов, принадлежащих одному и тому же типу. В уравнении (4) мы используем оператор среза, чтобы извлечь эмбединги для узлов типа i .

Одним из преимуществ такого агрегирования информации внутри каждого слоя является то, что информация от разных типов узлов может полностью интегрироваться друг с другом. Интегрированная информация затем подается на следующий слой для передачи. Еще одно преимущество заключается в том, что THGblock обновляет эмбединги всех узлов, а не пакета узлов на каждой итерации, чтобы избежать проблемы устаревания памяти [19, С. 28].

С помощью полученной информации на каждом уровне мы обновляем итоговые эмбединги узлов на временной отметке $t-1$, агрегируя все промежуточные эмбединги. То есть:

$$H_i^{t-1} = AGGREGATE(H_i^0, H_i^1, \dots, H_i^L), \quad (5),$$

где L – количество уровней свертки. Функция агрегации AGGREGATE может представлять собой любую функцию, описанную в [5, С. 538]. Здесь мы используем среднее агрегирование для финального обновления узла.

Для дальнейшего объединения временной информации в полученные эмбединги H_i^{t-1} на временной отметке $t-1$, воспользуемся гэйтинг-механизмами для контроля и управления потоком информации:

$$H_i^t = \text{UPDATE}(H_i^{t-1}, Z_i^t), \quad (6),$$

где Z_i^t – это векторы исходных признаков описаний i -типа элементов в порядке на временной отметке t . Здесь UPDATE – обучаемая функция обновления, например, рекуррентная нейронная сеть, такая как LSTM или GRU. В данной работе воспользуемся сетью GRU, поскольку она имеет меньше количество гэитов и так же эффективна, как и LSTM. Обратим внимание, что мы используем нормализованную матрицу смежности, построенную на основе исторических данных заказов, чтобы избежать утечки данных на этапе обучения модели. Однако на этапе вывода мы используем нормализованную матрицу смежности, построенную на основе как обучающих, так и тестовых данных [15, С. 1030].

При индуктивной настройке предлагаемый для использования в рамках настоящего исследования модуль THGblock может объединять информацию от соседей в пределах одного заказа x_i , чтобы изучить эмбединги невидимых узлов. В крайнем случае все узлы в x_i никогда не наблюдаются во время обучения, что означает, что никакая информация о соседях не может быть использована. THGblock использует GRU для обновления эмбедингов узлов из необработанных входных данных. Следовательно, необработанные характеристики будут использоваться для оценки времени доставки такого заказа.

1. Входные данные: набор заказов X , связанное время доставки u и входные векторы признаков Z .
- 2: Построение гетерогенного графа G из X .
3. Инициализируем скрытые эмбединги H^0 узлов v & c помощью Xavier.
4. Извлекаем двудольные подграфы с типом узла i и j $G_{i,j}$ из G и скрытые эмбединги H^0 из H^0
5. for X^t in X do
6. for l in $[1, ..., L]$ do
7. for i in $[1, ..., n]$ do
8. for j in $[1, ..., n]$ do
9. $H_{i,j}^l = \hat{A}_{i,j} H_{i,j}^{l-1}, (3)$
10. end for
11. end for
12. for i in $[1, ..., n]$ do
13. $H_i^l = AGGREGATE(\{H_{i,j}^l[:N_i] \})$
14. end for
15. end for
16. for i in $[1, ..., n]$ do
17. $H_i^{t-1} = AGGREGATE(H_i^0, H_i^1, ..., H_i^L)$
18. $H_i^t = UPDATE(H_i^{t-1}, Z_i^t)$
19. end for
20. $\hat{y}^t = Tblock(H^t, Z^t)$
21. $L = MAE(y^t, \hat{y}^t)$
22. $L.backward()$
23. end for

⇒ загружаем батч
⇒ послонное
распространение

⇒ двудольная
упрощенная
графовая
сверточная сеть

⇒ уравнение (4)

⇒ цикл
распространения
конечного слоя

⇒ уравнение (5)

⇒ уравнение (6)

⇒ предсказание
ETA с помощью
компонента
модели Tblock
⇒ потери

⇒ обратное
распространение

Рисунок 2 – Алгоритм (псевдокод) модуля прогнозирования сроков доставки ETAnow

2.2.2.2. Структурный компонент создаваемой модели машинного обучения Tblock

Стандартный трансформер предназначен для предсказания слов, а предложение рассматривается как полносвязный граф. В нашем случае заказ на доставку аналогичен предложению в обработке естественного языка (NLP), содержащему множество элементов. Таким образом, задача прогнозирования ETA в значительной степени может поддаваться использованию трансформеров. Представляется, что на таких полносвязных графах можно обучать большие модели таких трансформеров.

Более того, каждое слово обращается к каждому другому слову в предложении в архитектуре трансформера, что позволяет обеспечить высокоуровневые взаимодействия между словами. В NLP предсказание слова в предложении может меняться в зависимости от контекста, других слов и т.д. Аналогично, в нашем сценарии предсказание ETA зависит не только от необработанных признаков перевозчиков, пунктов отправления, пунктов назначения, но также от высокоуровневых взаимодействий между этими признаками.

Однако стандартный трансформер предполагает, что слова кодируются в одном и том же пространстве эмбедингов. На практике узлы в гетерогенном графе могут отличаться по размеру признаков, так как они принадлежат к разным типам. Таким образом, для включения исходных признаков в трансформер, возникает необходимость изменения формы пространства признаков. То есть:

$$e_i = f(z_i, h_i^t) \quad (7),$$

где z_i – это исходный признак узла i , h_i^t – это выходной эмбединг узла i из THGblock, а f – это функция мэппинга, позволяющая выравнивать узлы одинаковой размерности.

Представляется, что в таком случае интуитивно понятной функцией для сопоставления признаков из разных пространств разных размерностей в одно

пространство могло бы быть использование многослойного персептрона (MLP). Однако это вводит дополнительные параметры в модель, поэтому мы применяем заполнение нулями для каждого узла заказа. Наконец, мы преобразуем заказ в $x^i \in R^{n \times W}$, где n – количество элементов в заказе, а W – размерность заполненного нулями вектора признаков для каждого элемента.

Для сбора информации на уровне заказа, мы добавляем обучаемый токен заголовка, аналогичный BERT [8, С. 26] и ViT [10, С. 101]. При реализации воспользуемся MLP с одним скрытым слоем для обучения токенов. При позиционном кодировании, существующие графовые трансформеры используют различные методы для сохранения информации о расстоянии между узлами [46, С. 1242]. Наша разделенная архитектура графа использует стандартное позиционное кодирование трансформера, потому что в нашем сценарии заказ на доставку x_q может быть естественно представлен в виде последовательности элементов. Токен заголовка e_q^{header} , позиционные эмбединги e_q^{pos} и эмбединги отображаемых узлов e_q из x_q затем передаются в варианты трансформеров. Наконец, к выходным данным мы используем нормализацию слоя (Layer Normalization, LN) и одноуровневый многослойный персептрон (Multilayer perceptron, MLP) для прогнозирования ETA, как показано на рисунке 3.

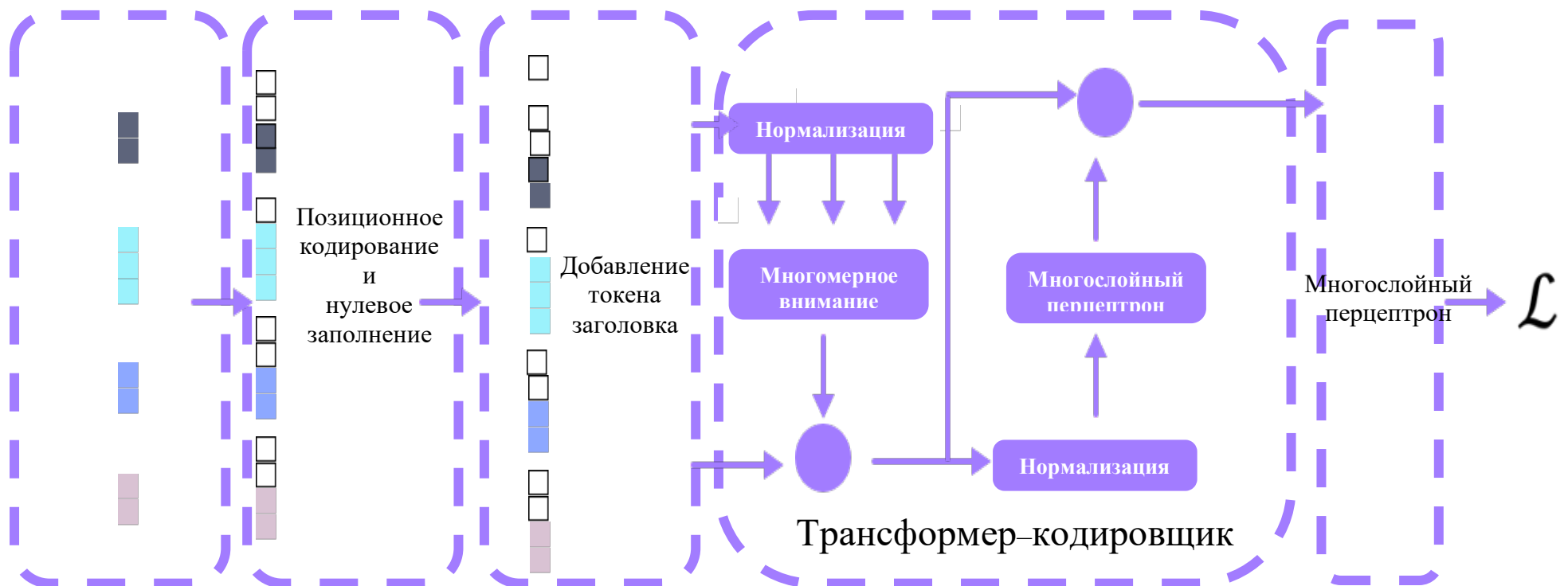


Рисунок 3 – Архитектура компонента Tblock

Формально, мы обозначаем исходный признак запроса заказа x_q как $z_q = (z_q^1, z_q^2, \dots, z_q^n)$, где z_q^i – это исходный признак i -го элемента в заказе. Прогнозирование ЕТА заказа x_q на основе трансформер-подобных моделей можно представить следующим образом:

$$\begin{aligned} e_q &= f(z_q, h_q^t), \\ \hat{e}_q^{header} &= \text{Transformer}(e_q^{header}, e_q^{pos}, e_q), \quad (8) \\ \hat{y}_q &= \text{MLP}\left(\text{LN}(\hat{e}_q^{header})\right) \end{aligned}$$

2.2.2.3. Обучение модели машинного обучения ETAKnow

Для обучения предложенной модели ETAKnow мы используем Adam [19, С. 36] в качестве оптимизатора и среднюю абсолютную ошибку (Mean Absolute Error, MAE) в качестве функции потерь. Учитывая набор N сэмплов из датасета $y = (y_1, y_2, \dots, y_N)$, представляющих истину и соответствующие прогнозы $\hat{y}$, функция потерь на основе MAE будет определена следующим образом:

$$L = \text{MAE}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (9)$$

Подробности алгоритма оптимизации, используемого для обучения модели ETAKnow, приведены в алгоритме на рисунке 2.

$$L = \text{MAE}(y^t, \hat{y}^t)$$

Таким образом, в рамках данной главы обеспечена разработка прототипа модели машинного обучения, предлагаемой для имплементации в создаваемую систему управления в сфере логистики и транспорта, в том числе, формализована постановка задачи для создания необходимой модели, установлены математические и аналитические зависимости, осуществлен инжиниринг модуля прогнозирования сроков доставки, математически и алгоритмически описаны его компоненты, проведено обучение модели машинного обучения ETAKnow.

3 ТЕСТИРОВАНИЕ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ ETAKNOW

3.1 Экспериментальные наборы данных

Для целей проведения эксперимента мы используем крупный набор логистических данных об отправлениях грузов, собранных на одной из фулфилмент-платформ, чтобы оценить производительность модели машинного обучения ETAKnow и базовых моделей.

3.2 Описание данных и статистики

Экспериментальный набор данных содержит 10490 записей от 4 перевозчиков, с 1 мая 2023 года по 31 марта 2024 года.

3.3 Методы и метрики для сравнения

Для оценки эффективности предлагается сравнить реализуемую в рамках настоящего исследования модель ETAKnow с следующими базовыми моделями:

1) Линейная регрессия (Linear regression, LR) моделирует время доставки заказов, обучая линейную регрессионную модель на основе исходных характеристик заказов. Для оценки воспользуемся инструментарием LR в scikit-learn [26, С. 2829].

2) XGBoost [2, С. 791] – это модель регрессии на основе градиентного бустинга деревьев решений. Входные данные такие же, как в модели LR. В данном случае воспользуемся XGBRegressor в xgboost с MAE в качестве метрики оценки.

3) MURAT [23, С. 1696] учитывает структуры дорожной сети и знания о пространственно-временных предпочтениях для оценки времени

перемещения OD (origin-destination). Информация о дорогах в нашем наборе данных недоступна, поэтому граф OD строится на основе адреса отправителя и получателя. Здесь мы предлагаем использовать как ResNet18, так и глубокие нейронные сети прямого распространения (Deconvolutional Neural Network, DNN), чтобы сопоставить обученные представления с целевыми задачами. В нашем сценарии мы обнаружили, что DNN работает лучше, чем ResNet18. Результаты, представленные в нашем исследовании, основаны на DNN.

4) xDeepFM [24, С. 1760] изучает как низкочастотные, так и высокочастотные взаимодействия функций как явным, так и неявным способом в рамках экстремально глубокой факторизационной машинной модели. В нашем сценарии мы преобразуем задачу рекомендации в задачу регрессии для прогнозирования ETA.

5) TEMPrel [32, С. 11] – это метод на основе ранее упомянутого OD-ориентированного подхода для оценки времени перемещения. Он усредняет масштабированное время перемещения всех соседних маршрутов с похожими начальной и конечной точками для прогнозирования ETA.

3.4 Сравнение производительности

Для традиционных моделей машинного обучения мы использовали полные данные обучения для ETA-прогнозирования в LR, XGBoost и TEMPrel. Другие модели обучаются пакетами размером 8192. Мы отслеживаем потери на валидационном наборе данных и применяем раннюю остановку, если не происходит улучшения после 100 последовательных эпох. Учитывая сходимость модели, общее количество эпох для обучения фиксировано на 1000. Также мы инициализируем латентные эмбединги H^0 узлов в графе G методом Ксавье [14, С. 252], настраиваем количество слоев L в модуле HGblock в размере $\{1, 2, 3, 4, 5\}$ и размер эмбединга D узлов в размере $\{16, 32, 64, 128, 256\}$. Реализация нашей модели осуществляется в PyTorch [25,

С. 8031]. Исходный код модели будет опубликован в сети Интернет по адресу: <https://github.com/vin-57/ETAcknow>.

Следуя [23, С. 1703], мы оценивали нашу модель и все базовые модели по метрикам средней абсолютной ошибки (Mean Absolute Error, MAE) и средней абсолютной относительной ошибки (Mean Absolute Relative Error, MARE). Определения MAPE и MARE приведены ниже:

$$MAPE(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (10)$$

$$MARE(y, \hat{y}) = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{\sum_{i=1}^N y_i}, \quad (11)$$

Аналогично MAE, определённого в уравнении (9), y обозначает истинное время доставки отправления, $\hat{y}$ обозначает предсказанное время доставки, а N обозначает количество сэмплов в тестовых данных.

Таблица 3 представляет собой сравнение производительности предложенной модели ETAcknow с базовыми моделями на выбранном наборе данных.

Таблица 3: Сравнение производительности на предсказании ETA между нашей моделью и базовыми моделями

Метрика	LR	XGBoost	MURAT	xDeepFM	TEMPrel	ETAcknow	Улучшение (%)
MAE (часы)	10,64	8,14	8,63	<u>8,05</u>	10,20	7,32	9,06
MAPE (%)	25,91	17,18	19,01	<u>15,98</u>	24,11	13,94	12,77
MARE (%)	23,02	17,38	18,34	<u>17,33</u>	21,99	15,19	8,82

Мы отмечаем лучшие результаты по каждой метрике жирным шрифтом, а вторые по значимости – подчёркнутым. Улучшение (%) рассчитывается между лучшим результатом, достигнутым базовыми моделями, и результатом нашей модели.

Как видно из Таблицы 3, результаты по трём метрикам показывают, что предложенная модель способна улучшить лучший базовый результат более чем на 9%.

С учетом того, что модуль ETAnow состоит из двух компонентов, такая разделенная архитектура нашей модели позволяет нам выполнять прогнозирование ETA на каждом компоненте независимо от другого. В ходе исследования в данной части было установлено, что TBlock (блок на основе трансформера) достигает лучшей производительности по сравнению с THeGBlock (темпорально-гетерогенным графовым блоком) в прогнозировании ETA. Поскольку TBlock работает с межзаказными взаимодействиями, он дополнительно исследует взаимодействия более высокого порядка между элементами заказа. Вместе с тем, именно комбинация Tblock и THGblock позволяет достичь значительного улучшения по сравнению с базовыми моделями.

4 РАЗРАБОТКА ПРОТОТИПА СИСТЕМЫ УПРАВЛЕНИЯ В СФЕРЕ ЛОГИСТИКИ И ТРАНСПОРТА, ОСНОВАННОЙ НА ТЕХНОЛОГИИ ОБРАБОТКИ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ML И ЭКОНОМИЧЕСКАЯ ОЦЕНКА РЕАЛИЗАЦИИ ПРОЕКТА

4.1 Краткое описание создаваемого прототипа системы

В рамках дальнейшего исследования планируется разработка прототипа системы бронирования грузоперевозок и оптимизации доставки с расчетным модулем прогнозирования сроков доставки. Предполагается, что прототип использует разработки в области технологий искусственного интеллекта, в частности, комплексного графового трансформера для прогнозирования ЕТА.

4.2 Формализация архитектуры создаваемого цифрового решения

Создаваемый программный продукт представляет собой цифровую платформу, которая позволяет осуществлять онлайн-бронирование грузоперевозок и предсказывать сроки доставки. Архитектура платформы состоит из трех основных элементов: клиентская часть, интеллектуальное ядро и партнерская часть (рисунок 4):

1) Клиентская часть

Клиентская часть платформы предоставляет пользователям удобный интерфейс для выбора и заказа логистических услуг. Она выполняет следующие ключевые функции:

- регистрация и авторизация пользователей – пользователи могут создавать учетные записи, входить в систему и управлять своими профилями;
- поиск и бронирование – интерфейс позволяет пользователям искать доступные грузоперевозки, сравнивать варианты и бронировать их;
- отслеживание заказов – пользователи могут отслеживать статус своих заказов в реальном времени;

- уведомления – система отправляет уведомления о статусе заказа, изменениях маршрута и сроках доставки.

Эта часть взаимодействует с интеллектуальным ядром для получения оптимальных предложений по маршрутам и срокам доставки, а также с партнерской частью для организации выполнения заказов и обеспечения необходимых сервисов;

2) Интеллектуальное ядро:

Интеллектуальное ядро платформы является непосредственным «мозгом системы». Оно обрабатывает данные о заказах, маршрутах, транспортных средствах и других факторах, чтобы предложить пользователям оптимальные варианты доставки. Интеллектуальное ядро включает следующие модули:

- модуль прогнозирования сроков доставки – использует данные о текущих условиях и исторические данные для предсказания времени доставки с высокой точностью, используя технологии машинного обучения, включающего алгоритмы комплексного графового трансформера;

- модуль обмена данными – обеспечивает связь между различными частями платформы, а также с внешними системами для плавного и эффективного выполнения заказов;

3) Партнерская часть:

Партнерская часть платформы позволяет интегрировать сторонние сервисы и партнеров для расширения функционала и покрытия. Основные функции включают:

- интеграция с перевозчиками и логистическими сервисами – позволяет расширить выбор транспортных средств и маршрутов для пользователей;

- подключение к каналам продаж и фулфилмент сервисам – автоматизирует процессы заказа и исполнения, улучшая координацию и сокращая время доставки;

- интеграция с сервисами таможенного оформления и страхования грузов – обеспечивает безопасность и легальность транспортировки;

- реализация биллинга: упрощает процесс оплаты и учет услуг, предоставляя прозрачные и удобные механизмы для финансовых операций.

Архитектура платформы обеспечивает комплексное решение для управления логистическими процессами, предлагая пользователям удобный интерфейс, оптимизированные решения и широкие возможности в интеграции и расширении функционала.

Взаимодействие между клиентской частью, интеллектуальным ядром и партнерской частью обеспечивает:

- гибкость и масштабируемость – возможность адаптации под растущие требования и расширение функционала за счет добавления новых модулей и сервисов;

- высокую точность и эффективность – за счет использования передовых технологий машинного обучения и алгоритмов оптимизации;

- улучшение пользовательского опыта – благодаря удобному интерфейсу и надежным сервисам, обеспечивающим высокое качество обслуживания.

Эта структура позволяет создавать систему, способную эффективно управлять логистическими процессами, обеспечивая высокую точность прогнозирования и оптимизацию всех этапов доставки.

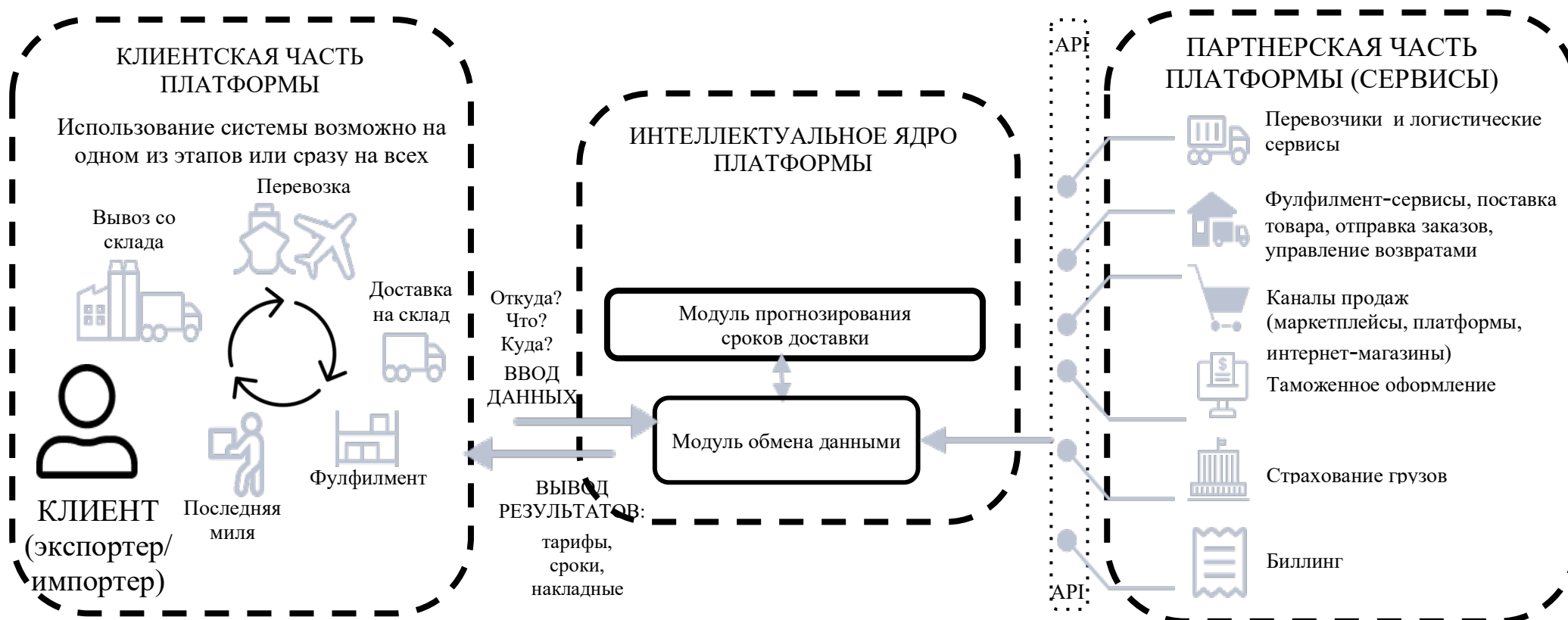


Рисунок 4 – Базовая технология (архитектура прототипа) цифрового продукта

4.3 Основные планируемые функции и характеристики прототипа

На основании приведенного в Приложении А технического задания, разработан личный кабинет платформы, основные элементы платформы. Интерфейс платформы разработан с помощью специализированного сервиса Figma.com и размещен в сети Интернет [47].

Помимо функционала личного кабинета, к основным планируемым функциям прототипа относятся:

- интерфейс импорта из соответствующих автоматизированных систем заказчика данных о заказах и товарных остатках на маркетплейсах;
- интерфейс импорта из соответствующих автоматизированных систем заказчика данных о товарных остатках на складах;
- загрузка исторических данных по срокам доставки грузов по локациями в формате .CSV;
- выбор параметров, по которым необходимо получить прогноз;
- выбор горизонта прогнозирования срока доставки грузов;
- расчет сроков доставки грузов на срок до 12 месяцев с возможностью учета особо загруженных дней, недель и месяцев.

Планируется, что прототип продукта в объеме указанной функциональности, будет обладать следующими характеристиками по скорости вычислений:

1) скорость расчета прогнозов для набора данных о 10490 доставках (по историческим данным с глубиной в 1 год):

- на 60 дней вперед составляет 85%;
- на 30 дней вперед составляет 95%;
- на 7 дней вперед составляет 70%;

2) скорость расчета прогнозов приводится для следующей конфигурации:

- система Ubuntu 22.04 LTS;
- процессор – 4 vCPU;

– оперативная память – 32 Гб;

3) при проведении испытаний в окружении, близком к реальному, были обеспечены следующие показатели точности расчетов Продукта:

Таблица 4 – Показатели точности продукта

Срок прогноза, дней	Ошибка при расчетах «по среднему», %	Ошибка при расчетах методами машинного обучения, %	Повышение точности, %
60	26,4%	17,7%	+32%
30	18,2%	15,1%	+17%
7	15,2%	1,2%	+14%

4.4 Стек разработки создаваемой системы

Для обеспечения базовых функций модели машинного обучения были применены следующие библиотеки:

– NumPy – библиотека языка Python, добавляющая поддержку больших многомерных массивов и матриц, вместе с большой библиотекой высокоуровневых (и очень быстрых) математических функций для операций с этими массивами;

– Pandas – библиотека языка Python для обработки и анализа данных. Работа с данными строится поверх библиотеки NumPy, предоставляющей специальные структуры данных и операции для манипулирования числовыми таблицами и временными рядами;

– Matplotlib – библиотека языка Python для визуализации данных двумерной и трехмерной графикой;

– StatsModels – библиотека языка Python для оценки множества различных моделей, а также проведения статистических тестов и исследования данных;

– DateTime – модуль Python, предоставляющий классы для обработки времени и даты разными способами;

– TQDM – модуль Python, предназначен для быстрого и расширяемого внедрения индикаторов выполнения во внешние интерфейсы программ, что обеспечивает визуальную индикацию хода вычислений или передачи данных в рамках прогнозирования сроков доставки; а также другие библиотеки.

Разработка прототипа основного продукта планируется на следующем стеке:

1) Frontend:

TypeScript (TS) – добавляет статическую типизацию в JavaScript, что позволяет уменьшить количество ошибок на этапе разработки и повысить стабильность кода. Это особенно важно при разработке сложных приложений, таких как системы управления логистикой;

React – популярная библиотека для создания пользовательских интерфейсов. Она позволяет разрабатывать компонентные и легко поддерживаемые интерфейсы, что ускоряет процесс разработки и упрощает масштабирование приложения;

Apollo Client – предоставляет удобные инструменты для работы с GraphQL, что позволяет эффективно управлять состоянием и запросами данных на клиентской стороне. Это особенно полезно в системе, где требуется частое взаимодействие с бэкэндом для получения и обновления информации о заказах и маршрутах;

Material-UI – библиотека компонентов для React, которая позволяет быстро создавать красивые и консистентные пользовательские интерфейсы в соответствии с Material Design. Это ускоряет процесс разработки и улучшает пользовательский опыт.

2) Backend:

Node.js (Express) – позволяет использовать JavaScript для серверной части, что обеспечивает единый язык разработки как на клиентской, так и на

серверной стороне. Express – минималистичный и гибкий фреймворк для Node.js, который ускоряет разработку серверных приложений;

Apollo Server – предоставляет мощные инструменты для создания GraphQL API. Он упрощает создание и управление схемами данных, а также интеграцию с различными источниками данных, что важно для обработки запросов в системе бронирования и оптимизации доставки;

GraphQL – позволяет клиентам запрашивать именно те данные, которые им нужны, и ничего лишнего. Это повышает эффективность передачи данных и снижает нагрузку на сеть. В системе логистики и транспорта это важно для быстрой и эффективной работы с большими объемами данных;

MySQL – реляционная база данных, известная своей надежностью, производительностью и масштабируемостью. Она подходит для хранения структурированных данных о заказах, маршрутах, транспортных средствах и других элементах системы.

Преимущества использования данного стека заключаются в следующем:

- скорость разработки и внедрения – использование современных инструментов и библиотек (React, Material-UI, Apollo) позволяет быстро разрабатывать и внедрять новые функции, обеспечивая высокую скорость реагирования на требования рынка;

- масштабируемость – архитектура на базе Node.js и MySQL позволяет легко масштабировать приложение как горизонтально, так и вертикально, что важно для системы с большим количеством пользователей и данных;

- эффективность и производительность – GraphQL и Apollo обеспечивают эффективное управление запросами и состоянием, что критично для производительности системы в реальном времени;

- надежность и поддержка: Выбор проверенных технологий (TypeScript, Node.js, MySQL) гарантирует надежность системы и наличие широкой поддержки со стороны сообщества разработчиков.

Таким образом, предложенный стек технологий обеспечивает оптимальное сочетание гибкости, производительности и масштабируемости

для разработки прототипа системы управления в сфере логистики и транспорта.

4.5 Проблемы, которые решает создаваемая система управления

Создаваемый продукт предназначен для онлайн загрузки транспортных средств и бронирования соответствующих грузоперевозок путем покупки лейблов на конкретную перевозку, также оптимизации доставки по стоимости за счет использования принципа LTL-перевозок (от англ. Less than Truckload, частичная загрузка грузового автомобиля) путем открытого в режиме реального времени формирования сборных грузов и онлайн-бронирования грузовых мест, а также оптимизации доставки по времени за счет объективного и независимого от заверений перевозчиков прогнозирования посредством соответствующего модуля системы сроков доставки грузов с учетом исторических данных и методов машинного обучения.

В силу специфики маркетинга представленного решения предполагается использование продукта, в первую очередь, для компаний в сфере транспортной экспедиции и торговли, поэтому в настоящей работе описано применение продукта для них.

Информация о динамике продаж используется компаниями для управления всей своей деятельностью: расходами на маркетинг, управлением производством и складскими запасами, логистикой. Компаниям важно поддерживать минимальные остатки по всей товаропроводящей цепи, оптимальный уровень производства, но не допускать временного отсутствия товаров на складах и маркетплейсах, в том числе при колебаниях спроса, например, сезонного или возникающего на фоне повышенных промоционных активностей маркетплейсов.

Заказ товара на производстве необходимо делать за 2-4 месяца в зависимости от товара. На сегодня прогноз по стране в целом многие компании осуществляют выпуск продукции с приемлемой точностью (80-

90%), но по регионам (с учетом нестабильной логистики) точность прогнозов может быть существенно ниже – 40-60%. Это регулярно приводит либо к временному отсутствию товаров на складах и маркетплейсах.

Объем продаж популярных товаров может исчисляться миллиардами рублей в месяц. Поэтому прогноз сроков доставки таких товаров при их экспорте, импорте или внутрироссийской доставке позволит компаниям оптимизировать производственные и сбытовые бюджеты на сотни тысяч рублей в месяц.

В настоящее время у большинства компаний прогнозирование и загрузка транспортных средств осуществляется в 1С или аналогичных системах, при этом зачастую используются коэффициенты, определяемые компетентными сотрудниками экспертно. Технологии искусственного интеллекта используются на рынке точно, но вызывают большой интерес у всех причастных к процессу планирования деятельности логистических и торговых компаний сотрудников.

4.6 Экономическая оценка реализации проекта

Проект по разработке прототипа системы управления в сфере логистики и транспорта представляет собой сложное и многоаспектное мероприятие, требующее значительных инвестиций на этапе разработки и дальнейшего внедрения. Затраты на разработку включают в себя расходы на заработную плату квалифицированных специалистов, приобретение необходимого технического оборудования и программного обеспечения, а также обучение и консультации для команды разработчиков. Административные расходы, такие как аренда офиса и бухгалтерское сопровождение, также занимают значительную долю в структуре затрат.

Ожидаемая прибыль проекта формируется из нескольких ключевых источников. Во-первых, это плата за подписку, которую пользователи будут вносить ежемесячно или ежегодно за доступ к платформе. Во-вторых, проект

предусматривает получение комиссий с каждого заказа, забронированного через систему. Дополнительные доходы могут быть получены за счет продажи премиум-функций, таких как расширенная аналитика или приоритетное обслуживание, а также интеграционных сборов за подключение внешних систем и сервисов, таких как таможенные службы или страховые компании.

Рентабельность проекта оценивается на основе соотношения затрат и доходов. Анализ показывает, что высокий спрос на цифровые решения в логистике и значительные операционные выгоды от использования платформы создают благоприятные условия для привлечения большого числа клиентов. Автоматизация процессов бронирования и оптимизация маршрутов способствуют снижению затрат, что повышает привлекательность системы для пользователей. Заключение долгосрочных контрактов с крупными логистическими компаниями обеспечивает стабильный доход и финансовую устойчивость проекта.

Срок окупаемости проекта, согласно прогнозам, составляет от 1,5 до 2 лет с момента запуска. Такой срок обусловлен высоким спросом на цифровые логистические решения и значительными преимуществами, которые платформа предоставляет своим пользователям.

Финансовая устойчивость проекта поддерживается за счет диверсификации источников доходов. Помимо основной прибыли от подписок и комиссий, проект предполагает получение дополнительных доходов от интеграционных сервисов и премиум-функций. Важным аспектом является привлечение инвесторов, что позволит обеспечить необходимое финансирование на начальных этапах разработки и запуска платформы.

Организационная структура проекта включает в себя проектную команду, состоящую из специалистов по разработке, маркетингу, продажам и обслуживанию клиентов. Установление партнерских отношений с логистическими компаниями, маркетплейсами и внешними сервисами является ключевым элементом для расширения функционала платформы. Эффективная маркетинговая стратегия, включающая продвижение

платформы через цифровые каналы и участие в отраслевых мероприятиях, также играет важную роль в успешной реализации проекта.

Особое внимание необходимо уделить макроэкономической обстановке, которая оказывает значительное влияние на развитие логистической отрасли. Восстановление логистики после глобальных кризисов, развитие новых международных транспортных коридоров и политика импортозамещения, включая разработку отечественного программного обеспечения, создают дополнительные возможности для внедрения инновационных решений. В условиях увеличения объемов международных грузоперевозок и необходимости повышения их эффективности, цифровые платформы, обеспечивающие оптимизацию маршрутов и точное прогнозирование сроков доставки, становятся особенно актуальными.

Подробная проработка вопросов экономических аспектов проекта осуществлялась в рамках мероприятий программы «Стартап как диплом». Основные финансово-экономические, маркетинговые и организационные параметры проекта представлены в Паспорте проекта в Приложении Б.

ЗАКЛЮЧЕНИЕ

В ходе исследования было выявлено, что неотъемлемой частью современных систем в сфере транспорта и логистики являются инструменты оптимизации, использующие модели машинного обучения. Указанное обстоятельство требует от исследователей в рассматриваемой предметной области поиска эффективных способов обработки и анализа большого объема данных и выявления большого количества признаков, учет которых позволил бы создать наиболее точные модели прогнозирования.

На современном этапе применения методов машинного обучения в логистике и транспорте основное внимание уделяется интеграции искусственного интеллекта для повышения точности прогнозов и оптимизации процессов. Модели машинного обучения позволяют анализировать исторические данные и текущие условия, выявлять скрытые закономерности и принимать обоснованные решения, что значительно улучшает работу логистических систем и сокращает затраты.

В данной работе был проведен ряд экспериментов с использованием различных алгоритмов, и исходя из полученных результатов следует, что использование графовых нейронных сетей (GCN) и трансформеров для прогнозирования сроков доставки позволяет достичь более высокой точности по сравнению с традиционными методами. Эксперименты показали, что модели, включающие эти современные подходы, значительно улучшают результаты предсказаний.

В нашем исследовании было выявлено, что применение комплексного графового трансформера (IGT) в системе бронирования грузоперевозок и оптимизации доставки позволяет не только повысить точность прогнозирования сроков доставки, но и улучшить общую эффективность логистических процессов. Это достигается за счет интеграции высокоэффективных алгоритмов машинного обучения, которые учитывают как исторические данные, так и текущие условия.

Дальнейшее исследование в этой области может быть направлено на улучшение моделей машинного обучения путем включения большего количества факторов, влияющих на сроки доставки, а также на разработку методов обеспечения уверенности в прогнозах (предоставление интервалов доверия). Важно также продолжить интеграцию различных внешних систем и сервисов, чтобы обеспечить комплексный подход к управлению логистикой.

Автор полагает, что дальнейшее развитие и внедрение подобных систем в логистических компаниях может существенно повысить эффективность управления цепями поставок, сократить операционные затраты и улучшить качество обслуживания клиентов. Реализация таких решений в рамках программы "Стартап как диплом" показывает, что подобные инновационные проекты могут успешно внедряться и масштабироваться, обеспечивая значительное улучшение бизнес-процессов в сфере логистики.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Artificial Intelligence in Manufacturing and Supply Chain Market Analysis 2023-2029: Key Players, Estimated Size and Share, Opportunities, and Challenges [Электронный ресурс] // Digital Journal. – URL: <https://www.digitaljournal.com/pr/news/artificial-intelligence-in-manufacturing-and-supply-chain-market-analysis-2023-2029-key-players-estimated-size-and-share-opportunities-and-challenges#ixzz822WKBncW> (дата обращения: 04.04.2024) – Текст : электронный.
2. Chen T. Xgboost: A scalable tree boosting system / T. Chen // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – 2016. – №1. – 785-794 p.
3. Cheng K. Skeleton-based action recognition with shift graph convolutional network / K. Cheng // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2020. – №1. – 183-192 p.
4. Cho K. Learning phrase representations using RNN encoder-decoder for statistical machine translation / K. Cho // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). – 2014. – №1. – 1724-1734 p.
5. Corso G. Principal neighbourhood aggregation for graph nets / G. Corso // Proceedings of the 34th International Conference on Neural Information Processing Systems. – 2020. – №34. – 534-542 p.
6. Cui R. Sooner or later? Promising delivery speed in online retail / R. Cui // Promising Delivery Speed in Online Retail. – 2020. – №6. – 232-239 p.
7. Defferrard M. Convolutional neural networks on graphs with fast localized spectral filtering / M. Defferrard // Advances in neural information processing systems. – 2016 – №29. – 88-96 p.
8. Devlin J. Bert: Pre-training of deep bidirectional transformers for language understanding / J. Devlin // Conference of the North American Chapter of the Association for Computational Linguistics. – 2019. – №5. – 21-29 p.

9. Doole M. Estimation of traffic density from drone-based delivery in very low level urban airspace / M. Doole // Journal of Air Transport Management. – 2020. – №88. – 862-893 p.
10. Dosovitskiy A. An image is worth 16x16 words: Transformers for image recognition at scale / A. Dosovitskiy // International Conference on Learning Representations. – 2021. – №2. – 99-117 p.
11. Dwivedi V. P. generalization of transformer networks to graphs / V.P. Dwivedi // Neural Networks. – 2020. – №1. – 106-118 p.
12. Fang X. Constgat: Contextual spatial-temporal graph attention network for travel time estimation at Baidu Maps / X. Fang // Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. – №26. – 2020. – 2697-2705 p.
13. Fu K. Compacteta: A fast inference system for travel time prediction / K. Fu // Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. – №26. – 2020. – 3337-3345 p.
14. Glorot X. Understanding the difficulty of training deep feedforward neural networks / X. Glorot // Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. JMLR Workshop and Conference Proceedings. – №6. – 2010. – 249-256 p.
15. Hamilton W. L. Inductive representation learning on large graphs / W. L. Hamilton // Proceedings of the 31st International Conference on Neural Information Processing Systems. – №4. – 2017. – 1025-1035 p.
16. Hong H. HetETA: Heterogeneous Information Network Embedding for Estimating Time of Arrival / H. Hong // Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. – №26. – 2020. – 2444-2454 p.
17. Hu J. Stochastic origin-destination matrix forecasting using dual-stage graph convolutional, recurrent neural networks / J. Hu // 2020 IEEE 36th International Conference on Data Engineering (ICDE). IEEE. – №2. – 2020. – 1417-1428 p.

18. Hu Z. Heterogeneous graph transformer / Z. Hu // Proceedings of The Web Conference 2020. – №1. – 2020. – 2704-2710 p.
19. Kazemi S. M. Representation Learning for Dynamic Graphs: A Survey / S. M. Kazemi // Journal of Machine Learning Research. – №70. – 2020. – 1-73 p.
20. Kingma D. P. A method for stochastic optimization / D. P. Kingma // International Conference on Learning Representations. – №6. – 2015. – 34-48 p.
21. Kipf T. N. Semi-supervised classification with graph convolutional networks / T. N. Kipf // International Conference on Learning Representations. – №4. – 2017. – 45-50 p.
22. Li Q. Deeper insights into graph convolutional networks for semi-supervised learning / Q. Li // Thirty-Second AAAI Conference on Artificial Intelligence. – №1. – 2018. – 64-68 p.
23. Li Y. Multi-Task Representation Learning for Travel Time Estimation / Y. Li // Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Association for Computing Machinery – №24. – 2018.– 1695-1704 p.
24. Lian J. xDeepFM: Combining explicit and implicit feature interactions for recommender systems / J. Lian // Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. – №24. – 2018.– 1754-1763 p.
25. Paszke A. Pytorch: An imperative style, high-performance deep learning library / A. Paszke // Advances in neural information processing systems. – №32. – 2019.– 8026-8037 p.
26. Pedregosa F. Scikit-learn: Machine learning in Python / F. Pedregosa // Journal of Machine Learning Research. – №12. – 2011.– 2825-2830 p.
27. Ragesh R. heterogeneous graph convolutional networks for text classification / R. Ragesh // Proceedings of the 14th ACM International Conference on Web Search and Data Mining. – №14. – 2021.– 860-868 p.

28. Su Y. Characterizing the environmental impact of packaging materials for express delivery via life cycle assessment / Y. Su // Journal of Cleaner Production. – №274. – 2020. – 122-161 p.
29. Vaswani A. Attention is all you need / A. Vaswani // Advances in neural information processing systems. – №1. – 2017. – 5998-6008 p.
30. Veličković P. Graph attention networks / P. Veličković // Proceedings of the Sixth International Conference on Learning Representations. – №6. – 2018. – 21-29 p.
31. Vladimir Z. Electronic commerce: structures and issues / Z. Vladimir // International Journal of Electronic Commerce. – №1. – 1996. – 3-23 p.
32. Wang H. A Simple Baseline for Travel Time Estimation Using Large-Scale Trip Data / H. Wang // ACM Trans. Intell. Syst. Technol. – №2. – 2019. – 10-15 p.
33. Wang J. Fraudster detection via graph convolutional networks in online app review system / J. Wang // Companion Proceedings of The 2019 World Wide Web Conference. – №2. – 2019. – 310-316 p.
34. Wang Z. Learning to estimate the travel time / Z. Wang // Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. – №24. – 2018. – 858-866 p.
35. Wu F. Simplifying graph convolutional networks / F. Wu // International conference on machine learning. PMLR. – №1. – 2019. – 6861-6871 p.
36. Wu F. A spatial-temporal sequential neural network model for estimating time of arrival in package delivery system / F. Wu // Proceedings of the AAAI Conference on Artificial Intelligence. – №33. – 2019. – 774-781 p.
37. Yao L. Graph convolutional networks for text classification / L. Yao // Proceedings of the AAAI conference on artificial intelligence. – №33. – 2019. – 7370-7377 p.
38. Yun S. Graph transformer networks / S. Yun // Advances in Neural Information Processing Systems. – №32. – 2019. – 11983-11993 p.

39. Zhou D. A data-driven graph generative model for temporal interaction networks / D. Zhou // Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. – №26. – 2020. – 401-411 p.
40. Zhou X. Layer-refined Graph Convolutional Networks for Recommendation / X. Zhou // Neural Networks. – №7. – 2022. – 96-99 p.
41. Zhou X. SelfCF: A Simple Framework for Self-supervised Collaborative Filtering / X. Zhou // Neural Networks. – №7. – 2021. – 87-90 p.
42. Zhou X. Bootstrap Latent Representations for Multimodal Recommendation / X. Zhou // Neural Networks. – №4. – 2022. – 16-31 p.
43. Zhu L. Order Fulfillment Cycle Time Estimation for On-Demand Food Delivery / L. Zhu // Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. – №26. – 2020. – 2571-2580 p.
44. Zhu Z. Hgcn: A heterogeneous graph convolutional network-based deep learning model toward collective classification / Z. Zhu // Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. – №26. – 2020. – 1161-1171 p.
45. Zhou X. Inductive Graph Transformer for Delivery Time Estimation / X. Zhou // Neural Networks. – №2. – 2022. – 17-21 p.
46. Zeng J. Graph Neural Networks for Recommendation: Learning from User Feedback / J. Zeng // IEEE Transactions on Neural Networks and Learning Systems. – №34. – 2023. – 1234-1245 p.
47. Zesty interface of web-portal [Электронный ресурс] // Figma. – URL: <https://www.figma.com/file/pIJOCGaZwqafenfgGApjUR/LogiAnalytics?type=design&node-id=0%3A1&mode=design&t=e945vKJ15TiIWwaq-1> (дата обращения: 04.04.2024). – Текст : электронный.

Техническое задание на разработку личного кабинета платформы

1 Введение

1.1 Общее описание продукта, название, идея

Разрабатываемый продукт – сквозная цифровая платформа, которая позволяет осуществлять онлайн-бронирование грузоперевозок и оптимизацию доставки по стоимости и по времени.

1.2. Целевая аудитория

Поставщики товаров (розничные и мелкооптовые экспортеры и импортеры), фулфилмент-сервисы, перевозчики, экспедиторы.

1.3. Какие цели стоят перед проектом

Разработка прототипа.

1.4. Ценность. Какую бизнес-ценность (business value) несет цифровой продукт

Создаваемый цифровой продукт позволит пользователям снизить затраты на доставку грузов до 30%, повысить эффективность поиска подрядчика и построения логистических цепочек на 60% исходя из того, что система будет работать в режиме «одного окна» и использовать технологии машинного обучения.

1.5. Структура проекта. Категории, подкатегории, страницы, окна:

ПК / Планшет / Мобильный:

1.5.1. Страница Авторизация

1.5.2. Страница Регистрация

1.5.3. Рор-уп – Завершение регистрации

1.5.4. Страница Восстановление пароля

1.5.5. Рор-уп – Восстановление пароля – данные отправлены на почту

1.5.6. Страница – Контактная информация (Личные данные)

1.5.7. Рор-уп – Выход из личного кабинета (“Вы действительно хотите выйти?”)

- 1.5.8. Страница Отправления: создать
- 1.5.9. Pop-up – Отправление создано
- 1.5.10. Страница Отправления: черновики
- 1.5.11. Страница Отправления: история отправок
- 1.5.12. Страница Адресная книга
- 1.5.13. Страница Создание контакта

2. Технические параметры проекта

2.1. Условия разработки. Ведение проекта осуществлялось под git-ом <https://github.com/>. Каждый разработчик работает на своей ветке. Результаты после повторной проверки сливались поэтапно в master. Ведение разработки осуществлялось на dev. сервере, после проводился перенос на прод основного домена (сервер).

2.2. Верстка: параметры представлены ниже, в разделе “Верстка”.

2.3. Язык проекта. Личный кабинет выполнен на русском языке.

2.4. Состав и содержание осуществленных работ по вводу сайта (личного кабинета) в эксплуатацию. На хостинг перенесена программа (сайт), сверстаный шаблон дизайна, структура и наполнение базы данных. VDS соответствует требованиям: CPU: 1x3.2 МГц ; DDR4: 8 Gb; SSD: 40 Gb; Сеть: 1 Gbit/s.

3. Выполнение работ по проекту

3.1. Формирование плана работы, формирование технического задания, а также выполнение прототипа.

3.2. Дизайн (подробнее – ниже, в разделе “Дизайн”).

3.3. Верстка (подробнее – ниже, в разделе “Верстка”).

3.4. Программирование front-end на React и back-end на NodeJS, интеграция с разработанной моделью машинного обучения (подробнее – п.6).

Продолжение ПРИЛОЖЕНИЯ А

3.5. Тестирование личного кабинета на тестовом сервере.

3.6. Бекап готового проекта и перенос на сервер.

3.7. Настройка SSL-сертификата

4. Дизайн

4.1. Макет дизайна информационной системы выполнен на основе прототипов, разработанных автором данной работы в Figma:

<https://www.figma.com/file/pIJOcGaZwqafenfgGApjUR/LogiAnalytics?type=design&node-id=0%3A1&mode=design&t=e945vKJ15TiIWwaq-1>

4.2. Страницы дизайна выполнены согласно описанных в структуре проекта (п.1.5.)

4.3. Страницы макета выполнялись в режиме Т&М, пока не были учтены все замечания научного руководителя. Ориентир времени, затраченного на выполнение – от 4х часов на страницу.

4.4. На страницах макета были обязательно отражены:

4.4.1. Страница Авторизация: предусмотрена возможность пользователя зайти на сайт под логином/паролем, через российскую соцсеть VK.

4.4.2. Страница Регистрация: предусмотрена возможность пользователя зарегистрироваться, принять условия обработки персональных данных.

4.4.3. Рор-уп – Завершение регистрации – оповещение с возможностью повторно отправить письмо на почту или вернуться к авторизации.

4.4.4. Страница Восстановление пароля – ввод почтового адреса, куда отправить ссылку для восстановления пароля.

4.4.5. Рор-уп – Восстановление пароля – данные отправлены на почту – оповещение пользователя с возможностью вернуться к авторизации.

Продолжение ПРИЛОЖЕНИЯ А

4.4.6. Страница Контактная информация (Личные данные) – возможность ввода личных данных в редактируемые поля.

4.4.7. Рор-уп – Выход из личного кабинета (Вы действительно хотите выйти?)

4.4.8. Страница Отправления: создать – ввод данных:

4.4.8.1. получателя, отправителя; возможность выбора из адресной книги существующих Клиентов;

4.4.8.2. информации о грузе: текстовые поля для ввода, предусмотрена возможность прикрепить документ;

4.4.8.3. информация об упаковке и дополнительном страховании;

4.4.8.4. информация о дате отправления – выбор из календаря; время отправления – с двумя вариантами выбора; текстовое поле для ввода инструкции курьеру;

4.4.8.5. размещены ссылки на уведомления “условия перевозки”, “запрещенные грузы”,

4.4.8.6. предусмотрена возможность создания заказа, отмены заказа и создания черновика.

4.4.9. Рор-уп – Отправление создано с возможностью редактировать отправление.

4.4.10. Страница Отправления: черновики – разместить сохраненные заполненные заказы с возможностью возобновить заказ из черновика.

4.4.11. Страница Отправления: история – история заказов, которые можно отсортировать по Статусу (статусы у каждого заказа присвоены), Дате, (другой вариант по наименованию колонки); поиск по заказам; возможность заказ Изменить, Отменить, Отправить повторно, Скопировать, Распечатать.

4.4.12. Страница Адресная книга – отображение карточек контактов, создание и редактирование контактов.

4.4.13. Страница Создание контакта – редактируемые поля для ввода данных.

5. Верстка

5.1. Кроссбраузерность:

Обеспечена кроссбраузерность в нижеперечисленных браузерах (версии браузеров на момент создания проекта верстки): Google Chrome, Mozilla Firefox, Opera, Yandex Browser, Edge, Safari (iOS).

5.2. Retina экраны:

В связи с прототипируемым характером цифрового продукта на данной стадии использовалась графика без применения дополнительных манипуляций по улучшению качества отображения на retina-ready экранах.

5.3. Респонсив:

Разработаны макеты респонсива в разрешениях на 1920px и на 320px

5.4. Валидация:

Предусмотрена валидация форм, содержащих адрес

5.5. Parallax:

С учетом того, что на мобильных устройствах параллакс обычно отключают (это является хорошей практикой в пользу производительности), то в рамках настоящего проекта на сайте предусмотрен эффект параллакса с применением обычного способа, реализованный с помощью стилей, который не работает на мобильных устройствах.

5.6. Формы:

Отправку форм на email в текущей стадии продукта не предусмотрена.

5.7. Макет:

Блоки в мобильных макетах перестраиваются согласно логике их потока в десктопных макетах. Перестройка блоков, которые меняют свое расположение на респонсиве, не предусмотрена.

5.8. Шрифты:

В макете используются шрифт Century Gothic.

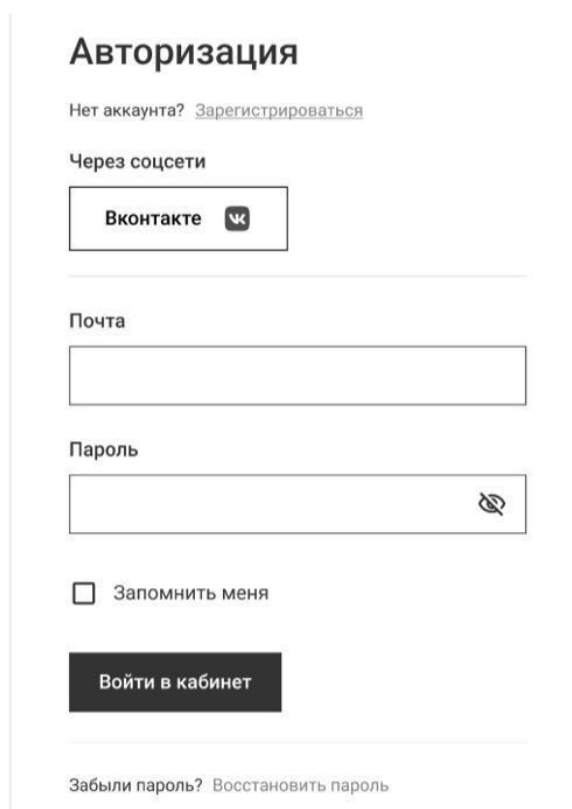
5.9. Оптимизация скорости загрузки сайта:

В рамках проекта обеспечена валидность кода, а также его лаконичность, оптимизация изображений, которые находятся в макете. Вместе с тем не гарантировано то, что веб-страница будет показывать высокие результаты по показателям скорости загрузки, например google pagespeed или gtmetrix, так как это зависит не только от верстки, а по большей части – от настроек хостинга, от кеширования на стороне сервера и т.п.

5.10. Pixel-Perfect:

По умолчанию сделана верстка с допустимой погрешностью от макета 5 пикселей (px).

6. Функционал



The screenshot shows a web form for authorization. At the top, the title 'Авторизация' is displayed. Below it, there is a link 'Нет аккаунта? Зарегистрироваться'. A section titled 'Через соцсети' contains a button for 'Вконтакте' with its logo. Below this, there are input fields for 'Почта' and 'Пароль'. The password field has an eye icon for toggling visibility. A checkbox labeled 'Запомнить меня' is located below the password field. A dark button labeled 'Войти в кабинет' is positioned below the checkbox. At the bottom, there are links for 'Забыли пароль?' and 'Восстановить пароль'.

Рисунок 1 – Форма авторизации

Продолжение ПРИЛОЖЕНИЯ А

6.1. Функционал страницы Авторизация:

6.1.1. Авторизация через соцсети (Вконтакте)

6.1.2. Поля авторизационных данных (почта и пароль)

6.1.3. Запомнить логин и пароль

Авторизация

Нет аккаунта? [Зарегистрироваться](#)

Через соцсети

Вконтакте 

Почта

Пароль 

☐ Запомнить меня

Войти в кабинет

[Забыли пароль?](#) [Восстановить пароль](#)

Рисунок 1 – Форма авторизации

6.2. Функционал страницы Регистрация:

6.2.1. Ввод персональных данных

6.2.2. Согласие с условиями обработки персональных данных

6.2.3. Отправка письма на введенную почту для подтверждения регистрации

Регистрация

Уже есть аккаунт? [Авторизироваться](#)

Имя

Почта

Пароль

☐ Я согласен с условиями обработки [персональных данных](#)

Зарегистрироваться

Рисунок 2 – Форма регистрации

6.3. Функционал pop-up Завершение регистрации:

6.3.1. Возврат к авторизации

6.3.2. Отправка письма на ранее указанную почту

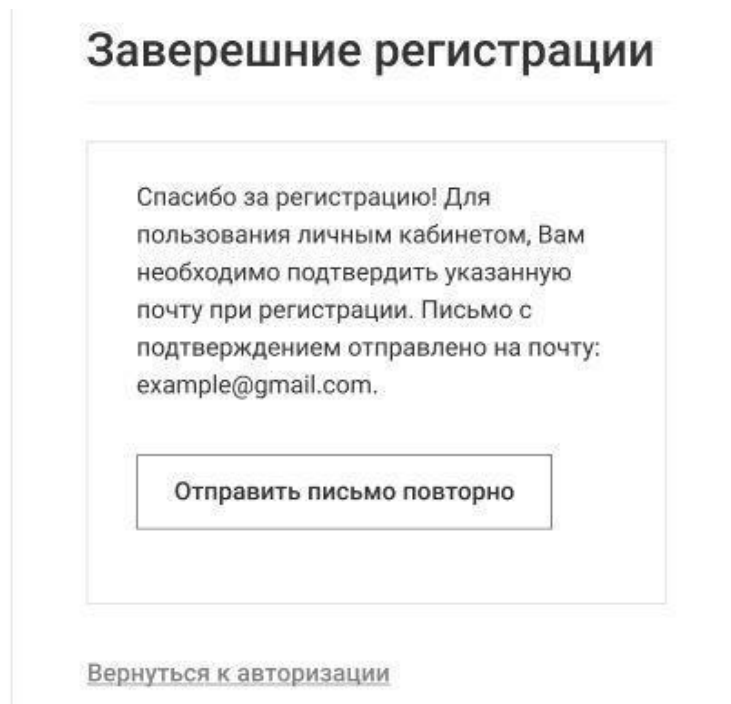


Рисунок 3 – Оповещение завершения регистрации

6.4. Функционал pop-up Восстановление пароля:

6.4.1. Ввод почты

6.4.2. Отправка письма для восстановления пароля на указанную почту

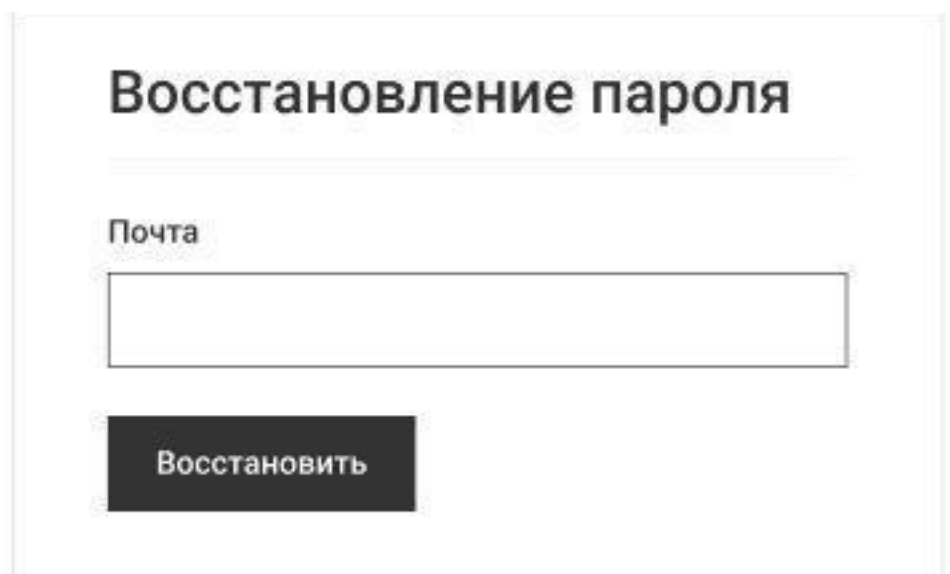


Рисунок 4 – Форма восстановления пароля

6.5. Функционал pop-up Восстановление пароля (оповещение):

6.5.1. Отправка данных для восстановления пароля на указанную почту

6.5.2. Возврат к авторизации

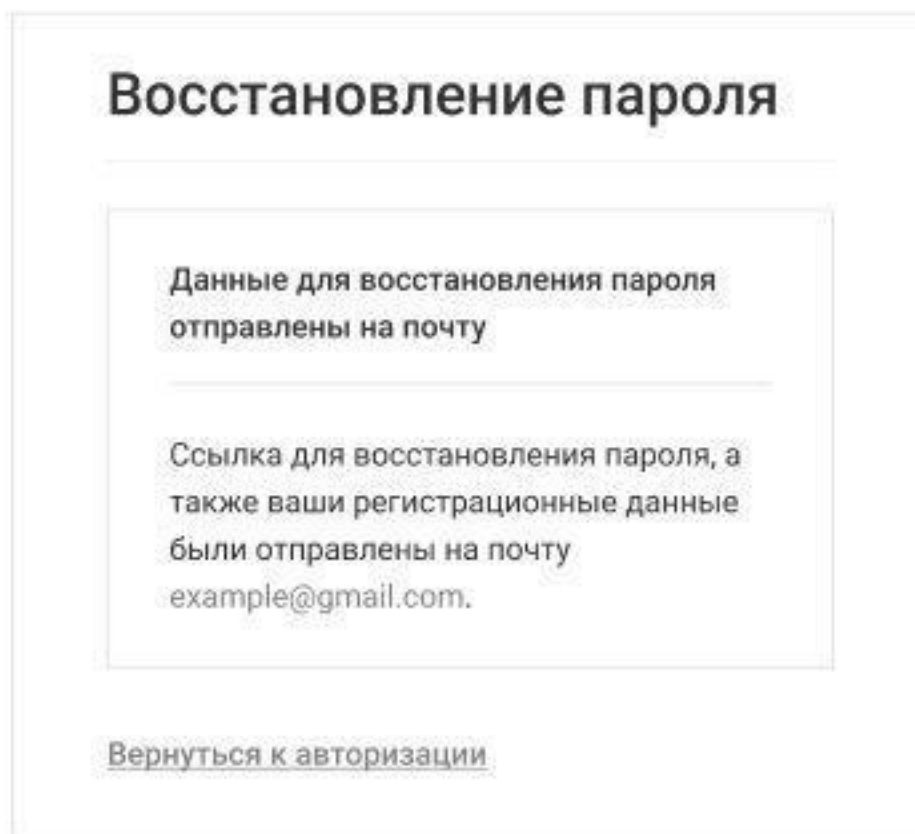


Рисунок 5 – Оповещение о восстановлении пароля

6.6. Функционал страницы Личные данные:

6.6.1. Ввод личных данных в редактируемые поля

6.6.2. Сохранение данных

6.6.3. Возможность редактирования личных данных

6.6.4. Выход из личного кабинета

The screenshot shows a web application interface with a teal sidebar on the left and a main content area. The sidebar contains a menu with the following items: 'Личные данные' (highlighted), 'Отправления', 'Адресная книга', 'Расчет стоимости', 'Лицевой счет', and 'Настройки'. The main content area is titled 'Личные данные' and contains several input fields for personal information: 'Фамилия', 'Имя', 'Отчество', 'Компания', 'Почта', and 'Телефон'. Each field is represented by a horizontal text box. At the bottom of the form is a dark button labeled 'Сохранить'.

Рисунок 6 – Страница личных данных

6.7. Функционал страницы Оформление отправления (Создать):

The screenshot shows a web application interface for the 'Shipping Form' page. The teal sidebar on the left has a menu with the following items: 'Личные данные', 'Отправления' (highlighted), 'Создать', 'Черновики', 'История', 'Отследить', 'Адресная книга', 'Расчет стоимости', 'Лицевой счет', and 'Настройки'. The main content area is titled 'Оформление отправления' and contains a section '1. Контактная информация'. This section includes a dropdown menu labeled 'Выбрать из адресной книги' with 'Пункт' selected. Below this are two columns of input fields. The left column is for the 'Отправитель' (Sender) and the right column is for the 'Получатель' (Recipient). Each column contains fields for 'Фамилия', 'Имя', 'Отчество', 'Компания', 'Почта', and 'Телефон'.

Рисунок 7 – Страница оформления отправления

Продолжение ПРИЛОЖЕНИЯ А

6.7.1. Ввод личных данных отправителя и получателя в редактируемые поля

6.7.2. Выбор и заполнение данных из адресной книги

6.7.3. Возможность редактирования личных данных

6.7.4. Заполнение информации о грузе:

6.7.4.1. описание груза – текстовое поле с подфономком описания

6.7.4.2. код товара – текстовое поле

6.7.4.3. кол-во, штук – текстовое поле

6.7.4.4. валюта

6.7.4.5. вес нетто/брутто – текстовое поле, кг

6.7.4.6. где был произведен – текстовое поле, кг

6.7.4.7. вывод информационных окошек с пояснением (текстовое поле)

6.7.4.8. прочие сборы – текстовое поле с подфономком описания

6.7.4.9. стоимость сборов – текстовое поле

6.7.4.10. валюта

6.7.4.11. Общий вес нетто = количество штук * вес нетто, кг

6.7.4.12. Общий вес брутто = количество штук * вес брутто, кг

6.7.4.13. Общая стоимость товаров = стоимость наименования * количество .

6.7.4.14. Добавить еще одно наименование – выпадает точно такой же блок с полями для ввода данных об еще одном грузе.

6.7.4.15. Если товаров несколько – то блок с общим весом нетто, брутто и общей стоимостью считается один раз и составляет: Общий вес нетто = сумме общих весов нетто, Общий вес брутто = сумме общих весов брутто, Общая стоимость товаров = сумме общих стоимостей товаров.

2. Информация о грузе

Документы Грузы

Создать инвойс У меня есть инвойс

Опишите содержание груза

Код отправителя

☒ Дополнительная страховка

Оценочная стоимость груза, ₽

Рисунок 8 – Форма информации о грузе

6.7.5. Заполнение информации об упаковке:

6.7.5.1. Ввод данных – текстовые поля (Количество – строго в шт, Вес – строго в кг, Длина/Ширина/Высота – строго в см)

6.7.5.2. Добавить место – вывод такого же блока с полями для ввода.

3. Упаковка

Введите количество, шт Вес, кг

Длина, см Ширина, см Высота, см

Добавить место +

Рисунок 9 – Форма информации об упаковке

6.7.6. Завершение (функционал):

6.7.6.1. Ввод даты – ввод вручную и/или выбор из календаря

6.7.6.2. Выбор варианта отправления (инпут)

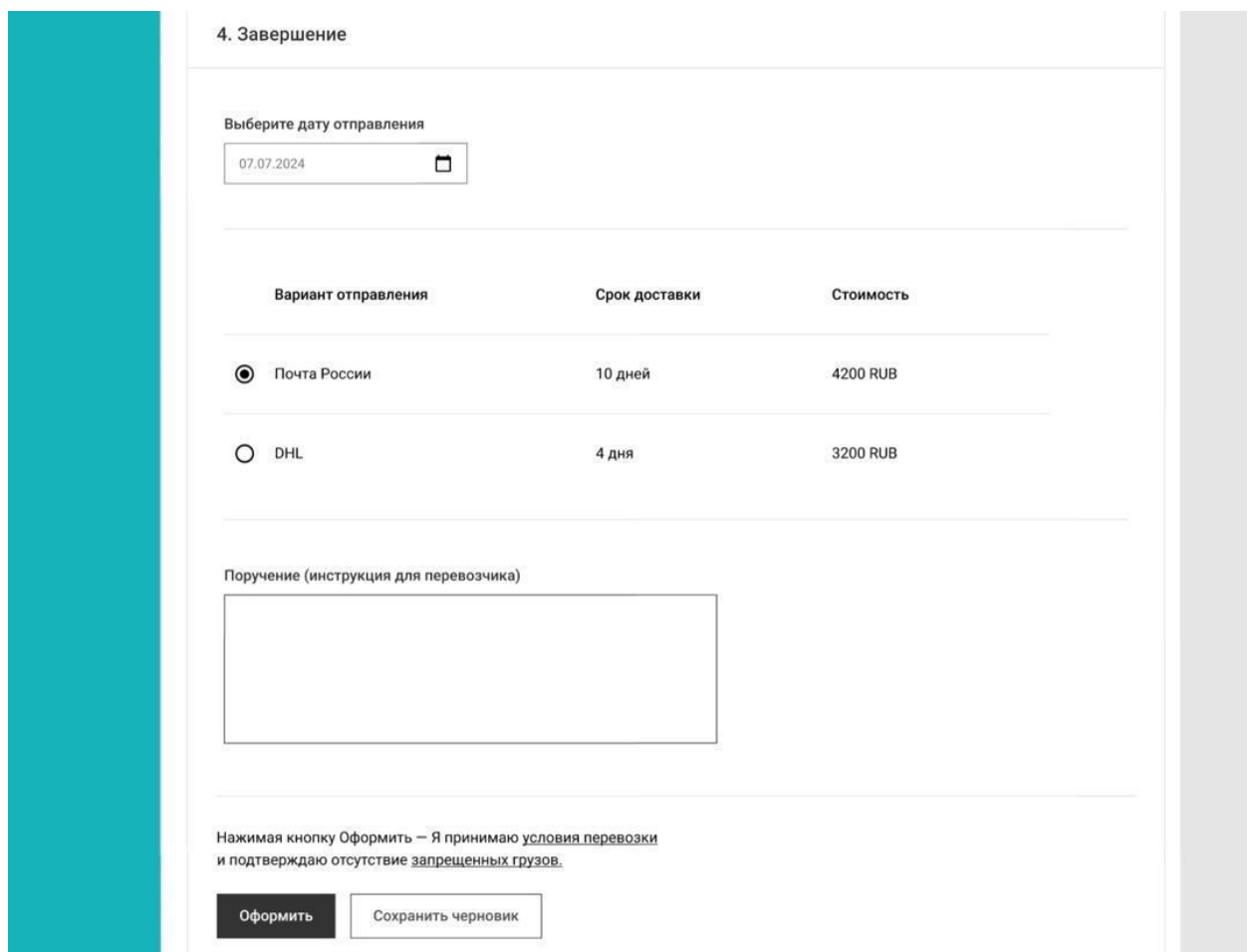
Продолжение ПРИЛОЖЕНИЯ А

6.7.6.3. Текстовое поле – описание

6.7.6.4. Сохранить черновик – при клике – отправка заказа в черновики

6.7.6.5. Оформить: при клике на кнопку – генерация документов: инвойс и csv-файл, поля которых заполняются автоматически из внесенных выше.

6.7.6.6. Отправка сгенерированных документов на почту.



4. Завершение

Выберите дату отправления

07.07.2024

Вариант отправления	Срок доставки	Стоимость
☒ Почта России	10 дней	4200 RUB
☐ DHL	4 дня	3200 RUB

Поручение (инструкция для перевозчика)

Нажимая кнопку Оформить – Я принимаю условия перевозки и подтверждаю отсутствие запрещенных грузов.

Оформить Сохранить черновик

Рисунок 10 – Форма завершения оформления отправления

ПРИЛОЖЕНИЕ Б

Паспорт стартап-проекта

КРАТКАЯ ИНФОРМАЦИЯ О СТАРТАП-ПРОЕКТЕ		
1	Название стартап-проекта* 	Разработка прототипа системы управления в сфере транспорта и логистики, основанной на технологии обработки данных с использованием ML
2	Тема стартап-проекта* <i>Указывается тема стартап-проекта в рамках темы акселерационной программы, основанной на Технологических направлениях в соответствии с перечнем критических технологий РФ, Рынках НТИ и Сквозных технологиях.</i>	Информационные технологии в B2B сегменте
3	Технологическое направление в соответствии с перечнем критических технологий РФ*	Информационные технологии
4	Рынок НТИ	Технет
5	Сквозные технологии	Технологии обработки и передачи данных: Искусственный интеллект, включая технологии машинного обучения
6	Фокусная тематика из перечня ФСИ https://fasie.ru/programs/programma-start/fokusnye-tematiki.php)	Нейрокомпьютерные технологии и эволюционные алгоритмы

Продолжение ПРИЛОЖЕНИЯ Б

КОМАНДА ПРОЕКТА

Информация о лидере и участниках стартап-проекта							
Лидер стартап-проекта*							Иванов Иван Александрович
Команда стартап-проекта (участники стартап-проекта, которые работают в рамках акселерационной программы)							
№	Unti ID	Leader ID	ФИО	Роль в проекте	Тел., почта	Должность (при наличии)	Опыт и квалификация (краткое описание)
1			Иванов Иван Александрович	СЕО			10 лет в управлении проектами, 2 года в транспорте и логистике, 2 года в инженерии машинного обучения, студент магистратуры по программе "Инженерия машинного обучения"
2			Смирнов Антон Викторович	СТО			9 лет в управлении проектами, 6 лет в ИТ, 2 года в инженерии машинного обучения, студент магистратуры по программе "Инженерия машинного обучения"
			Иванов Андрей Александрович	финансист			5 лет в сфере проектного финансирования и банкинга, 3 года в логистике и управлении стартапами
3			Юничева Мария Дмитриевна	маркетолог			5 лет в сопровождении медиа-проектов, 3 года в логистике

Продолжение ПРИЛОЖЕНИЯ Б

Опыт взаимодействия с другими институтами развития	
Участвовал ли кто-либо из членов проектной команды в «Акселерационно-образовательных интенсивах по формированию и преакселерации команд»	+
Участвовал ли кто-либо из членов проектной команды в программах «Диагностика и формирование компетентностного профиля человека / команды»	+
Перечень членов проектной команды, участвовавших в программах Leader ID и АНО «Платформа НТИ»	—
Участие в программе «Стартап как диплом»	+
Участие в образовательных программах повышения предпринимательской компетентности и наличие достижений в конкурсах АНО «Россия – страна возможностей»	—
Для исполнителей по программе УМНИК	—
Номер контракта и тема проекта по программе «УМНИК»	—
Роль лидера по программе «УМНИК» в заявке по программе «Студенческий стартап»	—

РАЗРАБОТКА И РЕАЛИЗАЦИЯ СТАРТАП-ПРОЕКТА		
7	Аннотация проекта* <i>Указывается краткая информация (не более 1000 знаков, без пробелов) о стартап-проекте (краткий реферат проекта, детализация отдельных блоков предусмотрена другими разделами Паспорта): цели и задачи проекта, ожидаемые результаты, области применения результатов, потенциальные потребительские сегменты</i>	Реализуется проект по созданию продукта, представляющий собой интеллектуальную систему, которая позволяет осуществлять онлайн-бронирование отправок сборных грузов и прогнозировать сроки доставки
Характеристика проблемы, на решение которой направлен стартап-проект		
8	Описание проблемы* <i>Необходимо детально описать проблему, которая приводит к необходимости разработки и внедрения нового продукта/решения</i>	За последние годы на рынке трансграничной торговли обострились следующие проблемы: — неустойчивость цепочек поставок и увеличение сроков доставки грузов — рост стоимости международных перевозок и снижение российского несырьевого экспорта По указанным причинам компании теряют клиентов, время и деньги. И это только прямые убытки, не считая испорченной репутации, штрафов за нарушение сроков и снижения удовлетворенности клиентов

Продолжение ПРИЛОЖЕНИЯ Б

9	Какая часть проблемы решается (может быть решена) * <i>Необходимо детально раскрыть вопрос, поставленный в соответствии потребностями потребителей, описав, какая часть проблемы или вся проблема решается с помощью стартап-проекта</i>	Описанные пользовательские боли будут реализованы в рамках функциональности создаваемой системы
10	«Держатель» проблемы, его мотивации и возможности решения проблемы с использованием продукции* <i>Необходимо детально описать взаимосвязь между выявленной проблемой и потенциальным потребителем (см. пункты продукт, проблему, каналы)</i>	Как малые и средние предприятия, так и крупные корпорации, работающие в различных отраслях, где логистика играет важную роль в их бизнес-процессах. Категории потенциальных потребителей: участники рынка электронной коммерции, производственные компании, транспортные компании
11	Каким способом будет решена проблема* <i>Необходимо описать детально, как именно образом ваши товары и услуги помогут потребителям справиться с проблемой</i>	Будет создана интеллектуальная система
12	Оценка потенциала «рынка» и рентабельности бизнеса* <i>Необходимо привести кратко обоснование сегмента и доли рынка, потенциальные возможности для масштабирования бизнеса, а также детально раскрыть информацию</i>	Целевой рынок: логистические, производственные, розничные и дистрибьюторские компании. Размер рынка: 1,6 млрд. руб. *, растущий. * Доклад НИУ ВШЭ "Цифровая трансформация: ожидания и реальность" 2022 г.
Базовая бизнес-идея		
13	Какой продукт (товар/ услуга/ устройство/ ПО/ технология/ процесс и т.д.) будет продаваться* <i>Указывается максимально понятно и емко информация о продукте, лежащем в основе стартап-проекта, благодаря реализации которого планируется получать основной доход</i>	Интеллектуальная система, которая позволяет осуществлять онлайн-бронирование перевозок грузов и прогнозировать сроки доставки

Продолжение ПРИЛОЖЕНИЯ Б

14	Какую и чью (какого типа потребителей) проблему решает* <i>Указывается максимально и емко информация о проблеме потенциального потребителя, которую (полностью или частично) сможет решить ваш продукт</i>	Проблемы, связанные с логистикой, возникающие у следующих категорий потенциальных потребителей: участники рынка электронной коммерции, производственные компании, транспортные компании
15	Потенциальные потребительские сегменты* <i>Указывается краткая информация о потенциальных потребителях с указанием их характеристик (детализация предусмотрена в части 3 данной таблицы): для юридических лиц – категория бизнеса, отрасль, и т.д.; для физических лиц – демографические данные, вкусы, уровень образования, уровень потребления и т.д.; географическое расположение потребителей, сектор рынка (B2B, B2C и др.)</i>	Целевой рынок: логистические, производственные, розничные и дистрибьюторские компании. Размер рынка: 1,6 млрд. руб.*, растущий. * Доклад НИУ ВШЭ "Цифровая трансформация: ожидания и реальность" 2022 г.
16	На основе какого научно-технического решения и/или результата будет создан продукт (с указанием использования собственных или существующих разработок) * <i>Указывается необходимый перечень научно-технических решений с их кратким описанием для создания и выпуска на рынок продукта</i>	Собственное решение
17	Бизнес-модель* <i>Указывается кратко описание способа, который планируется использовать для создания ценности и получения прибыли, в том числе, как планируется выстраивать отношения с потребителями и поставщиками, способы привлечения финансовых и иных ресурсов, какие каналы продвижения и сбыта продукта планируется использовать и развивать, и т.д.</i>	лицензии; комиссии от транзакций; плата за дополнительные услуги

Продолжение ПРИЛОЖЕНИЯ Б

18	Партнеры (поставщики, продавцы) <i>Указывается информация о Вашем представлении о партнерах/ поставщиках/продавцах на момент выхода предприятия на самоокупаемость, т.е. о том, как может быть.</i>	перевозчики, Ассоциация «Цифровая логистика и транспорт»
19	Коллектив (характеристика будущего предприятия) <i>Указывается информация о составе коллектива (т.е. информация по количеству, перечню должностей, квалификации), который Вы представляете на момент выхода предприятия на самоокупаемость. Вероятно, этот состав шире и(или) будет отличаться от состава команды по проекту, но нам важно увидеть, как Вы представляете себе штат созданного предприятия в будущем, при переходе на самоокупаемость</i>	СЕО, СТО, финансист, маркетолог и т.д.
20	Основные конкуренты* <i>Кратко указываются основные конкуренты (не менее 5)</i>	Agora Freight Магистрали
21	Ценностное предложение* <i>Формулируется объяснение, почему клиенты должны вести дела с вами, а не с вашими конкурентами, и с самого начала делает очевидными преимущества ваших продуктов или услуг</i>	мы снижаем затраты на доставку грузов до 30% * повышаем эффективность поиска подрядчика и построения логистических цепочек на на 60-70% * исходя из того, что работаем в режиме “одного окна” и используем нейронные сети,
22	Обоснование реализуемости (устойчивости) бизнеса (конкурентные преимущества (включая наличие уникальных РИД, действующих промышленных партнеров, доступ к ограниченным ресурсам и т.д.); дефицит, дешевизна, уникальность и т.п.)* <i>Приведите аргументы в пользу реализуемости бизнес-идеи, в чем ее полезность и востребованность продукта по сравнению с другими продуктами на рынке, чем обосновывается потенциальная прибыльность бизнеса, насколько будет бизнес устойчивым</i>	Устойчивость бизнеса обусловлена трендом цифровизации логистики

Продолжение ПРИЛОЖЕНИЯ Б

	Характеристика будущего продукта	
23	Основные технические параметры, включая обоснование соответствия идеи/задела тематическому направлению (лоту)* <i>Необходимо привести основные технические параметры продукта, которые обеспечивают их конкурентоспособность и соответствуют выбранному тематическому направлению</i>	Соответствие потребностям пользователей, отказоустойчивость, искусственный интеллект в основе части функциональных возможностей
24	Организационные, производственные и финансовые параметры бизнеса* <i>Приводится видение основателя (-лей) стартапа в части выстраивания внутренних процессов организации бизнеса, включая партнерские возможности</i>	Требуемые инвестиции – 20 млн. руб. Безубыточность – 36 месяцев Возврат инвестиций – 48 месяцев
25	Основные конкурентные преимущества* <i>Необходимо привести описание наиболее значимых качественных и количественных характеристик продукта, которые обеспечивают конкурентные преимущества в сравнении с существующими аналогами (сравнение по стоимостным, техническим параметрам и проч.)</i>	Аналогичные сервисы в России: Agora freight – цифровой экспедитор Недостатки российского аналога: – Очень небольшое количество предлагаемых вариантов. В некоторые направления и вовсе отсутствуют какие-либо варианты перевозки. – Нет прозрачной стоимости доставки. – Высокие цены среди доступных вариантов перевозки. Аналогичные иностранные сервисы: Searates, Freightos, Flexport Основной недостаток – не оказывают сервиса в России. Сравнив существующие решения на рынке России и в мире мы имеем следующее преимущество: во-первых, это оптимизация маршрутов по времени и стоимости на основе предобученной нейронной сети с учетом возможных видов транспорта, которые уже загружены в нашу систему во-вторых, интеграции с перевозчиками, маркетплейсами, фулфилмент сервисами (в том числе зарубежными), а также возможность интеграции с системами заказчика Тем самым мы можем предположить, что наша система на порядок лучше систем конкурентов, и позволяет нашим клиентам меньше тратить на логистику и больше зарабатывать.

Продолжение ПРИЛОЖЕНИЯ Б

26	Научно-техническое решение и/или результаты, необходимые для создания продукции* <i>Описываются технические параметры научно-технических решений/результатов, подтверждающие/обосновывающие достижение характеристик продукта, обеспечивающих их конкурентоспособность</i>	Создано предварительное MVP, востребованность подтверждена рынком
27	«Задел». Уровень готовности продукта TRL <i>Необходимо указать максимально емко и кратко, насколько проработан стартап-проект по итогам прохождения акселерационной программы (организационные, кадровые, материальные и др.), позволяющие максимально эффективно развивать стартап дальше</i>	На данный момент проведены и реализованы: Анализ рынка посредством анкетирования и интервьюирования целевой аудитории (200 анкет, 30 интервью). Маркетинговая стратегия выхода на рынок и привлечения аудитории. Финансовая модель, базирующаяся на unit-экономике. Дизайн-система интеллектуальной системы. Аналитическое системное и бизнес описание реализуемой функциональности. Механизм авторизации на развернутой сетевой инфраструктуре.
28	Соответствие проекта научным и(или) научно-техническим приоритетам образовательной организации/региона заявителя/предприятия*	Соответствует
29	Каналы продвижения будущего продукта* <i>Необходимо указать, какую маркетинговую стратегию планируется применять, привести кратко аргументы в пользу выбора тех или иных каналов продвижения</i>	Продвижение: – Мессенджеры – E-mail рассылки – Социальные сети – Сарафанное радио – Контекстная реклама – SEO-оптимизация – Партнерские программы
30	Каналы сбыта будущего продукта* <i>Указать какие каналы сбыта планируется использовать для реализации продукта и дать кратко обоснование выбора</i>	– Продажи на сайте – Заключение контрактов на отраслевых мероприятиях – Клиенты (реферальные программы) – Партнеры (реферальные программы) – Участие в программе «Мой экспорт»

Продолжение ПРИЛОЖЕНИЯ Б

Финансовый план реализации проекта (планирование доходов и расходов на реализацию проекта)		
31	Доходы:	28 334 875 (3 года)
32	Расходы:	20 679 231 (3 года)
33	Источники привлечения ресурсов для развития стартап-проекта после завершения договора гранта и обоснование их выбора (грантовая поддержка Фонда содействия инновациям или других институтов развития, привлечение кредитных средств, венчурных инвестиций и др.)	Собственные средства, меры поддержки, частные инвестиции / инвестиционные фонды
Характеристика будущего предприятия (результат стартап-проекта, плановые оптимальные параметры (на момент выхода предприятия на самоокупаемость))		
34	Объем реализации продукции (в натуральных единицах) <i>Указывается предполагаемый Вами объем реализации продукции на момент выхода предприятия на самоокупаемость, т.е. Ваше представление о том, как может быть осуществлено</i>	2025 год: 300 клиентов; 50 подключенных перевозчиков 5 000 транзакций; 2026 год: 1 000 клиентов; 200 подключенных перевозчиков; 20 000 транзакций; 2027 год: 3 000 клиентов; 500 подключенных перевозчиков; 40 000 транзакций
35	Доходы (в рублях) <i>Указывается предполагаемый Вами объем всех доходов (вне зависимости от их источника, например, выручка с продаж и т.д.) предприятия на момент выхода предприятия на самоокупаемость, т.е. Ваше представление о том, как это будет достигнуто.</i>	2025 год: 3 млн. руб., 2026 год: 5 млн. руб., 2027 год: 20 млн. руб.
36	Расходы (в рублях) <i>Указывается предполагаемый Вами объем всех расходов предприятия на момент выхода предприятия на самоокупаемость, т.е. Ваше представление о том, как это будет достигнуто</i>	2025 год: 5 млн. руб., 2026 год: 10 млн. руб., 2027 год: 5 млн. руб.,
37	Планируемый период выхода предприятия на самоокупаемость <i>Указывается количество лет после завершения гранта</i>	30 месяцев

Продолжение ПРИЛОЖЕНИЯ Б

ПЛАН ДАЛЬНЕЙШЕГО РАЗВИТИЯ СТАРТАП-ПРОЕКТА		
38	Календарный план проекта*	Проверка рынка – 2 месяца MVP – 6 месяцев Развитие – 16 месяцев
39	Перечень планируемых работ с детализацией**	Стабилизация бизнес-модели Написание аналитических постановок Тестирование функциональности системы Размещение MVP на площадки спроса Насыщение системы контентом Исследование рынка (4 тысячи анкет, 100 интервью) Построение CJM Формирование ценностных предложений для сегментов аудитории Формирование рекламного кабинета (определение эффективности рекламных каналов) Проведение а-в тестирования интерфейса Проведение а-в тестирования интерфейса "Проектирование UI-кита, Проектирование экранных форм" Объединение экранных форм в интерактивный прототип Построение информационного портала (тильда или аналоги) Разработка механизма авторизации Разработка механизма конструктора маршрутов Разработка механизма уведомлений Системное тестирование (юнит + интеграция + нагрузочное) Подготовка серверных площадок под вывод в продакшн Исправление ошибок в работе системы

Продолжение ПРИЛОЖЕНИЯ Б

	<i>Вопросы на период грантовой поддержки и максимально прогнозируемый срок, но не менее 2-х лет после завершения договора гранта)</i>	
40	Формирование коллектива	4 месяца
41	Функционирование юридического лица	< 1 месяца
42	Выполнение работ по разработке продукции с использованием результатов научно-технических и технологических исследований (собственных и/или легитимно полученных или приобретенных), включая информацию о создании MVP и (или) доведению продукции до уровня TRL 3 и обоснование возможности разработки MVP / достижения уровня TRL 3 в рамках реализации договора гранта	8 месяцев
43	Выполнение работ по уточнению параметров продукции, «формирование» рынка сбыта (взаимодействие с потенциальным покупателем, проверка гипотез, анализ информационных источников и т.п.)	2 месяца
44	Организация производства продукции	< месяца
45	Реализация продукции	6 месяцев

*Календарный план проекта**

№ этапа	Название этапа календарного плана	Длительность этапа, мес	Стоимость, руб.
1	Проверка рынка	2	2 000 000
2	Создание MVP	6	3 000 000
3	Развитие функциональных возможностей	16	10 000 000

Продолжение ПРИЛОЖЕНИЯ Б

*Перечень планируемых работ с детализацией***

ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РАБОТ С ДЕТАЛИЗАЦИЕЙ					
Этап 1 (длительность – 2 месяца)					
Задача	Оценка, fte	Собственный капитал	FTE, руб.	Дополнительные вложения, руб.	Итого, стоимость работы
1	2	3	4	5	6
Открытие юр. Лица	40	40	0,00		0,00
Стабилизация бизнес-модели	80	80	0,00		0,00
Написание аналитических постановок	160	0	240 000,00		240 000,00
Тестирование функциональности системы	40	0	60 000,00		60 000,00
Выкладка MVP на площадки спроса	40	0	60 000,00		60 000,00
Насыщение системы контентом	40	0	60 000,00	200 000,00	260 000,00
Исследование рынка (4 тысячи анкет, 100 интервью)	32	0	48 000,00	400 000,00	448 000,00
Построение CJM	64	0	96 000,00		96 000,00
Формирование ценностных предложений для сегментов аудитории	40	0	60 000,00		60 000,00
Формирование рекламного кабинета (определение эффективности рекламных каналов)	40	0	60 000,00	280 000,00	340 000,00
Проведение a-b тестирования интерфейса	64	0	96 000,00		96 000,00

Продолжение ПРИЛОЖЕНИЯ Б

1	2	3	3	5	6
Проведение а-в тестирования интерфейса	8	0	12 000,00		12 000,00
Проектирование UI-кита, Проектирование экранных форм	160	0	240 000,00		240 000,00
Объединение экранных форм в интерактивный прототип	16	0	24 000,00		24 000,00
Построение информационного портала	40	0	60 000,00		60 000,00
Разработка механизма авторизации	64	0	96 000,00		96 000,00
Разработка механизма конструктора маршрутов	80	0	120 000,00		120 000,00
Разработка механизма уведомлений	48	0	72 000,00		72 000,00
Системное тестирование (юнит + интеграция + нагрузочное)	96	0	144 000,00		144 000,00
Подготовка серверных площадок под вывод в продакшн	32	0	48 000,00	100 000,00	148 000,00
Исправление ошибок в работе системы	64	0	96 000,00		96 000,00
Этап 2 (длительность – 10 месяцев)					
Наименование работы	Описание работы	Стоимость	Результат		
Развитие	Реализация прочих функциональных возможностей	10 000 000	Продукт удовлетворяет запрос рынка, содержит функции монетизации, привлек более 300 пользователей		