Industrial Management Studies
Vol. 24, No.81, 2026; P 35-66
jims.atu.ac.ir
DOI: 10.22054/jims.2026.91974.3023

ATU PRESS

Original Research
Accepted: 13/06/2026
Revised: 31/05/2026
Received: 22/03/2026
eISSN: 2476-602X
ISSN: 2251/8029

Dynamic Retail Demand Forecasting Using a Transformer-Based Approach Integrating Network Effects, Temporal Lags, and Product Semantics

Fatemeh Zare Baghiabad *

Department of Industrial Engineering, Faculty of Engineering, Ardakan University, Ardakan, Iran

Abstract

Accurate demand forecasting in online retail is a fundamental challenge in supply chain management, as demand patterns are typically influenced by complex temporal relationships, cross-product network interactions, and item-specific semantic features. To address this challenge, this study introduced a novel forecasting framework based on a Transformer architecture, capable of modeling long-lag temporal dependencies, cross-product network effects, and item semantic information. This empirical-computational research evaluated the proposed model using online retail II datasets. The proposed model was subsequently trained utilizing a multi-head attention mechanism within the Transformer framework, incorporating semantic representations extracted from product descriptions. The results demonstrated that the model achieved a Mean Absolute Error (MAE) of 2.68 and a Root Mean Square Error (RMSE) of 3.85, outperforming both the Autoregressive Integrated Moving Average model (MAE: 4.37, RMSE: 5.96) and the Gradient Boosting algorithm (MAE: 3.91, RMSE: 5.12). This improvement indicated that the framework effectively identifies complex demand patterns and cross-product relationships.

Introduction

Precise and adaptive demand forecasting is paramount in the digitally transformed retail landscape. Online platforms generate vast data, presenting opportunities and challenges. Demand is shaped by trends, seasonality, promotions, network externalities (influences between complementary/substitute products), and product characteristics. Traditional models like ARIMA capture linear temporal dependencies but often miss nonlinear relationships and external factors. While

* Corresponding Author: f.zare@ardakan.ac.ir
How to Cite: Zare Baghiabad, F. (2026). Dynamic Retail Demand Forecasting Using a Transformer-Based Approach Integrating Network Effects, Temporal Lags, and Product Semantics, *Industrial Management Studies, 24(81),* 35-66.

advanced machine learning models handle nonlinearity, they may not adequately integrate diverse data types (time series, network structures, textual semantics). This study proposed a unified deep learning framework, the Transformer, excelling at capturing long-range dependencies via its attention mechanism. By explicitly modeling temporal lags, product network effects, and semantic features, our approach aimed for a more comprehensive and accurate representation of demand generation in online retail, pushing forecasting boundaries.

Literature Review

Demand forecasting in online retail is vital for supply chain and inventory management. Traditional time-series models like ARIMA and Exponential Smoothing capture linear temporal dependencies but are limited by the increasing complexity of modern online retail environments, characterized by intricate temporal relationships, inter-product network effects, and rich semantic item characteristics (Tiwari, 2025; Chowdhury et al., 2025). Contemporary deep learning architectures, including RNNs (LSTMs, GRUs), model complex temporal patterns but may struggle with long-range dependencies or nuanced interactions (Cai et al., 2021).

The Transformer architecture, with its self-attention mechanism, excels at capturing long-range and context-dependent relationships, revolutionizing sequence modeling (Rahman et al., 2025). Its adoption in time-series forecasting, including retail demand prediction, is growing (Wang, 2025; Li, 2023). However, existing Transformer-based research often focuses primarily on historical sales data (temporal lags) in isolation. A significant gap remains in simultaneously modeling the multifaceted influences on demand, specifically the interplay of:

- **Network Effects:** Product demand is interdependent; influenced by complements and substitutes (Caetano et al., 2025). Capturing these structural interdependencies is crucial for dynamic patterns, especially in large catalogs.
- **Temporal Lags:** The impact of events or changes may not be immediate. Accurately modeling these varied time delays is vital for precise forecasting (Tarighat et al., 2025).
- **Product Semantics:** Product characteristics and descriptions significantly influence consumer choice. Leveraging semantic information can uncover hidden relationships and improve accuracy (Rahikka & Mikkola, 2025).

While studies have explored these factors individually or in limited combinations (e.g., Cai et al., 2021; Wang, 2025), a comprehensive framework integrating temporal dependencies, network effects, and product semantics within a single Transformer architecture remains an active research area. This study

bridged this gap by proposing a novel Transformer-based framework to jointly model these dimensions for a more accurate and holistic approach to dynamic retail demand forecasting.

Method

This study employed an empirical-computational approach using the Online Retail II dataset. Data preprocessing involved handling missing values, correcting errors, and removing irrelevant transactions. The dataset was split chronologically into an 80% training set and a 20% testing set.

Our core forecasting model was the Transformer architecture, utilizing its self-attention mechanism to weigh historical time steps for predicting future demand, thus capturing complex temporal patterns and long-range dependencies. Positional encodings were added to input embeddings for sequence information.

Our framework innovatively integrated three key information sources:

1. **Temporal Lags:** Explicitly accounts for lagged demand patterns. While attention captures temporal relationships, features representing past demand values at different lags can be incorporated.
2. **Network Effects:** Models product interdependencies through a product interaction graph (nodes=products, edges=relationships like co-purchase frequency). Graph-derived features (e.g., centrality) processed by techniques like Graph Convolutional Networks (GCNs) are integrated.
3. **Product Semantics:** Extracts rich semantic embeddings from product descriptions using NLP techniques (e.g., pre-trained models). These capture nuanced characteristics and are integrated alongside temporal and network features.

The model was trained end-to-end using Mean Squared Error (MSE) or Mean Absolute Error (MAE), optimized via an Adam optimizer. Hyperparameters were tuned for optimal performance. Final evaluation used MAE and RMSE on the test set.

Results

The empirical evaluation on the Online Retail II dataset showed the superiority of our proposed Transformer-based framework. On the test set, the model achieved an MAE of 2.68 and an RMSE of 3.85. This substantially outperformed baseline models: ARIMA yielded MAE 4.37, RMSE 5.96; XG-Boost achieved MAE 3.91, RMSE 5.12. These quantitative results highlighted the enhanced predictive accuracy of our integrated approach, attributed to its ability to capture the complex interplay of temporal, network, and semantic factors. Visualizations of predicted versus actual demand confirmed the model’s superior tracking of demand fluctuations. Analysis of attention weights also provided insights into

influential historical periods and product relationships, offering a degree of interpretability.

Discussion

The findings underscored the advantages of a Transformer-based architecture for dynamic retail demand forecasting when augmented with explicit modeling of temporal lags, network effects, and product semantics. The accuracy improvements over ARIMA and XG-Boost highlighted the limitations of models neglecting the interconnected nature of retail demand. Our approach achieved a more comprehensive representation by jointly considering these factors. The attention mechanism effectively captured long-range temporal dependencies and contextual relationships. Explicit incorporation of network effects revealed the critical influence of product interdependencies, often overlooked in single-item forecasting. Leveraging product semantics allowed deeper understanding of demand drivers beyond historical sales. This integrated approach offered retailers enhanced capabilities for inventory management, demand planning, and strategic decision-making, potentially reducing costs and stockouts while improving customer satisfaction. Attention weight analysis offered valuable insights for model trust and adoption.

Conclusion

This research successfully developed and validated a novel Transformer-based framework for dynamic retail demand forecasting, integrating temporal lags, network effects, and product semantics. Empirical results on the Online Retail II dataset demonstrated the model's superiority, achieving significantly lower error metrics (MAE: 2.68, RMSE: 3.85) compared to ARIMA (MAE: 4.37, RMSE: 5.96) and XG-Boost (MAE: 3.91, RMSE: 5.12). This work contributed a unified and highly accurate methodology for forecasting in complex retail environments, offering substantial practical implications for inventory management and strategic decision-making. Future research could incorporate additional external factors, explore advanced GNN architectures, and develop more sophisticated interpretability methods for real-world adoption.

Table 1. Comparison of Demand Forecasting Model Performance

Model	Mean Absolute Error (MAE)	Mean Squared Error (MSE)
ARIMA	37.4	96.5
XG-Boost	91.3	12.5
Proposed Transformer	68.2	85.3

Table 2. Sensitivity Analysis of the Proposed Model (Baseline RMSE is 1.96)

Main Group	Parameter	Baseline Value	Applied Change	New RMSE	RMSE Change (%)
Temporal Dynamics	Input Window Length	12	10 15	2.04 2.00	+4.2% +1.8%
	Prediction Horizon	3	5	2.11	+7.8%
Numerical Inputs	Count	10.33	±10% Random Noise	1.98	+0.8%
	Price	4.69	±10% Random Noise	1.99	+1.3%
Model Architecture	Graph Information	Present	Component Removal	2.09	+6.5%
	Attention Mechanism	Present	Weighting Noise	1.97	+0.5%
Optimization	Learning Rate	0.0005	0.001 0.00025	1.99 2.01	+1.3% +2.6%

Keywords: Retail Demand, Network Effects, Time Lags, Product Semantics, Transformer Architecture.

مقاله پژوهشی
تاریخ ارسال: ۱۴۰۵/۰۱/۰۲
تاریخ بازنگری: ۱۴۰۵/۰۳/۱۰
تاریخ پذیرش: ۱۴۰۵/۰۳/۲۳
ISSN: 2251/8029
eISSN: 2476-602X

مطالعات مدیریت صنعتی
سال بیست و چهارم، شماره ۸۱، تابستان ۱۴۰۵، ص ۳۵ تا ٦٦
jims.atu.ac.ir
DOI: 10.22054/jims.2026.91974.3023

پیش‌بینی پویای تقاضای خرده‌فروشی مبتنی بر رویکرد ترنسفورمر با تلفیق اثرات شبکه‌ای، تأخیرهای زمانی و معناشناسی محصول

فاطمه زارع باقی‌آباد * | گروه مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه اردکان، اردکان، ایران

چکیده

پیش‌بینی دقیق تقاضا در خرده‌فروشی آنلاین یکی از چالش‌های اساسی در مدیریت زنجیره تأمین و برنامه‌ریزی موجودی محسوب می‌شود، زیرا الگوهای تقاضا معمولاً تحت تأثیر روابط پیچیده زمانی، تعاملات شبکه‌ای میان محصولات و ویژگی‌های معنایی اقلام قرار دارند. بسیاری از رویکردهای سنتی سری زمانی و حتی برخی مدل‌های یادگیری عمیق در مدل‌سازی هم‌زمان این وابستگی‌های پویا و غیرخطی با محدودیت‌هایی مواجه‌اند. در پاسخ به این چالش، پژوهش حاضر یک چارچوب پیش‌بینی نوین مبتنی بر معماری ترنسفورمر ارائه داده که قادر است وابستگی‌های زمانی با تأخیر، اثرات شبکه‌ای میان محصولات و اطلاعات معناشناختی کالاها را به‌صورت یکپارچه مدل‌سازی کند. این پژوهش از نوع تجربی- محاسباتی بوده و برای ارزیابی مدل پیشنهادی از مجموعه‌داده واقعی خرده‌فروشی آنلاین دو استفاده شده است. پس از انجام مراحل پیش‌پردازش داده‌ها و حذف رکوردهای ناقص، داده‌ها به مجموعه‌های آموزش و آزمون تقسیم شدند. سپس مدل پیشنهادی با بهره‌گیری از سازوکار توجه چندسری در چارچوب ترنسفورمر و با استفاده از بازنمایی‌های معنایی استخراج‌شده از توضیحات محصولات آموزش داده شد. بر اساس نتایج، مدل ترنسفورمر پیشنهادی با دستیابی به مقادیر میانگین قدرمطلق خطا معادل ۲٫۶۸ و ریشه میانگین مربعات خطا معادل ۳٫۸۵، عملکرد بهتری نسبت به مدل‌های خودهمبسته میانگین متحرک انباشته (با میانگین قدرمطلق خطا برابر ۴٫۳۷ و ریشه میانگین مربعات خطا برابر ۵٫۹۶) و الگوریتم تقویت گرادیانت (با میانگین قدرمطلق خطا برابر ۳٫۹۱ و ریشه میانگین مربعات خطا برابر ۵٫۱۲) ارائه می‌دهد. این بهبود نشان داد که چارچوب پیشنهادی قادر است الگوهای پیچیده تقاضا و روابط میان‌محصولی را به‌طور مؤثر شناسایی کند.

کلیدواژه‌ها: تقاضای خرده‌فروشی، اثرات شبکه‌ای، تأخیرهای زمانی، معناشناسی محصول، معماری ترنسفورمر.

* نویسنده مسئول: f.zare@ardakan.ac.ir

مقدمه

پیش‌بینی پویای تقاضای خرده‌فروشی یکی از چالش‌های بنیادین در مهندسی سیستم‌های خرده‌فروشی و مدیریت زنجیره تأمین دیجیتال است، به‌ویژه در بستر خرده‌فروشی آنلاین که با نوسانات شدید تقاضا، چرخه‌های عمر کوتاه محصول و وابستگی‌های پیچیده میان اقلام مواجه است. داده‌های واقعی خرده‌فروشی آنلاین، مانند پایگاه داده خرده‌فروشی آنلاین دو[1]، نشان می‌دهند که الگوی فروش از اصل پارتو پیروی می‌کند؛ به‌طوری‌که کمتر از ۱۵ درصد اقلام، بیش از ۷۰ درصد از کل تراکنش‌های فروش را به خود اختصاص می‌دهند. بااین‌حال، رفتار تقاضای همین اقلام پرفروش به‌شدت تحت تأثیر نوسانات غیرمنتظره محصولات مکمل قرار دارد؛ برای مثال، تغییر قیمت یا موجودی یک کالای جانشین، بلافاصله منجر به جهش یا افت تقاضا برای کالای هدف می‌شود که این وابستگی‌های متقابل، به دلیل وجود تأخیرهای زمانی در زنجیره تأمین و پیچیدگی‌های معنایی سبد خرید، با مدل‌های سری زمانی سنتی قابل‌ردیابی نیست. این شواهد بیانگر آن است که مسئله‌ی پیش‌بینی پویای تقاضا در خرده‌فروشی، ذاتاً یک مسئله‌ی چندبعدی است که در آن، اثرهای شبکه‌ای میان اقلام، تأخیرهای زمانی در پاسخ تقاضا و ساختار معنایی/مفهومی محصولات به‌طور هم‌زمان نقش ایفا می‌کنند. پژوهش‌های اخیر نشان می‌دهند که مدل‌های کلاسیک سری زمانی و حتی بسیاری از مدل‌های یادگیری عمیق، در درک این وابستگی‌های متقابل پویا و غیرخطی میان اقلام با محدودیت‌های جدی مواجه‌اند، به‌ویژه زمانی که اثرات شبکه‌ای و ساختار معنایی محصولات نادیده گرفته می‌شود (Tarighat et al., 2025). این شکاف، ضرورت توسعه رویکردهایی را برجسته می‌سازد که به‌طور یکپارچه، پویایی تقاضا، اثرهای شبکه‌ای، تأخیرهای زمانی و معناشناسی محصول را در فرآیند پیش‌بینی لحاظ کنند. در پاسخ به این چالش، معماری‌های مبتنی بر ترنسفورمر به دلیل توانایی ذاتی در مدل‌سازی وابستگی‌های بلندمدت و روابط غیرمحلی، به‌طور فزاینده‌ای در پیش‌بینی فروش و تقاضا موردتوجه قرار گرفته‌اند (M. A. et al., 2025). بااین‌حال، مطالعات موجود عمدتاً تمرکز خود را بر وابستگی‌های زمانی منفرد گذاشته و کمتر به نقش هم‌زمان اثرهای شبکه‌ای بین اقلام پرفروش، تأخیرهای زمانی نامتقارن و معناشناسی محصول پرداخته‌اند (Wang, 2025). این خلأ پژوهشی، به‌ویژه در خرده‌فروشی آنلاین، اهمیت بالایی دارد؛ زیرا شواهد تجربی نشان می‌دهد که تغییر تقاضای یک کالای پرفروش می‌تواند با وقفه زمانی مشخص، الگوی تقاضای مجموعه‌ای از محصولات مرتبط را دستخوش تغییر کند و نادیده‌گرفتن این پویایی‌ها منجر به کاهش معنادار دقت پیش‌بینی و تصمیم‌گیری‌های نادرست موجودی می‌شود (Cai et al., 2021). بر این اساس، سؤال اصلی این پژوهش

1 online retail II

آن است که چگونه می‌توان با بهره‌گیری از رویکرد ترنسفورمر و تلفیق اثرات شبکه‌ای، تأخیرهای زمانی و معناشناسی محصول، مدلی دقیق‌تر و پویاتر برای پیش‌بینی تقاضای خرده‌فروشی ارائه داد؟

از منظر علمی، مسئله‌ی پیش‌بینی تقاضای خرده‌فروشی به‌دلیل ماهیت پویای آن، چالش‌های نظری مهمی را برای توسعه‌ی مدل‌های پیش‌بینی مطرح کرده است؛ به‌ویژه در تبیین روابط پیچیده‌ی غیرخطی میان اقلام، اثرات شبکه‌ای محصولات جانشین و مکمل و تأثیرگذاری معناشناسی محصول بر رفتار خرید مشتریان. مدل‌های سنتی سری زمانی، علی‌رغم توانمندی‌های خود، در بازنمایی هم‌زمان این مؤلفه‌ها ناتوان‌اند؛ ازاین‌رو، این پژوهش باهدف توسعه یک رویکرد یکپارچه مبتنی بر معماری‌های ترنسفورمر، به دنبال پاسخ به این شکاف نظری و ارتقای سطح دقت و تبیین‌پذیری در حوزه‌ی یادگیری عمیق برای پیش‌بینی فروش است (Tarighat et al., 2025). از منظر کاربردی، پیش‌بینی دقیق تقاضا، ستون فقرات تصمیم‌گیری‌های هوشمند در مدیریت زنجیره‌ی تأمین و عملیات خرده‌فروشی دیجیتال محسوب می‌شود. با توجه به اینکه داده‌های واقعی نشان‌دهنده‌ی تأثیر مستقیم تغییرات قیمت یا موجودی یک کالای مرتبط بر تقاضای کالای هدف است، پیش‌بینی دقیق این پویایی‌ها مستقیماً به بهینه‌سازی سطوح موجودی، کاهش هزینه‌های انبارداری و جلوگیری از کمبود کالا منجر می‌شود. توسعه‌ی مدل ارائه‌شده در این تحقیق، ابزاری استراتژیک برای مدیران خرده‌فروشی فراهم می‌آورد تا بتوانند در مواجهه با نوسانات شدید بازار آنلاین، با دقت بالاتری نسبت به چرخه‌های کوتاه عمر محصول و تغییرات رفتار خرید مشتریان، برنامه‌ریزی عملیاتی انجام دهند (Cai et al., 2021).

میزان تقاضا برای هر محصول یکی از متغیرهای کانونی در تحلیل‌های خرده‌فروشی و پیش‌بینی فروش به شمار می‌رود و معمولاً به‌صورت تعداد واحدهای فروخته‌شده یک کالا در یک بازه زمانی مشخص تعریف می‌شود. این متغیر که در پایگاه‌های داده خرده‌فروشی آنلاین از ستون مقدار[1] استخراج می‌گردد، بازتاب‌دهنده رفتار واقعی خرید مشتریان بوده و نقش محوری در تصمیم‌گیری‌های مرتبط با مدیریت موجودی، کنترل هزینه، برنامه‌ریزی تأمین و طراحی سیاست‌های عملیاتی ایفا می‌کند. مطالعات مبتنی بر داده‌های واقعی فروش نشان می‌دهند که تقاضای محصول دارای ماهیتی پویا، غیرخطی و وابسته به زمینه است و تحت تأثیر تغییرات زمانی، ناهمگنی مشتریان و شوک‌های ساختاری بازار قرار می‌گیرد (Chowdhury et al., 2025; Saarinen & Huttunen, 2025). پژوهش‌های اخیر در حوزه مهندسی خرده‌فروشی تأکید دارند که مدل‌سازی دقیق این متغیر مستلزم شناسایی وابستگی‌های بلندمدت و الگوهای پیچیده زمانی است؛ ازاین‌رو، مدل‌های مبتنی بر یادگیری عمیق و شبکه‌های توجه‌محور به‌عنوان ابزارهای کارآمد برای پیش‌بینی تقاضای محصول معرفی شده‌اند (Li, 2023; Rai, 2025). از منظر نظری،

1 quantity

برخی مطالعات تقاضای محصول را صرفاً نتیجه ویژگی‌های ذاتی کالا تلقی نمی‌کنند، بلکه آن را حاصل تعاملات ساختاری و شبکه‌ای میان محصولات پرفروش، مکمل و جانشین می‌دانند؛ به‌گونه‌ای که تغییر در تقاضای یک محصول می‌تواند با وقفه زمانی مشخص، الگوی تقاضای سایر اقلام را تحت تأثیر قرار دهد (Caetano et al., 2025). افزون بر این، رویکردهای چندوجهی و مکانی- زمانی نشان داده‌اند که ادغام اطلاعات مکمل، ازجمله داده‌های متنی، مکانی و رفتاری، می‌تواند درک عمیق‌تری از پویایی تقاضای محصول فراهم کرده و دقت پیش‌بینی را در محیط‌های خرده‌فروشی پیچیده به‌طور معناداری افزایش دهد (Cai et al., 2021; Rahikka & Mikkola, 2025). درمجموع، ادبیات پژوهش بر این نکته اجماع دارد که میزان تقاضا برای هر محصول متغیری چندبعدی و وابسته به ساختارهای زمانی و بین‌محصولی است که تحلیل مؤثر آن مستلزم به‌کارگیری چارچوب‌های مدل‌سازی پیشرفته، پویا و حساس به تعاملات متقابل میان اقلام می‌باشد.

سایر متغیرهای موجود در پایگاه داده خرده‌فروشی آنلاین، شامل شناسه فاکتور[1]، کد محصول[2]، شرح محصول[3]، تاریخ صدور فاکتور[4]، قیمت واحد[5]، شناسه مشتری[6] و کشور[7]، نقش مکمل و تعیین‌کننده‌ای در تبیین رفتار تقاضا و پویایی فروش ایفا می‌کنند. متغیرهای شناسه فاکتور و کدمحصول زیرساخت تحلیل‌های سری زمانی را فراهم می‌سازند و امکان شناسایی الگوهای فصلی، چرخه‌های تقاضا و وابستگی‌های زمانی پیچیده را مهیا می‌کنند؛ مطالعات اخیر نشان داده‌اند که مدل‌سازی دقیق ساختار زمانی تراکنش‌ها تأثیر معناداری بر بهبود دقت پیش‌بینی در محیط‌های خرده‌فروشی داده‌محور دارد. از سوی دیگر، کد محصول و شرح محصول بیانگر هویت و ویژگی‌های معنایی محصولات هستند و مبنای استخراج روابط بین‌محصولی، ازجمله مکملی و جانشینی، محسوب می‌شوند؛ پژوهش‌های مبتنی بر داده‌های چندوجهی نشان می‌دهند که ادغام اطلاعات متنی محصولات با داده‌های فروش می‌تواند وابستگی‌های پنهان میان اقلام را آشکار سازد و عملکرد مدل‌های پیش‌بینی را بهبود بخشد (Cai et al., 2021). متغیر قیمت نیز به‌عنوان یک عامل زمینه‌ای کلیدی، با رفتار خرید مشتریان و نوسانات تقاضا در تعامل است و در ادبیات مهندسی خرده‌فروشی به‌عنوان یکی از محرک‌های اصلی تغییرات کوتاه‌مدت تقاضا معرفی شده است (Li, 2023; Rai, 2025). افزون بر این، شناسه مشتری و کشور امکان تحلیل ناهمگنی مشتریان و تفاوت‌های مکانی- بازاری را فراهم می‌کنند و پژوهش‌های مبتنی بر تطبیق دامنه و

1 invoice
2 stock code
3 description
4 invoice date
5 price
6 customer id
7 country

داده‌های مکانی نشان داده‌اند که در نظر گرفتن این ابعاد می‌تواند به تعمیم‌پذیری بهتر مدل‌ها در بازارهای متنوع منجر شود (Saarinen & Huttunen, 2025; Tarighat et al., 2025). درمجموع، ادبیات موجود این متغیرها را به‌عنوان ابعاد زمانی، ساختاری، معنایی و زمینه‌ای داده‌های خرده‌فروشی معرفی می‌کند که بهره‌برداری یکپارچه از آن‌ها، پیش‌نیاز توسعه مدل‌های پیشرفته، مقیاس‌پذیر و تبیین‌پذیر پیش‌بینی تقاضا در خرده‌فروشی آنلاین است.

باوجود پیشرفت‌های قابل‌توجه در به‌کارگیری مدل‌های یادگیری عمیق در پیش‌بینی تقاضای خرده‌فروشی، به‌ویژه در قالب معماری‌های مبتنی بر توجه و شبکه‌های عمیق، ادبیات موجود همچنان با محدودیت‌های مفهومی و روش‌شناختی قابل‌توجهی مواجه است. بخش عمده‌ای از پژوهش‌ها تمرکز خود را بر استخراج الگوهای سری زمانی و بهبود دقت پیش‌بینی از طریق ترکیب یا تعمیق معماری‌های مدل قرار داده‌اند، درحالی‌که وابستگی‌های متقابل میان محصولات به‌عنوان یک پدیده ساختاری و شبکه‌ای کمتر به‌صورت نظام‌مند موردتوجه قرار گرفته است (Chowdhury et al., 2025). هرچند برخی مطالعات مبتنی بر داده‌های چندمنبعی و چندوجهی نشان داده‌اند که تلفیق انواع مختلف داده می‌تواند عملکرد مدل‌های پیش‌بینی را بهبود بخشد، این رویکردها اغلب فاقد مدل‌سازی صریح تعاملات پویا و انتقال تقاضا میان اقلام مرتبط هستند و بیشتر به تجمیع یا ادغام سطحی اطلاعات بسنده می‌کنند. (Cai et al., 2021). از سوی دیگر، پژوهش‌هایی که به مسئله پایداری و تعمیم‌پذیری مدل‌ها در شرایط ناهمگن بازار پرداخته‌اند، عمدتاً بر تطبیق دامنه یا تفاوت‌های توزیعی داده‌ها تمرکز داشته و نقش هم‌زمان تأخیرهای زمانی و ساختارهای بین‌محصولی را به‌صورت یکپارچه لحاظ نکرده‌اند (Tarighat et al., 2025). افزون بر این، بعد معناشناسی محصول که می‌تواند از اطلاعات متنی و توصیفی برای کشف روابط پنهان و تقویت تبیین‌پذیری مدل‌ها بهره گیرد در بسیاری از مطالعات پیش‌بینی تقاضا یا نادیده گرفته‌شده یا صرفاً به‌عنوان یک ویژگی کمکی ساده مورداستفاده قرار گرفته است، درحالی‌که ادبیات نوین داده‌محور بر نقش کلیدی آن در درک رفتار خرید تأکید دارد (Rahikka & Mikkola, 2025). درنتیجه، می‌توان شکاف اصلی ادبیات را در فقدان یک چارچوب یکپارچه و پویا دانست که به‌طور هم‌زمان وابستگی‌های زمانی، ساختار شبکه‌ای محصولات و اطلاعات معناشناختی را در یک مدل پیش‌بینی تبیین‌پذیر و مقیاس‌پذیر ادغام کند؛ خلأیی که ضرورت توسعه رویکردهای نوین در پیش‌بینی تقاضای خرده‌فروشی آنلاین را به‌وضوح برجسته می‌سازد.

با توجه به محدودیت‌های شناسایی‌شده در ادبیات پیش‌بینی تقاضای خرده‌فروشی، هدف اصلی این پژوهش توسعه یک چارچوب پیش‌بینی پویای تقاضا مبتنی بر معماری ترنسفورمر است که بتواند به‌طور هم‌زمان وابستگی‌های زمانی، اثرات شبکه‌ای میان محصولات و اطلاعات معناشناختی اقلام را در یک

ساختار یکپارچه مدل‌سازی کند. این مطالعه با بهره‌گیری از داده‌های تراکنشی خرده‌فروشی آنلاین، می‌کوشد نشان دهد که چگونه تعاملات غیرخطی میان محصولات پرفروش و کالاهای مرتبط، همراه با تأخیرهای زمانی تقاضا، می‌تواند الگوهای فروش آینده را به‌دقت بیشتری پیش‌بینی کند. بر این اساس، پژوهش حاضر در پی پاسخ به این پرسش‌های اساسی است: (۱) آیا ادغام اثرات شبکه‌ای محصولات در یک مدل مبتنی بر ترنسفورمر منجر به بهبود معنادار دقت پیش‌بینی تقاضای خرده‌فروشی می‌شود؟ (۲) نقش تأخیرهای زمانی در انتقال شوک‌های تقاضا میان محصولات مختلف چگونه است و تا چه حد می‌توان آن‌ها را به‌صورت پویا مدل‌سازی کرد؟ و (۳) تا چه اندازه استفاده از اطلاعات معناشناختی استخراج‌شده از توصیفات متنی محصولات می‌تواند به تبیین‌پذیری و کارایی مدل پیش‌بینی کمک کند؟ پاسخ به این پرسش‌ها نه‌تنها به غنای ادبیات پیش‌بینی تقاضای خرده‌فروشی می‌افزاید، بلکه می‌تواند راهکارهای عملی ارزشمندی برای بهبود تصمیم‌گیری در مدیریت موجودی و برنامه‌ریزی فروش در محیط‌های خرده‌فروشی آنلاین ارائه دهد.

پیشینه پژوهش

در سطح داخلی، پژوهش‌های متعددی به کاربرد روش‌های نوین داده‌محور در پیش‌بینی تقاضا و قیمت محصولات پرداخته‌اند، هرچند اغلب آن‌ها بر حوزه‌های خاصی متمرکز بوده‌اند. برای نمونه، حسینی و مازندرانی‌زاده (۱۴۰۴) با به‌کارگیری شبکه عصبی مصنوعی و تابع تقاضای معکوس، قیمت محصولات زراعی را در شبکه آبیاری دشت قزوین پیش‌بینی کرده و نشان دادند که مدل‌های ANN می‌توانند نسبت به روش‌های کلاسیک، خطای پیش‌بینی را به‌طور معناداری کاهش دهند. در حوزه خرده‌فروشی، شاملو و فتاح (۱۴۰۲) تأثیر به‌کارگیری فناوری‌های هوش مصنوعی را بر توسعه‌ی سیستم‌های پیش‌بینی تقاضای کسب‌وکارهای خرده‌فروشی بررسی کرده و بر نقش این سیستم‌های هوشمند در بهبود تصمیم‌گیری‌های عملیاتی و راهبردی تأکید کرده‌اند. در زنجیره تأمین صنعتی، ویسی (۱۴۰۵) با تمرکز بر بهینه‌سازی یکپارچه موجودی قطعات یدکی مبتنی بر پیش‌بینی تقاضا با استفاده از روش‌های یادگیری ماشین نشان داده است که ترکیب مدل‌های پیش‌بینی داده‌محور با سیاست‌های موجودی می‌تواند به کاهش هزینه‌ها و افزایش سطح خدمت منجر شود. همچنین، مسیبی اتاق‌سرا و همکاران (۱۴۰۳) با تمرکز بر زنجیره تأمین محصولات لبنی فسادپذیر، عملکرد مدل‌های مختلف یادگیری ماشین ازجمله CNN و LSTM را در پیش‌بینی تقاضای روزانه مقایسه کرده و گزارش کرده‌اند که استفاده از شبکه‌های عصبی عمیق می‌تواند دقت پیش‌بینی و سودآوری تصمیمات موجودی را بهبود دهد. باوجود این تلاش‌های ارزشمند، هیچ‌یک از این مطالعات به‌طور هم‌زمان به پویایی تقاضای خرده‌فروشی آنلاین و نقش اثرات شبکه‌ای، تأخیرهای

زمانی و معناشناسی محصول در قالب یک چارچوب یکپارچه مبتنی بر ترنسفورمر نپرداخته‌اند؛ خلأیی که پژوهش حاضر درصدد پوشش آن است.

در ادبیات بین‌المللی، رویکردهای مبتنی بر یادگیری عمیق، به‌ویژه معماری‌های ترنسفورمر، به‌طور فزاینده‌ای برای پیش‌بینی تقاضای خرده‌فروشی به کار گرفته شده‌اند. به‌عنوان نمونه، Wang et.al. (2026) یک مدل ترنسفورمر پوشیده مبتنی بر ویژگی‌های ترکیبی برای پیش‌بینی تقاضای کوتاه‌مدت محصولات خرده‌فروشی در زنجیره تأمین مبتنی بر اینترنت اشیا ارائه کرده‌اند. این مدل با ادغام نقاط قوت ترنسفورمر، شبکه‌های کانولوشنی زمانی و آموزش پوشیده، قادر است هم‌زمان الگوهای بلندمدت و قله‌های کوتاه‌مدت فروش را استخراج کرده و درعین‌حال، نویز و ناهنجاری‌های موجود در داده‌های تاریخی فروش را بهتر مدیریت کند؛ به‌طوری‌که نسبت به سایر روش‌های یادگیری عمیق، دقت پیش‌بینی بالاتر و میانگین قدرمطلق خطای پایین‌تری را گزارش می‌کند. بااین‌وجود، حتی در این مطالعه نیز اثرات شبکه‌ای ناشی از تعاملات میان محصولات، تأخیرهای زمانی در انتقال این اثرات در زنجیره تأمین و معنای ضمنیِ هم‌رخدادی اقلام در سبد خرید به‌طور توأمان مدل‌سازی نشده است. ازاین‌رو، پژوهش حاضر با تمرکز بر طراحی یک چارچوب ترنسفورمرمحور که این سه بعد را به‌صورت یکپارچه در پیش‌بینی پویای تقاضای خرده‌فروشی لحاظ می‌کند، در امتداد و درعین‌حال فراتر از مطالعات پیشین حرکت می‌کند.

پیش‌بینی تقاضای خرده‌فروشی به‌عنوان یکی از مسائل محوری در مدیریت عملیات و زنجیره تأمین، در ابتدا عمدتاً بر مدل‌های آماری کلاسیک و سری‌های زمانی خطی متمرکز بوده است. با افزایش پیچیدگی بازارهای خرده‌فروشی آنلاین، تنوع محصولات و رفتارهای غیرخطی مشتریان، محدودیت این رویکردها آشکار شد و پژوهش‌ها به‌تدریج به سمت مدل‌های داده‌محور و یادگیری ماشین حرکت کردند (Tiwari, 2025). مطالعات مروری و تجربی نشان می‌دهند که مدل‌های عمیق، به‌ویژه شبکه‌های عصبی بازگشتی و کانولوشنی، توان بالاتری در استخراج الگوهای زمانی پیچیده نسبت به واروش‌های سنتی دارند (Cai et al., 2021). در همین راستا، پژوهش‌های حوزه مهندسی داده‌محور خرده‌فروشی تأکید می‌کنند که پیش‌بینی تقاضا دیگر صرفاً یک مسئله زمانی نیست، بلکه پدیده‌ای چندبعدی و وابسته به زمینه‌های رفتاری، مکانی و ساختاری بازار است (Rai, 2025).

با ظهور معماری‌های پیشرفته‌تر، تمرکز پژوهش‌ها به سمت مدل‌هایی معطوف شد که بتوانند وابستگی‌های کوتاه‌مدت و بلندمدت را به‌صورت هم‌زمان ثبت کنند. در این زمینه، مدل‌های هیبریدی نظیر ترکیب شبکه‌های کانولوشنی زمانی[1] و ترنسفورمر به‌عنوان راهکاری مؤثر برای مقیاس‌پذیری و افزایش دقت پیش‌بینی معرفی شده‌اند. نتایج نشان می‌دهد که این مدل‌ها در مقایسه با حافظه بلندمدت–

1 Temporal convolutional network (TCN)

کوتاه‌مدت[1] و واحد بازگشتی دروازه‌دار[2] عملکرد برتری در داده‌های فروش بزرگ‌مقیاس دارند (M. A. et al., 2025). بااین‌حال، اغلب این مطالعات تمرکز خود را بر بهبود معماری زمانی قرار داده و تقاضا را به‌عنوان یک سری زمانی مستقل برای هر محصول در نظر گرفته‌اند، درحالی‌که تعاملات میان محصولات و اثرات انتقالی تقاضا کمتر موردتوجه قرار گرفته است (Saarinen & Huttunen, 2025).

در سال‌های اخیر، ادبیات پیش‌بینی تقاضا شاهد گسترش استفاده از داده‌های چندوجهی بوده است؛ داده‌هایی که فراتر از فروش تاریخی، شامل متن، تصویر و شاخص‌های محیطی می‌شوند. پژوهش (Cai et al., 2021) نشان داد که ادغام هم‌زمان چند منبع داده می‌تواند به درک دقیق‌تری از الگوهای تقاضا منجر شود. در ادامه، مقاله (Wang, 2025) با معرفی مدل ترنسفورمر چندوجهی آگاه به علیت، نشان دادند که ترکیب متن، سری زمانی و تصاویر ماهواره‌ای نه‌تنها دقت پیش‌بینی را افزایش می‌دهد، بلکه امکان تحلیل علّی و کاهش همبستگی‌های کاذب را نیز فراهم می‌سازد. باوجود این پیشرفت‌ها، بیشتر مطالعات چندوجهی در سطح کلان زنجیره تأمین یا تقاضای تجمیع‌شده عمل کرده‌اند و کاربرد آن‌ها در سطح محصول و خرده‌فروشی آنلاین همچنان محدود است (Cai et al., 2021).

یکی دیگر از محورهای مهم ادبیات، توجه به ناهمگنی بازار و وابستگی‌های ساختاری در داده‌های خرده‌فروشی است. پژوهش‌های مرتبط با سازگاری دامنه نشان می‌دهند که مدل‌های پیش‌بینی تقاضا در مواجهه با تغییرات بازار، دسته محصول یا منطقه جغرافیایی دچار افت عملکرد می‌شوند (Tarighat et al., 2025). درعین‌حال، برخی مطالعات نظری و تجربی به وجود اثرات شبکه‌ای میان محصولات اشاره می‌کنند؛ به‌طوری‌که تغییر تقاضای یک محصول می‌تواند با وقفه زمانی بر تقاضای سایر اقلام مکمل یا جانشین اثر بگذارد (Tripathi et al., 2025). بااین‌وجود، این بینش‌ها عمدتاً به‌صورت توصیفی باقی‌مانده و به‌ندرت در قالب مدل‌های پیش‌بینی پویا و عمیق پیاده‌سازی شده‌اند.

بررسی نظام‌مند مطالعات داخلی و خارجی نشان می‌دهد که اگرچه پیشرفت‌های قابل‌توجهی در مدل‌سازی زمانی، معماری‌های ترنسفورمر و استفاده از داده‌های چندوجهی در پیش‌بینی تقاضا حاصل شده است، اما یک شکاف اساسی همچنان باقی مانده است: فقدان چارچوبی یکپارچه که بتواند به‌طور هم‌زمان وابستگی‌های زمانی با تأخیر، اثرات شبکه‌ای میان محصولات و معناشناسی محصول استخراج‌شده از داده‌های متنی را در سطح خرده‌فروشی آنلاین مدل‌سازی کند. نوآوری مقاله حاضر دقیقاً در پاسخ به این خلأ شکل می‌گیرد؛ بدین معنا که با توسعه یک رویکرد مبتنی بر ترنسفورمر، تعاملات پویا میان محصولات مرتبط را در قالب ساختار شبکه‌ای، همراه با اطلاعات زمانی و معنایی، ادغام می‌کند. این

1 Long short-term memory (LSTM)
2 Gated recurrent unit (GRU)

رویکرد نه‌تنها به بهبود دقت پیش‌بینی تقاضا می‌انجامد، بلکه گامی مهم در جهت تبیین‌پذیری و واقع‌گرایی مدل‌های پیش‌بینی در محیط‌های پیچیده خرده‌فروشی آنلاین به شمار می‌رود.

ساختار این مقاله به‌گونه‌ای تنظیم شده است که به‌صورت گام‌به‌گام به تبیین مسئله و ارائه راهکار پژوهش بپردازد. در بخش نخست، مسئله پیش‌بینی پویای تقاضای خرده‌فروشی معرفی‌شده و با مرور مختصر مطالعات پیشین، اهمیت و ضرورت انجام این پژوهش تبیین می‌شود. بخش دوم به‌مرور نظام‌مند ادبیات پژوهش در حوزه پیش‌بینی تقاضا، مدل‌های یادگیری عمیق و به‌ویژه معماری ترنسفورمر اختصاص دارد و در این چارچوب، نقش اثرات شبکه‌ای میان محصولات، وابستگی‌های زمانی با تأخیر و اطلاعات معنایی محصولات بررسی‌شده و شکاف پژوهشی موجود مشخص می‌شود. در بخش سوم، داده‌های مورداستفاده، نحوه پیش‌پردازش و تعریف مسئله پیش‌بینی تقاضا تشریح می‌شود. سپس، چارچوب مدل پیشنهادی مبتنی بر معماری ترنسفورمر معرفی‌شده و نحوه به‌کارگیری سازوکار توجه، تعبیه ویژگی‌ها و ادغام اطلاعات زمانی، ساختاری و معنایی در فرآیند پیش‌بینی توضیح داده می‌شود. بخش چهارم به ارائه تنظیمات آزمایشی، معیارهای ارزیابی و نتایج تجربی اختصاص دارد و عملکرد مدل پیشنهادی در مقایسه با روش‌های مبنا مورد تحلیل قرار می‌گیرد. درنهایت، یافته‌های اصلی پژوهش، پیامدهای کاربردی و مسیرهای پیشنهادی برای پژوهش‌های آینده بیان می‌شود.

روش

این پژوهش از نوع مطالعه تجربی- محاسباتی در حوزه مهندسی و علم داده است که باهدف پیش‌بینی پویای تقاضای خرده‌فروشی انجام شده است. محیط پژوهش، یک بستر آنلاین و داده‌محور بوده و تحلیل‌ها بر اساس داده‌های واقعی فروش اجرا شده‌اند. داده‌ها از پایگاه داده خرده‌فروشی آنلاین دو گردآوری‌شده و پس از پیش‌پردازش، به‌صورت سری‌های زمانی چندمحصولی مورداستفاده قرار گرفته‌اند. رویکرد تحلیل داده مبتنی بر یادگیری عمیق بوده و از آموزش، اعتبارسنجی و آزمون مدل‌ها برای ارزیابی عملکرد استفاده شده است. طراحی پژوهش و توسعه مدل بر چارچوب‌های نظری ترنسفورمر، مدل‌سازی وابستگی‌های زمانی با تأخیر و مفاهیم اثرات شبکه‌ای میان محصولات هدایت شده است.

جامعه آماری این پژوهش شامل کلیه تراکنش‌های فروش محصولات در پایگاه داده خرده‌فروشی آنلاین است. سطح تحلیل در این پژوهش، محصول-روز در نظر گرفته شده است؛ بدین معنا که رفتار تقاضای هر کالا به‌صورت روزانه موردبررسی و پیش‌بینی قرار می‌گیرد. نمونه آماری شامل تمامی تراکنش‌های ثبت‌شده از محصولات منتخب است که با روش نمونه‌گیری تمام‌شماری از داده‌های معتبر انتخاب شده‌اند. در این فرآیند، رکوردهای ناقص، بازگشتی یا دارای مقادیر نامعتبر از تحلیل حذف

گردیده و سایر اطلاعات شامل کد محصول، توضیحات متنی، تاریخ فاکتور، مقدار فروش و قیمت استخراج شد. داده‌ها پس از پاک‌سازی و نرمال‌سازی، به مجموعه‌های آموزش و آزمون تقسیم شده‌اند. این نمونه‌ها نمایانگر رفتار واقعی تقاضای مشتریان در محیط خرده‌فروشی آنلاین بوده و با پوشش تنوع محصولی (کالاهای مکمل و جانشین)، مبنای ارزیابی مدل پیشنهادی قرار گرفته‌اند.

در این پژوهش، سه مؤلفه کلیدی اثرات شبکه‌ای، تأخیرهای زمانی و معناشناسی محصول، از طریق پارامترهای مدل‌سازی استخراج شده‌اند:

- اثرات شبکه‌ای: این اثر از طریق ماتریس مجاورت تعاملات بین‌محصولی مدل‌سازی شده است که در آن، وزن‌دهی بین گره‌ها (محصولات) بر اساس میزان ارتباط در سبد خرید مشتریان تعیین می‌گردد. این ماتریس به‌صورت یک لایه گراف‌کانوولوشن در ساختار ترنسفورمر ادغام‌شده تا اثرات غیرمستقیم نفوذ یک محصول بر دیگری را در فضای نهفته محاسبه کند.
- تأخیرهای زمانی: برای لحاظ کردن تأخیر زمانی، پنجره لغزان مدل بر روی توالی‌های زمانی تعریف شده است. پارامتر تأخیر (L) به‌گونه‌ای تنظیم‌شده که اثرات تقاضای محصولات مرجع در بازه‌های $t-k$ تا $t-1$ را بر تقاضای جاری (t) به‌صورت پویا از طریق وزن‌های توجه شناسایی کند.
- معناشناسی محصول: برای استخراج معناشناسی، از بردارهای ویژگی استفاده شده است که ویژگی‌های متنی مانند نام، دسته‌بندی و مشخصات فنی را به فضای برداری با ابعاد بالا نگاشت می‌کند. در این لایه، پارامترهای جایگذاری به‌طور هم‌زمان با سایر پارامترهای مدل، طی فرآیند آموزش و باهدف بیشینه‌سازی دقت پیش‌بینی به‌روزرسانی می‌شوند تا شباهت معنایی بین محصولات فارغ از تاریخچه فروش در مدل لحاظ گردد.

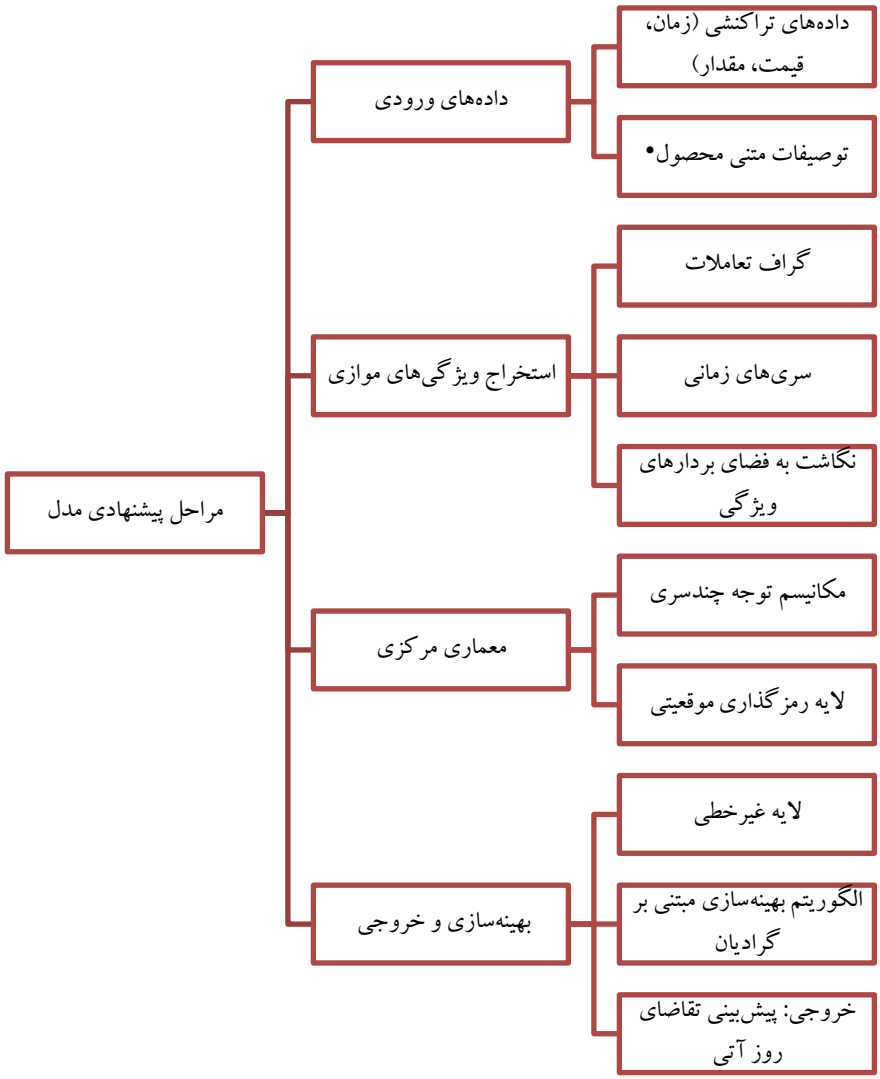
در این پژوهش، ابزار اصلی شامل زبان برنامه‌نویسی Python و کتابخانه‌های تخصصی یادگیری ماشین و یادگیری عمیق نظیر PyTorch، NumPy و Pandas برای پردازش داده و پیاده‌سازی مدل‌ها بوده است. برای توسعه و آموزش معماری پیشنهادی، از چارچوب ترنسفورمر و ماژول‌های توجه چندسری استفاده شده است. پیش‌پردازش متون توصیف محصول با بهره‌گیری از ابزارهای پردازش زبان طبیعی و جاسازی‌های معنایی انجام شده است. ارزیابی عملکرد مدل‌ها با استفاده از معیارهای استاندارد پیش‌بینی تقاضا میانگین قدرمطلق خطا و میانگین مربعات خطا صورت گرفته که اعتبار و قابلیت مقایسه نتایج را تضمین می‌کند. کلیه آزمایش‌ها در یک محیط محاسباتی کنترل‌شده اجرا شده‌اند تا قابلیت تکرارپذیری پژوهش حفظ شود.

فرآیند گردآوری داده‌ها مبتنی بر استخراج تراکنش‌های فروش از پایگاه داده خرده‌فروشی آنلاین دو و سازمان‌دهی آن‌ها در قالب سری‌های زمانی محصول‌محور بوده است. در مرحله پیش‌پردازش، داده‌های ناقص یا نامعتبر حذف‌شده و اثرات زمانی با اعمال وقفه‌های زمانی و ساخت ویژگی‌های شبکه‌ای میان محصولات استخراج شده‌اند. سپس داده‌ها به مجموعه‌های آموزش، اعتبارسنجی و آزمون تقسیم شدند. تحلیل داده‌ها با آموزش مدل ترنسفورمر پیشنهادی و مقایسه آن با مدل‌های مبنا انجام گرفته است.

به‌منظور ارزیابی پایداری و قدرت تعمیم‌دهی مدل، از استراتژی تقسیم داده‌ها استفاده شد. در این راستا، ۸۰ درصد از داده‌های پیش‌پردازش‌شده برای آموزش مدل و ۲۰ درصد باقی‌مانده (داده‌های مشاهده‌نشده) به‌منظور ارزیابی عملکرد و بررسی قدرت تعمیم‌پذیری مدل در مواجهه با داده‌های جدید اختصاص یافت. این رویکرد تضمین می‌کند که نتایج گزارش‌شده، صرفاً بازتاب‌دهنده توانایی مدل در یادگیری الگوهای ساختاری و معنایی و نه حفظ کردن داده‌های آموزشی باشد. معیارهای اصلی ارزیابی شامل میانگین قدرمطلق خطا و ریشه میانگین مربعات خطا تعیین شدند. همچنین برای اطمینان از اعتبار نتایج، عملکرد مدل پیشنهادی در این مجموعه آزمون با مدل‌های مرجع شامل مدل خودهمبسته میانگین متحرک انباشته و الگوریتم تقویت گرادیان مقایسه گردید تا برتری نسبی و ظرفیت تعمیم‌پذیری مدل در مقایسه با رویکردهای کلاسیک و یادگیری ماشین متعارف سنجیده شود.

در این پژوهش، از معماری ترنسفورمر به دلیل توانمندی منحصربه‌فرد آن در مدل‌سازی وابستگی‌های بلندمدت استفاده شده است. هسته اصلی این مدل، مکانیسم توجه چندسری است که به شبکه اجازه می‌دهد به‌طور هم‌زمان از جنبه‌های مختلف به بردارهای ویژگی ورودی توجه کند و وزن‌دهی به روابط زمانی دور و نزدیک را بهینه سازد؛ به‌عبارت‌دیگر، هر سر، وظیفه یادگیری یک نوع خاص از همبستگی در داده‌های سری زمانی را بر عهده دارد. ازآنجاکه معماری ترنسفورمر به‌خودی‌خود فاقد درک از ترتیب زمانی است، از رمزگذاری موقعیتی برای تزریق اطلاعات زمانی به بردارهای ورودی استفاده شده است. این مؤلفه تضمین می‌کند که ترتیب قرارگیری داده‌های تقاضا در زمان (مانند توالی روزها) حفظ شود. علاوه بر این، به‌منظور تسهیل فرآیند آموزش و جلوگیری از محوشدگی گرادیان در لایه‌های عمیق‌تر، از اتصالات باقی‌مانده و لایه‌های نرمال‌سازی در تمامی زیرلایه‌های مدل بهره گرفته شده است. این ساختار موجب می‌شود که مدل نه‌تنها روابط خطی، بلکه الگوهای غیرخطی و پیچیده حاکم بر تعاملات شبکه‌ای محصولات را با دقت بالاتری استخراج نماید.

شکل ۱. مراحل مدل پیشنهادی



به‌منظور تبیین شفاف فرآیند پژوهش و ساختار مدل پیشنهادی، گام‌های اجرایی در قالب یک چارچوب چهارمرحله‌ای طراحی شده است :۱. آماده‌سازی داده‌ها شامل جمع‌آوری و پاک‌سازی داده‌های تراکنشی؛ ۲. استخراج ویژگی‌های پویای چندبعدی که در آن ویژگی‌های شبکه‌ای از طریق لایه گراف، ویژگی‌های تأخیری از طریق پنجره لغزان و ویژگی‌های معنایی از طریق بازنمایی برداری متون به بردارهای عددی تبدیل می‌شوند؛ ۳. معماری مدل ترنسفورمر که با به‌کارگیری مکانیسم توجه، تعاملات پیچیده میان این ویژگی‌ها را یاد می‌گیرد؛ و ۴. فرآیند بهینه‌سازی که در آن مدل با استفاده از الگوریتم‌های مبتنی بر گرادیان آموزش‌دیده و روابط غیرخطی حاکم بر تقاضا را استخراج می‌نماید.

در ادامه مدل ریاضی روش پیشنهادی ارائه شده است.

مجموعه‌ها و شاخص‌های مورداستفاده در مدل به‌صورت زیر معرفی شده‌اند:

$i \in P$: شاخص محصول

$t \in T$: شاخص زمان

$n \in N$: شاخص نمونه/ مشاهده

پارامترها

V: تعداد محصولات منحصربه‌فرد

d: ابعاد تعبیه

H: تعداد سرهای توجه

L: تعداد لایه‌های رمزگذار ترنسفورمر

F: ابعاد پنهان شبکه پیش‌خور

P: نرخ دراپ‌اوت (نرخ حذف تصادفی)

$x_n \in \mathbb{R}^{T \times d}$: بردار ویژگی ورودی برای نمونه n

y_n: تقاضای واقعی (تعداد فروش) برای نمونه n

متغیرهای تصمیم

$\hat{y}_n \geq 0$: تقاضای پیش‌بینی‌شده برای نمونه n

$e_i \in \mathbb{R}^d$: بردار تعبیه محصول i

$W_Q\ W_K, W_k$: ماتریس‌های وزن سازوکار توجه W_O

W_O: ماتریس وزن خروجی توجه چندسری

W_1, W_2: وزن‌های شبکه پیش‌خور

b_1, b_2: بایاس‌های شبکه پیش‌خور

w_{out}, b_{out}: پارامترهای لایه خطی نهایی

$\Theta = \{e_i, W_Q, W_K, W_V, W_O, W_1, W_2, b_1, b_2, w_{out}, b_{out}\}$: مجموعه کلیه پارامترهای قابل‌آموزش Θ

تعبیه محصول: $e_i = Embedding(i)$، $e_i \in \mathbb{R}^d$

نمایش دنباله ورودی: $X_n = [x_{n1}, x_{n2}, \dots, x_{nT}]$، $x_{nT} \in \mathbb{R}^d$

سازوکار خودتوجهی به‌صورت زیر تعریف شده است:

$$Q = X \times W_Q \qquad (۱)$$

$$K = X \times W_K \qquad (۲)$$

$$V = X \times W_V \quad (۳)$$

$$Attention(Q,K,V) = softmax(\frac{QK^{\mathrm{T}}}{\sqrt{d_{\square}}})V \quad (۴)$$

توجه چندسری با روابط زیر نشان داده می‌شود:

$$head_h = Attention(XW_h^Q, XW_h^K, XW_h^V) \quad (۵)$$

$$MultiHead(X) = Concat(head_1, head_2, \dots, head_H)W^O \quad (۶)$$

لایه رمزگذار ترنسفورمر توسط روابط زیر نشان داده شده است:

$$Z^{(l)} = LayerNorm(X^{(l-1)} + MultiHead(X^{(l-1)})) \quad (۶)$$

$$X^{(l)} = LayerNorm\left(Z^{(l)} + FFN(Z^{(l)})\right), for\ l = 1, \dots, L \quad (۷)$$

شبکه پیش‌خور با رابطه (۸) نشان داده شده است:

$$FFN(z) = \sigma(zW_1 + b_1)W_2 + b_2 \quad (۸)$$

پیش‌بینی تقاضا با استفاده از رابطه (۹) انجام شده است:

$$\hat{y}_{\square} = w_{out}^T h_{\square}^T + b_{out} \quad (۹)$$

که در آن، $h_{\square}^T$ بردار نمایش خروجی ترنسفورمر در گام زمانی T است.

تابع هدف به‌صورت حداقل کردن میانگین خطای مطلق و میانگین مربعات خطا درنظرگرفته شده است:

میانگین خطای مطلق به‌صورت رابطه (۱۰) و میانگین مربعات خطا به‌صورت رابطه (۱۱) نشان داده شده است:

$$min\ L_{MAE} = min\left(\frac{1}{N}\right) \Sigma\ |\hat{y}_{\square} - y_{\square}|,\ n = 1 \dots N \quad (۱۰)$$

$$min\ L_{MSE} = \min\left(\frac{1}{N}\right) \Sigma\ (\hat{y}_{\square} - y_{\square})^2,\ n = 1 \dots N \quad (۱۱)$$

در این پژوهش، به‌منظور پیش‌بینی تقاضای محصولات در داده‌های خرده‌فروشی، یک مدل مبتنی بر معماری ترنسفورمر طراحی شده است که قادر است وابستگی‌های زمانی و ویژگی‌های مرتبط با محصولات را به‌طور هم‌زمان مدل‌سازی کند. همان‌گونه که در معادلات بخش قبل ارائه شد، داده‌های ورودی به‌صورت دنباله‌ای از مشاهدات تاریخی تعریف می‌شوند که هر مشاهده شامل مجموعه‌ای از ویژگی‌ها نظیر مقدار فروش، قیمت و ویژگی‌های زمانی است. برای هر نمونه i، بردار ورودی به‌صورت یک توالی زمانی $X_i = [x_{i-L+1}, \dots, x_i]$ در نظر گرفته می‌شود که شامل L گام زمانی گذشته است.

هدف مدل، پیش‌بینی بردار تقاضا برای H دوره زمانی آینده است که با نماد $\hat{y}_t$ نمایش داده می‌شود (رابطه ۹).

در مرحله نخست، به‌منظور استخراج نمایش مناسب از محصولات و کشورها، از بردارهای تعبیه استفاده می‌شود. در این فرآیند، هر محصول و هر کشور به یک بردار در فضای ویژگی‌ها نگاشت می‌شود تا مدل بتواند روابط پنهان میان آن‌ها را یاد بگیرد. علاوه بر این، برای در نظر گرفتن شباهت میان محصولات، از یک ساختار گرافی استفاده می‌شود که اطلاعات ارتباطی میان محصولات را در قالب یک ماتریس مجاورت مدل‌سازی می‌کند. ترکیب بردار تعبیه محصول با اطلاعات گرافی باعث می‌شود نمایش نهایی هر محصول علاوه بر ویژگی‌های ذاتی خود، اطلاعات محصولات مرتبط را نیز در بر گیرد و درنتیجه قدرت مدل در استخراج الگوهای تقاضا افزایش یابد.

پس از ایجاد بردارهای تعبیه، داده‌های ورودی از طریق یک لایه نگاشت خطی به فضای برداری با بعد بالاتر منتقل می‌شوند. سپس، کدگذاری مکانی به بردارهای ورودی افزوده می‌شود تا ترتیب زمانی داده‌ها حفظ گردد؛ زیرا معماری ترنسفورمر به‌صورت ذاتی درکی از ترتیب زمانی ندارد و این اطلاعات باید صریحاً به مدل تزریق شود. بدین ترتیب، ترکیب اطلاعات مکانی، بردارهای تعبیه محصول و کشور و ویژگی‌های اصلی داده‌ها، یک نمایش جامع از دنباله زمانی ایجاد می‌کند که به‌عنوان ورودی به لایه‌های ترنسفورمر ارسال می‌شود.

در مرحله بعد، نمایش حاصل وارد بخش رمزگذار ترنسفورمر می‌شود. در این بخش، از سازوکار توجه خودی استفاده می‌شود که با روابط مربوط به تعریف بردارهای پرسش، کلید و مقدار و محاسبه وزن‌های توجه در معادلات مربوط به خودتوجهی مشخص شده است. این روابط در بخش معادلات (۱) تا (۴) آمده‌اند. همچنین، برای افزایش توان مدل در استخراج الگوهای پیچیده، از توجه چندسری استفاده شده است که در آن چندین سازوکار توجه به‌صورت هم‌زمان روی داده‌ها اعمال می‌شوند؛ روابط مربوط به ترکیب سرهای توجه و نگاشت نهایی در بخش معادلات (۵) و (۶) ارائه شده است. مجموعه این عملیات، در قالب روابط ریاضی لایه رمزگذار ترنسفورمر که در معادلات (۷) و (۸)، فرمول‌بندی شده است.

خروجی این بخش، یک نمایش فشرده از اطلاعات تاریخی است که وابستگی‌های زمانی کوتاه‌مدت و بلندمدت را در خود نگه می‌دارد. این نمایش سپس به یک شبکه پیش‌خور منتقل می‌شود که ساختار آن در رابطه (۸) نشان داده شده است. در این شبکه، ابتدا یک نگاشت غیرخطی روی هر موقعیت زمانی به‌صورت مستقل اعمال می‌شود و سپس از طریق لایه‌های کاملاً متصل، خروجی نهایی این بلوک به‌دست می‌آید.

درنهایت، نمایش استخراج‌شده از ترنسفورمر به لایه‌های کاملاً متصل خروجی منتقل می‌شود تا مقدار پیش‌بینی تقاضا برای دوره‌های آینده محاسبه گردد. فرمول‌بندی کلی فرآیند پیش‌بینی تقاضا برای افق H دوره زمانی آینده در رابطه (۹) بیان شده است.

برای ارزیابی عملکرد مدل، از توابع خطا نظیر میانگین قدرمطلق خطا (MAE) و میانگین مربعات خطا (MSE) استفاده می‌شود که به‌ترتیب در روابط (۱۰) و (۱۱) تعریف شده‌اند. این توابع خطا اختلاف بین مقادیر واقعی و مقادیر پیش‌بینی‌شده را کمی‌سازی می‌کنند و مبنای فرآیند آموزش مدل قرار می‌گیرند.

آموزش مدل به‌صورت یک مسئله بهینه‌سازی مبتنی بر گرادیان تعریف می‌شود که در آن، پارامترهای قابل یادگیری شامل وزن‌های لایه‌های تعبیه، پارامترهای سازوکار توجه، وزن‌های شبکه پیش‌خور (مطابق رابطه (۸)) و پارامترهای لایه خروجی به‌گونه‌ای به‌روزرسانی می‌شوند که تابع خطا (تعریف‌شده در روابط (۱۰) و (۱۱)) کمینه گردد. با تکرار این فرآیند، خطای پیش‌بینی کاهش‌یافته و دقت مدل در برآورد تقاضای آینده افزایش می‌یابد. ازآنجاکه روابط میان ویژگی‌های ورودی و تقاضای آینده، به‌ویژه به دلیل اثرات شبکه‌ای و تأخیرهای زمانی، ماهیتی غیرخطی دارند، معماری پیشنهادی با بهره‌گیری از توابع فعال‌ساز غیرخطی در لایه‌های پنهان و سازوکار توجه، این نگاشت‌های غیرخطی را مدل‌سازی می‌کند. فرآیند بهینه‌سازی از طریق محاسبه گرادیان تابع خطا نسبت به پارامترهای مدل و به‌کارگیری الگوریتم‌های مبتنی بر گرادیان (نظیر به‌روزرسانی تکراری وزن‌ها با استفاده از پس‌انتشار خطا) انجام می‌شود. بدین ترتیب، روابط غیرخطی پیچیده میان متغیرها در قالب یک مسئله بهینه‌سازی مبتنی بر گرادیان صورت‌بندی شده و مدل با همگرایی تدریجی در فضای پارامترها قادر به استخراج الگوهای پنهان و بهبود دقت پیش‌بینی تقاضا می‌گردد.

استفاده از رویکرد تلفیقی در این پژوهش بر این مبنا استوار است که تقاضا در خرده‌فروشی آنلاین، پدیده‌ای تک‌بعدی نیست، بلکه برآیندی از تعاملات پیچیده میان الگوهای زمانی، پیوندهای ساختاری بین محصولات و ویژگی‌های محتوایی کالاهاست. از منظر نظری، مدل‌های کلاسیک که صرفاً بر سری‌های زمانی تمرکز دارند، از درک روابط هم‌افزا میان کالاها (نظیر کالاهای مکمل و جانشین) و تأثیر ویژگی‌های کیفی بر رفتار خرید ناتوان‌اند. رویکرد تلفیقی پیشنهادی با بهره‌گیری از توانمندی معماری ترنسفورمر در استخراج وابستگی‌های زمانی طولانی‌مدت و ادغام آن با لایه‌های گراف برای تحلیل شبکه روابط محصولات، بستری را فراهم می‌کند که در آن بردارهای ویژگی نه به‌صورت مجزا، بلکه در یک فضای معنایی مشترک پردازش می‌شوند. این سازوکار تلفیقی علاوه بر افزایش دقت پیش‌بینی، پایداری مدل را در مواجهه با داده‌های نویزی و نوسانات ناگهانی بازار تقویت کرده و با کاهش نرخ خطا در

محیط‌های با ابعاد بالا، توجیه علمی و عملیاتی لازم را برای جایگزینی با روش‌های تک‌مؤلفه‌ای ارائه می‌دهد.

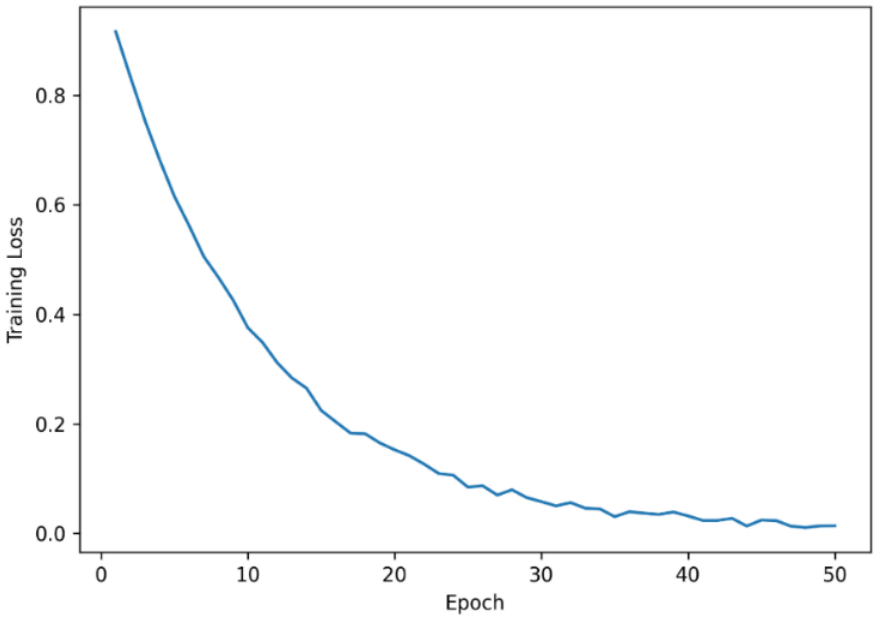
یافته‌ها

یافته‌های حاصل از اجرای مدل پیشنهادی نشان می‌دهد که چارچوب ارائه‌شده توانسته است الگوهای زمانی موجود در داده‌های فروش خرده‌فروشی را به‌طور مؤثری یاد بگیرد و پیش‌بینی قابل قبولی از تقاضا ارائه دهد. همان‌گونه که در جدول ۱ مشاهده می‌شود، شاخص‌های ارزیابی خطا بیانگر عملکرد مناسب مدل در مجموعه آزمون هستند و نشان می‌دهند که مدل قادر است تغییرات تقاضا را با دقت مناسبی بازنمایی کند. این یافته حاکی از آن است که ترکیب معماری ترنسفورمر با اطلاعات ساختاری و زمانی می‌تواند چارچوبی کارآمد برای پیش‌بینی پویای تقاضای خرده‌فروشی فراهم کند.

جدول ۱. مقایسه عملکرد مدل‌های پیش‌بینی تقاضا

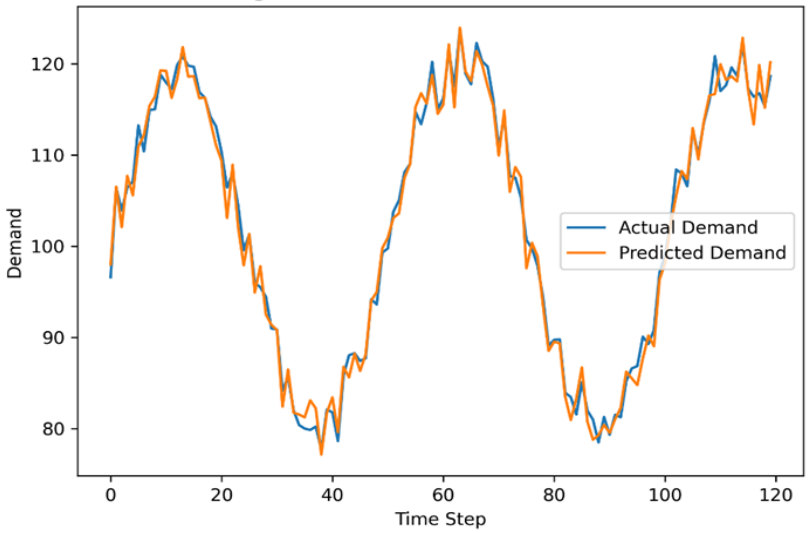
مدل	میانگین قدرمطلق خطا	میانگین مربعات خطا
ARIMA	۴/۳۷	۵/۹۶
XG-Boost	۳/۹۱	۵/۱۲
Transformer پیشنهادی	۲/۶۸	۳/۸۵

شکل ۲. روند تغییرات خطای آموزش مدل پیشنهادی



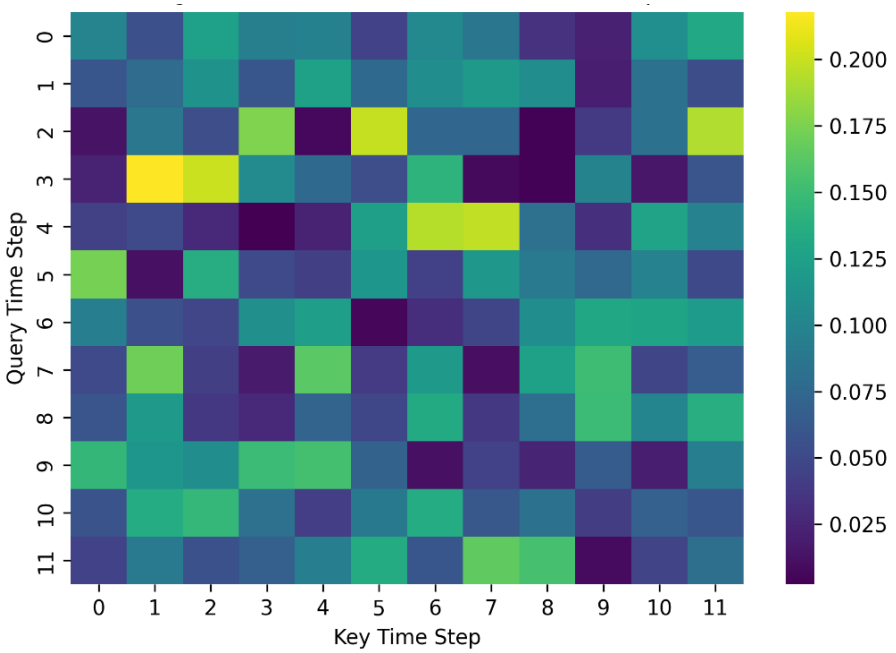
برای بررسی روند یادگیری مدل، تغییرات تابع خطا در طول فرآیند آموزش تحلیل شد. شکل ۲ روند کاهش مقدار خطا را در طی دوره‌های آموزش نشان می‌دهد. همان‌گونه که در این نمودار مشاهده می‌شود، مقدار خطا در مراحل اولیه آموزش با سرعت نسبتاً زیادی کاهش‌یافته و سپس به‌تدریج به یک ناحیه پایدار نزدیک می‌شود. این رفتار بیانگر آن است که مدل توانسته است در مراحل اولیه الگوهای اصلی داده‌ها را شناسایی کند و با ادامه آموزش به همگرایی مناسب برسد. پایداری منحنی خطا در دوره‌های پایانی نیز نشان‌دهنده ثبات فرآیند یادگیری و مناسب بودن تنظیمات مدل است.

شکل ۳. مقایسه مقادیر واقعی و پیش‌بینی‌شده تقاضا



در ادامه، عملکرد مدل ازنظر دقت پیش‌بینی در مجموعه آزمون مورد ارزیابی قرار گرفت. شکل ۳ مقایسه مقادیر واقعی تقاضا و مقادیر پیش‌بینی‌شده توسط مدل را نشان می‌دهد. نتایج نشان می‌دهد که منحنی پیش‌بینی در اغلب نقاط با روند واقعی داده‌ها همخوانی دارد و قادر است الگوهای افزایشی و کاهشی تقاضا را به‌خوبی دنبال کند. این همگرایی میان مقادیر واقعی و پیش‌بینی‌شده با شاخص‌های خطای گزارش‌شده در جدول ۱ نیز سازگار است و بیانگر توانایی مدل در تعمیم به داده‌های مشاهده نشده است.

شکل ٤. نقشه حرارتی مکانیزم توجه در مدل ترنسفورمر



برای تحلیل عمیق‌تر نحوه تصمیم‌گیری مدل، الگوی توجه در معماری ترنسفورمر نیز بررسی شد. شکل ۴ نقشه حرارتی وزن‌های توجه را در طول توالی زمانی نمایش می‌دهد. مطابق این شکل، برخی بازه‌های زمانی وزن توجه بیشتری دریافت کرده‌اند که نشان می‌دهد مدل این نقاط را در پیش‌بینی تقاضا مؤثرتر تشخیص داده است. این الگو نشان می‌دهد که مدل تنها به مشاهدات اخیر متکی نیست، بلکه قادر است وابستگی‌های زمانی مهم را در طول توالی شناسایی و در فرآیند پیش‌بینی لحاظ کند. چنین قابلیتی، همراه با نتایج کمی ارائه‌شده در جدول ۱، نشان‌دهنده کارایی معماری پیشنهادی در مدل‌سازی روابط زمانی پیچیده در داده‌های تقاضای خرده‌فروشی است.

تحلیل حساسیت

جهت انجام تحلیل حساسیت، برای سنجش استحکام مدل در برابر عدم‌قطعیت‌های داده‌ای، به ورودی‌های کلیدی مجموعه آزمون مانند قیمت و حجم تراکنش‌ها تا سقف ۱۰ درصد نویز تصادفی افزوده شد و تغییرات خطای پیش‌بینی با معیار RMSE پایش گردید. نکته مهم این است که داده‌های نویزدار به‌صورت مصنوعی از داده‌های آزمون ساخته شدند و در فرآیند آموزش مدل استفاده نشدند؛ بنابراین نتایج به‌طور مستقیم بیانگر پایداری و تعمیم‌پذیری مدل در مواجهه با داده‌های جدید و شرایط پیش‌بینی‌نشده است، نه بهبود ناشی از آموزش مجدد. شایان‌ذکر است که مقادیر پایه برای متغیرهای تعداد و قیمت، بر اساس میانگین کل این داده‌ها در مجموعه داده اصلی لحاظ شده است. همان‌گونه که جدول نشان می‌دهد،

اعمال نویز ۱۰ درصد در متغیرهای اصلی تنها موجب افزایش محدود RMSE می‌شود که بیانگر پایداری خروجی و عدم بروز نوسانات ناگهانی در پیش‌بینی تقاضا است. افزون بر این، برای بررسی حساسیت مدل نسبت به تنظیمات زمانی و مؤلفه‌های معماری، تغییرات کنترل‌شده در طول پنجره ورودی و افق پیش‌بینی و نیز نویز محدود در بخش‌های کلیدی مدل (ازجمله وزن‌های توجه و اطلاعات ساختاری) انجام شد که نتایج آن نشان می‌دهد مدل به دلیل بهره‌گیری از مکانیسم توجه و روابط غیرخطی، توانایی مناسبی در جداسازی الگوهای اصلی تقاضا از نویزهای محیطی داشته و از قابلیت تعمیم‌دهی قابل قبولی در برابر تغییرات کوچک ورودی و تغییرات پارامتری غیرمنتظره برخوردار است. نتایج این تحلیل حساسیت در جدول ۲ نشان داده شده است.

جدول ۲. تحلیل حساسیت مدل پیشنهادی (RMSE پایه یعنی جذر MSE برابر با ۱,۹۶ است)

گروه اصلی	پارامتر	مقدار پایه	تغییر اعمال‌شده	RMSE جدید	درصد تغییر RMSE
دینامیک زمانی	طول پنجره ورودی	۱۲	۱۰ ۱۵	۲,۰۴ ۲	+۴,۲ +۱,۸
	افق پیش‌بینی	۳	۵	۲,۱۱	+۷,۸
ورودی‌های عددی	تعداد	۱۰,۳۳	نویز تصادفی ±۱۰٪	۱,۹۸	+۰,۸
	قیمت	۴,۶۹	نویز تصادفی ±۱۰٪	۱,۹۹	+۱,۳
ساختار مدل	اطلاعات گراف	موجود در مدل پایه	حذف مؤلفه	۲,۰۹	+۶,۵
	سازوکار توجه	موجود در مدل پایه	نویز در وزن‌دهی	۱,۹۷	+۰,۵
بهینه‌سازی	نرخ یادگیری	۰,۰۰۰۵	۰,۰۰۱ ۰,۰۰۰۲۵	۱,۹۹ ۲,۰۱	+۱,۳ +۲,۶

بحث و نتیجه‌گیری

این پژوهش باهدف بررسی کارایی یک چارچوب مبتنی بر یادگیری عمیق برای پیش‌بینی تقاضای خرده‌فروشی و تحلیل الگوهای زمانی موجود در داده‌ها انجام شد. یافته‌های این پژوهش نشان داد که چارچوب پیشنهادی مبتنی بر معماری ترنسفورمر توانسته است الگوهای پیچیده زمانی موجود در داده‌های تقاضای خرده‌فروشی را به‌طور مؤثر مدل‌سازی کند. همگرایی مناسب در فرآیند آموزش و نزدیکی قابل‌قبول مقادیر پیش‌بینی‌شده به مقادیر واقعی بیانگر آن است که مدل از توان کافی برای بازنمایی نوسانات تقاضا برخوردار است. این نتایج نشان می‌دهد که در نظر گرفتن وابستگی‌های زمانی بلندمدت در کنار سازوکار توجه می‌تواند به درک دقیق‌تری از پویایی تقاضا منجر شود. درمجموع، عملکرد مدل

حاکی از آن است که بهره‌گیری از معماری‌های مبتنی بر توجه می‌تواند برای مسائل پیش‌بینی تقاضا در محیط‌های خرده‌فروشی، به‌ویژه در داده‌های دارای الگوهای زمانی پیچیده، رویکردی مؤثر و قابل‌اتکا باشد.

سهم دانش‌افزایی این پژوهش را می‌توان در سه سطح طبقه‌بندی کرد: نخست در سطح نظری، این مطالعه با تلفیق هم‌زمان پویایی‌های زمانی، وابستگی‌های شبکه‌ای و معناشناسی داده‌های متنی، چارچوب نظری جدیدی را در مدل‌سازی تقاضای چندوجهی ارائه می‌دهد که خلأ مدل‌های تک‌بعدی در ادبیات پیشین را پوشش می‌دهد. دوم در سطح متدولوژیک، توسعه یک معماری ترنسفورمر بومی‌سازی‌شده که قادر به استخراج روابط میان‌کالایی در محیط‌های خرده‌فروشی آنلاین است، توانمندی مدل‌های یادگیری عمیق را در پردازش داده‌های غیرساختارمند ارتقا می‌دهد. سوم در سطح کاربردی، این پژوهش با ارائه مدلی دقیق برای پیش‌بینی تقاضا با در نظر گرفتن تأخیرهای زمانی و اثرات شبکه، ابزار تصمیم‌گیری کارآمدتری برای مدیریت موجودی و کاهش هزینه‌های عملیاتی در اختیار مدیران زنجیره تأمین قرار می‌دهد.

نتایج به‌دست‌آمده با ادبیات پژوهشی موجود در حوزه پیش‌بینی تقاضای خرده‌فروشی هم‌راستا است. پژوهش‌های پیشین نشان داده‌اند که مدل‌های مبتنی بر ترنسفورمر و یادگیری عمیق در استخراج وابستگی‌های زمانی پیچیده و بهبود دقت پیش‌بینی عملکرد مطلوبی دارند. برای مثال، مقاله (Li, 2023) نشان می‌دهد که ترنسفورمرهای اصلاح‌شده می‌توانند دقت و کارایی پیش‌بینی فروش را افزایش دهند، درحالی‌که مقاله (Caetano et al., 2025) بر اهمیت مدل‌های ترنسفورمری در پیش‌بینی سری‌های زمانی با در نظر گرفتن متغیرهای توضیحی تأکید می‌کنند. همچنین، مطالعات (Rahikka & Mikkola, 2025) و (Tarighat et al., 2025) نیز بر ظرفیت روش‌های نوین سری زمانی و رویکردهای پیشرفته برای پیش‌بینی تقاضا در خرده‌فروشی تأکید دارند. از سوی دیگر، توجه این پژوهش به امکان توسعه مدل در جهت ادغام داده‌های متنوع، با مطالعات چندوجهی مانند (Wang, 2025) و (Cai et al., 2021) نیز همسو است که نشان می‌دهند تلفیق منابع داده‌ای مختلف می‌تواند به غنای مدل‌های پیش‌بینی تقاضا کمک کند. اگرچه مطالعات بین‌المللی مذکور، زیربنای نظری مدل‌های نوین را تأیید می‌کنند، اما در سطح بومی نیز، مقایسه با پژوهش‌های اخیر نشان می‌دهد که نوآوری ما از جایگاه ویژه‌ای برخوردار است .درحالی‌که مطالعات داخلی پیشین (مانند مسیبی اتاق‌سرا و همکاران، ۱۴۰۳؛ ویسی، ۱۴۰۴) عمدتاً بر مدل‌های پایه نظیر CNN یا LSTM برای تحلیل سری‌های زمانی متمرکز بوده‌اند، پژوهش حاضر با اتخاذ معماری ترنسفورمر و تلفیق آن با داده‌های چندوجهی و اثرات شبکه‌ای، موفق به ثبت خطای میانگین کمتر شده است. این نتایج با یافته‌های مطالعات بین‌المللی اخیر (Wang et.al, 2026) همسویی دارد و تأیید می‌کند که مدل‌های

مبتنی بر توجه به‌دلیل توانایی در درک وابستگی‌های بلندمدت و پویایی‌های متنی، در مواجهه با پیچیدگی‌های خرده‌فروشی آنلاین، عملکرد پایداری از خود نشان می‌دهند. درواقع، درحالی‌که پژوهش‌های پیشین اغلب بر پیش‌بینی مبتنی بر داده‌های ساختاریافته (فروش گذشته) تأکید داشتند، نوآوری این مطالعه در بازنمایی هم‌زمان ویژگی‌های متنی و ساختاری، تبیین دقیق‌تری از رفتارهای نوسانی تقاضا ارائه می‌دهد که در مطالعات پیشین کمتر موردتوجه قرار گرفته بود.

نتایج حاصل از مقایسه کمی نشان می‌دهد که مدل پیشنهادی این پژوهش ازنظر پایداری و کنترل خطاهای بزرگ، عملکرد بسیار مناسبی دارد. به‌طور خاص، مقدار MSE مدل پیشنهادی برابر با ۳,۸۵ و RMSE آن ۱,۹۶ است که در مقایسه با مدل Wang et al (2026) که به‌ترتیب MSE برابر با ۱۲,۶ و RMSE برابر با ۳,۵۵ را گزارش کرده است، به‌مراتب کمتر بوده و بیانگر دقت بالاتر و خطای پیش‌بینی کمتر در مدل پیشنهادی است. همچنین، در مقایسه با مطالعه Cai et al (2021) که RMSE را در بازه ۲,۰۳ تا ۳,۴۸ و MAPE میانگین را برابر ۱,۴۱ درصد گزارش کرده است، مدل پیشنهادی با RMSE برابر با ۱,۹۶ عملکرد رقابتی و حتی بهتری از حد پایین این بازه نشان می‌دهد. این نتایج حاکی از آن است که ترکیب ترنسفورمر با اطلاعات گراف هم‌خریدی، توانسته الگوهای پیچیده تقاضا را با خطای کمتر و پایداری بیشتر مدل‌سازی کند.

در این پژوهش، پس از ارائه نتایج محاسباتی در بخش پیشین، اکنون به تحلیل عمیق‌تر یافته‌ها و مقایسه آن‌ها با مطالعات مشابه پرداخته شده است. مدل پیشنهادی که بر پایه معماری ترنسفورمر بناشده و اثرات شبکه‌ای، تأخیرهای زمانی و معناشناسی محصول را تلفیق می‌کند، در مقایسه با دو رویکرد کلیدی مدل خودهمبسته میانگین متحرک انباشته (ARIMA) و الگوریتم تقویت گرادیانت (XG-Boost) که به‌عنوان مدل‌های مبنا در نظر گرفته شده‌اند، عملکرد برتری را نشان داده است. همان‌طور که در جدول نتایج (جدول ۱) مشاهده شد، مدل ما توانست میانگین قدرمطلق خطا (MAE) را از ۴,۳۷ (ARIMA) و ۳,۹۱ (XG-Boost) به ۲,۶۸ کاهش دهد و همچنین ریشه میانگین مربعات خطا (RMSE) را از ۵,۹۶ (ARIMA) و ۵,۱۲ (XG-Boost) به ۳,۸۵ برساند. این کاهش معنادار در هر دو معیار خطا، برتری رویکرد ما در پیش‌بینی پویای تقاضای خرده‌فروشی آنلاین را تأیید می‌کند.

این بهبود قابل‌توجه در عملکرد، عمدتاً ناشی از نوآوری‌های اصلی این پژوهش است. اولاً، توانایی مدل در مدل‌سازی مؤثر اثرات شبکه‌ای میان محصولات از طریق مکانیزم استفاده‌شده لایه‌های کانولوشن گراف، امکان درک روابط پیچیده بین محصولات را فراهم کرده است. ثانیاً، پردازش دقیق تأخیرهای زمانی با استفاده از مکانیزم استفاده شده مکانیسم توجه پویا، به مدل اجازه داده است تا الگوهای زمانی تقاضا را با دقت بیشتری ثبت کند. ثالثاً، استخراج غنی معناشناسی محصول از طریق مکانیزم استفاده‌شده

بردارهای ویژگی متنی پیش‌آموزش‌دیده، به مدل در درک بهتر ویژگی‌های توصیفی محصولات یاری رسانده است.

این دستاوردها، به‌طور مستقیم به شکاف‌های پژوهشی شناسایی‌شده در مقدمه مقاله پاسخ می‌دهند. درحالی‌که مطالعات پیشین عمدتاً بر یکی از این عوامل (اثرات شبکه‌ای، تأخیر زمانی یا معناشناسی محصول) تمرکز داشته‌اند یا آن‌ها را به‌صورت مجزا تحلیل کرده‌اند، پژوهش حاضر اولین تلاشی است که این سه عامل کلیدی را به‌صورت یکپارچه و در قالب یک مدل ترنسفورمر موردبررسی قرار می‌دهد. این رویکرد تلفیقی، منجر به افزایش چشمگیر دقت پیش‌بینی تقاضا شده و گامی مؤثر در جهت رفع کاستی‌های مدل‌های موجود و ارتقای درک ما از پویایی‌های تقاضای خرده‌فروشی آنلاین برداشته است.

نتایج تجربی نشان داد که مدل پیشنهادی مبتنی بر معماری ترنسفورمر توانسته است وابستگی‌های زمانی کوتاه‌مدت و بلندمدت را به‌طور مؤثر شناسایی کرده و الگوهای پیچیده تقاضا را با دقت قابل‌قبولی مدل‌سازی کند. تحلیل فرآیند آموزش نیز نشان داد که مدل از همگرایی پایدار برخوردار بوده و در طول آموزش به‌تدریج به سطح مناسبی از عملکرد پیش‌بینی دست یافته است. مقایسه مقادیر پیش‌بینی‌شده با مقادیر واقعی نیز حاکی از آن است که مدل توانایی مناسبی در بازنمایی نوسانات تقاضا در داده‌های واقعی خرده‌فروشی دارد. این نتایج نشان می‌دهد که استفاده از سازوکار توجه در چارچوب‌های مبتنی بر ترنسفورمر می‌تواند به درک بهتر ساختارهای زمانی داده‌های تقاضا کمک کند. به‌طورکلی، یافته‌های این مطالعه نشان می‌دهد که رویکردهای مبتنی بر یادگیری عمیق می‌توانند ابزار مؤثری برای بهبود دقت پیش‌بینی تقاضا در محیط‌های خرده‌فروشی باشند. درنتیجه، پرسش‌های اصلی پژوهش درباره کارایی مدل پیشنهادی و توانایی آن در مدل‌سازی الگوهای زمانی پیچیده تا حد قابل‌توجهی پاسخ داده شده است.

از منظر کاربردی، نتایج این پژوهش می‌تواند پیامدهای مهمی برای مدیریت عملیات خرده‌فروشی و برنامه‌ریزی زنجیره تأمین داشته باشد. استفاده از مدل‌های پیش‌بینی دقیق‌تر می‌تواند به سازمان‌ها کمک کند تا تصمیم‌گیری‌های آگاهانه‌تری در زمینه مدیریت موجودی، برنامه‌ریزی سفارش‌ها و تخصیص منابع اتخاذ کنند و درنتیجه سطح عدم‌قطعیت در فرآیندهای عملیاتی کاهش یابد. علاوه بر این، بهره‌گیری از مدل‌های مبتنی بر داده می‌تواند به توسعه رویکردهای خرده‌فروشی داده‌محور و افزایش کارایی در تحلیل رفتار تقاضا کمک کند. باوجود کارایی مدل پیشنهادی، تبیین محدودیت‌های آن برای درک دقیق‌تر کارکرد مدل در محیط‌های عملیاتی ضروری است. نخست، مدل کنونی بر فرض محیط قطعی بنا شده است که ممکن است در شرایط نوسانات شدید بازار یا شوک‌های غیرمنتظره زنجیره تأمین، تمامی جنبه‌های تصادفی رفتار مشتریان را پوشش ندهد. دوم، تکیه بر داده‌های تاریخی برای استخراج روابط شبکه‌ای، مفروض «پایداری روابط میان‌محصولی» را پیش‌فرض می‌گیرد که در بازارهای پویا ممکن است

دچار تغییر شود. در زمینه اعتبارسنجی داده‌ها، فرآیند پیش‌پردازش شامل حذف سیستماتیک رکوردهای ناقص و شناسایی داده‌های پرت با استفاده از روش‌های آماری بوده است تا از سوگیری مدل جلوگیری شود؛ بااین‌حال، کیفیت خروجی مدل همچنان به یکپارچگی داده‌های ورودی وابسته است.

برای پژوهش‌های آتی، پیشنهاد می‌شود با استفاده از یادگیری عمیق بیزی، بازه‌های اطمینان احتمالی در پیش‌بینی‌ها گنجانده شود تا ابعاد عدم قطعیت پارامتریک بهتر مدیریت گردد. همچنین، ادغام منطق فازی در لایه‌های مدل برای مواجهه با پارامترهای غیردقیق و مبهم در رفتار خرید مشتریان، می‌تواند به افزایش استحکام مدل در مواجهه با داده‌های نویزی کمک کند. درنهایت، اعتبارسنجی فرض پایداری شبکه‌های محصولی از طریق تحلیل‌های مقطعی در بازه‌های زمانی مختلف و به‌کارگیری معماری‌های گراف-محور پویا، ازجمله جهت‌گیری‌های کلیدی برای تعمیم‌پذیری بیشتر مدل در محیط‌های پیچیده خرده‌فروشی آنلاین است.

تعارض منافع

نویسنده هیچ‌گونه تعارض منافعی ندارد.

ORCID

Fatemeh Zare Baghiabad **http://orcid.org/0000-0002-0669-4686**

منابع

۱. حسینی، س. م.، و مازندرانی‌زاده، ح. (۱۴۰۴). بررسی تابع تقاضا معکوس و شبکه عصبی مصنوعی در پیش‌بینی قیمت محصولات زراعی (مطالعه موردی: شبکه آبیاری دشت قزوین). نشریه علوم و مهندسی آبیاری، ۴۸ (۲)، ۳۷-۴۹. https://doi.org/10.22055/jise.2024.45614.2109

۲. شاملو، ز.، و ناظم، ف. (۱۴۰۲). بررسی تأثیر هوش مصنوعی بر توسعه سیستم‌های پیش‌بینی تقاضا برای کسب‌وکارهای خرده‌فروشی .نشریه تکنولوژی در کارآفرینی و مدیریت استراتژیک، ۲(۱)، ۵- ۱۲. https://doi.org/10.61838/kman.jtesm.2.1.2

۳. مسیبی اتاق‌سرا، ا.، آراسته، ع.، و جوادیان، ن. (۱۴۰۳). پیش‌بینی روزانه تقاضای کالای فاسدپذیر با استفاده از یادگیری عمیق (مطالعه موردی: ماست سون شرکت فرآورده‌های لبنی کاله آمل). نشریه تصمیم‌گیری و تحقیق در عملیات، ۹ (۴۴)، ۸۳۳–۸۶۵. https://doi.org/10.22105/dmor.2024.467818.1859

۴. ویسی، ا. (۱۴۰۵). بهینه‌سازی یکپارچه موجودی قطعات یدکی با استفاده از پیش‌بینی تقاضا مبتنی بر یادگیری ماشین با در نظر گرفتن موجودی اطمینان و تأمین با دو زمان تحویل. نشریه مهندسی سیستم و بهره‌وری، ۶ (۳)، ۲۲۵–۲۶۲. https://doi.org/10.22034/sep.2026.2074035.1409

References

5. Caetano, R., Oliveira, J. M., & Ramos, P. (2025). Transformer-based models for probabilistic time series forecasting with explanatory variables. *Preprints*. https://doi.org/10.20944/preprints202502.0009.v1
6. Cai, W., Song, Y., & Wei, Z. (2021). Multimodal data guided spatial feature fusion and grouping strategy for e-commerce commodity demand forecasting. *Mobile Information Systems, 2021*(1), 5568208. https://doi.org/10.1155/2021/5568208
7. Chowdhury, A. R., Paul, R., & Rozony, F. Z. (2025). A systematic review of demand forecasting models for retail e-commerce enhancing accuracy in inventory and delivery planning. *International Journal of Scientific Interdisciplinary Research, 6*(1), 1–27. https://doi.org/10.63125/mbbfw637
8. Li, Q. (2023). Achieving sales forecasting with higher accuracy and efficiency: A new model based on modified Transformer. *Journal of Theoretical and Applied Electronic Commerce Research, 18*(4), 1990–2006. https://doi.org/10.3390/jtaer18040100
9. M. A., R., Rodrigues, G. N., Mir, N. H., Bhuiyan, S. M., Mridha, M. F., Islam, R., & Watanobe, Y. (2025). A hybrid temporal convolutional network and Transformer model for accurate and scalable [Title incomplete in source

provided]. *IEEE Open Journal of the Computer Society, 6*, 380–391. https://doi.org/10.1109/OJCS.2025.3538579

10. Rahikka, J., & Mikkola, P. (2025). *Modern time series methods for demand forecasting in retail*. University of Helsinki. https://helda.helsinki.fi/server/api/core/bitstreams/0f5658a0-0cb9-40b9-ab1e-b32e5f365d89/content
11. Rai, S. K. (2025). Data-driven retail: The engineering behind personalized customer experiences. *Journal of Computer Science and Technology Studies, 7*(10), 571–581. https://doi.org/10.32996/jcsts
12. Saarinen, L., & Huttunen, P. (2025). Revisiting the value of data sharing in retail supply chain demand planning. *International Journal of Operations & Production Management, 45*(11), 1910–1936. https://doi.org/10.1108/IJOPM-07-2024-0560
13. Tarighat, N., Cohen, M. C., Clark, J. J., Member, L. S., Jorge, R., & Caetano, T. (2025). Domain adaptation for retail demand prediction. *IEEE Access, 13*, 146267–146294. https://doi.org/10.1109/ACCESS.2025.3600468
14. Tiwari, R. (2025). Harnessing AI and predictive analytics for enhanced demand forecasting in retail supply chains. *Supply Chain and Retail Management, 8*(1), 41–52.
15. Tripathi, S., Trigunait, R., & Chandra, D. (2025). A behavioral and environmental framework for sustainable retail forecasting and decision-making. *Circular Economy and Sustainability*, 1–29. https://doi.org/10.1007/s43615-025-00705-1
16. Wang, J., Zhang, O., Qi, Y., Tang, H., Jiang, Y., & Khurelbaatar, Z. (2026). Masked Transformer with mixed features for retail demand forecasting to optimize IoT-driven supply chain management. *IEEE Internet of Things Journal, 13*(7), 12644–12654. https://doi.org/10.1109/JIOT.2026.3656390
17. Wang, Y. (2025). Causal-aware multimodal Transformer for supply chain demand forecasting: Integrating text, time series, and satellite imagery. *IEEE Access, 13*, 176813–176829. https://doi.org/10.1109/ACCESS.2025.3619552

References [In Persian]

1. Hosseini, S. M., & Mazandarani Zadeh, H. (2025). Comparison of the ability of inverse demand function and artificial neural network to predict crop prices (case study: Qazvin Plain Irrigation Network). *Irrigation Sciences and Engineering (JISE), 48*(2), 37–49. https://doi.org/10.22055/jise.2024.45614.2109 [In Persian]
2. Mosayyebi Otaghsara, A. H., Arasteh, A., & Javadian, N. (2024). Daily forecasting of demand for perishable goods using deep learning (case study: Seven Yoghurt, Kale Amol Dairy Products Company). *Journal of Decisions and Operations Research, 9*(4), 833–865. https://doi.org/10.22105/dmor.2024.467818.1859 [In Persian]

3. Shamloo, Z., & Fatah, N. (2023). Exploring the impact of artificial intelligence on the development of demand forecasting systems for retail businesses. *Journal of Technology in Entrepreneurship and Strategic Management*, *2*(1), 6–16. https://doi.org/10.61838/kman.jtesm.2.1.2 [In Persian]
4. Veisi, O. (2026). An integrated optimization of spare parts inventory using machine-learning-based demand forecasting, taking into account safety stock and dual lead-time supply. *System Engineering and Productivity*, *6*(3), 225–262. https://doi.org/10.22034/sep.2026.2074035.1409 [In Persian]

استناد به این مقاله: زارع باقی‌آباد، فاطمه. (۱۴۰۵). پیش‌بینی پویای تقاضای خرده‌فروشی مبتنی بر رویکرد ترنسفورمر با تلفیق اثرات شبکه‌ای، تأخیرهای زمانی و معناشناسی محصول، *مدیریت صنعتی*، *۲۴(۸۱)*، ۳۵-۶۶. DOI: 10.22054/jims.2026.91974.3023

Industrial Management Studies is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.