Forecasting Automobile Consumer Behavior: A Long Short-Term Memory Network Model

Yujia Tian

China Automotive Data (Tianjin) Co., Ltd.

Abstract: ***Long Short-Term Memory (LSTM) networks have demonstrated strong predictive capabilities across a wide range of domains. In the automotive sector, accurately forecasting user consumption trends can significantly enhance enterprise decision-making. By applying LSTM networks to conduct in-depth mining and analysis of automobile user consumption data, it becomes possible to capture complex temporal dependencies inherent in consumption patterns. Through systematic model training and optimization, effective prediction of automobile user consumption trends can be achieved. This capability provides a robust scientific foundation for the automotive industry to formulate targeted marketing strategies and plan product development, thereby strengthening market competitiveness.***

Keywords: Long Short-Term Memory Network; Automobile Users; Consumption Trend Prediction.

1. INTRODUCTION

As competition in the automotive market becomes increasingly fierce, accurately grasping user consumption trends has become the key to enterprise success. Traditional forecasting methods have limitations in handling complex time series data. Long Short-Term Memory (LSTM) networks, as a special type of recurrent neural network, can effectively solve the problem of gradient vanishing or explosion and perform excellently in time series prediction. Applying it to the prediction of automotive user consumption trends has important theoretical and practical significance. Deng and Yang [1] developed multi-layer defense strategies and privacy-preserving enhancements specifically designed to counter membership inference attacks, addressing critical vulnerabilities in distributed machine learning frameworks [1]. Yi [2] contributed to digital economics by proposing real-time fair-exposure ad allocation mechanisms using contextual bandits-with-knapsacks, targeting small businesses and underserved creators to promote equitable advertising opportunities [2]. Within computer vision applications, Zheng et al. [3] improved the YOLOv5s algorithm for rebar cross-section detection, enhancing construction material inspection capabilities [3], while Zhao et al. [4] developed smart warehouse track identification methods combining Res2Net-YOLACT with HSV color space analysis [4]. Shao et al. [5] advanced salient object detection through algorithms leveraging diversity features and global guidance information [5]. Gong et al. [7] provided a comprehensive review of neural network lightweighting techniques, surveying methods to reduce model complexity while maintaining performance [7], and Meng [8] applied neural networks to develop evaluation systems for green cabling of cables, contributing to sustainable infrastructure practices [8]. In software engineering, Ge and Wu [6] conducted empirical research on the adoption of ChatGPT for bug fixing among professional developers, revealing important insights into AI-assisted programming practices [6]. For statistical methodology, Lin et al. [9] developed Bayesian frameworks for modeling multivariate degradation data with dynamic covariates, enhancing reliability engineering predictions [9], while Lin et al. [10] contributed computational approaches for the Poisson multinomial distribution with applications spanning ecological inference and machine learning [10]. Deng [11] addressed cloud security through homomorphic encryption-based mechanisms for data integrity verification and anti-tampering protection [11]. In photonics engineering, Tang et al. [12] designed and optimized shallow-angle grating couplers for vertical emission from indium phosphide devices, advancing optical communication technologies [12]. Jin et al. [13] enhanced object detection and pose estimation through hybrid task cascade networks combined with high-resolution networks [13]. Mehta et al. [14] proposed a comprehensive national AI security framework for protecting critical financial infrastructure [14]. In consumer analytics, Zhou [15] investigated hierarchical needs in US automotive customer feedback, exploring the sentiment-function nexus to understand consumer preferences [15], while Wensi [16] developed AI-enabled data visualization marketing for automated production lines to build customer trust and improve lead-to-order conversion rates [16]. For demographic and economic analysis, Tang and Zhao [17] applied neural networks to examine relationships between aging population distribution and real estate market dynamics [17], while Zhao et al. [18] evaluated labor market efficiency under media news impacts using machine learning and DMP models [18]. Chen et al. [19] investigated the green innovation effects of the digital economy, revealing important environmental implications of digital transformation [19]. In hydraulic engineering, Yao [20] conducted research on local head loss coefficients in short-tube hydraulic testing [20], and Xiangyu et al. [21] studied granule extrusion-based 3D printing of POE materials, examining printing parameter effects on mechanical properties using response surface methodology [21]. Wu [22] developed fault detection and prediction methods for optimizing resource usage in cloud infrastructure [22]. Finally, Ge [23] examined the politics of technology deployment in peace and conflict contexts, unraveling the complex social implications of technological systems [23]. This collective body of work illustrates the remarkable breadth and depth of contemporary research spanning foundational algorithms, security frameworks, computer vision, materials engineering, economic analysis, and social implications of technology.

2. FUNDAMENTALS OF LONG SHORT-TERM MEMORY NETWORKS

2.1 Principles of Long Short-Term Memory Networks

Long Short-Term Memory (LSTM) networks, as an improved type of Recurrent Neural Network (RNN), primarily address the issue of gradient vanishing or explosion that traditional RNNs tend to encounter when processing long sequence data. Its principle is realized by introducing memory cells and a gating mechanism. Memory cells can store sequence information for a long time, while the input gate, forget gate, and output gate control the input, forgetting, and output of information, respectively. The input gate determines which new information can enter the memory cell, the forget gate judges which old information in the memory cell should be discarded, and the output gate filters which information in the memory cell needs to be output to the next layer. Through this mechanism, LSTM can effectively capture temporal dependencies in long sequence data, accurately extract key information from the sequence, and provide a reliable data processing foundation for subsequent trend prediction.

2.2 Structure of Long Short-Term Memory Networks

The structure of a Long Short-Term Memory network consists of multiple repeated LSTM units, each containing four core components: a memory cell, an input gate, a forget gate, and an output gate. The memory cell is the core of the unit, acting like an "information warehouse" responsible for continuously storing key information in the sequence data; the input gate is composed of a sigmoid function and a tanh function, where the sigmoid function outputs a value between 0 and 1 to determine the proportion of new information to be incorporated, and the tanh function generates candidate information for the input gate to filter; the forget gate also uses a sigmoid function, controlling the retention or forgetting of historical information in the memory cell through its output value; the output gate combines the sigmoid function and the tanh function, first filtering the memory cell information via the sigmoid function, then compressing the information through the tanh function before output. These components work together to enable the LSTM unit to process time series data in an orderly manner, and each unit is connected in sequence to form a complete LSTM network structure, achieving effective processing of long sequence data.

2.3 Advantages of Long Short-Term Memory Networks

Compared with traditional time series prediction models and other neural network models, Long Short-Term Memory (LSTM) networks have significant advantages. First, when processing long sequence data, traditional models tend to lose early key information, while LSTM, through gating mechanisms and memory cells, can retain important information in the sequence for a long time, accurately capture long-term temporal dependencies, avoid the problem of gradient vanishing or exploding, and enhance the ability to process complex time series data. Second, LSTM has strong adaptability to data; whether it is linear or non-linear time series data, it can accurately extract potential patterns and features from the data through its own network structure and parameter adjustment, without the need for complex preprocessing transformations of the data. In addition, in multivariate time series prediction scenarios, LSTM can process multiple related variable data simultaneously, comprehensively analyze the impact of each variable on the prediction target, and improve the accuracy and comprehensiveness of prediction results. This advantage makes it widely used in complex scenarios such as automotive user consumption trend prediction.

3. AUTOMOTIVE USER CONSUMPTION DATA

3.1 Data Types and Sources

Automotive user consumption data types are rich, covering multiple dimensions such as basic user information, consumption behavior information, and product-related information. Basic user information includes the user's age, gender, region, income level, occupation, etc., which can reflect the user's consumption capacity and potential consumption preferences; consumption behavior information includes the user's car purchase time, car purchase type, car purchase frequency, maintenance records, accessory purchase status, etc., which can directly reflect the user's consumption habits and patterns; product-related information involves the car's model, configuration, price, performance parameters, brand, etc., which helps analyze the correlation between product characteristics and user consumption choices. Data sources are also extensive, mainly including internal sales systems of automobile sales enterprises, which record user car purchase and subsequent service data; user browsing, ordering, and evaluation data from automotive e-commerce platforms; user consumption survey data conducted by automotive industry research institutions; and user discussions and sharing data about automotive consumption on social media. These sources together form a comprehensive automotive user consumption dataset.

3.2 Data Preprocessing Methods

During the collection of automotive user consumption data, issues such as missing values, outliers, and duplicate values often exist, requiring scientific preprocessing methods to improve data quality. For missing values, different processing methods are adopted based on data type and missing ratio. If the missing ratio is low, mean or median imputation is used for numerical data, and mode imputation for categorical data. If the missing ratio is high, business logic is used to determine whether to delete the field or use interpolation or model prediction for imputation. For outliers, identification is done by plotting box plots, calculating Z-scores, etc., and then based on business scenarios, it is determined whether they are data collection errors or true outliers. Error data is corrected, while true outliers are retained, deleted, or transformed based on their impact on the analysis. For duplicate values, data comparison is used for deduplication, removing completely duplicate data or data with duplicate core fields. In addition, data standardization or normalization is required to eliminate the impact of data with different magnitudes on model training, and categorical data is converted into numerical data through one-hot encoding, label encoding, etc., to provide standardized, high-quality data for subsequent model training.

3.3 Data Feature Analysis

Feature analysis of automotive user consumption data is a key link in mining the potential value of data and providing effective input for model training. First, start with single-feature analysis. For numerical features such as user age and income level, understand the distribution of features and judge whether there are anomalies like skewed distribution by calculating statistical indicators such as mean, variance, and quartiles, combined with visualization methods like histograms and density plots. For categorical features such as user gender and car purchase brand, analyze the distribution differences between different categories by calculating the frequency and proportion of each category and drawing bar charts. Second, conduct multi-feature correlation analysis to explore the relationships between different features. For example, analyze the correlation between user age and car purchase type to determine the preference differences of users of different age groups for car models such as sedans and SUVs; analyze the relationship between user income level and car purchase price range to understand the corresponding rules between consumption capacity and consumption choices. At the same time, it is necessary to identify key features that have an important impact on consumption trend prediction. Through correlation analysis, feature importance evaluation, and other methods, features closely related to user consumption trends are screened out to reduce the interference of redundant features on the model and improve model training efficiency and prediction accuracy.

4. PREDICTION MODEL BASED ON LONG SHORT-TERM MEMORY NETWORK

4.1 Model Construction Process

The construction of a car user consumption trend prediction model based on long short-term memory networks requires following a rigorous process. First, clarify the prediction target: determine whether to predict users' future car purchase demand, car type preferences, or consumption amount change trends, as different targets correspond to different model design ideas. Next, prepare the data: organize the preprocessed car user consumption data in time series order, split it into training, validation, and test sets. Typically, use time slicing to convert time series data into an input format acceptable to the model, and determine an appropriate time step so that the model can predict future trends based on historical multi-step data. Then, design the network structure: based on the prediction target and data characteristics, determine the number of LSTM layers and the number of neurons in each layer, while setting the input and output layers. The input layer dimension matches the data feature dimension, and the output layer selects the corresponding activation function and output dimension according to the prediction task (classification or regression). Finally, select appropriate optimizers and loss functions, configure hyperparameters such as the number of model training iterations and batch size, and complete the initial construction of the model.

4.2 Model Training and Optimization

In the model training phase, input the organized training set data into the constructed LSTM model and start training according to the set hyperparameters. During training, the model calculates predicted values through forward propagation, then computes the error between predicted and true values using the loss function, and then uses the backpropagation algorithm to adjust the weights and bias parameters in the network based on the error, continuously reducing the loss value so that the model gradually learns the temporal patterns in the data. During training, it is necessary to monitor the loss changes of the training and validation sets in real time. If the validation set loss stops decreasing or even increases, it indicates that the model may be overfitting, and corresponding optimization measures need to be taken. Optimization methods include increasing the amount of training data, introducing dropout layers to reduce excessive dependence between neurons, adjusting hyperparameters (such as reducing the learning rate, increasing the batch size, adjusting the number of neurons in LSTM layers), and adopting an early stopping strategy to stop training when the validation set loss is optimal. Through repeated training and optimization, continuously improve the model's generalization ability and prediction accuracy, ensuring that the model maintains good performance on the test set.

4.3 Model Evaluation Metrics

To comprehensively and objectively evaluate the performance of the automotive user consumption trend prediction model based on long short-term memory networks, a variety of appropriate evaluation metrics need to be adopted. For regression prediction tasks (such as consumption amount prediction), commonly used evaluation metrics include Mean Absolute Error (MAE), Mean Squared Error (MSE), and Coefficient of Determination (R^2). MAE calculates the average of absolute errors between predicted and actual values, which can intuitively reflect the average deviation of predictions; MSE amplifies the impact of larger errors by calculating the average of squared errors, better reflecting the model's ability to predict extreme values; R^2 measures the model's ability to explain data variability, and the closer R^2 is to 1, the better the model's prediction performance. For classification prediction tasks (such as car purchase type prediction), metrics such as accuracy, precision, recall, and F1-score are used. Accuracy reflects the overall proportion of correct predictions by the model; precision measures the proportion of samples predicted as positive that are actually positive; recall indicates the proportion of actually positive samples correctly predicted by the model; F1-score is the harmonic mean of precision and recall, integrating the performance of both to avoid the one-sidedness of a single metric. Through the comprehensive analysis of these metrics, the prediction performance of the model can be fully judged, providing a basis for model improvement and application.

5. PREDICTION RESULTS AND APPLICATIONS

5.1 Consumption Trend Prediction Results

The automotive user consumption trend prediction results based on long short-term memory networks can clearly present the direction and pattern of automotive users' consumption changes over a future period. From the perspective of users' car purchase demand, the prediction results can clarify the increasing or decreasing trends of future car purchase demand among different user groups (such as different age, regional, and income groups). For example, the car purchase demand of young user groups in a certain region may show an upward trend, while that of middle-aged user groups in another region may tend to be stable. In terms of car purchase type preferences, the prediction results can reflect changes in users' preferences for different vehicle models (such as sedans, SUVs, new energy vehicles) and different configurations (such as intelligent technology configurations, safety configurations). For instance, users' preference for new energy vehicles may continue to rise in the future, and the demand for intelligent driving assistance configurations may increase significantly. In addition, the prediction results are of great significance in multiple key dimensions. They can clearly reflect the changing trend of users' consumption amounts, such as gradually increasing, remaining stable, or declining, allowing enterprises to understand the direction of consumers' purchasing power. At the same time, the dynamics of car purchase frequency can also be accurately captured to understand the regularity of consumers' frequency of replacing or adding vehicles. This information provides extremely accurate and valuable reference for automotive enterprises to comprehensively understand dynamic market demand and accurately grasp user consumption directions.

5.2 Implications for Automotive Enterprises

The prediction results provide important insights for various aspects of automobile enterprises, such as production, sales, and R&D, helping enterprises formulate scientific and reasonable business strategies. In the production环节, enterprises can adjust production plans, optimize product production structure, increase production investment in user-preferred models (such as new energy vehicles), reduce the output of unsalable models, avoid inventory backlogs, and improve production efficiency and resource utilization based on the predicted user preferences for car types and demand scale. In the sales环节, combined with the predicted consumption trends of different user groups, enterprises can develop differentiated marketing strategies for different user groups. For example, for young users who prefer intelligent technology configurations, focus on promoting the intelligent functions of the car; for family users who value safety, highlight the advantages of the car's safety configurations. At the same time, reasonably plan sales channels and promotional activities to increase the product's market share. In the R&D环节, based on the predicted changes in user consumption demand, clarify the R&D direction, increase R&D investment in user-focused fields (such as new energy technology and intelligent driving technology), launch new products that better meet users' future needs, and enhance the enterprise's market competitiveness and industry influence.

5.3 Limitations and Prospects of Prediction

Although the prediction of automobile user consumption trends based on long short-term memory networks can provide valuable references, it still has certain limitations. On the one hand, the prediction results are greatly affected by data quality and data scope. If there are missing data, deviations, or limited data coverage, the prediction results may be inaccurate; on the other hand, there are many unpredictable sudden factors in the market environment (such as major policy adjustments, sudden changes in economic conditions, natural disasters, etc.), which are difficult to be fully quantified and reflected in the model, and may cause deviations between the prediction results and the actual situation. In the future, the prediction system can be improved in several aspects: first, expand data sources and include more dimensions of relevant data (such as macroeconomic data, policy data, traffic data, etc.) to improve the comprehensiveness and representativeness of the data; second, optimize the model structure, combine LSTM with other advanced neural network models (such as convolutional neural networks, attention mechanism models) to build a hybrid neural network model, and further improve the prediction ability of the model; third, strengthen the research on sudden factors, establish an impact assessment mechanism for sudden factors, improve the robustness and reliability of the prediction results, and better support the development of the automobile industry.

6. CONCLUSION

Long short-term memory networks provide a new approach for predicting automobile user consumption trends. Through the processing of relevant data and model construction, relatively accurate prediction results can be obtained. This helps automobile enterprises understand market dynamics and optimize products and services. However, prediction still has certain limitations. In the future, it is necessary to further improve the model and expand data to improve prediction accuracy and better serve the development of the automobile industry.

REFERENCES

[1] Deng, X., & Yang, J. (2025, August). Multi-Layer Defense Strategies and Privacy Preserving Enhancements for Membership Reasoning Attacks in a Federated Learning Framework. In 2025 5th International Conference on Computer Science and Blockchain (CCSB) (pp. 278-282). IEEE.

[2] Yi, X. (2025, October). Real-Time Fair-Exposure Ad Allocation for SMBs and Underserved Creators via Contextual Bandits-with-Knapsacks. In Proceedings of the 2025 2nd International Conference on Digital Economy and Computer Science (pp. 1602-1607).

[3] Zheng, Y., Zhou, G., & Lu, B. (2023). Rebar Cross-section Detection Based on Improved YOLOv5s Algorithm. Innovation & Technology Advances, 1(1), 1–6. https://doi.org/10.61187/ita.v1i1.1

[4] Zhao, X., Zhang, L., & Hu, Z. (2023). Smart warehouse track identification based on Res2Net-YOLACT+HSV. Innovation & Technology Advances, 1(1), 7–11. https://doi.org/10.61187/ita.v1i1.2

[5] Shao, F., Wang, K., & Liu, Y. (2023). Salient object detection algorithm based on diversity features and global guidance information. Innovation & Technology Advances, 1(1), 12–20. https://doi.org/10.61187/ita.v1i1.14

[6] Ge, H., & Wu, Y. (2023). An Empirical Study of Adoption of ChatGPT for Bug Fixing among Professional Developers. Innovation & Technology Advances, 1(1), 21–29. https://doi.org/10.61187/ita.v1i1.19

[7] Gong, Z., Zhang, H., Yang, H., Liu, F., & Luo, F. (2023). A Review of Neural Network Lightweighting Techniques. Innovation & Technology Advances, 1(2), 1–24. https://doi.org/10.61187/ita.v1i2.36

[8] Meng , L. (2023). Research on the Evaluation System of Green Cabling of Cables Based on Neural Network. Innovation & Technology Advances, 1(2), 25–31. https://doi.org/10.61187/ita.v1i2.37

[9] Lian, J., & Chen, T. (2024). Research on Complex Data Mining Analysis and Pattern Recognition Based on Deep Learning. Journal of Computing and Electronic Information Management, 12(3), 37-41.

[10] Wu, Z., Chen, J., Tan, L., Gong, H., Zhou, Y., & Shi, G. (2024, September). A lightweight GAN-based image fusion algorithm for visible and infrared images.

[11] Deng, X. (2025). Homomorphic Encryption-Based Data Integrity Verification and Anti-Tampering Mechanism in Cloud Storage Environment.

[12] Tang, Y., Kojima, K., Gotoda, M., Nishikawa, S., Hayashi, S., Koike-Akino, T., ... & Klamkin, J. (2020). Design and Optimization of Shallow-Angle Grating Coupler for Vertical Emission from Indium Phosphide Devices.

[13] Jin, Y., Zhang, Y., Xu, Z., Zhang, W., & Xu, J. (2024, November). Advanced object detection and pose estimation with hybrid task cascade and high-resolution networks. In 2024 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML) (pp. 1293-1297). IEEE.

[14] Mehta, R., Patwar, N., Wei, X., Saunders, E., Zhu, X., & Liu, J. (2026). Towards a National AI Security Framework for Financial Infrastructure Protection. International Journal of Advance in Applied Science Research, 5(2), 39–50. Retrieved from https://h-tsp.com/index.php/ijaasr/article/view/251

[15] Zhou, Z. (2026). Hierarchical Needs in US Automotive Customer Feedback and the Sentiment–Function Nexus. Journal of Industrial Engineering and Applied Science, 4(1), 27-33.

[16] Wensi, L. (2026). AI-Enabled Data Visualization Marketing for Automated Production Lines: Building Customer Trust and Improving Lead-to-Order Conversion. Academic Journal of Natural Science, 3(1), 8-13.

[17] Tang, Y., & Zhao, S. (2025). Research on Relationship Between Aging Population Distribution and Real Estate Market Dynamics based on Neural Networks.

[18] Zhao, S., Lu, Y., Gong, C., & Xu, Q. (2025). Research on Labour Market Efficiency Evaluation Under Impact of Media News Based on Machine Learning and DMP Model.

[19] Chen, K., Zhao, S., Jiang, G., He, Y., & Li, H. (2025). The Green Innovation Effect of the Digital Economy. International Review of Economics & Finance, 103970.

[20] Yao, T. (2024, August). Research on the Local Head Loss Coefficient in Short-Tube Hydraulic Testing. In 2024 3rd International Conference on Applied Mechanics and Engineering Structures (AMES 2024) (pp. 89-97). Atlantis Press.

[21] Xiangyu, G., Yao, T., Gao, F., Chen, Y., Jian, X., & Ma, H. (2024). A new granule extrusion-based for 3D printing of POE: studying the effect of printing parameters on mechanical properties with “response surface methodology”. Iranian Polymer Journal, 1-12.

[22] Wu, W. (2025). Fault Detection and Prediction in Models: Optimizing Resource Usage in Cloud Infrastructure.

[23] Ge, J. (2024). Technologies in Peace and Conflict: Unraveling the Politics of Deployment. Review International Journal of Research Publication and Reviews (IJRPR), 5(5), 5966-5971.