

УДК 004.8

РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ПРОГНОЗИРОВАНИЯ ПОТРЕБИТЕЛЬСКОГО СПРОСА С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ

Логутов Илья Павлович, магистрант, направление подготовки 09.04.02 Информационные системы и технологии, Оренбургский государственный университет, Оренбург
e-mail: logytov@mail.ru

Глотова Марина Ивановна, кандидат педагогических наук, доцент, доцент кафедры математических методов и моделей в экономике, Оренбургский государственный университет, Оренбург
e-mail: glotova_marina@rambler.ru

Аннотация. В условиях высокой динамики рынка и роста объема данных актуальность разработки точных систем прогнозирования потребительского спроса значительно возрастает. Целью данной работы является создание интеллектуальной системы на основе рекуррентной нейронной сети (RNN), способной учитывать сложные временные зависимости в данных. В исследовании применяются методы предобработки данных, включая нормализацию и устранение шумов, а также оптимизация гиперпараметров модели для повышения точности прогнозирования. Для интерпретации результатов используются методы LIME и SHAP, что позволяет выявить ключевые факторы спроса. Научная новизна работы заключается в комбинации RNN с методами объяснимого ИИ и алгоритмом подбора гиперпараметров, обеспечивающей не только высокую точность, но и прозрачность прогнозов, что значительно повышает практическую значимость интеллектуальной системы для организаций, занимающихся торговой деятельностью.

Ключевые слова: RNN, прогнозирование спроса, LIME, SHAP, подбор гиперпараметров.

Для цитирования: Логутов И. П., Глотова М. И. Разработка интеллектуальной системы прогнозирования потребительского спроса с помощью нейронной сети // Шаг в науку. – 2025. – № 4. – С. 31–36.

DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR FORECASTING CONSUMER DEMAND USING A NEURAL NETWORK

Logutov Ilya Pavlovich, postgraduate student, training program 09.04.02 Information Systems and Technologies, Orenburg State University, Orenburg
e-mail: logytov@mail.ru

Glotova Marina Ivanovna, Candidate of Pedagogical Sciences, Associate Professor, Associate Professor of the Department of Mathematical Methods and Models in Economics, Orenburg State University, Orenburg
e-mail: glotova_marina@rambler.ru

Abstract. In the face of highly dynamic markets and growing data volumes, the development of accurate consumer demand forecasting systems is becoming increasingly important. The goal of this study is to create an intelligent system based on a recurrent neural network (RNN) capable of accounting for complex temporal dependencies in data. The study utilizes data preprocessing methods, including normalization and noise removal, as well as optimization of model hyperparameters to improve forecasting accuracy. LIME and SHAP methods are used to interpret the results, enabling the identification of key demand drivers. The scientific novelty of this study lies in the combination of RNN with explainable AI methods and a hyperparameter selection algorithm, which ensures not only high accuracy but also forecast transparency, significantly increasing the practical value of the intelligent system for retail organizations.

Key words: RNN, demand forecasting, LIME, SHAP, hyperparameter selection.

Cite as: Logutov, I. P., Glotova, M. I. (2025) [Development of an intelligent system for forecasting consumer demand using a neural network]. *Shag v nauku* [Step into science]. Vol. 4, pp. 31–36.

Современные рынки характеризуются высокой степенью неопределённости, что делает задачу точного прогнозирования потребительского спроса особенно актуальной. Прогнозирование потребительского спроса является критически важным направлением для бизнеса, занимающегося торговлей, стремящегося минимизировать затраты и максимально удовлетворить потребности клиентов. Точные прогнозы спроса позволяют компаниям эффективно управлять запасами, планировать производство, разрабатывать маркетинговые стратегии и улучшать обслуживание клиентов.

При современном уровне автоматизации, каждая организация имеет большие хранилища данных о ведении собственной деятельности. С точки зрения прогнозирования, данные об истории продаж, динамике спроса и цены являются полезными [2].

Традиционные методы, основанные на статистических подходах, часто не учитывают сложные временные зависимости, скрытые закономерности и нелинейные взаимосвязи между признаками. Применение рекуррентных нейронных сетей, обладающих способностью учитывать последовательную природу данных, представляет собой перспективное направление в решении задачи прогнозирования.

Прогноз потребительского спроса имеет следующие черты: прогноз всегда ошибочен, при оценке ошибок прогнозы являются более полными, для большего набора имен прогноз более точный, прогноз более точный за более короткий период времени [1]. Точность прогноза зависит от используемой модели и самих данных [5].

Целью данной работы является описание процесса разработки интеллектуальной системы прогнозирования потребительского спроса на основе рекуррентной нейронной сети, обеспечивающей не только высокую точность прогноза, но и прозрачность модели для конечного пользователя.

На сегодняшний день существует значительное количество исследований, посвящённых применению методов машинного обучения в задачах прогнозирования. Рассматриваются подходы как статистические методы, такие как авторегрессия (Auto-regression, AR), авторегрессия интегрированных скользящих средних (Auto-regression of integrated moving average, ARIMA), сезонная ARIMA (Seasonal ARIMA, SARIMA), векторные авторегрессионные модели (Vector autoregression, VAR) и нейронные сети [6].

Рекуррентные нейронные сети (RNN) – это класс нейронных сетей, предназначенный для обработки последовательных данных, таких как временные ряды. Основная особенность RNN заключается в том, что они сохраняют информацию о предыдущих шагах последовательности, что позволяет учитывать

исторические данные [3]. RNN способны улавливать сложные временные зависимости, что особенно важно в задаче прогнозирования спроса, зависящего от сезонных и поведенческих факторов.

Одним из недостатков нейронных сетей в том числе и RNN является важность подбора гиперпараметров и сложность в интерпретации полученной модели и прогноза.

Некорректность значений гиперпараметров, например, таких как порог возбуждения и коэффициент ослабления, может приводить к перегрузке RNN в процессе прогнозирования либо, наоборот, к чрезмерному затуханию и недостатку возбуждающего потенциала нейронов для ассоциативного вызова новой информации, что негативно сказывается на точности прогнозирования [4].

Однако при всей эффективности RNN, в большинстве публикаций акцент делается исключительно на достижении высокой точности предсказаний. Между тем, в реальных бизнес-условиях пользователям важно не только получить точный прогноз, но и понимать, как и почему был получен тот или иной результат. Интерпретируемость моделей приобретает всё большее значение в условиях принятия решений на основе ИИ. Это позволяет повысить доверие к системе, облегчить процесс внедрения в практику и обеспечить соблюдение требований нормативного регулирования.

Существует ряд методов, обеспечивающих интерпретацию сложных моделей, включая SHAP и LIME. Тем не менее, их применение к рекуррентным архитектурам остаётся относительно редким. Причина в том, что такие модели обладают внутренними состояниями и работают с последовательностями, что усложняет интерпретацию. Таким образом, разработка подхода, сочетающего преимущества RNN с интерпретируемостью, является актуальной задачей современной прикладной аналитики.

Краткий алгоритм работы LIME:

- выбор примера данных: выбирается конкретный пример данных, для которого требуется объяснить предсказание модели;
- создание случайных изменений: генерируются искусственные данные путем случайного изменения признаков вокруг выбранного примера;
- обучение локальной модели: на сгенерированных данных обучается локальная модель (например, линейная регрессия), которая предсказывает результаты, аналогичные исходной модели в данной точке данных;
- анализ весов модели: интерпретация предсказания основывается на анализе весов или коэффициентов, полученных в локальной модели. Это позволяет определить важность различных признаков для данного предсказания.

LIME является мощным инструментом для обеспечения объяснимости моделей машинного обучения на уровне индивидуальных предсказаний.

Краткий алгоритм метода SHAP:

- создание интерпретируемых обобщений сложных внутренних представлений нейронной сети, таких как визуализация активаций и анализ весов;
- разработка механизмов уведомления пользователей о процессе принятия решений моделью, включая визуальные и текстовые отчеты;
- включение интерактивных компонентов, которые позволяют пользователям задавать вопросы модели и получать объяснения по конкретным аспектам её работы;
- проведение анализа устойчивости и чувствительности модели к изменениям входных данных, что помогает выявить слабые места и области для улучшения.

Функционал метода SHAP:

- визуализация активаций: показ, как различные слои нейронной сети реагируют на определенные входные данные;
- объяснение важности признаков: определе-

ние, какие входные признаки наиболее влияют на решение модели;

- уведомления о процессах: информирование пользователей о том, как модель пришла к определенному выводу, включая информацию о промежуточных шагах и использованных данных [35];
- интерактивные интерфейсы: возможность для пользователей задавать вопросы и получать ответы от модели в интерактивном режиме;
- анализ чувствительности: изучение, как изменения входных данных влияют на результаты модели, что помогает в обнаружении потенциальных уязвимостей и предвзятостей.

Подробнее опишем важные этапы работы интеллектуальной системы прогнозирования потребительского спроса.

Описание данных. Для корректной работы модели важно обеспечить высокое качество данных. Использовался набор данных, содержащий историю продаж ритейл-компании за 3 года (2019–2021). В таблице 1 отражены такие признаки, как дата, цена, размер скидки, количество продаж. Ниже приведён пример фрагмента данных.

Таблица 1. Пример фрагмента обучающего набора данных

Продукт	Дата	Цена	Размер_скидки	Количество_продаж
молоко	01.01.2020	80	0	226
молоко	02.01.2020	80	0	227
молоко	03.01.2020	80	0	224
молоко	04.01.2020	80	0	207
молоко	05.01.2020	80	0	211

Источник: разработано авторами

Далее из этих данных выявляются такие признаки как: день недели, дата месяца.

Предобработка входных данных. Для корректной работы модели важно обеспечить высокое качество данных. Набор данных включал значения цены, размера скидки, дня недели, календарной даты и индикатора праздничного дня и значение спроса на продукцию. Признаки были нормализованы методом стандартизации. Пропущенные значения обрабатывались на основе медианных значений либо методом линейной интерполяции по времени. Также проводилась проверка на выбросы, что критично при обучении моделей на временных рядах.

Разбиение набора данных. Для предотвращения утечки информации и сохранения временной структуры выборка делилась хронологически: 80% данных

использовалось для обучения, 20% – для тестирования. Дополнительно применялась временная кросс-валидация, при которой модель обучалась на определённом временном отрезке и тестировалась на следующем, имитируя реальные условия прогнозирования в будущем.

Подбор гиперпараметров. Гиперпараметры модели (размер скрытого состояния, число слоёв, learning rate, размер батча, коэффициенты регуляризации) подбирались методом перебора. В качестве метрик оценки использовались MAE, MSE и Коэффициент детерминации (R^2). Главным преимуществом коэффициента детерминации перед мерами, основанными на ошибках, является его инвариантность к масштабу данных.

Фрагмент результата подбора значений гиперпараметров показан в таблице 2.

Таблица 2. Результат подбора гиперпараметров

Lauer	Units	Activation	Dropout	Types	L_rate	Batch	MAE	MSE	RBI
1	2	3	4	5	6	7	8	9	10
3	150	relu	0.2	GRU	0.001	64	0.0291	0.00285	0.8333
2	100	relu	0.2	GRU	0.001	64	0.0290	0.00286	0.8186
2	150	relu	0.2	GRU	0.001	32	0.0300	0.00288	0.7732
3	100	relu	0.1	GRU	0.001	32	0.0298	0.00288	0.6545

Источник: разработано авторами

Как следует из таблицы, нейронная сеть с параметрами: количество слоёв 3, количество нейронов 150, функция-активатор Relu, тип нейронной сети GRU, скорость обучения 0.001, размером пакета 64 по метрикам MSE и R^2 наиболее точная. Для предотвращения переобучения применялась стратегия ранней остановки (early stopping), позволяющая прекратить обучение при отсутствии улучшения на валидационной выборке.

Обучение модели. Модель была реализована с использованием TensorFlow. Архитектура включала ранее подобранные параметры. Модель обучалась до стабилизации ошибки на валидационной выборке.

Прогнозирование пользовательских данных. Реализован пользовательский интерфейс, позволяющий загружать данные в формате CSV. Система автоматически обрабатывает входные данные, выполняет прогноз и отображает результаты.

Интерпретация модели. SHAP (SHapley Additive explanation) показывают, насколько выбранный признак изменил результат прогнозирования. Идея такого подхода основана на принципе, используемом в теории игр

для определения того, насколько каждый игрок при совместной игре способствует ее успешному исходу [7]. В контексте прогнозирования потребительского спроса это позволяет визуализировать причины спроса.

LIME (Local Interpretable Model-agnostic Explanations) генерирует локальные объяснения для отдельных прогнозов, выделяя ключевые признаки (например, цена, день недели), которые повлияли на предсказание в конкретный момент времени. Это особенно полезно для анализа аномалий или пиков спроса.

На рисунке 1 показан график SHAP анализа всех признаков.

Полученные данные интерпретируются следующим образом:

- значения слева от центральной вертикальной линии – это негативное влияние, справа – позитивное влияние;
- толщина линии прямо пропорциональна количеству точек наблюдения;
- чем краснее точки, тем большее значение имеет признак в этой точке.

Общий SHAP анализ

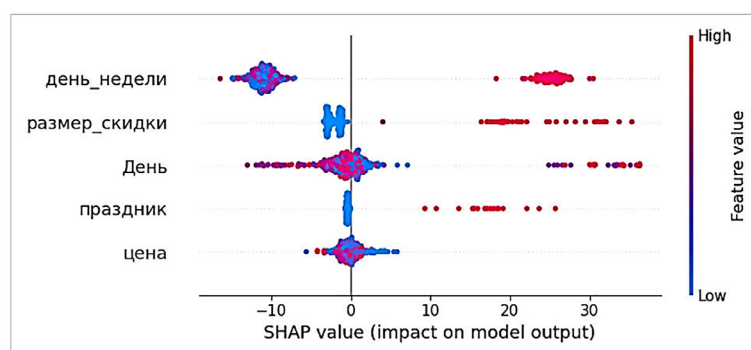


Рисунок 1. Результаты общего SHAP анализа

Источник: разработано авторами

Как видно из рисунка, наиболее яркое разделение демонстрирует день недели, это указывает на то, что

в определенные дни спрос выше, чем в остальные дни, таким образом, можно сделать предположение,

что речь идёт о выходных днях.

Размер скидки, играет роль, только когда скидка есть, тогда значение скидки оказывает большую роль на спрос.

Исходя из рисунка, большинство дней не оказывают сильного влияния на значения спроса, но в определенные дни значения спроса сильно колеблются как в большую, так и в меньшую сторону.

Праздник оказывает влияние на значение спро-

са только если есть наличие праздничного дня, тогда фактор влияния имеет высокую роль, в остальные не праздничные дни этот фактор никак не влияет на спрос.

Значение цены показывает, что нет сильного разброса в значении спроса при изменении цен.

LIME анализ строится для уже готового прогноза с использованием пользовательских данных.

На рисунках 2, 3 будут показаны результаты LIME анализа двух дней.

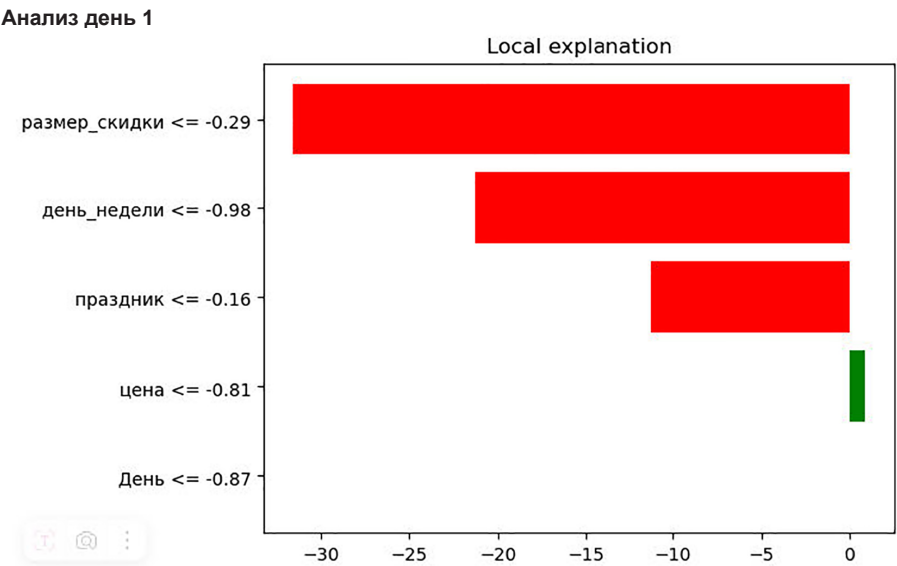


Рисунок 2. Результаты LIME анализа день 1
Источник: разработано авторами

Как видно из рисунка, позитивным фактором спроса выступает только цена, негативные факто-

ры – это отсутствие скидки, день недели, и отсутствие праздника.

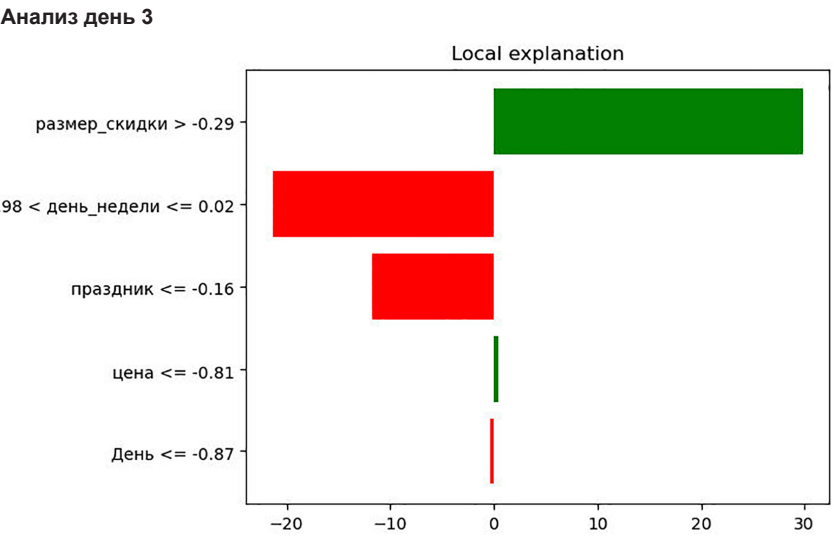


Рисунок 3. Результаты LIME анализа день 2
Источник: разработано авторами

Как видно из рисунка, позитивные факторы – это размер скидки, цена, негативными факторами выступают: будний день, отсутствие праздника, дни, не связанные с зарплатой или выходными.

Таким образом была проведена интерпретация влияния параметров как на модель прогнозирования, так и на сам прогноз, наиболее значимыми признаками оказались день недели, день месяца, размер скидки, цена не сильно влияет на спрос данного продукта.

Разработанная модель продемонстрировала высокую точность прогнозирования благодаря подбору гиперпараметров RNN, способность учитывать широкий спектр признаков, влияющих на спрос. Интерпретация с помощью SHAP показала, важность факторов на спрос. LIME позволил понять поведение модели

в частных случаях, что особенно важно при работе с отдельными товарами и днями продаж.

Практическая значимость системы заключается не только в способности давать достаточно точные прогнозы, но и в предоставлении бизнесу прозрачных оснований для принятия решений. Это позволяет использовать систему в качестве инструмента поддержки управления закупками, логистикой и маркетингом.

Разработанная интеллектуальная система на основе рекуррентной нейронной сети обладает высоким потенциалом применения в бизнес-практике. Она учитывает сложные временные зависимости и предоставляет возможности интерпретации как на уровне всей модели, так и на уровне отдельного прогноза. Комбинация точности и объяснимости делает её особенно ценной в условиях принятия управленческих решений.

Литература

1. Амирханова П. М. Методы прогнозирования спроса // Вестник науки. – 2020. – Т. 4, № 4(25). – С. 40–42.
2. Жиратков Д. Н., Липинский Л. В. О нейросетевых моделях в управлении запасами // Актуальные проблемы авиации и космонавтики. – 2015. – Т. 1, № 11. – С. 305–307.
3. Латыпова Р. Р. Прогнозирование макроэкономических показателей на основе рекуррентной нейронной сети // Известия Санкт-Петербургского государственного экономического университета. – 2025. – № 1(151). – С. 104–108.
4. Милосердов Д. И. Усовершенствованная система нейросетевого прогнозирования групп временных рядов с непрерывным обучением // Информационно-управляющие системы. – 2024. – № 1(128). – С. 20–30. – <https://doi.org/10.31799/1684-8853-2024-1-20-30>.
5. Пилипенко А. Ю. Прогнозирование спроса на товары средствами машинного обучения // StudNet. – 2022. – Т. 5, № 2. – URL: <https://elibrary.ru/item.asp?id=48038580>.
6. Сюндюкова Е. В. Прогнозирование розничного спроса с использованием нейронных сетей и макроэкономических переменных // Экономика и качество систем связи. – 2025. – № 1(35). – С. 122–131.
7. Фирюлина М. А., Каширина И. Л. Описание процесса прогнозирования проблемных состояний с применением ансамблевых методов машинного обучения // Инженерный вестник Дона. – 2022. – № 4(88). – С. 34–46.

Статья поступила в редакцию: 13.05.2025; принята в печать: 13.10.2025.

Автор прочитал и одобрил окончательный вариант рукописи.