

**Автономная некоммерческая организация высшего образования
«Университет Иннополис»**

**АННОТАЦИЯ
НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ
(БАКАЛАВРСКУЮ РАБОТУ)
ПО НАПРАВЛЕНИЮ ПОДГОТОВКИ
09.03.01 – «ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА»**

**НАПРАВЛЕННОСТЬ (ПРОФИЛЬ) ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ
«ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА»**

Тема

**Использование ИИ-алгоритмов в бизнес-процессах
индустриальных компаний**

Выполнил

Будник Георгий Васильевич

подпись

Иннополис, Innopolis, 2025

Оглавление

1	Введение	4
2	Обзор литературы	5
3	Методология	8
4	Реализация и результаты	11
5	Выводы	13
	Список использованной литературы	14

Аннотация

В данной работе рассматриваются пути цифровой трансформации компаний промышленного сектора с использованием алгоритмов Искусственного интеллекта (ИИ) и без его участия. Работа рассматривает оптимизацию двух ключевых показателей: времени работы с клиентом и среднее количество позиций товаров в заказе (SKU).

Для сокращения времени работы с клиентом предлагается внедрить обновленный бизнес-процесс оформления штрихкарт, которые будут заменять печать документов для выдачи и оплаты товаров. Для увлечения среднего количества SKU в заказе клиента рассматривается использование вспомогательной рекомендательной системы для продавцов-консультантов. Исследование включает сравнение рекомендательных систем, основанных на классическим методах и глубоком обучении, а также анализ и репрезентацию входных данных.

Все эксперименты и анализ данных проводились в рамках крупнейшей строительной компании на Дальнем Востоке. Выпускная квалификационная работа состоит из N страниц, включая 22 рисунка, 2 таблицы и 40 использованных источников.

Глава 1

Введение

Внедрение Искусственного интеллекта (ИИ) в бизнес-процессы промышленных компаний, которые занимают 25% мирового ВВП [1], становится ключевым фактором цифровой трансформации. Цель исследования — оценить влияние ИТ-решений на эффективность работы, рассматривая два ключевых показателя: На сколько можно сократить время оформления заказа с помощью механизма штрихкарт (**RQ1**)? На сколько можно увеличить среднее число SKU в заказе при помощи рекомендательной системы для продавцов (**RQ2**)?

В течение года Я занимал позицию руководителя цифровой трансформации в крупнейшей строительной розничной сети на Дальнем Востоке России. В рамках этой дипломной работы мне был предоставлен доступ к коммерческим данными компании.

В ходе работы был проведён анализ лучших практик цифровизации, подготовлены и очищены данные пользователей, реализованы и протестированы ИТ-решения, а также произведена оценка их эффективности в реальных условиях. Все материалы доступны в [открытом репозитории](#).

Глава 2

Обзор литературы

RQ1 – Оптимизация времени обслуживания клиента. Современные исследования демонстрируют различные способы сокращения времени взаимодействия с клиентом. Самые часто используемые из них:

1. **RFID (Radio Frequency IDentification) технологий:** В исследовании [2] продемонстрирована «умная корзина», которая на основании RFID-меток на товарах определяет общую сумму покупки, отправляя запрос на сервер для подтверждения кассиром. По результатам время обслуживание в среднем уменьшилось от 5 до 10 минут на покупателя. Компания Uniqlo [3] внедрила аналогичный принцип в кассы самообслуживания, сократив время обслуживания на 50%
2. **Функциональные упрощения интерфейса:** Благодаря изменению алгоритма заполнения шаблона формы нефтегазовые специалисты смогли сократить время заполнения заявки с 25 до 5 минут [4]

Однако, среди всех решений механизм штрихкодов признан наиболее универсальным и доступным способом оптимизации в промышленном сек-

торе [5]. Стандарт EAN-13 [6] позволяет легко и надёжно идентифицировать продукцию и связанные с ней операции, а контролировать операции позволяют специализированные системы планирования (Enterprise resource planning, ERP).

В исследованиях [5], [7] показано, как через штрихкарты можно связать данные о строительных материалах с базой данных и автоматизировать расчёты. Такой подход позволил найти расхождения в 17.9% в нескольких статьях расхода проекта, что превышает операционную выгоду относительно других решений. Общая архитектура таких систем представлена как цикл: генерация кода, привязка к сущности, верификация и удаление устаревших записей.

RQ2 – Увеличение количества SKU в заказе. Множество исследований подчеркивают, что персонализация является ключевым направлением увеличения продаж и расширения состава заказа [8]. Персонализированные рекомендации позволяют повышать лояльность и точность рекомендаций.

Как отмечается в работе [9], важно учитывать бизнес-ограничения, корректность рекомендаций и особенности предметной области. Например, при продаже стройматериалов точность рекомендаций может быть менее критична, чем в медицине.

Рекомендательные системы строятся на машинном обучении и анализе взаимодействий пользователей с товарами. Классические подходы делятся на:

- **Контентные (СВ)** – используют характеристики товаров и формируют профиль пользователя на основе схожих товаров.

- **Коллаборативные (CF)** – основаны на матрице взаимодействий. Бывают memory-based (поиск схожих пользователей или товаров) и model-based (матричные разложения и нейросети).
- **Гибридные (Hybrid)** – объединяют СВ и CF, чтобы устранить недостатки, например, такие как холодный старт [9].

Методы глубокого обучения обеспечили качественный скачок в точности рекомендаций [10]:

- **Transformer-модели**, использующие bidirectional context для обработки последовательностей взаимодействий, как в BERT4Rec [11].
- **AutoRec** – автоэнкодеры для восстановления пользовательских предпочтений.
- **GANRec** – генерация новых пар пользователь–товар для обучения моделей, несмотря на нестабильность обучения.

Современные бизнес-решения всё чаще используют гибридные подходы, совмещающие эмбединги, кластеризацию и трансформеры, что позволяет достичь высокой точности рекомендаций при сохранении их интерпретируемости.

Глава 3

Методология

RQ1 — Постановка задачи и Интеграция механизма внедрения штрих-карт. Время обслуживания клиента в строительном магазине можно представить как сумму трёх этапов:

$$t_{\text{checkout}} = t_{\text{order}} + t_{\text{services}} + t_{\text{completion}}, \quad (3.1)$$

где t_{order} — оформление товара продавцом в ERP программе, t_{services} — регистрация и печать сопутствующих услуг (доставка, упаковка, лояльность), $t_{\text{completion}}$ — оплата на кассе и выдача товара.

Оптимизируем путь клиента, интегрировав механизм многоразовых ID-карт с уникальными штрихкодами, позволяющий отказаться от бумажных документов и ускорить работу на кассе.

Алгоритм работы механизма. После оформления заказа продавец сканирует штрихкарту, привязывая её к заказу. Клиент получает штрихкарту вместо печатных документов. Продавец отправляет клиента на кассу, оформляя складские документы. Клиент отдаёт штрихкарту кассиру, который сканирует её. ERP программа автоматически открывает нужный заказ.

После оплаты карта очищается и возвращается продавцу.

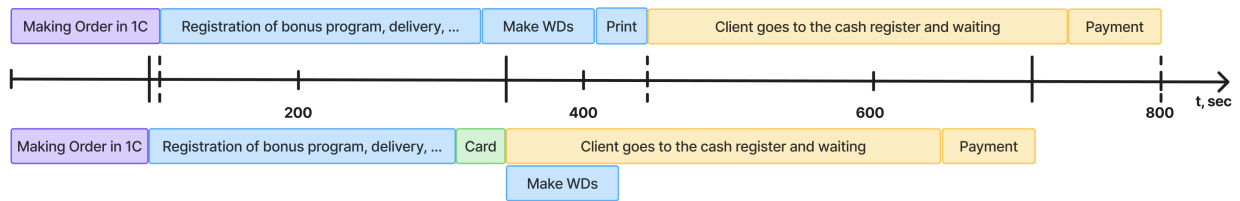


Рис. 3.1: Изменения пути клиента до/после внедрения механизма штрихкарт

В качестве исходных данных использован датасет из трёх флагманских магазинов, содержащий **59,998** заказов до очистки за период трёх месяцев (Июль - Сентябрь 2024) до внедрения и трёх месяцев после (Октябрь - Декабрь 2024). После удаления дубликатов и фильтрации выбросов методом межквартильного размаха (IQR) [12] осталось **59,491** наблюдений, пригодных для достоверного анализа. Это обеспечило корректное сравнение фаз до и после внедрения.

RQ2 — Постановка задачи и Интеграция рекомендательной системы.

Целью задачи является внедрение алгоритма, способного в реальном времени предлагать клиенту дополнительные товары, исходя из текущего состава корзины, введённой продавцом. Основное ограничение — отсутствие данных о клиенте до завершения оформления заказа, что требует работы модели в условиях холодного старта [9].

Пусть $B = i_1, i_2, \dots, i_n$ — текущая корзина клиента, содержащая n товарных позиций (SKU). Рекомендательная система должна сформировать ранжированный список $\hat{R}(B) = r_1, r_2, \dots, r_k$ из k наиболее релевантных товаров, опираясь на глобальную историю покупок других клиентов.

Для обучения и тестирования использован датасет из **665,733** заказов оформленных во Владивостоке, которые были выгружены из ERP про-

граммы за ближайшие два года. После удаления нерелевантных данных и выбросов по методу межквартильного размаха (IQR) [12] осталось **492,999** заказов и **12,881** уникальных товаров. Разметка данных для валидации была выполнена разбиением каждого тестового заказа O_i на две половины: B_i — известная часть корзины, и T_i — скрытые позиции, по которым проверяется качество рекомендаций.

Вычисление метрик проводилось по стандартным показателям Recall@20 и nDCG@20:

$$\text{Recall} \bullet k = \frac{|\hat{R}(B_i) \cap T_i|}{|T_i|}, \quad \text{nDCG} \bullet k = \frac{1}{\text{IDCG} \bullet k} \sum_{j=1}^k \frac{\mathbb{I}[r_j \in T_i]}{\log_2(j+1)} \quad (3.2)$$

Сформированный датасет $D_{RQ2} = O_1, O_2, \dots, O_N$ использовался в различных моделях — от классических (Global Top, ALS, Item2Item) [9] до глубоких (BERT4Rec) [11], включая гибридные подходы с использованием эмбедингов на товары (MiniLM, SBERT) и кластеризацией (DBSCAN, HDBSCAN), основанной на пользовательских взаимодействиях с понижением размерности через технологию UMAP [13]. Полученные карты плотностей использовались для кластерно-осознанной системы ранжирования.

Глава 4

Реализация и результаты

RQ1 — Внедрение механизма штрихкарт. По результатам анализа **59,491 заказов** в трёх флагманских магазинах удалось достичь сокращения времени обслуживания клиента на **6.8%–8.6%**, что составляет **52–59 секунд** на одного клиента в зависимости от магазина.

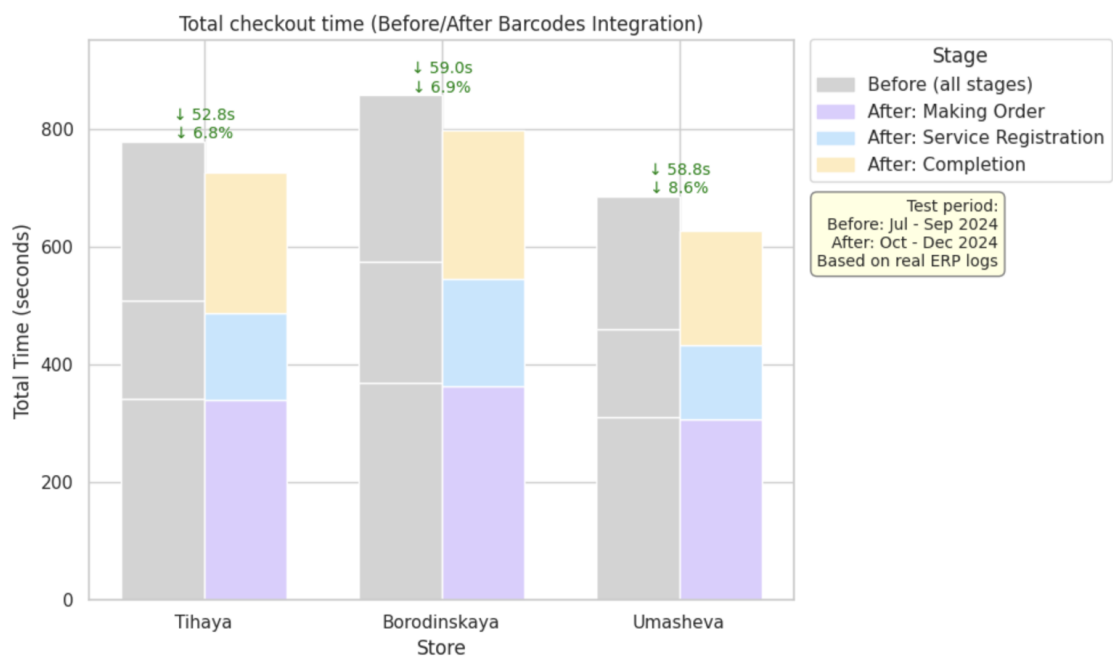


Рис. 4.1: Сравнение времени обслуживания до/после внедрения

Основной эффект связан со снижением значений t_{order} и $t_{\text{completion}}$, оформление складских документов ведется параллельно оплате клиента, а кассир не тратит много времени для сверки конкретного заказа клиента, благодаря уникальному ID штрихкарты. Продавцами и директором магазина отмечено снижение нагрузки на персонал, упрощение кассовых процессов и возможность структурного анализа работы сотрудников за счёт логирования операций по ID штрихкарты.

RQ2 — Интеграция рекомендательной системы. Система рекомендаций, основанная на архитектуре **BERT4Rec + HDBSCAN**, показала наивысшие метрики: $n\text{DCG}@20 = 0.2473$, $\text{Recall}@20 = 0.3790$. Анализ показал, что модель превосходит другие подходы за счёт учёта контекста внутри заказа и кластеризации товаров. Сравнив результаты 306 заказов за три дня до внедрения (14 - 16 февраля 2025) и 296 заказов за три дня после (17 - 19 февраля 2025), Среднее количество SKU выросло с **9.784 до 10.534**, что соответствует приросту на **7.66%**:

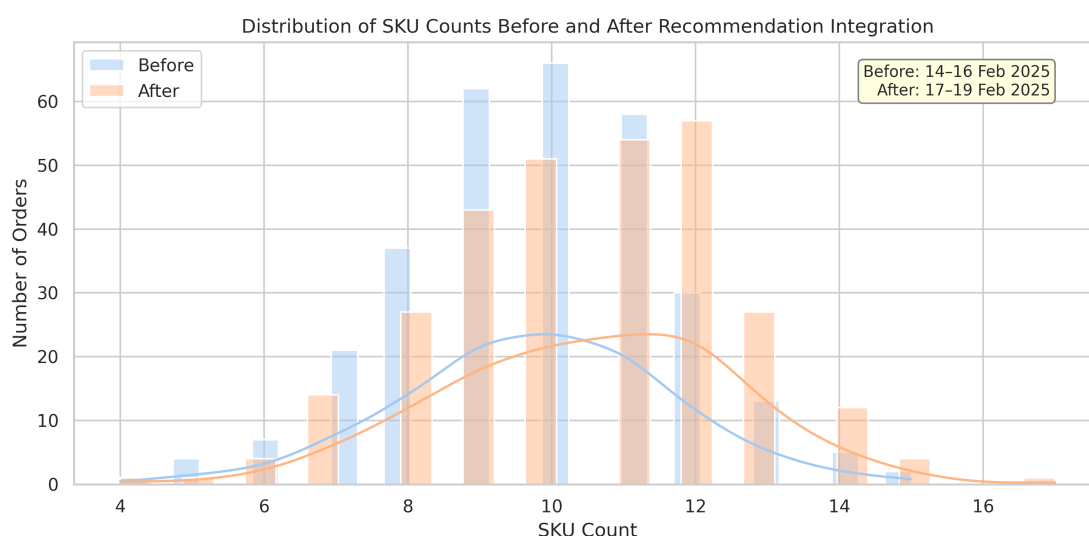


Рис. 4.2: Распределение количества товаров в заказе до/после внедрения

Глава 5

Выводы

В ходе работы было реализовано два ИТ-решения, направленные на оптимизацию ключевых показателей. Внедрение штрихкارت позволило сократить путь клиента **в среднем на 7.43% (примерно -57 секунд)**, а рекомендательная система, основанная на модели BERT4Rec+HDBSCAN, увеличила среднее количество SKU в заказе **на 7.66% (примерно +0.75 позиции)**. Эксперименты прошли апробацию и были одобрены руководством компании «Помощник». На основании экспертизы Исполнительного директора:

- Общая оценка оптимизации расходов внедрения штрихкارت (RQ1) оценивается в размере **78,050,420 рублей в год** во всей строительной розничной сети
- Общая оценка потенциального дохода от внедрения рекомендательной системы (RQ2) оценивается в размере **42,673,540 рублей в год** во всей строительной розничной сети

Результаты подтверждают актуальность цифровой трансформации в промышленности и перспективность применения ИИ в индустрии.

Список использованной литературы

- [1] World Bank, *World Development Indicators*, <https://wdi.worldbank.org/table/4.2>, 2023.
- [2] «Data Transmission Using RFID System on Smart Shopping Carts for Checkout Process Efficiency in Supermarket at Indonesia,» *Procedia Computer Science*, т. 179, с. 674—681, 2021, Accessed: May 22, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050921001101>.
- [3] Japan Insides, *Uniqlo Self Checkout is 50% Faster in Japan With RFID Technology*, <https://japaninsides.com/uniqlo-self-checkout-is-50-faster-in-japan-with-rfid-technology-39569>, 2023.
- [4] IEOM Society, *Optimization of Order Processing Procedures to Reduce Order Execution Time*, <https://index.ieomsociety.org/index.cfm/article/view/ID/11421>, 2022.
- [5] International Journal of Engineering and Technology, *A Development of Practical Cost Control System for Small and Medium-sized Contractors*, <https://www.ijetch.org/papers/358-P10048.pdf>, 2012.

- [6] International Journal of Science and Research (IJSR), *Generating EAN-13 Standard Barcodes*, <https://www.ijsr.net/archive/v3i6/MDIwMTQ0OTg=.pdf>, 2012.
- [7] M.-C. Tsai и C.-Y. Yang, *Mobile Construction Supply Chain Management Using PDA and Bar Codes*, <https://ir.lib.nycu.edu.tw/bitstream/11536/13559/1/000229288500002.pdf>, 2005.
- [8] ResearchGate, *The Role of Artificial Intelligence in Personalized Marketing Strategies*, <https://shorturl.at/AN9gS>, 2024.
- [9] F. Ricci и L. Rokach, «Recommendation Systems: Algorithms, Challenges, Metrics, and Business Opportunities,» *Applied Sciences*, т. 10, № 21, с. 7748, 2020. DOI: [10.3390/app10217748](https://doi.org/10.3390/app10217748).
- [10] ResearchGate, *Sequence Recommendation based on Deep Learning*, https://www.researchgate.net/publication/369048688_Sequence_Recommendation_based_on_Deep_Learning, 2023.
- [11] F. Sun, J. Liu, J. Wu и др., «BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer,» в *Proc. 28th ACM Int. Conf. on Information and Knowledge Management (CIKM)*, 2019, с. 1441—1450.
- [12] «Using the Interquartile Range in Infection Prevention and Control Research,» *Infection Prevention in Practice*, 2024. url: <https://www.sciencedirect.com/science/article/pii/S2590088924000015>.
- [13] R. Gonzalez-Marquez и D. Kobak, «Learning Representations of Learning Representations,» в *Data-centric Machine Learning Research (DMLR) Workshop at ICLR 2024*, 2024. url: <https://arxiv.org/pdf/2404.08403>.