
МЕТОДОЛОГИЯ И АРХИТЕКТУРА ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ПРОДАЖ В МАГАЗИНАХ РОЗНИЧНОЙ СЕТИ

Шибиченко Михаил Иванович,

Аспирант, Московский финансово-юридический университет МФЮА
Россия, Москва
shibmish@yandex.ru

Павлов Валерий Анатольевич,

кандидат экономических наук, доцент кафедры информационных систем и технологий и автоматизации в строительстве, ФГБОУ ВО «Национальный исследовательский Московский государственный строительный университет»
Россия, Москва
Testerpav@gmail.com

Аннотация

В статье рассматривается прогнозирование спроса в розничных сетях критически важно для оптимизации логистики и управления запасами. Отмечено, что традиционные статистические методы зачастую не справляются с учетом большого количества сложных нелинейных факторов, влияющих на продажи. Промо-акции, макроэкономические показатели, сезонность и конкурентная среда, - на их основе современные методологии глубокого обучения строят высокоточные прогнозные модели, способные выявить скрытые паттерны в многомерных временных рядах. Выявлено, что грамотное сочетание строгой методологии и передовых архитектур позволяет значительно повысить точность прогнозов, снижая затраты на логистику.

Ключевые слова: прогнозирование продаж, розничная торговля, глубокое обучение, временные ряды, LSTM, машинное обучение, искусственный интеллект, управление запасами

DEEP LEARNING METHODOLOGY AND ARCHITECTURE FOR PREDICTING SALES IN RETAIL STORES

Mikhail I. Shibichenko,

postgraduate student
Moscow University of Finance and Law MFUA
Russia, Moscow
shibmish@yandex.ru

Valery A. Pavlov,

PhD in Economics, Associate Professor of the
Department of Information Systems and Technologies and Automation in
Construction
National Research Moscow State University of Civil Engineering

Russia, Moscow
Testerpav@gmail.com

ABSTRACT

This article examines demand forecasting in retail chains, which is critical for optimizing logistics and inventory management. It notes that traditional statistical methods often fail to account for the large number of complex nonlinear factors influencing sales. Promotions, macroeconomic indicators, seasonality, and the competitive environment—all of these factors—are used by modern deep learning methodologies to build highly accurate forecast models capable of identifying hidden patterns in multivariate time series. It is found that a well-designed combination of rigorous methodology and advanced architectures significantly improves forecast accuracy, reducing logistics costs.

Keywords: sales forecasting, retail, deep learning, time series, LSTM, machine learning, artificial intelligence, inventory management

Глубокое обучение — это могущественная ветвь искусственного интеллекта, которая, перестав быть прерогативой узкого круга ученых, стала двигателем цифровой трансформации. Эта дисциплина основывается на двух фундаментальных столпах: методологии и архитектуре. Их диалектическое единство и составляет суть успеха любого проекта в области глубокого обучения. Цель работы заключается в анализе методологии и архитектуры глубокого обучения для прогнозирования продаж на примере розничных сетевых магазинов. Задачами работы стали теоретический анализ методологии глубокого обучения, а также, - особенности внедряемых в архитектуру глубокого обучения для решения задач сетевых магазинов моделей. Данная обзорная статья призвана проанализировать: какие новые проблемы появились в сфере прогнозирования продаж с развитием искусственного интеллекта.

Методология глубокого обучения — это стратегический каркас, который определяет весь жизненный цикл проекта. Это не просто набор технических приемов, а целостный подход к решению проблемы. Архитектура же является тактическим воплощением этой стратегии [2, с. 36]. Это конкретный, инженерный выбор того, как будут соединены между собой нейроны внутри искусственной нейронной сети для наилучшего выполнения поставленной задачи. Можно провести аналогию со строительством дома: методология — это генеральный план, включающий цели, сроки, бюджет и философию строительства. А вот архитектура — это конкретные чертежи, определяющие расположение входных групп, арок, окон или витражей, которые и делают дом уникальным.

Ключевым аспектом методологии является ее цикличность. Процесс редко бывает линейным; он - как спираль, где каждый виток приближает модель к идеалу. Все начинается с фундаментального этапа, — сбора и подготовки данных. Данные должны репрезентативными, очищенными от шума и корректно размеченными [4, с. 125]. Следом наступает этап проектирования, - то есть, выбора архитектуры модели. Здесь методология диктует необходимость выбора не самой сложной, а наиболее адекватной задачи архитектуры.

Цикл продолжается обучением модели. На этом этапе методология предъявляет строгие требования к мониторингу процесса: отслеживание функции потерь на тренировочном и валидационном наборах данных позволяет избежать ключевых проблем.

Проблем обычно две: переобучение, - когда модель идеально запоминает примеры, но не способна к обобщению, и недообучение, - когда она оказывается слишком примитивной для решения задачи. Оценка производительности на тестовом наборе дают в итоге количественную оценку работы алгоритма [3].

Но и на этом методологический цикл не заканчивается. Развертывание модели, ее постоянный мониторинг на предмет «дрейфа» данных (когда распределение входных данных постепенно меняется, в следствии чего модель теряет свою актуальность), последующее обслуживание, – все это неотъемлемые части современной методологии MLOps.

Архитектура глубокого обучения является решением современных проблем. Архитектура глубокого обучения эволюционировала от простых полносвязных сетей, которые плохо справлялись с сохранением пространственной и временной структуры, к сложным специализированным структурам.

Например, сверточные нейронные сети (CNN) произвели революцию в компьютерном зрении [4, с. 120]. Их особенность заключается в идее локальных рецептивных полей и разделяемых весов: сверточные слои сканируют изображение небольшими «окнами» (фильтрами), извлекая низкоуровневые признаки [4, с. 119]. Рекуррентные нейронные сети (RNN) были созданы для работы с последовательностями, где контекст и порядок имеют ключевое значение. То есть: для работы с текстом, речью, временными рядами. Их ключевая особенность в наличии «памяти», которая хранит информацию о предыдущих элементах последовательности для обработки текущего. Однако классические RNN страдали от проблемы исчезающего градиента. Это мешало им запоминать долгосрочные зависимости. Более продвинутые архитектуры, - LSTM (Long Short-Term Memory), GRU (Gated Recurrent Unit), - решили эту проблему с помощью специальных механизмов – «вентиров».

Но настоящим прорывом последних лет стала архитектура «трансформер»: отказавшись от рекуррентности в пользу механизма внимания («attention mechanism»), «трансформеры» научились взвешивать и оценивать важность каждого элемента последовательности относительно всех других. При этом, расстояние между нами не имеет значение. Это позволило проводить параллельные вычисления и обучать модели на огромных объемах данных (что раньше не было возможно). «Трансформеры» легли в основу больших языковых моделей, - таких как BERT и GPT.

Прогнозирование продаж в ритейле с помощью глубокого обучения, – это не просто задача по обучению нейронных сетей работать с учетом обновленных данных [3]. Это комплексная дисциплина, где научная строгость методологии управления данными работает в симбиозе с архитектурами LSTM и «Трансформеров». Их применение позволяет менеджерам перейти к проактивному стратегическому планированию: когда решение о закупке, акции и ценообразовании подкреплено точным, многомерным прогнозом будущего спроса.

Методология в данном контексте – это дорожная карта, ведущая к прогнозной модели, интегрированной в бизнес-процессы. Это алгоритм, основанный на подходах вроде CRISP-DM (Cross-Industry Standard Process for Data Mining). Внедрение глубокого обучения в бизнес начинается не с написания кода, а с глубокого погружения в бизнес-контекст: что именно мы прогнозируем?; продажи в штуках или в денежном выражении?; на уровне товарной категории, SKU (артикула) или конкретного магазина?; какой горизонт прогнозирования, – день, неделя или месяц? И ответы на эти вопросы определяют непосредственно всю архитектуру программного продукта.

Выбор архитектуры – это выбор двигателя для корабля прогнозирования. Для задач прогнозирования продаж (по своей сути являющихся многомерными временными рядами)

наиболее эффективными считаются рекуррентные нейронные сети (RNN). Точнее, их усовершенствованная версия, — сети с долгой краткосрочной памятью (LSTM) [2, с. 34].

LSTM — это архитектура, специально созданная для работы с последовательностями, со встроенным решением проблемы исчезающего градиента. Именно это мешало мешала обычным RNN запоминать долгосрочные зависимости. Ее гениальность, — в наличии механизма «вентиров», который решает, какую информацию забыть, а какую, — сохранить в памяти и использовать в текущей работе. В контексте прогнозирования продаж это означает, что LSTM может «научиться» помнить, что пик продаж мороженого был две недели назад во время аномальной жары (долгосрочная зависимость). И, в то же время, — учитывать, что вчера была запущена акция «два по цене одного» (краткосрочное воздействие). Сеть анализирует всю историю продаж вместе с сопутствующими факторами, выявляя сложные, нелинейные паттерны, незаметные для человека и линейных моделей [2, с. 32].

Более современным и мощным подходом становится использование архитектуры «трансформер», которая произвела революцию в обработке естественного языка. Ее ключевой компонент: механизм внимания (attention). Он позволяет модели научиться «внимательно» относиться к разным участкам временного ряда. Например, при прогнозе продаж на Новый год модель может научиться уделять максимальное внимание данным за аналогичный период прошлого года, а не данным за прошлую неделю. Это обеспечивает гибкость в учете сезонных и событийных особенностей.

Нет так давно крупнейший мировой ритейлер Walmart поставил перед собой амбициозную цель, — предугадывать желания миллионов покупателей в каждом из своих магазинов. Традиционные методы прогнозирования, справлявшиеся с общими трендами, пасовали перед необходимостью учитывать уникальность спроса на каждый конкретный продукт в конкретной локации. Неточность выливалась в прямые убытки: пустые полки в разгар распродаж, горы списанного скоропортящегося товара. Компании требовался принципиально иной подход, способный уловить взаимосвязь между погодой за окном, праздником в календаре, рекламной листовкой в почтовом ящике, — и внезапной тягой к лимонаду у жителей конкретного района [1].

Решением программистов стало внедрение архитектуры глубокого обучения, способной к долгосрочной памяти. Команда Walmart Labs сделала ставку на рекуррентные нейронные сети типа LSTM. Их ключевое преимущество заключалось в умении не просто анализировать данные, но и запоминать контекст, выявляя сложные, неочевидные временные зависимости. В отличие от статических моделей, такая сеть могла понять, что аномальная жара в городе — это не разовый всплеск продаж воды, а устойчивый тренд на неделю вперед. Или что праздничные распродажи в преддверии Рождества имеют свою специфику в США, эта специфика будет повторяться из года в год. Однако для стран с мусульманским населением все будет не так очевидно.

Модель «питалась» петабайтами разнородных данных. В нее загружалась не только история продаж каждого товара, но и внешние сигналы: прогноз погоды, календарь государственных и школьных праздников, макроэкономическая статистика, график маркетинговых активностей. Нейросеть училась взвешивать влияние каждого фактора, самостоятельно определяя, что для прогноза продаж грилей в одном городе температура воздуха важнее, чем скидка в десять процентов, а для продаж зонтов в городе другого региона, — все наоборот.

Результатом стало то, что точность прогнозовкратно выросла. Что привело к ощутимым бизнес-результатам. Сократились потери от списаний, — система научилась с высокой точностью предсказывать необходимое количество скоропортящихся товаров. На полках практически перестали заканчиваться популярные товары в пиковые дни спроса, —

а это повысило лояльность клиентов и увеличило выручку. И компания перешла от реагирования к предвосхищению. Глубокая нейросеть превратила прогнозирование из рутинной статистической процедуры в интеллектуальный инструмент, способный предугадывать капризы потребителя за много дней вперед. Этот пример наглядно показывает, что в современном ритейле побеждает тот, кто не просто собирает большие данные, а умеет добывать из них правильную информацию с помощью глубокого понимания их контекста.

В современном ритейле точное предсказание спроса определяет эффективность всей цепочки создания стоимости: от управления запасами и логистики до планирования маркетинговых активностей. Практическая значимость проведенного исследования в том, что оно на примере показывает, что традиционные статистические методы, - такие как ARIMA, - часто оказываются несостоятельными перед лицом сложности данных современной розничной сети. На сцену выходит глубокое обучение: раздел искусственного интеллекта, предлагающий не просто инструменты, а целый уникальный подход к прогнозированию. Его эффективность, как показывает пример сетевого магазина Walmart, всецело зависит от симбиоза двух основ: строгой, выверенной методологии и корректно выбранной архитектуры нейронной сети.

Список литературы:

1. How Walmart Is Using Machine Learning AI, IoT And Big Data To Boost Retail Performance, - URL: <https://www.forbes.com/sites/bernardmarr/2017/08/29/how-walmart-is-using-machine-learning-ai-iot-and-big-data-to-boost-retail-performance/> (дата обращения: 12.09.2025).
2. Анафиев, А. С. Обзор подходов к решению задач оптимизации гиперпараметров для алгоритмов машинного обучения / А. С. Анафиев, А. С. Карюк // ТВИМ. - 2022. - № 2. - С. 30-37.
3. Рогулин Р.С. Обзор прикладных основ использования аналитики данных и машинного обучения в прогнозировании спроса // ЭСГИ. 2023. №3 (39). С. 115-126.
4. Лясковской, Е. А. Прогнозирование спроса на рынке дорожностроительной техники с использованием инструментов интеллектуального анализа данных / Е. А. Лясковской // Вестник ЮУрГУ. Серия: Компьютерные технологии, управление, радиоэлектроника. - 2022. - Т. 22, № 3. - С. 117-131.

References:

1. How Walmart Is Using Machine Learning: AI, IoT, and Big Data to Boost Retail Performance, - URL: <https://www.forbes.com/sites/bernardmarr/2017/08/29/how-walmart-is-using-machine-learning-ai-iot-and-big-data-to-boost-retail-performance/> (Accessed: 12.09.2025).
2. Anafiyev, A. S. Review of approaches to solving hyperparameter optimization problems for machine learning algorithms / A. S. Anafiyev, A. S. Karyuk // TVIM. - 2022. - No. 2. - pp. 30-37.
3. Rogulin R.S. Review of applied fundamentals of using data analytics and machine learning in demand forecasting // ESGI. 2023. No. 3 (39). P. 115-126.

4. Lyaskovskaya, E. A. Forecasting demand in the road construction equipment market using data mining tools / E. A. Lyaskovskaya // Bulletin of SUSU. Series: Computer technologies, control, radio electronics. – 2022. – Vol. 22, No. 3. – P. 117–131.