PRESS
PINNACLE SCIENCE
Transactions on Computational and Scientific Methods | Vo. 4, No. 3, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Multi-Source Data Integration and Sales Forecasting for Chain Retail Enterprises Using ETL Techniques

Zouxiaowei Ma
Columbia University, New York, USA
zouxiaowei.ma@outlook.com

Abstract: Chain retail enterprises accumulate a large amount of business data from sales, stores, promotions, customer flow, and time environment in their daily operations. However, differences in structure, inconsistent field definitions, and complex information relationships between different data sources affect the stability and accuracy of store sales forecasting. To address this issue, this paper constructs an ETL data fusion and store sales forecasting method for chain retail enterprises. This method first extracts, cleans, transforms, and loads sales records and store attribute information from the Rossmann Store Sales dataset, performing preprocessing operations such as missing value handling, outlier correction, categorical feature encoding, and continuous variable standardization. Based on this, static store attributes, calendar features, promotion status, customer flow information, and recent sales signals are integrated into a structured feature representation, and the time dependency in store sales changes is preserved through historical window construction. Subsequently, a sales forecasting model is built based on the fused daily store data to estimate future sales revenue. Comparative experimental results show that the proposed method performs well in terms of RMSE, MAE, MAPE, and R^2, effectively reducing prediction errors and enhancing model fitting ability. Research shows that ETL data fusion can provide a more complete and reliable data foundation for retail sales forecasting, and the constructed methods can provide data support for chain retail enterprises' inventory management, replenishment planning, promotion arrangements, and operational decisions.

Keywords: ETL data fusion; store sales forecasting; chain retail; historical window modeling

1. Introduction

As chain retail enterprises continue to expand their operations, the number of stores, product categories, membership groups, and sales channels increases, resulting in data generated during daily operations that is characterized by diverse sources, complex structures, frequent updates, and dense correlations[1,2]. Sales figures, inventory records, member consumption data, promotional activities, store attributes, regional environment, and holiday information collectively form the crucial data foundation for retail business analysis. Because this data is typically scattered across different business systems and management platforms, differences in data format, statistical definitions, and update cycles directly impact the accuracy of subsequent data analysis and business decisions. Therefore, how to standardize the extraction, cleaning, transformation, and integration of multi-source retail data has become a critical issue for chain retail enterprises to improve their data management capabilities and business analysis levels.

In the chain retail scenario, store sales are influenced by a combination of factors, including product structure, pricing strategies, promotional rhythms, inventory status, customer traffic changes, consumption habits, and the regional economic environment, exhibiting significant dynamism and uncertainty. Traditional methods relying on manual experience or single historical sales figures are insufficient to fully capture the differences between stores and to promptly reflect the impact of external factors on sales fluctuations[3]. When a company has a large number of stores and a wide variety of products, sales forecasting results not only affect procurement plans and inventory allocation, but also further impact supply chain scheduling, staff scheduling, marketing efforts, and capital turnover efficiency. Therefore, developing a sales forecasting methodology for chain retail enterprises is crucial for improving their refined operational capabilities.

ETL data processing provides fundamental support for sales forecasting in chain retail enterprises. By uniformly extracting multi-source heterogeneous data, controlling data quality, standardizing fields, handling missing values, identifying outliers, and integrating business features, the impact of noise in the raw data on the forecast results can be effectively reduced, and the correlation between different business data can be enhanced[4]. Compared to analyzing only single sales data, the fused data can more comprehensively reflect the store's operating status, product sales patterns, and changes in the external environment, thus providing a more complete and reliable input foundation for the sales forecasting model. Therefore, ETL data fusion is not only a data preprocessing step, but also a crucial bridge connecting enterprise business systems with intelligent predictive analytics.

Research on ETL data fusion and store sales forecasting for chain retail enterprises has strong theoretical value and practical application significance. From a theoretical perspective, this research helps explore the organization and integration mechanisms of multi-source business data in retail forecasting tasks, providing methodological references for data-driven forecasting in complex business scenarios[5]. From an application perspective, accurate store sales forecasting can help companies identify sales trends in advance, optimize replenishment decisions and inventory structure, reduce problems such as stockouts, overstocking, and resource waste, and provide data support for promotional planning and store operation adjustments. Therefore, this research can promote the transformation of chain retail enterprises from experience-driven to data-driven approaches, and is of great significance for improving operational efficiency, enhancing market responsiveness, and driving the digital transformation of retail.

2. BackGround

The IT infrastructure of chain retail enterprises is typically supported by multiple business systems, including sales management, inventory management, membership management, supply chain management, and financial management. These systems differ in their development time, data standards, business definitions, and storage methods, leading to problems such as data fragmentation, inconsistent fields, data redundancy, delayed updates, and inconsistent data quality when conducting business analysis[6]. ETL, as a key process in enterprise data integration, can extract, transform, and load data scattered across different business systems, creating analyzable, traceable, and reusable data resources within a unified data environment. For chain retail enterprises, ETL not only handles basic data processing but also affects the reliability of subsequent data modeling, indicator calculations, and business decisions.

Store sales forecasting is a crucial task in retail enterprise management. Its core lies in identifying patterns in sales changes from historical sales records and related business factors, and making reasonable estimates of future sales trends. As retail operations shift from single-channel offline sales to multi-channel collaborative operations, sales changes are influenced by factors such as store location, product mix, membership structure, promotional strategies, inventory status, and seasonal cycles, making data relationships more complex[7]. Traditional statistical methods have limitations in handling complex nonlinear relationships and multivariate interactions. In contrast, data fusion-based forecasting methods can more fully utilize internal operational data and external auxiliary information, providing more detailed data for sales assessment at the store level. Therefore, constructing a sales forecasting research framework around ETL data fusion can provide crucial support for data governance and intelligent operations in chain retail enterprises.

3. Datasets and Dataset Preprocessing

3.1 Dataset

This study selects the Rossmann Store Sales dataset as the data source for predicting store sales in chain retail enterprises. This dataset contains historical sales records from 1115 stores, which can comprehensively reflect the sales changes of different stores over time in a chain retail scenario. The data includes fields such as store number, date, sales amount, customer traffic, business status, promotion status, holiday information, and school holiday information, as well as store attribute information such as store type, product mix, distance to competing stores, and promotion cycle. Because this dataset possesses time-series characteristics, store-specific characteristics, and business environment characteristics, it can support the fusion and modeling of multi-source retail data and provide a rich data foundation for store sales prediction tasks.

3.2 Dataset preprocessing

The data preprocessing stage first involves linking and integrating sales data and store attribute data, and standardizing date formats, field types, and data encoding methods. Missing values, outliers, and invalid sales records are cleaned to reduce the impact of noisy data on subsequent modeling. Subsequently, categorical features are numerically transformed, and continuous variables such as sales revenue, customer traffic, and competitive distance are standardized to form structured data suitable for sales forecasting tasks. The statistical analysis of this dataset is shown in Figure 1.

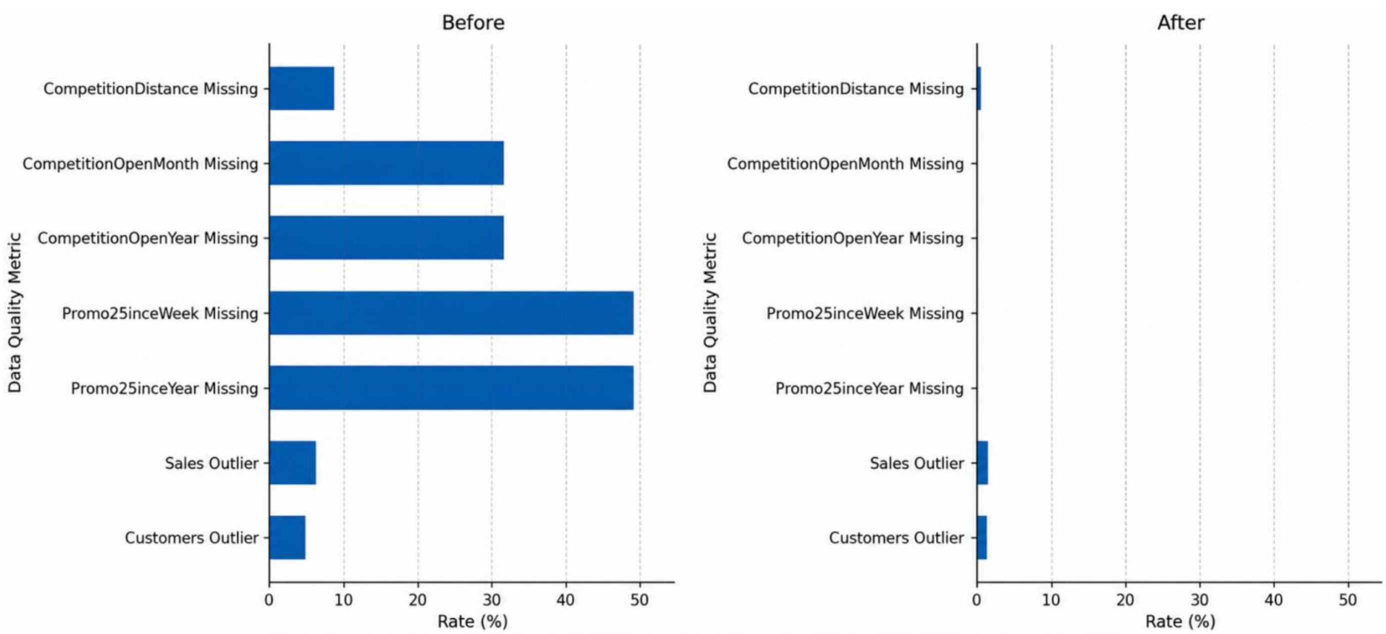

Figure 1. Dataset Statistical Analysis

As shown in Figure 1, the Rossmann Store Sales dataset had significant data quality issues before preprocessing. Specifically, fields related to the promotion start week, promotion start year, and competitor store opening dates had a high proportion of missing values, and sales figures and customer traffic also contained outliers. After data cleaning, missing value handling, and outlier correction, all data quality indicators significantly decreased, indicating that the preprocessing effectively reduced incompleteness and noise in the original data. The preprocessed data has a more regular distribution, providing a more stable and reliable structured data foundation for subsequent ETL data fusion and store sales forecasting.

4. Methodology

A method for store level sales forecasting must first transform heterogeneous retail records into a unified analytical representation rather than treating sales, calendar, promotion, customer traffic, and store attributes as isolated tables. The raw retail information is therefore represented as a set of relational sources $\mathcal{S} = \{S_1, S_2, \ldots, S_K\}$, in which each source contributes a different business view of the same chain retail system, and the extraction stage preserves the temporal index, store identifier, operating status, promotion state, holiday attributes, and store specific descriptors required for downstream prediction. Its overall model architecture is shown in Figure 2.

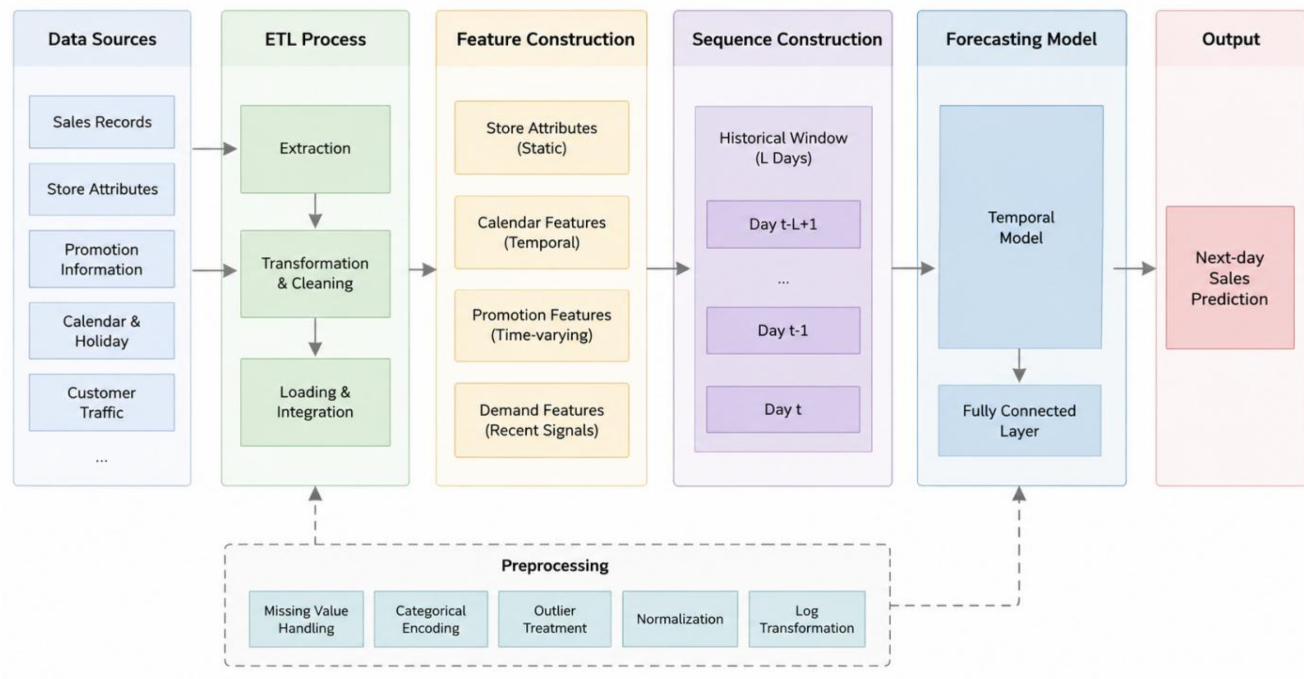

Figure 2. Overall model architecture

To reduce semantic inconsistency across sources, each field is mapped into a normalized schema through a transformation function that aligns naming conventions, data types, missing value rules, and business meanings before feature construction.

$$\boldsymbol{z}_{i,t} = \Phi_{etl}\left(S_1, S_2, \ldots, S_K; i, t\right)$$

Here, $\boldsymbol{z}_{i,t}$ denotes the fused record of store i on day t, while Φ_{etl} expresses the complete ETL transformation that converts multi-source retail records into a consistent store day feature unit. This formulation is meaningful because sales forecasting in chain retail enterprises depends not only on historical sales but also on whether the corresponding business context has been consistently extracted and integrated. To further ensure that the fused record reflects both stable store attributes and time-varying operational signals, the feature vector is organized as a structured combination of static, temporal, promotional, and demand-related components.

$$\boldsymbol{x}_{i,t} = \left[\boldsymbol{a}_i, \boldsymbol{c}_t, \boldsymbol{p}_{i,t}, \boldsymbol{d}_{i,t}\right]$$

In this expression, $\boldsymbol{a}_i$ represents store-level attributes such as store type, assortment, and competition distance, $\boldsymbol{c}_t$ captures calendar information such as weekday, month, and holiday status, $\boldsymbol{p}_{i,t}$ encodes promotion-related variables, and $\boldsymbol{d}_{i,t}$ contains demand observations such as customer traffic and recent sales signals. Such a representation makes the prediction task closer to real retail decision-making, because each sales value is interpreted under a specific operational condition rather than as an independent numerical point.

Preprocessing is introduced to improve the reliability of feature learning while maintaining the original business structure of the dataset. Missing values are handled according to their field semantics, categorical fields are converted into learnable numerical representations, and continuous variables are scaled to reduce the dominance of large magnitude attributes such as sales revenue, customer traffic, and competitive distance. Instead of applying a uniform transformation to all fields, numerical variables are standardized after data cleaning so that the model can compare heterogeneous retail indicators under a shared numerical scale.

$$\widetilde{x}_{i,t}^{(m)} = \frac{x_{i,t}^{(m)} - \mu_m}{\sigma_m + \epsilon}$$

Through this normalization, the mth continuous feature in the store day record (i,t) is adjusted by the feature mean μ_m and standard deviation σ_m, while the small constant ϵ avoids numerical instability during calculation. This step is important because retail data often contains variables with different units and ranges, and uncontrolled scale differences may cause the prediction model to overemphasize high-magnitude fields while weakening the contribution of smaller but meaningful indicators. For the forecasting target, sales values are transformed with a logarithmic function to reduce skewness and improve the stability of regression learning.

$$y_{i,t} = log\left(Sales_{i,t} + 1\right)$$

After this transformation, $y_{i,t}$ serves as the target value corresponding to the original sales amount of store i on day t, and the logarithmic form helps represent both low sales and high sales observations within a smoother value range.

Temporal dependence is modeled by constructing a store-specific observation window, because the sales of a chain retail store are usually affected by recent demand patterns, promotion continuity, weekly cycles, and short-term customer fluctuations. Rather than relying only on the current day record, the prediction input is formed by combining the latest L days of fused features for each store, allowing the model to observe local sales dynamics and contextual changes before making a forecast.

$$\boldsymbol{H}_{i,t} = \left[\boldsymbol{x}_{i,t-L+1}, \boldsymbol{x}_{i,t-L+2}, \ldots, \boldsymbol{x}_{i,t}\right]$$

Within this temporal matrix, $\boldsymbol{H}_{i,t}$ describes the historical feature sequence of store i ending at day t, and L controls the length of the historical window used to capture short-term retail behavior. This design is useful because the same promotion or holiday signal may have different effects depending on the previous sales state, customer volume, and recent store activity. The forecasting function then estimates future sales by learning nonlinear relations between the temporal feature matrix and the target value.

$$\widehat{y}_{i,t+1} = f_\theta\left(\boldsymbol{H}_{i,t}\right)$$

Here, f_θ denotes the trainable forecasting model parameterized by θ, and $\widehat{y}_{i,t+1}$ represents the predicted log-scale sales value for the next day. By learning from fused historical sequences, the model can jointly consider store heterogeneity, calendar periodicity, promotion influence, and recent demand variation.

Optimization is defined around the difference between predicted sales and observed sales in the transformed target space, so that the model focuses on minimizing prediction deviation across stores and dates while maintaining stable training behavior. A regularized regression objective is adopted to control model complexity and reduce the risk of overfitting to store-specific noise or temporary sales fluctuations.

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{(i,t)\in\Omega} \left(y_{i,t+1} - \widehat{y}_{i,t+1}\right)^2 + \lambda \parallel \theta \parallel_2^2$$

Under this objective, Ω denotes the collection of valid store-day samples, N is the number of training records, and λ balances the prediction error term with the parameter regularization term inside a single learning criterion. The final prediction can be converted back to the original sales scale when business interpretation is required, which makes the method compatible with practical retail decision-making such as replenishment planning, promotion evaluation, and store operation adjustment. Overall, the proposed methodological framework connects ETL-based data fusion with temporal sales forecasting, enabling chain retail data to be processed as an integrated decision-oriented representation rather than as fragmented operational records.

5. Experimental Results and Analysis

5.1 Experimental setup

The experiment was conducted under a unified model training configuration to ensure that the predictive model could learn the patterns of store sales changes under stable optimization conditions. The model training employed the AdamW optimizer, combined with learning rate decay, weight regularization, and Dropout mechanisms to control the parameter update process, thereby reducing the risk of overfitting to local noise and short-term fluctuations. The main hyperparameter settings are shown in Table 1; these parameters collectively determine the model's training step size, network capacity, time-dependent modeling range, and optimization stability.

Table 1: Detailed hyperparameter settings

Hyperparameter	Setting
Optimizer	AdamW
Initial learning rate	0.001
Batch size	64
Training epochs	100
Weight decay	0.0001
Dropout rate	0.2
Hidden dimension	128
Number of temporal layers	2
Historical window length	30
Gradient clipping threshold	1.0

5.2 Experimental Results and Analysis

To verify the applicability of the constructed ETL data fusion and store sales forecasting method in chain retail operation scenarios, research related to retail demand forecasting, sales time series modeling, and multi-source operational feature fusion in recent years was selected as a comparison object. The relevant methods cover machine learning, deep sequence modeling, probabilistic prediction, and hierarchical time series prediction, reflecting the differences in model structure, feature utilization, and time-series dependency modeling in store sales forecasting tasks from different perspectives. Table 2 shows the comparison of different methods under a unified evaluation index.

Table 2: Comparative results of related methods and the proposed method

Method	RMSE	MAE	MAPE	R^2
Carbonneau et al.[8]	36.84	24.71	15.43	0.82
Pavlyshenko et al.[9]	33.58	22.46	13.75	0.84
Bandara et al.[10]	31.92	21.37	12.96	0.86
Salinas et al.[11]	30.64	20.58	12.31	0.87
Makridakis et al.[12]	29.35	19.64	11.72	0.88
Saha et al.[13]	28.47	18.91	11.15	0.89
Challu et al.[14]	27.63	18.22	10.68	0.90
Ours	24.18	15.84	8.96	0.94

Table 2 shows that the proposed algorithm demonstrates stronger error control capabilities in store sales forecasting. Lower RMSE, MAE, and MAPE indicate that the model can more accurately characterize local fluctuations and overall deviations in sales revenue changes, reducing the distance between predicted values and actual sales. A higher R^2 indicates that the model has better explanatory power for store sales revenue changes and can more fully learn the temporal patterns, store differences, and business influencing factors contained in the sales data.

Combined evaluation metrics show that the proposed method exhibits better overall performance in error reduction and improved fitting ability. These results demonstrate that after ETL data fusion, feature construction, and time series modeling, the model can more effectively utilize multi-source retail information and enhance its ability to express complex sales changes. Therefore, the proposed algorithm can provide more stable and reliable data analysis support for sales forecasting in chain retail enterprises.

To further analyze the role of different components in the method for predicting store sales, this paper focuses on ablation settings for three key stages: ETL data fusion, structured feature construction, and historical window sequence construction. Table 3 shows the prediction performance under different settings. By gradually removing the core modules in the method, the impact of each stage on error control and fitting ability can be observed, thereby verifying the necessity of the complete framework in multi-source retail data modeling.

Table 3. Ablation study results of the proposed method

Ablation Setting	RMSE	MAE	MAPE	R^2
w/o ETL Data Fusion	31.76	20.88	12.64	0.86
w/o Structured Feature Construction	29.84	19.32	11.52	0.88
w/o Historical Window Construction	28.37	18.41	10.91	0.89
Ours	24.18	15.84	8.96	0.94

Table 3 shows that the proposed algorithm has better store sales prediction performance under the complete structure. Removing ETL Data Fusion makes it difficult for the model to fully integrate multi-source data such as sales records, store attributes, promotional information, and time context, leading to a decrease in the completeness and consistency of input information. Removing Structured Feature Construction weakens the model's ability to express static store features, time features, promotional features, and demand features, making it difficult to fully characterize the multi-factor relationships in retail business. Removing Historical Window Construction results in insufficient utilization of recent sales status and time dependencies, making the prediction process more susceptible to fluctuations in daily data. The complete method can simultaneously retain data fusion, structured feature expression, and historical window modeling capabilities, thus exhibiting more stable performance in error control and fitting ability, indicating that the proposed algorithm has good overall effectiveness.

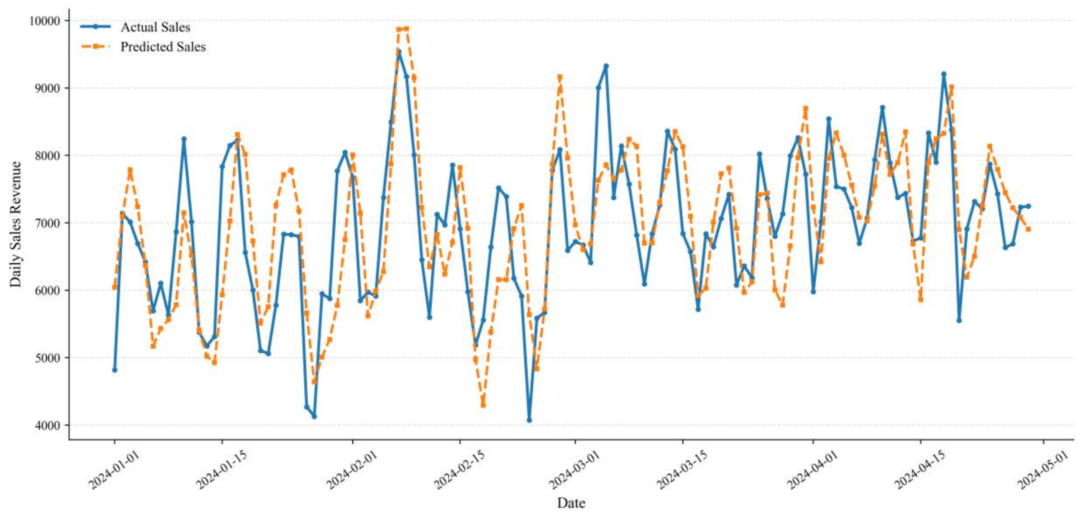

Figure 3. Experimental results comparing actual and predicted values

The algorithm proposed in this paper can fit the actual changes in store sales well. The predicted curve and the actual sales curve maintain a high degree of consistency in the overall fluctuation position and the direction of local changes, as shown in Figure 3. Even when faced with fluctuations in sales data caused by promotions, changes in customer traffic, and cyclical operational factors, the model can still maintain relatively stable predictive performance, indicating that ETL data fusion and historical window modeling can effectively enhance the model's ability to utilize complex retail business information. This result further demonstrates that the proposed method has good adaptability and application value in store sales forecasting tasks, and can provide reliable data support for the sales planning, inventory arrangement, and operational decisions of chain retail enterprises.

The algorithm proposed in this paper maintains a good consistency between predicted and actual values. The scatter distribution generally revolves around the ideal fitted line, indicating that the model has a relatively stable predictive ability for store samples across different sales ranges. Despite some fluctuations and dispersion in the sales data, the prediction results still closely approximate actual sales levels, demonstrating the model's robustness in handling complex retail data.

The proposed method effectively learns the mapping relationship between sales revenue and multi-source business characteristics, maintaining a good fit across a wide sales range, as shown in Figure 4. These results demonstrate that ETL data fusion, structured feature construction, and historical window modeling can collectively enhance the model's ability to express changes in store sales, providing reliable methodological support for chain retail enterprises in sales forecasting and operational decision-making.

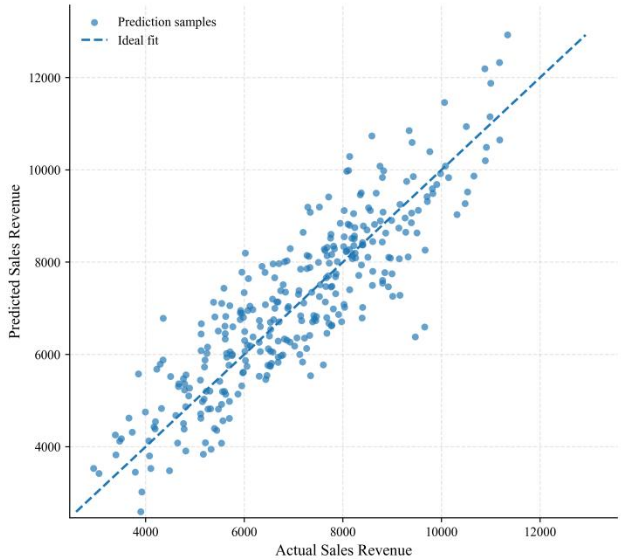

Figure 4. Predicting Scatter Experiment Results

6. Discussion

This paper addresses the challenges faced by chain retail enterprises, including fragmented multi-source business data, complex sales influencing factors, and reliance on data for store operational decisions. It constructs a store sales forecasting method based on ETL data fusion, demonstrating strong practical application value. By unifying and fusing sales records, store attributes, promotional information, time factors, and customer flow data, the method enhances enterprises' ability to perceive changes in store sales and provides more reliable data support for replenishment planning, inventory allocation, promotional arrangements, and operational management. This method helps reduce the uncertainty brought about by experience-based judgments, enabling chain retail enterprises to develop more refined data analysis processes in multi-store, multi-product, and multi-scenario operating environments, thereby improving operational efficiency and market responsiveness.

However, this research still has certain limitations. First, the modeling is primarily based on existing structured retail data and has not yet incorporated richer external factors such as weather, regional consumption levels, customer flow in commercial districts, and changes in the competitive environment. Therefore, there is still room for improvement in characterizing complex real-world scenarios. Secondly, the model mainly focuses on sales forecasting at the store level, and the analysis of single-item level, category level, and cross-regional store linkage is still insufficient. Future research can further combine more granular sales data and spatial relationship information to enhance the model's applicability and interpretability in different retail business scenarios.

7. Conclusion

This paper constructs a predictive research framework for ETL data fusion, addressing the needs of chain retail enterprises for multi-source data integration and store sales forecasting. This framework extracts, cleanses, transforms, and fuses sales records, store attributes, promotional information, time factors, and customer flow data in a unified manner, forming a more standardized, complete, and modelable structured data foundation. Furthermore, it combines feature construction and historical window modeling to characterize the business relationships and time dependencies in store sales changes. The results show that the proposed method effectively improves the accuracy and stability of store sales forecasting, demonstrating good error control and fitting capabilities in comparative experiments, indicating that multi-source data fusion plays a crucial supporting role in retail sales forecasting. This research not only provides chain retail enterprises with feasible data processing and modeling ideas for sales forecasting but also offers a methodological reference for the retail industry's shift from experience-driven management to data-driven decision-making. By more accurately grasping changes in store sales, enterprises can obtain a more reliable decision-making basis in areas such as inventory replenishment, merchandise allocation, promotional planning, personnel arrangement, and operational management, thereby improving resource allocation efficiency and market responsiveness.

Future research can further expand data sources and application scenarios based on the existing framework to enhance the applicability and promotional value of the model in complex retail environments. On the one hand, external factors such as weather changes, regional consumption levels, customer traffic in shopping districts, online and offline channel collaboration information, and changes in the competitive environment can be incorporated into the data fusion process, enabling the sales forecasting model to more comprehensively reflect the multidimensional influencing factors in real business scenarios. On the other hand, the research scope can be further expanded to more granular tasks such as single-product sales forecasting, category demand forecasting, regional store collaboration forecasting, and supply chain linkage optimization, thereby improving the model's support capabilities for different business levels. As the digital transformation of chain retail enterprises continues to deepen, ETL data fusion and intelligent forecasting methods will play a more important role in application areas such as business analysis, inventory management, precision marketing, and supply chain collaboration, providing continuous technical support for retail enterprises to achieve refined operations, intelligent management, and high-quality development.

References

[1] Y. Ensafi, S. H. Amin, G. Zhang, et al., "Time-series forecasting of seasonal items sales using machine learning-A comparative analysis," International Journal of Information Management Data Insights, vol. 2, no. 1, p. 100058, 2022.

[2] L. Eglite and I. Birzniece, "Retail sales forecasting using deep learning: Systematic literature review," Complex Systems Informatics and Modeling Quarterly, no. 30, pp. 53-62, 2022.

[3] C. Giri and Y. Chen, "Deep learning for demand forecasting in the fashion and apparel retail industry," Forecasting, vol. 4, no. 2, pp. 565-581, 2022.

[4] P. Ramos and J. M. Oliveira, "Robust sales forecasting using deep learning with static and dynamic covariates," Applied System Innovation, vol. 6, no. 5, p. 85, 2023.

[5] I. Vallés-Pérez, E. Soria-Olivas, M. Martínez-Sober, et al., "Approaching sales forecasting using recurrent neural networks and transformers," Expert Systems with Applications, vol. 201, p. 116993, 2022.

[6] R. S. Ahmed, M. Hasnain, M. H. Mahmood, et al., "Comparison of deep learning algorithms for retail sales forecasting," ICCK Transactions on Intelligent Systematics, vol. 1, no. 3, pp. 112-126, 2024.

[7] D. Eşki and T. Kaya, "Retail demand forecasting using temporal fusion transformer," in International Conference on Intelligent and Fuzzy Systems, Cham: Springer Nature Switzerland, 2024, pp. 165-170.

[8] R. Carbonneau, K. Laframboise, and R. Vahidov, "Application of machine learning techniques for supply chain demand forecasting," European Journal of Operational Research, vol. 184, no. 3, pp. 1140-1154, 2008.

[9] B. M. Pavlyshenko, "Machine-learning models for sales time series forecasting," Data, vol. 4, no. 1, p. 15, 2019.

[10] K. Bandara, C. Bergmeir, and S. Smyl, "Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach," Expert Systems with Applications, vol. 140, p. 112896, 2020.

[11] D. Salinas, V. Flunkert, J. Gasthaus, et al., "DeepAR: Probabilistic forecasting with autoregressive recurrent networks," International Journal of Forecasting, vol. 36, no. 3, pp. 1181-1191, 2020.

[12] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "M5 accuracy competition: Results, findings, and conclusions," International Journal of Forecasting, vol. 38, no. 4, pp. 1346-1364, 2022.

[13] P. Saha, N. Gudheniya, R. Mitra, et al., "Demand forecasting of a multinational retail company using deep learning frameworks," IFAC-PapersOnLine, vol. 55, no. 10, pp. 395-399, 2022.

[14] C. Challu, K. G. Olivares, B. N. Oreshkin, et al., "Nhits: Neural hierarchical interpolation for time series forecasting," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, no. 6, 2023, pp. 6989-6997.