

ПРОГНОЗИРОВАНИЕ РОЗНИЧНОГО СПРОСА С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ И МАКРОЭКОНОМИЧЕСКИХ ПЕРЕМЕННЫХ

Е.В. Сундюкова, Нижегородский государственный университет им. Н.И. Лобачевского, sundukova234k@gmail.com.

УДК 004.852:339.37

Аннотация. В связи с растущей конкуренцией между фирмами в условиях глобализации корпоративной среды и сложностью подходов к прогнозированию спроса появилось большое количество исследований по прогнозированию розничного спроса с использованием различных подходов. Однако большинство из них в значительной степени опирается на микропеременные в качестве исходных данных, тем самым игнорируя влияние макроэкономических условий на спрос домохозяйств на розничные товары. Разработанная модель, включающая внешнюю макроэкономическую информацию, превосходит модель, разработанную без макроэкономической информации, тем самым демонстрируя большой потенциал для применения в промышленности с улучшенными возможностями прогнозирования.

Ключевые слова: прогнозирование спроса; нейронные сети; *LSTM*; макроэкономика; искусственный интеллект.

FORECASTING RETAIL DEMAND USING NEURAL NETWORKS AND MACROECONOMIC VARIABLES

E.V. Syundyukova, Nizhny Novgorod State University N.I. Lobachevsky.

Annotation. Due to the growing competition between firms in a globalized corporate environment and the complexity of approaches to forecasting demand, a large amount of literature on forecasting retail demand using various approaches has appeared. However, the current literature relies heavily on macro variables as input data, thereby ignoring the impact of macroeconomic conditions on household demand for retail goods. As expected, the developed model incorporating this external macroeconomic information is superior to the model developed without this macroeconomic information, thereby demonstrating great potential for application in industry with improved forecasting capabilities.

Keywords: demand forecasting; neural networks; *LSTM*; macroeconomics; artificial intelligence.

Введение

Точное прогнозирование спроса имеет решающее значение для компании, поскольку оно напрямую влияет на ее финансовые показатели [1]. Прогнозирование будущего спроса помогает фирмам адаптировать свою деятельность к рыночному спросу и получить конкурентные преимущества в условиях высокой глобализации рынка. В настоящее время все большее число фирм переходят на передовые модели прогнозирования, стремясь улучшить управление цепочками поставок [2]. В связи с этим появляется большое количество исследований, посвященных прогнозированию будущего спроса. В литературе используются многочисленные статистические подходы, а также подходы, основанные на искусственном интеллекте (ИИ) и машинном обучении (*Machine learning, ML*), включая различные продвинутые алгоритмы с использованием нейронных сетей. При разработке моделей прогнозирования большинство

исследователей опирается на данные временных рядов о потребительском спросе на интересующие розничные товары, а также на некоторые другие важные переменные, такие как цена товара, характеристики магазина, любые специальные предложения или сделки, информация о каком-либо особом дне или событии и т.д. Однако фирмы осуществляют свою деятельность в макроэкономической среде, которая оказывает большое влияние на расходы домохозяйств. В современной литературе не используется это влияние для прогнозирования будущего потребительского спроса. Эмпирические данные свидетельствуют о том, что экономические условия существенно влияют на покупательское поведение домохозяйств и их расходы [3]. Таким образом, учет этих макроэкономических условий и факторов внешней среды в дополнение к получению исторических данных о потребительском спросе имеет решающее значение для прогнозирования потребительского спроса.

Данная проблема рассматривается в настоящем исследовании. Происходит дополнение временных рядов данных о потребительском спросе посредством включения следующих макроэкономических переменных: индекс потребительских цен (*Consumer price index, CPI*), индекс настроения потребителей (*Index of consumer sentiment, ICS*) и уровень безработицы как отражение экономической ситуации. Эти дополнительные данные используются для разработки многоуровневой модели глубоких нейронных сетей для прогнозирования будущего спроса. В частности, была реализована модель долгой краткосрочной памяти (*Long short-term memory, LSTM*), которая способна улавливать ненаблюдаемые нелинейные тенденции из наблюдаемых данных с помощью скрытых слоев и сохранять в памяти прошлую информацию. В данном исследовании удалось добиться достаточно низкой ошибки прогнозирования со среднеквадратичной ошибкой (*Root mean square deviation, RMSE*), равной 1,539. Более того, модель, разработанная с учетом этих макроэкономических условий, превосходит модель, разработанную без учета макроэкономической информации.

В исследовании дана интерпретация разработанной модели для предоставления эмпирических доказательств того, что экономическая среда оказывает сильное влияние на прогнозирование будущего спроса. Цель данного исследования – это демонстрация инструментов улучшения способности фирм прогнозировать будущий спрос на продукцию, что напрямую влияет на их финансовые показатели.

Подходы к прогнозированию розничного спроса

Из-за растущего интереса и желания фирм принимать более эффективные решения по цепочке поставок обширное количество исследовательских работ посвящено повышению точности прогнозирования с минимальной ошибкой, опираясь на различные подходы. Подходы к прогнозированию, реализованные в исследованиях, включают 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 auto-regression, VAR*), и т.д., а также продвинутые алгоритмы искусственного интеллекта с использованием нейронных сетей. В то же время используются различные гибридные модели прогнозирования, которые сочетают в себе несколько подходов к составлению прогнозов. Каждый подход имеет свои ограничения и преимущества и подходит для конкретного применения в конкретных ситуациях.

Ранние исследования по прогнозированию спроса в основном основывались на статистических подходах, таких как скользящее среднее (*Moving average, MA*),

ARIMA, *SARIMA*, *VAR* и т.д. Эти статистические подходы требуют стационарности и основаны на предположении, что отношения между переменными являются линейными по своей природе. В работе [4] внедрили различные модели *ARIMA* и выбрали подходящую модель, используя процедуру временных рядов Бокса-Дженкинса для прогнозирования спроса в пищевой компании с использованием данных временных рядов. Результаты этого исследования дают менеджерам некоторые рекомендации относительно управления цепочкой поставок продуктов. Исследователи в [5] использовали методы *ARIMA* и *SARIMA* для оценки будущего спроса Турции на первичные энергоносители в период с 2005 г. по 2020 г.

Хотя предполагается, что эти статистические подходы обеспечивают надежные результаты в случае линейных зависимостей, большинство данных временных рядов реального мира отображают нелинейные зависимости, и в этом случае предположение о линейных зависимостях нарушается. Методы искусственного интеллекта, такие как нейронные сети и различные древовидные методы, такие как случайные леса, методы опорных векторов (*Support vector machines, SVM*) и т.д., преодолевают это ограничение и могут фиксировать нелинейные связи между переменными. В нескольких исследованиях сравниваются эффективность статистических подходов с эффективностью методов компьютерного анализа. Эмпирические данные, полученные в ходе этих исследований, свидетельствуют о том, что модели, разработанные с использованием алгоритмов нейронных сетей, как правило, демонстрируют более высокую эффективность прогнозирования. Например, авторы [6] внедрили *ARIMA* и *LSTM* для прогнозирования будущего спроса на заказы, и они нашли доказательства того, что *LSTM* демонстрирует повышенную эффективность в случае краткосрочного прогнозирования. В [7] сравнили эффективность нейросетевой модели с *ARIMA* Бокса-Дженкинса, а также с регрессионными моделями, разработав модели для прогнозирования ежедневных максимальных уровней содержания озона. Результаты этого исследования также свидетельствуют о том, что модель нейронных сетей превосходит регрессионные модели и *ARIMA*. В [9] сравнили несколько статистических методов, таких как *MA* и *ARIMA*, с моделями нейронных сетей, используя базу данных инвентаризации компании *Panasonic Refrigeration Devices*, расположенной в Сингапуре. Результаты их исследования также подтверждают выводы аналогичных исследований о том, что модели, основанные на нейронных сетях, обеспечивают более высокую эффективность по сравнению с другими статистическими методами.

Алгоритмы нейронных сетей обладают определенными преимуществами. Эти модели являются непараметрическими и основаны на данных, поэтому они могут фиксировать любые нелинейные взаимосвязи без предварительного определения взаимосвязей между переменными в модели, что делает алгоритмы нейронных сетей более подходящими для работы с реальными данными, содержащими множество сложных нелинейных зависимостей. Учитывая эту гибкость и превосходную производительность, в современной литературе для прогнозирования приложений в основном используются модели на основе нейронных сетей. В исследовании [10] используют различные модели машинного обучения для прогнозирования спроса в ресторанах. В [11] соавторы разработали модель прогнозирования спроса с использованием модели нейронных сетей Холта-Винтера для розничной торговли. В [12] разрабатывают различные модели для прогнозирования будущего спроса на ликеро-водочную продукцию и находят доказательства значительного повышения точности *LSTM*-модели на основе нейронных сетей по сравнению с другими моделями. В работе [13] разработали модель на основе нейронных сетей для прогнозирования будущего розничного

спроса, используя данные, предоставленные *Walmart*, и предложили интерпретацию разработанной модели с использованием библиотеки *Shapely*. Существует также несколько других исследований, в которых разрабатываются модели для прогнозирования будущего спроса с использованием различных алгоритмов [14]. В нескольких исследованиях представлены систематические обзоры соответствующих статей в литературе [15].

В настоящем исследовании исторические данные о розничном спросе дополняются макроэкономической информацией. Используются расширенные данные для разработки модели *LSTM* для прогнозирования будущего спроса. В частности, использовались индекс потребительских цен *ICS* и уровень безработицы, а также данные о потребительском спросе во временных рядах для обучения модели. В большом количестве эмпирических исследований было задокументировано влияние этих выбранных макроэкономических переменных на объяснение спроса домашних хозяйств. Например, в исследовании [16] показано, что индекс потребительских цен, наряду с другими факторами, играет важную роль в объяснении спроса на воздушные перевозки в Нигерии. Аналогичные данные были представлены в работе [17]. Результаты исследования [18] свидетельствуют о том, что *ICS* полезен для прогнозирования будущих потребительских расходов. Аналогичным образом, [19] также подтверждает влияние безработицы на расходы домохозяйств на туризм. Результаты настоящего исследования однозначно подтверждают необходимость включения в модель информации о внешнеэкономических условиях, так как при включении этих данных в модель значение *RMSE* уменьшается.

Набор данных и методы анализа

Предварительная обработка данных и разработка функциональных возможностей

Данные, использованные в этом исследовании, представляют собой исторические данные о товарах, проданных в количестве 3000 наименований в десяти различных магазинах из трех штатов США (Калифорния, Техас и Висконсин) за пять лет. В каждом магазине в трех штатах есть одинаковый набор из 3000 товаров. В дополнение к историческим данным о проданном продукте, набор данных также включает идентификатор продукта, цену продукта, категорию продукта, отдел, информацию о магазине, любые рекламные акции, день недели и любые события в день продажи продукта. Эта информация содержится в трех различных наборах данных: исторические данные о продажах, календарные данные и данные о цене продукта. Набор исторических данных о продажах – это обширный набор данных, содержащий информацию о товарах, проданных в количестве 3000 наименований за период в 1800 дней, со столбцом для каждого уникального дня и строкой для каждого уникального продукта. Этот набор данных также содержит идентификатор продукта, категорию продукта и местоположение магазина. Набор данных календаря содержит информацию о специальных мероприятиях или праздниках, рекламных акциях, типе мероприятия и т.д. на каждый день в каждом штате. Набор данных о ценах на товары содержит данные о цене каждого товара в каждом магазине на каждый день. Макроэкономические переменные, рассматриваемые в этом исследовании, включают индекс потребительских цен *ICS* и уровень безработицы. Исторические данные по индексу потребительских цен и безработице собраны из базы данных Всемирного банка по показателям мирового развития (*World development indicators, WDI*), а исторические данные по *ICS* собраны с веб-сайта Мичиганского университета. Сначала происходит объединение набора данных календаря с

макроэкономическими данными в столбце «данные», чтобы получить объединенный набор данных со строкой для каждого дня. Затем этот объединенный набор данных объединяется с набором данных о ценах на продукты и наборами данных о продажах. Этот окончательный набор данных содержит информацию о продажах, ценах и рекламных акциях для каждого из 3000 товаров из каждого магазина в трех штатах за каждый из 1800 дней. Произведено ограничение набора данных последними 500 днями, чтобы избежать каких-либо старых тенденций, которые могут оказаться недействительными для текущих дней. Кроме того, в набор включаются скользящие средние значения, а также скользящее стандартное отклонение нескольких запаздывающих значений цен продажи как часть входных данных.

Нейронные сети

В данном исследовании использовался алгоритм *LSTM* для обучения модели прогнозирования. Нейронные сети состоят из сети узлов с несколькими слоями, такими как входные слои, скрытые слои и выходные слои. Каждый узел на входном и скрытом уровнях связан с каждым узлом на следующем уровне с определенным весом. Ряд узлов, соединенных друг с другом, способен распознавать сложные, нелинейные и ненаблюдаемые взаимосвязи, скрытые в данных. Каждый узел можно рассматривать как линейную регрессию с соответствующими весами. Начальные веса присваиваются в начале обучения. После каждого цикла прямого и обратного распространения эти веса корректируются путем минимизации функции затрат. Скорость корректировки определяется параметром, называемым скоростью обучения. Существует несколько типов алгоритмов нейронных сетей, и *LSTM* является одним из них. *LSTM* обладает памятью, которая зависит от наблюдений во времени, и, следовательно, подходит для работы с временными рядами. В исследовании использовался многослойный алгоритм нейронной сети *LSTM* для обучения модели.

Влияние макроэкономических переменных на модели прогнозирования спроса

Прежде чем перейти к разработке модели нейронной сети, в первую очередь, были исследованы данные во временной области, чтобы понять любые временные зависимости. Одной из проблем в текущем анализе является тот факт, что большинство отдельных временных рядов для продуктов в имеющемся наборе данных разрежены, и на основе этих данных трудно изучить какую-либо временную зависимость. Поэтому исследуется временная зависимость с использованием агрегированных данных, т. е. общей суммы продаж всех товаров из всех десяти магазинов в трех штатах, а не продажи отдельных товаров.

На рис. 1 и 2 показаны автокорреляция и частичная автокорреляция общего объема ежедневных продаж соответственно. Оценки автокорреляции показывают, что текущий общий объем продаж коррелирует с его значениями с запаздыванием более чем на 250 дней. Коэффициент автокорреляции отражает как прямое, так и косвенное влияние запаздывающих значений на текущее значение, тогда как частичная автокорреляция отражает только прямое влияние, тем самым устраняя любые косвенные эффекты. Оценки частичной автокорреляции также убедительно свидетельствуют о последовательной зависимости и демонстрируют, что текущий общий объем продаж напрямую коррелирует с запаздывающими значениями с задержкой до 40 дней. Наличие сильной временной зависимости с большими запаздывающими значениями указывает на то, что временные ряды в этом наборе данных обладают высокой информативностью при прогнозировании будущих значений. Низкое значение *RMSE* разработанной модели *LSTM* в этом

исследовании, согласуется с выводом о сильной последовательной корреляции в наборе данных.

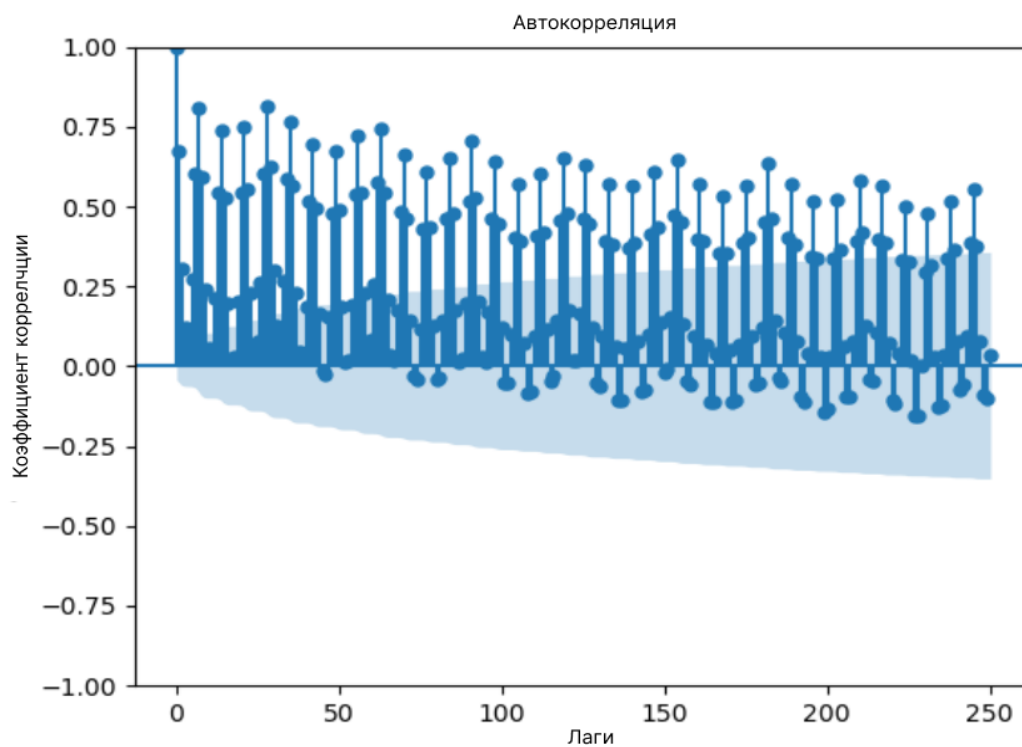


Рисунок 1

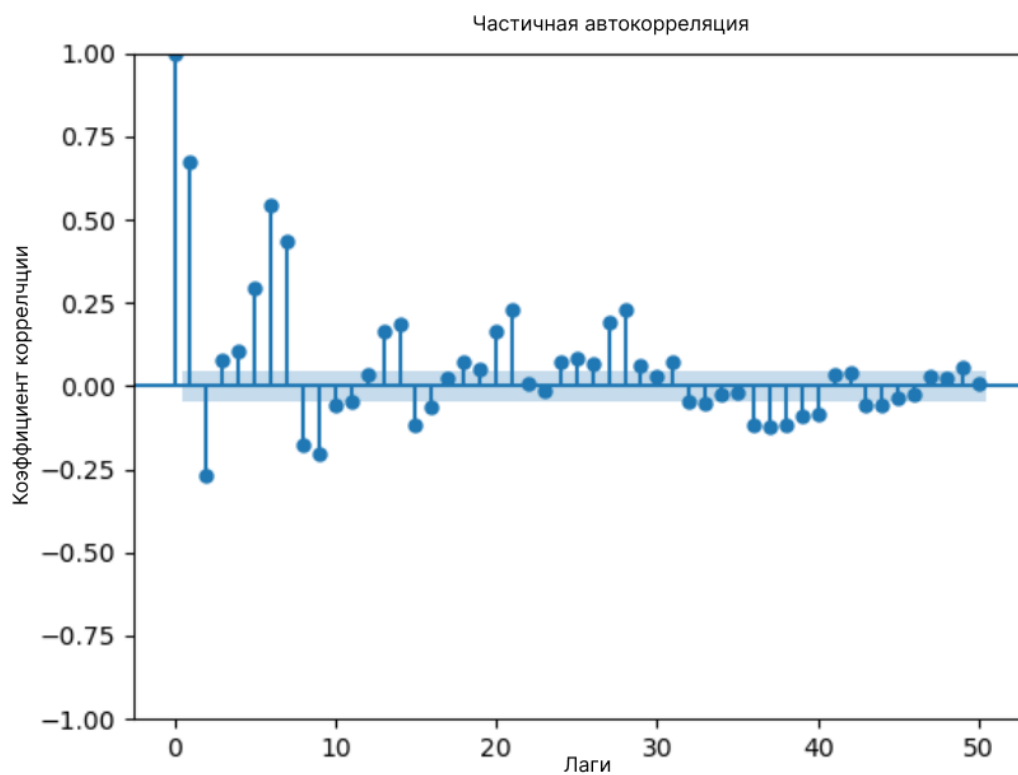


Рисунок 2

Результаты, представленные в каждой версии модели, основаны на одной и той же архитектуре модели *LSTM*, которая состоит из двух скрытых уровней и

предназначена для прогнозирования спроса на 28 будущих временных шагов. Первоначально модель была ограничена только значениями по штату Калифорния, а затем была расширена, посредством включения еще двух штатов: Техаса и Вайоминга.

Чтобы избежать каких-либо прошлых тенденций, отсутствующих в текущих данных, которые могут привести к нежелательным результатам, модели были переданы данные временных рядов за последние 500 дней для ее обучения. Базовая модель не включала цены на продукты, а затем была расширена с добавлением цен. Чтобы сравнить предлагаемую модель с описанной в упомянутых выше исследованиях, приводятся характеристики модели, включающей макроэкономические переменные, и модели без макропеременных.

Подробные сравнительные характеристики исходных моделей (по штату Калифорния) представлены в табл. 1. При включении внешних макроэкономических переменных модель демонстрирует превосходную производительность со значением $RMSE$, равным 1,926, в то время как модель без макропеременных имеет большее значение $RMSE$, равное 2,124. К аналогичному выводу можно прийти, если включить цены на продукты в базовые модели, которые не содержат цен на продукты. Модель $LSTM$ с макропеременными по-прежнему превосходит традиционную модель без макроэкономических переменных. Далее происходит расширение модели путем включения данных из всех трех штатов. Сравнительные характеристики расширенной модели (по штатам Калифорния, Техас, Вайоминг) представлены в табл. 2, и ее результаты аналогичны базовой модели. Модель с экономической информацией превосходит модели без экономической и экологической информации. Эти результаты согласуются с данными литературы, которые свидетельствуют о том, что внешняя экономическая среда может объяснить спрос домашних хозяйств. Во всех случаях разработанная модель прогнозирования, включающая те внешние переменные, которые могут объяснить спрос домашних хозяйств, превосходит модель, разработанную без учета этих внешних переменных.

Таблица 1.

Модель	$RMSE$	
	Без макропеременных	С макропеременными
$LSTM$ без цен на продукты	2,124	1,926
$LSTM$ с ценами на продукты	1,975	1,895

Таблица 2.

Модель	$RMSE$	
	Без макропеременных	С макропеременными
$LSTM$ с ценами на продукты	1,578	1,539

В задаче требуется определение характеристик, которые важны для прогнозирования будущих значений. Хотя, в целом, модели глубокого обучения обеспечивают более высокую эффективность прогнозирования по сравнению с обычными статистическими подходами, модели глубокого обучения считаются черными ящиками, которые не дают никакой интерпретации того, как делаются прогнозы. В данном исследовании использовалась библиотека *Shapely* для интерпретации модели $LSTM$ и изучения вклада каждой макроэкономической

характеристики наряду с другими характеристиками в прогнозирование будущего спроса. Значения *Shapely* используют подход теории совместных игр для обеспечения желаемых свойств и широко используются в литературе для объяснения моделей вычислительного интеллекта [4]. Линейная диаграмма значений важности признаков, представленная на рис. 3, подтверждает, что эти макроэкономические переменные действительно важны и вносят свой вклад в результаты моделирования. В частности, *ICS*, уровень безработицы и индекс потребительских цен являются второй, седьмой и девятой важными характеристиками, соответственно, из шестнадцати основных характеристик, включенных в модель.

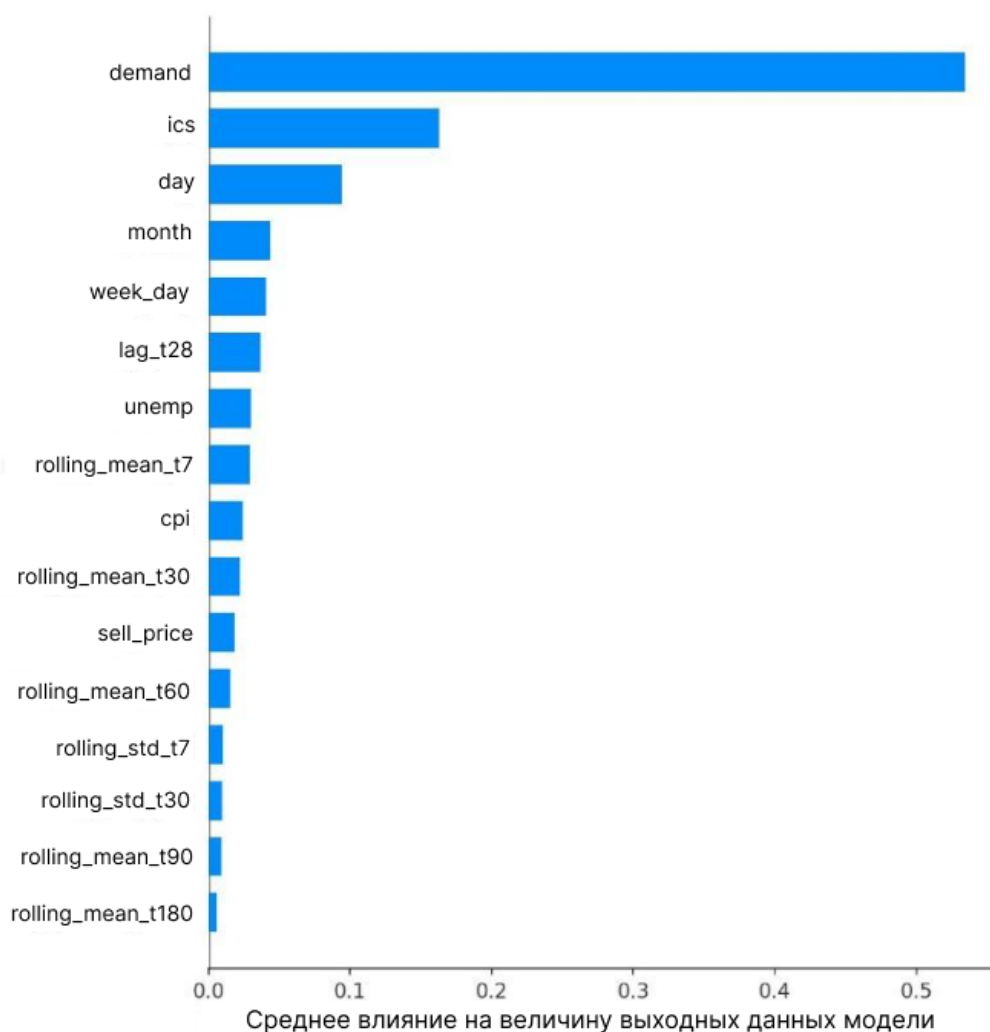


Рисунок 3

Заключение

В настоящем исследовании была разработана модель *LSTM* с включением макроэкономических переменных наряду с другими характеристиками, связанными со спросом, для прогнозирования розничного спроса с повышенной точностью. Основные выводы этого исследования демонстрируют, что включение макроэкономических переменных в модели улучшает эффективность прогнозирования. Были исследованы различные выборки с различными характеристиками. Также исследуется важность предложенных переменных по сравнению с другими соответствующими данными. Данное исследование дополняет существующую литературу, предлагая решение для улучшения

возможностей прогнозирования, тем самым создавая возможности для фирм по улучшению управления цепочками поставок и повышению финансовых показателей. Хотя результаты этого исследования основаны на модели нейронных сетей, пока неизвестно, применимы ли эти результаты к другим соответствующим методам машинного обучения и статистическим подходам. Будущие исследования могут быть направлены на изучение применимости этих результатов к более широкому набору соответствующих методов и подходов.

Литература

1. Aderamo A.J. Demand for Air Transport in Nigeria, *J Economics* 1, 2010. – № 1. – P. 123-131.
2. Wang C.C., Chien C.H., Trappey A. On the Application of ARIMA and LSTM to Predict Order Demand Based on Short Lead Time and On-Time Delivery Requirements. *Processes*, 2021. – № 9. – P. 1157.
3. Aamer A., Eka Yani L., Alan Priyatna I. Data analytics in the supply chain management: Review of machine learning applications in demand forecasting. *Operations and Supply Chain Management: An International Journal*, 2021. – № 14 (1). – P. 1-13.
4. Chen J., Koju W., Xu S., and Liu Z. Sales Forecasting Using Deep Neural Network and SHAP Techniques. *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, 2021. – P. 135-138.
5. Chung D., Lee C.G., Yang S. A hybrid machine learning model for demand forecasting: Combination of K-Means, Elastic-Net and Gaussian Process Regression. *International Journal of Intelligent Systems and Applications in Engineering*, ISSN 2147-6799, 2023.
6. Dou Z., Sun Y., Zhang Y., Wang T., Wu C., and Fan S. Regional Manufacturing Industry Demand Forecasting: A Deep Learning Approach. *Applied Sciences*, 2021. – № 11. – 6199 p.
7. Tanizaki T., Hoshino T., Shimmura T., Takenaka T. Demand forecasting in restaurants using machine learning and statistical analysis. *Procedia CIRP*, 2019. – № 79. – P. 679-683.
8. Ediger V.S. Akar S. ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy*, 2007. – № 35 (3). – P. 1701-1708.
9. Fattah J., Ezzine L., Aman Z., Moussami H.E., Lachhab A. Forecasting of demand using ARIMA model. *Engineering Business Management*, 2018. – № 10. – P. 1-9.
10. Gao Y-F, Liang Y-S, Liu Y, Zhan S-B and Ou Z-W. A neural-network-based forecasting algorithm for retail industry. *Proceedings of the eighth international conference on machine learning and Cybernetics*, Baoding, 2009.
11. Huth W.L., Eppright D.R., Taube P.M. The Indexes of Customer Sentiment and Confidence: Leading or Misleading Guide to Future Buyer Behavior. *Journal of Business Research*, 1994. – № 29 (3). – P. 199-206.
12. Icoz O., Var T., Kozaka M. Tourism Demand in Turkey. *Annals of Tourism Research*, 1998. – № 25. – P. 236-240.
13. Scholdra T.P., Wichmann J.R.K., Eisenbeiss M., Reinartz W.J. Households Under Economic Change: How Micro-and Macroeconomic Conditions Shape Grocery Shopping Behavior. *Journal of Marketing*, 2022. – № 86 (4). – P. 95-117.
14. Kandananond K.A. A comparison of various forecasting methods for autocorrelated time series. *International Journal of Engineering Business Management*, 2012. – № 4. – P. 4.

15. Lee I., Mangalaraj G. Big data analytics in supply chain management: A systematic literature review and research directions. *Big Data and Cognitive Computing*, 2022. – № 6. – P. 17.
16. Mitrea C.A, Lee CKM, Wu Z. A comparison between neural networks and traditional forecasting methods: a case study. *Int J Eng Bus Manage*, 2009. – № 1. – P. 11.
17. Prybutok V.R, Yi J., Mitchell D. Comparison of neural network models with ARIMA and regression models for prediction of Houston’s daily maximum ozone concentrations. *European Journal of Operational Research*, 2000. – № 122 (1). – P. 31-40.
18. Palkar A., Deshpande M., Kalekar S., Jaswal S. Demand Forecasting in Retail Industry for Liquor Consumption using LSTM. *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, IEEE Xplore Part Number: CFP20V66-ART*, 2020.
19. Obaid H.M.A. Factors determining housing demand in Saudi Arabia. *International Journal Economics and Financial Issues*, 2020. – № 10. – P. 150-157.