






Comparison of the Performance of a Three-Layer LSTM Model with Classical Algorithms in Sentiment Analysis of Persian Texts in the Automotive Industry

Lida Shojaei Barjoui¹, Ghasem Bakhshandeh^{1*} Mahmood Daniali Deh Houz²

¹ Department of Management, Ahv.C., Islamic Azad University, Ahvaz, Iran

² Department of Accounting, Iz.C., Islamic Azad University, Izeh, Iran

* Corresponding author email address: gh.bakhshandeh@iau.ac.ir

Received: 2026-03-07

Revised: 2026-05-15

Accepted: 2026-05-16

Initial Publish: 2026-05-19

Final Publish: 2027-01-01

Abstract

In the field of business, sentiment analysis is considered an effective tool for monitoring customer feedback, improving decision-making processes, and enhancing product quality. This study aimed to evaluate the efficiency of different sentiment analysis methods by comparing the performance of a deep learning model based on a three-layer Long Short-Term Memory (LSTM) recurrent neural network with four classical machine learning algorithms, including Linear Support Vector Machine (Linear SVM), Naive Bayes, Multilayer Perceptron (MLP), and Decision Tree. The research data consisted of real user reviews regarding domestic automobiles, which were utilized for model training and evaluation after standard preprocessing procedures. The findings indicated that among the classical algorithms, the Linear SVM model achieved the best performance with an accuracy of 0.90, followed by the MLP model with an accuracy of 0.88, whereas the Naive Bayes and Decision Tree models demonstrated the weakest performance. Compared with these methods, the three-layer LSTM model significantly outperformed the other approaches by achieving an accuracy of 95.4%. This superiority can be attributed to the capability of the LSTM architecture to learn temporal dependencies, capture the sequential structure of sentences, and extract deeper semantic relationships.

Keywords: *Sentiment analysis, LSTM, classical machine learning, automotive industry, Persian texts*

How to cite this article:

Shojaei Barjoui, L., Bakhshandeh, G., & Daniali Deh Houz, M. (2027). Comparison of the Performance of a Three-Layer LSTM Model with Classical Algorithms in Sentiment Analysis of Persian Texts in the Automotive Industry. *Management Strategies and Engineering Sciences*, 9(1), 1-11.

1. Introduction

The rapid expansion of digital communication technologies and online social platforms has generated unprecedented volumes of textual data that reflect public opinions, emotions, preferences, and behavioral intentions. In contemporary business environments, organizations increasingly rely on sentiment analysis techniques to extract actionable insights from customer-generated content, including online reviews, comments, social media posts, and product evaluations. Sentiment analysis, as a major branch of natural language processing (NLP), enables businesses to identify emotional orientations embedded within textual data and thereby improve strategic decision-making, customer relationship management, marketing intelligence,

and product development [1, 2]. The growing availability of unstructured textual data has consequently transformed sentiment analysis into one of the most influential applications of artificial intelligence and machine learning in business analytics [3, 4].

In recent years, the integration of artificial intelligence into business intelligence systems has significantly enhanced the capacity of organizations to process large-scale customer feedback in real time. Traditional survey-based methods are increasingly being replaced by automated analytical systems capable of continuously monitoring public sentiment across multiple digital channels [5, 6]. Sentiment analysis has become particularly important in industries characterized by high customer involvement and intense market competition, such as finance, retail,



transportation, and the automotive sector [7, 8]. In such environments, organizations seek to understand consumer attitudes toward products and services in order to improve competitiveness and optimize operational strategies.

The automotive industry represents one of the most dynamic domains for sentiment analysis applications because customer perceptions directly influence purchasing behavior, brand loyalty, and market reputation. Online automotive platforms and digital review systems provide customers with opportunities to express opinions regarding technical performance, safety, comfort, fuel efficiency, design quality, after-sales services, and driving experience. These large volumes of textual feedback create valuable datasets for analyzing customer satisfaction and predicting consumer trends [9, 10]. Moreover, the emergence of intelligent and automated vehicles has intensified research attention toward emotion-aware transportation systems and user-centered automotive technologies [11, 12]. Studies on emotional interactions within automated vehicles demonstrate that understanding passenger sentiment can contribute to safer and more adaptive transportation systems [13, 14].

The increasing complexity of textual data has motivated researchers to develop advanced computational methods for sentiment classification. Early sentiment analysis systems primarily relied on lexicon-based approaches and traditional machine learning algorithms such as Naive Bayes, Support Vector Machines (SVM), Decision Trees, and Multilayer Perceptrons (MLP). These models typically use handcrafted features and statistical representations such as Bag-of-Words or Term Frequency–Inverse Document Frequency (TF-IDF) vectors [15, 16]. Although classical machine learning algorithms have demonstrated acceptable performance in many classification tasks, they often struggle to capture contextual semantics, sequential dependencies, and complex linguistic structures in natural language [2, 17].

The limitations of traditional approaches have accelerated the adoption of deep learning architectures for sentiment analysis. Deep learning methods automatically learn hierarchical semantic representations from textual data and have demonstrated superior performance in numerous NLP tasks [3, 4]. Among these architectures, recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks have attracted considerable attention because of their capability to process sequential information and preserve long-term dependencies within texts [18, 19]. LSTM models are particularly effective for sentiment analysis because they can capture contextual relationships

between words and identify subtle emotional patterns embedded in sentences.

The growing effectiveness of LSTM architectures has led to their widespread application across multiple domains. In financial forecasting and stock market analysis, sentiment-aware LSTM models have been utilized to improve predictive accuracy by incorporating textual information from news articles and social media [19, 20]. Similarly, in digital marketing and consumer behavior prediction, deep learning models have demonstrated substantial capabilities in identifying purchasing intentions and recommendation patterns [21, 22]. In cybersecurity research, AI-driven sentiment and text analysis systems have also been employed to detect harmful online behaviors and cyber threats [16, 23]. These developments indicate that deep learning has evolved into a powerful analytical paradigm for extracting meaningful insights from complex textual environments.

Recent advances in large language models (LLMs) and transformer-based architectures have further expanded the capabilities of sentiment analysis systems. Models such as BERT and ChatGPT have demonstrated exceptional performance in contextual language understanding and semantic representation learning [4, 24]. Researchers have increasingly explored the integration of these advanced language models into business intelligence and customer analytics systems [5, 6]. Despite these advances, LSTM-based architectures remain highly relevant due to their computational efficiency, interpretability, and effectiveness in domain-specific sentiment analysis tasks, particularly in languages with limited annotated resources.

One of the major challenges in sentiment analysis research concerns aspect-level sentiment extraction. Unlike document-level classification, aspect-level sentiment analysis aims to identify specific product attributes or service components discussed within customer reviews and determine the sentiment associated with each aspect [25, 26]. In the automotive sector, customers often express different sentiments simultaneously regarding vehicle performance, fuel consumption, safety, comfort, and multimedia systems. Therefore, advanced sentiment analysis models must effectively capture these multidimensional emotional expressions. Graph-based deep learning methods and context-aware architectures have been proposed to address these challenges, particularly for aspect-specific sentiment representation [25, 26].

Another significant challenge involves multilingual sentiment analysis and low-resource languages. While

English-language sentiment analysis has achieved substantial progress, sentiment classification in languages such as Persian remains comparatively underexplored due to limited datasets, linguistic complexity, and insufficient NLP resources [21, 27]. Persian texts present additional challenges associated with morphological variation, colloquial expressions, informal writing styles, and orthographic inconsistencies. Consequently, conventional machine learning approaches may experience reduced performance when applied to Persian sentiment analysis tasks.

Recent studies have increasingly emphasized the effectiveness of deep learning approaches for Persian-language sentiment analysis. Research by Kazance et al. demonstrated that deep learning architectures significantly improve the classification accuracy of Persian user reviews compared with conventional machine learning algorithms [27]. Similarly, the integration of semantic embeddings and recurrent neural networks has shown promising results in extracting contextual information from Persian textual data [21]. These findings suggest that advanced neural architectures can effectively overcome some of the linguistic limitations associated with Persian NLP tasks.

In parallel with advances in sentiment analysis, machine learning applications have expanded rapidly in industrial and engineering domains. Predictive maintenance systems, intelligent monitoring frameworks, and AI-based optimization models increasingly rely on deep learning methods for pattern recognition and forecasting tasks [28, 29]. The successful application of LSTM networks in bearing life estimation and maintenance prediction illustrates the capability of recurrent architectures to model temporal dependencies in sequential datasets [29]. These achievements reinforce the suitability of LSTM models for complex analytical tasks involving time-dependent and context-sensitive information.

The importance of sentiment analysis has also intensified in the context of geopolitical and economic uncertainty. Public sentiment extracted from digital platforms can significantly influence financial markets, consumer confidence, and energy economics [30, 31]. Studies examining public reactions to geopolitical crises and fuel price fluctuations have demonstrated that machine learning-based sentiment analysis can provide valuable insights into market behavior and consumer perceptions [32, 33]. These developments are particularly relevant for the automotive industry, where fluctuations in fuel prices, environmental concerns, and technological innovation strongly affect

customer attitudes toward vehicles and transportation systems.

Despite extensive progress in sentiment analysis research, several limitations remain evident in existing studies. Many investigations primarily focus on English-language datasets or social media corpora, whereas fewer studies address Persian customer reviews within industrial contexts. Furthermore, comparative evaluations between deep learning models and classical machine learning algorithms remain limited, particularly in domain-specific datasets related to the automotive industry. Although transformer-based models have recently attracted substantial attention, there remains a need for systematic assessments of LSTM architectures relative to traditional algorithms such as Linear SVM, Naive Bayes, MLP, and Decision Trees in Persian sentiment analysis applications.

Additionally, the majority of existing research emphasizes overall sentiment classification accuracy without sufficiently examining the impact of class imbalance, preprocessing strategies, and feature representation techniques on model performance. Classical machine learning algorithms commonly rely on TF-IDF representations, whereas deep learning models typically employ distributed semantic embeddings such as FastText or Word2Vec. The comparative effectiveness of these approaches in Persian automotive review analysis remains insufficiently explored. Understanding the relative strengths and limitations of these models is essential for developing robust business intelligence systems capable of accurately analyzing customer feedback in real-world industrial environments.

Given the strategic importance of customer sentiment analysis in the automotive industry and the growing significance of Persian-language NLP applications, comprehensive comparative investigations are necessary to identify the most effective analytical frameworks for sentiment classification. Evaluating the performance of deep learning architectures alongside classical machine learning algorithms can contribute to both theoretical advancements and practical implementations in intelligent customer analytics systems. Therefore, the present study aims to compare the performance of a three-layer LSTM deep learning model with classical machine learning algorithms, including Linear SVM, Naive Bayes, MLP, and Decision Tree, in the sentiment analysis of Persian automotive review texts.

2. Methodology

This study was conducted as an extension of the authors' previous work in the field of sentiment analysis of Persian texts in the automotive industry. In the earlier study, the primary focus was on the design and evaluation of different Long Short-Term Memory (LSTM) recurrent neural network architectures. In the present study, the emphasis is placed on comparing the three-layer LSTM model with classical models and other algorithms. In the first stage, a dataset consisting of user reviews regarding automobiles available in the Iranian market was collected. These data were extracted from several specialized automotive websites, including Khodrobank, Car, Hamrah Mechanic, Asbe Bokhar, Pedal, Mashin, Bama, and several other platforms. After collection and initial cleaning, user comments were categorized into three labels: positive, negative, and neutral. In total, 4,348 comments were collected and utilized in this study for model training and evaluation. During the preprocessing stage, text-cleaning operations, including the removal of punctuation marks, elimination of unnecessary characters, normalization of Persian letters, and conversion of texts into an appropriate processing format, were performed. Subsequently, to convert text into numerical representations, the Term Frequency–Inverse Document Frequency (TF-IDF) method was applied in classical models, whereas FastText word embeddings were employed in neural network-based models.

In this study, several categories of models were examined for sentiment analysis. In the deep learning section, a recurrent neural network model based on a three-layer LSTM architecture was employed as the primary model. To compare performance, several classical machine learning

algorithms were also implemented, including Naive Bayes, Multilayer Perceptron (MLP), Decision Tree, and Linear Support Vector Machine (Linear SVM), using TF-IDF features. To ensure accurate model evaluation, the 10-fold cross-validation method was utilized. In this approach, the dataset was divided into ten subsets, where, in each iteration, one subset was used for testing and the remaining subsets were used for training. This process was repeated ten times, and the average results were reported as the final model performance.

To evaluate model performance, the metrics of Accuracy, Precision, Recall, and F1-Score were employed. These metrics were calculated based on the values of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN), thereby providing a comprehensive representation of model performance in data classification. Finally, the results of the different models were compared, and the effects of data balancing methods and deep learning architectures on the accuracy of sentiment analysis in Persian texts were examined.

3. Findings and Results

In this study, a dataset consisting of 4,348 textual user reviews in the automotive industry was analyzed. These reviews were labeled into three sentiment categories: positive, negative, and neutral. Among the collected data, 2,159 reviews were positive, 1,642 were negative, and 547 were neutral, indicating the presence of an initial class imbalance in the label distribution. For qualitative analysis, various aspects related to automobiles were extracted from the review texts and are presented in Table 1. The frequency of first-level qualitative topics is also reported in Table 2.

Table 1. Extracted Aspects from Automobile Review Texts

First-Level Category	Second-Level Components
Technical Performance	Gearbox, differential, injector, bearing, turbocharger, crankshaft, belt, valve, cylinder, spark plug, combustion, compressor, engine, acceleration, speed, power, torque, traction, engine RPM, handling, steering, stability, vibration
Fuel Consumption and Depreciation	Fuel consumption, fuel-efficient, high fuel consumption, gasoline, gas, depreciation, low depreciation, repair cost, oil, oil leakage, cylinder head gasket, radiator, filter, exhaust
Dimensions and Size	Body, chassis, cabin, fuel tank, tire, wheel, length, width, height, distance, interior cabin, trunk, front space, rear space, rear cabin
Design and Appearance	Appearance, design, style, beauty, color, metallic, hood, bumper, fender, headlight, grille, front view, rear view, sport design, rim, logo
Cabin Comfort Features	Cruise control, stability control, handbrake, ABS, airbag, seat belt, seat, seat heater, seat cooler, ventilation, air conditioner, heater, mirror, power window, sunroof, dashboard, console, cup holder
Braking System	Brake, brake pad, disc, oil seal, anti-lock braking system, ABS
Safety	Safety, secure, rollover, reinforced body, seat belt, airbag, chassis
After-Sales Services	Roadside assistance, dealership, repair shop, factory, build quality, assembly, parts quality, warranty, golden card, after-sales service, Saipa expert, contract, cancellation
Multimedia System	Audio, sound, speaker, sound system, monitor, touch screen, radio, media player, rear sensor, rear camera, multimedia
Multiplex System	Multiplex, alarm, adapter, battery, electrical connection, error, sensor, gauge, wiring, connector

Table 2. Frequency of Qualitative Topics (First-Level Categories) in Customer Reviews

No.	Category (First Level)	Frequency (Number of Reviews)	No.	Category (First Level)	Frequency (Number of Reviews)
1	Technical Performance	1,558	8	Braking System	478
2	Dimensions and Size	1,320	9	Multiplex System	375
3	Multimedia	886	10	Steering	315
4	Technical Specifications	823	11	Driver Assistance System	307
5	Cabin Comfort Features	793	12	Auxiliary Equipment	252
6	Appearance Specifications	605	13	Lighting System	214
7	Services	493	14	Cabin Safety	134
			15	Mirrors and Windows	134

To evaluate model performance, the 10-fold cross-validation method was employed. In this approach, the dataset was divided into ten subsets, and in each iteration, nine subsets were used for training while one subset was

used for testing. The average results of the ten iterations were reported as the final performance of the models.

The three-layer LSTM model was selected as the baseline model, and the results of its training and evaluation are illustrated in Figure 1.

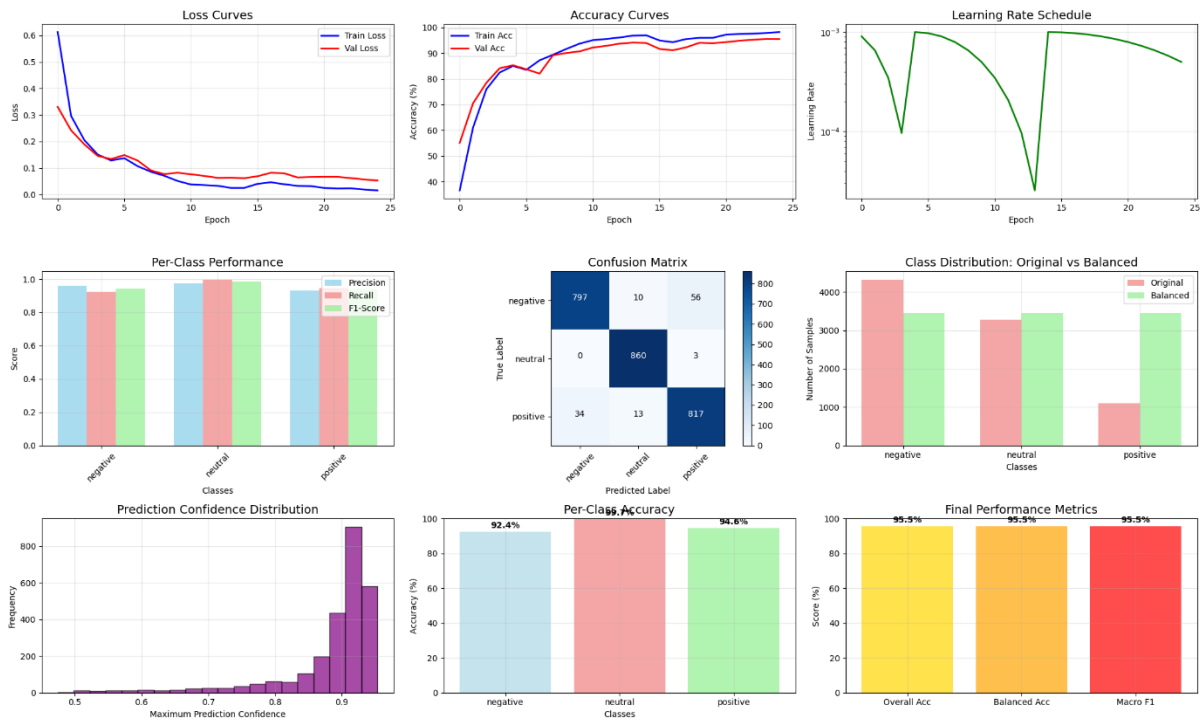


Figure 1. Training and Evaluation Diagrams of the Three-Layer LSTM Model

In Figure 1, the loss and accuracy curves initially demonstrate the learning process of the model over 25 training epochs. As observed, the consistent reduction in loss and the stable increase in accuracy indicate proper model convergence and the absence of severe overfitting. Furthermore, the learning rate adjustment curve illustrates how the learning rate changed over time, contributing to stability during the training process.

Additionally, the confusion matrix demonstrates the model’s performance in classifying the three labels of

“positive,” “negative,” and “neutral.” This matrix clearly illustrates the model’s success in correctly identifying each class as well as the classification errors. Most predictions are concentrated along the main diagonal, indicating that the model achieved satisfactory classification performance.

The model’s performance for each class was also evaluated based on the metrics of Precision, Recall, and F1-Score. The findings reveal that the LSTM model maintained an appropriate balance among the different evaluation

metrics and demonstrated acceptable performance across all three sentiment classes.

Moreover, the prediction confidence distribution presents the confidence level of the model in its outputs. The concentration of a substantial number of predictions at high-confidence levels suggests that the model performed classification tasks with relatively high certainty.

The class distributions before and after balancing are also presented in the figure, indicating that the original dataset suffered from significant imbalance, which was corrected in the balanced version. This issue had a substantial effect on improving model performance. Finally, the model's overall

evaluation metrics, including Accuracy, Balanced Accuracy, and Macro F1-Score, all reached approximately 95.5%, indicating the strong performance of the baseline model in sentiment classification.

In this section, the performance of the LSTM model was compared with four classical machine learning algorithms, including Linear SVM, Naive Bayes, MLP, and Decision Tree. The classification reports for each classical model are presented in Tables 3 to 6. These tables include the confusion matrices for the positive, negative, and neutral classes and illustrate the distribution of correct and incorrect predictions generated by each model.

Table 3. Classification Report of the Linear SVM Model (TF-IDF)

Actual \ Predicted	Positive	Negative	Neutral
Positive	1,870	180	109
Negative	210	1,360	72
Neutral	75	55	417

Table 4. Classification Report of the Naive Bayes Model (TF-IDF)

Actual \ Predicted	Positive	Negative	Neutral
Positive	1,760	260	139
Negative	285	1,280	77
Neutral	95	70	382

Table 5. Classification Report of the MLP Model (TF-IDF)

Actual \ Predicted	Positive	Negative	Neutral
Positive	1,690	320	149
Negative	330	1,180	132
Neutral	120	90	337

Table 6. Classification Report of the Decision Tree Model (TF-IDF)

Actual \ Predicted	Positive	Negative	Neutral
Positive	1,650	350	159
Negative	360	1,150	132
Neutral	140	95	312

The final evaluation results of the models using the 10-fold cross-validation method are reported in Table 7. According to this table, among the classical models, the Linear SVM model achieved the best performance with an

accuracy of 0.88. This was followed by the MLP model, whereas the Naive Bayes and Decision Tree models demonstrated comparatively weaker performance.

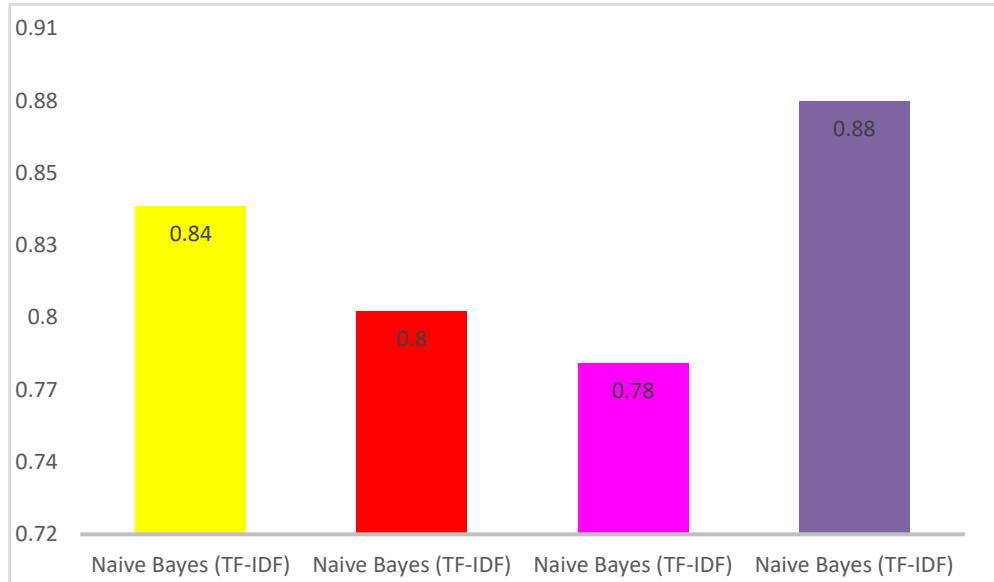
Table 7. Model Evaluation Results Using 10-Fold Cross-Validation and TF-IDF Features

Model	Accuracy	Precision (Weighted)	Recall (Weighted)	F1-Score (Weighted)
Naive Bayes (TF-IDF)	0.84	0.85	0.84	0.84
MLP (k = 5, TF-IDF)	0.80	0.81	0.80	0.80
Decision Tree (TF-IDF)	0.78	0.79	0.78	0.78
Linear SVM (TF-IDF)	0.88	0.89	0.88	0.88

This performance difference is clearly observable in Figure 2, which presents the comparative accuracy chart of the evaluated models. As illustrated, the LSTM model

demonstrated a substantial performance advantage compared with all TF-IDF-based models.

Figure 2. Comparative Accuracy Chart of the Evaluated Models



4. Discussion and Conclusion

The findings of the present study demonstrated that the three-layer LSTM model achieved superior performance in sentiment analysis of Persian automotive review texts compared with classical machine learning algorithms, including Linear SVM, Naive Bayes, MLP, and Decision Tree. The results indicated that the LSTM architecture achieved an overall accuracy of approximately 95.5%, while the best-performing classical model, Linear SVM, achieved an accuracy of 88%. Furthermore, Naive Bayes and Decision Tree demonstrated comparatively weaker performance across all evaluation metrics. These findings suggest that deep learning architectures, particularly recurrent neural network models based on LSTM, possess substantial capabilities for analyzing complex sentiment patterns in Persian textual data. The observed superiority of the LSTM model can be attributed to its ability to preserve sequential dependencies, capture contextual semantics, and learn hidden representations from large-scale textual datasets. These findings are consistent with previous studies emphasizing the effectiveness of deep learning approaches for sentiment classification tasks [18, 27].

One of the important findings of this study concerns the capability of the LSTM model to maintain balanced

performance across all three sentiment categories, including positive, negative, and neutral reviews. The confusion matrix results indicated that the majority of predictions were concentrated along the main diagonal, demonstrating strong classification consistency and reduced misclassification errors. This balanced performance is particularly important in practical business applications because customer reviews often contain ambiguous or mixed emotional expressions. Classical machine learning algorithms generally struggle to capture subtle semantic relationships within sentences because they rely primarily on handcrafted statistical features such as TF-IDF vectors. In contrast, LSTM networks process textual sequences dynamically and therefore can preserve contextual information throughout sentence structures. Similar conclusions have been reported in previous sentiment analysis studies, where recurrent neural networks demonstrated superior capability in extracting emotional dependencies and contextual semantics from user-generated content [1, 3].

The strong performance of the LSTM model observed in this study can also be explained through advances in deep learning-based representation learning. Traditional models such as Naive Bayes and Decision Trees generally assume feature independence or rely on relatively shallow decision boundaries, which limits their capability to process complex linguistic patterns. Persian language texts present additional

difficulties because of morphological diversity, informal expressions, and syntactic variations. Consequently, statistical feature engineering alone may not sufficiently capture the semantic structure of customer reviews. The use of FastText embeddings in the present study enabled the LSTM model to generate distributed semantic representations that preserve contextual similarities among words. These semantic embeddings, combined with recurrent memory mechanisms, significantly enhanced classification accuracy. Previous studies on Persian sentiment analysis similarly demonstrated that deep learning architectures outperform classical approaches because they effectively capture semantic and contextual information within low-resource languages [21, 27].

Another important finding relates to the impact of data balancing on model performance. The original dataset contained a considerable imbalance among positive, negative, and neutral classes, with positive reviews representing the majority category. After balancing the dataset, the classification performance of the LSTM model improved substantially, particularly in identifying minority classes. This finding highlights the importance of preprocessing strategies in sentiment analysis tasks. Imbalanced datasets can bias machine learning algorithms toward dominant categories, thereby reducing the reliability of classification results for minority classes. Deep learning models are particularly sensitive to data distribution because the training process optimizes parameters based on observed patterns within the dataset. Similar concerns regarding class imbalance have been discussed extensively in previous sentiment analysis research, where balancing methods and data augmentation strategies significantly improved predictive performance [1, 2].

The comparative results between LSTM and Linear SVM also provide valuable insights into the relative strengths of deep learning and classical machine learning methods. Linear SVM demonstrated the strongest performance among the classical models due to its capability to construct optimal decision boundaries in high-dimensional feature spaces. The acceptable performance of SVM-based models has been widely documented in sentiment analysis literature, particularly when combined with TF-IDF feature representations [15, 16]. However, despite its relatively strong performance, the Linear SVM model remained substantially less effective than the LSTM architecture. This difference can be explained by the inability of SVM-based approaches to model sequential dependencies and contextual relationships between words. Sentiment polarity often

depends on sentence structure, negation patterns, and semantic interactions that cannot be adequately represented through isolated statistical features. The superior results of the LSTM model therefore support the growing consensus that contextual representation learning is essential for advanced NLP applications [3, 4].

The weaker performance of Naive Bayes and Decision Tree models observed in this study is also consistent with earlier investigations. Naive Bayes assumes conditional independence among features, which rarely holds true in natural language data because words interact contextually to produce semantic meaning. Similarly, Decision Tree models are highly sensitive to feature sparsity and data fragmentation, which can negatively affect generalization performance in high-dimensional textual datasets. Consequently, both models exhibited relatively lower accuracy and greater classification errors compared with the LSTM architecture. Previous research similarly reported that shallow classifiers often experience reduced effectiveness in sentiment analysis tasks involving complex and context-dependent linguistic structures [2, 17].

The findings of the present study also reinforce the growing importance of sentiment analysis in business intelligence and customer analytics. Customer reviews provide organizations with valuable insights regarding product quality, technical performance, user satisfaction, and consumer expectations. In the automotive industry, customer perceptions concerning safety, fuel efficiency, design, multimedia systems, and after-sales services directly influence purchasing decisions and brand reputation. Therefore, accurate sentiment classification systems can assist automotive companies in identifying customer priorities and improving strategic decision-making processes. Similar applications of sentiment analysis in marketing and consumer behavior prediction have been emphasized in previous studies, which demonstrated that AI-driven analytics significantly enhance customer relationship management and market forecasting capabilities [21, 22].

The present findings are also aligned with recent developments in AI-driven personalization and large language model applications. Contemporary NLP research increasingly focuses on context-aware architectures capable of understanding semantic relationships within large textual corpora. Although transformer-based models such as BERT and ChatGPT have recently demonstrated exceptional performance in sentiment analysis tasks, LSTM architectures continue to provide strong predictive

performance while maintaining computational efficiency and architectural simplicity [6, 24]. The strong performance of the LSTM model in this study suggests that recurrent neural networks remain highly effective for domain-specific sentiment analysis, particularly in languages with limited annotated datasets and computational resources.

Another important implication of the findings concerns the applicability of sentiment analysis for broader industrial and economic forecasting systems. Previous studies demonstrated that sentiment extracted from textual data can influence financial markets, consumer confidence, and industrial performance indicators [8, 19]. Similarly, public opinion regarding fuel prices, geopolitical conflicts, and technological innovation has been shown to affect automotive market dynamics and transportation preferences [32, 33]. Consequently, sentiment analysis systems capable of accurately processing Persian customer reviews can contribute to predictive business analytics and strategic market intelligence in the automotive sector.

The results also support recent interdisciplinary applications of deep learning models across engineering and industrial domains. Studies on predictive maintenance, fault diagnosis, and intelligent monitoring systems increasingly employ LSTM architectures because of their capability to model temporal dependencies and sequential patterns [28, 29]. The successful implementation of LSTM networks in both engineering prediction tasks and sentiment classification demonstrates the versatility of recurrent neural architectures for handling complex data structures. These findings therefore contribute to the broader literature on AI-driven industrial analytics and intelligent decision-support systems.

In addition, the present study contributes to the limited body of research on Persian-language sentiment analysis within the automotive industry. Most previous investigations have primarily focused on English-language corpora or social media datasets, whereas comparatively fewer studies have addressed Persian customer reviews in industrial contexts. The findings therefore provide empirical evidence supporting the effectiveness of deep learning methods for analyzing Persian textual data. Furthermore, the comparative framework employed in this study offers a comprehensive evaluation of both classical and deep learning approaches under identical experimental conditions, thereby enhancing the reliability and practical relevance of the findings.

One limitation of the present study concerns the size and diversity of the dataset. Although 4,348 customer reviews

were analyzed, the dataset was restricted to reviews collected from selected Iranian automotive platforms and may not fully represent the entire diversity of customer opinions in the automotive market. Additionally, the dataset contained class imbalance, which required balancing procedures during preprocessing. Another limitation involves the exclusive focus on textual data without incorporating multimodal information such as images, audio, or behavioral indicators. Furthermore, only one deep learning architecture based on LSTM was examined, whereas more advanced transformer-based models such as BERT or GPT-based systems were not evaluated in the comparative analysis.

Future studies are recommended to investigate larger and more diverse Persian-language datasets collected from multiple social media platforms, e-commerce systems, and automotive forums. Researchers may also examine the integration of transformer-based architectures, attention mechanisms, and hybrid deep learning frameworks to further improve sentiment classification accuracy. In addition, future investigations could explore aspect-level sentiment analysis in greater depth by examining customer opinions regarding specific automotive components such as safety systems, fuel efficiency, or multimedia technologies. The incorporation of multimodal data and explainable artificial intelligence methods may also contribute to the development of more transparent and interpretable sentiment analysis systems.

From a practical perspective, the findings of this study highlight the importance of implementing AI-driven sentiment analysis systems within the automotive industry and broader business environments. Automotive companies can utilize advanced sentiment analysis models to continuously monitor customer satisfaction, identify emerging product issues, and optimize marketing strategies based on real-time consumer feedback. Managers and decision-makers may also employ these systems to improve after-sales services, enhance product quality, and strengthen customer engagement. Moreover, the successful application of deep learning models in Persian sentiment analysis demonstrates the potential for developing localized intelligent business analytics systems tailored to regional linguistic and cultural characteristics.

Authors' Contributions

Authors equally contributed to this article.

Acknowledgments

Authors thank all participants who participate in this study.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethical Considerations

All procedures performed in this study were under the ethical standards.

References

- [1] M. Birjali, M. Kasri, and A. Beni-Hssane, "A Comprehensive Survey on Sentiment Analysis: Approaches, Challenges and Trends," *Knowledge-Based Systems*, vol. 226, p. 107134, 2021.
- [2] M. Rodriguez-Ibanez, A. Casanez-Ventura, F. Castejon-Mateos, and P. M. Cuenca-Jimenez, "A Review on Sentiment Analysis from Social Media Platforms," *Expert Systems with Applications*, p. 119862, 2023.
- [3] W. Khan, A. Daud, K. Khan, S. Muhammad, and R. Haq, "Exploring the Frontiers of Deep Learning and Natural Language Processing: A Comprehensive Overview of Key Challenges and Emerging Trends," *Natural Language Processing Journal*, vol. 4, p. 100026, 2023.
- [4] R. Patil and V. Gudivada, "A Review of Current Trends, Techniques, and Challenges in Large Language Models (LLMs)," *Applied Sciences*, vol. 14, p. 2074, 2024.
- [5] F. Sudirjo, K. Diantoro, J. A. Al-Gasawneh, H. K. Azzaakiyyah, and A. M. A. Ausat, "Application of ChatGPT in Improving Customer Sentiment Analysis for Businesses," *Jurnal Teknologi dan Sistem Informasi Bisnis*, vol. 5, pp. 283-288, 2023.
- [6] J. Chen, "When Large Language Models Meet Personalization: Perspectives of Challenges and Opportunities," *World Wide Web*, vol. 27, p. 42, 2024.
- [7] E. Georgiadou, S. Angelopoulos, and H. Drake, "Big Data Analytics and International Negotiations: Sentiment Analysis of Brexit Negotiating Outcomes," *International Journal of Information Management*, vol. 51, p. 102048, 2020.
- [8] S. S. Mehrkian and H. Davari-Ardakani, "An Integrated Model of Sentiment Analysis and Quantitative Index Data for Predicting Stock Market Trends: A Case Study of Tehran Stock Exchange," *Expert Systems with Applications*, vol. 269, p. 126298, 2025.
- [9] P. N. Lopez *et al.*, "Detection of Occupant Emotion in Automated Vehicles under Different Driving Conditions," *Transportation Research Procedia*, vol. 72, pp. 3917-3924, 2023.
- [10] B. Palomares-Nicolas *et al.*, "Detection of Occupant Emotion in Automated Vehicles under Different Driving Conditions," *Transportation Research Procedia*, vol. 72, pp. 3917-3924, 2023.
- [11] M. Oehl, A. A. Pape, K. Ihme, M. Vukelic, E. Bosch, and M. Braun, "Emotions in the Age of Automated Driving," 2019.
- [12] C. Nadri *et al.*, "Empathic Vehicle Design: Use Cases and Design Directions from Two Workshops," 2022, pp. 1-7.
- [13] T. Y. Kim, H. Ko, and S. H. Kim, "Data Analysis for Emotion Classification Based on Bio-Information in Self-Driving Vehicles," *Journal of Advanced Transportation*, 2020.
- [14] I. I. Barosan, "CarESP: An Emotional Vehicle with Stress, Personality and Embodiment of Emotions," 2020.
- [15] R. Bose, R. K. Dey, S. Roy, and D. Sarddar, "Sentiment Analysis on Online Product Reviews," Springer, 2020, pp. 559-569.
- [16] T. Ahmed, M. Rahman, S. Nur, A. Islam, and D. Das, "Deployment of Machine Learning and Deep Learning Algorithms in Detecting Cyberbullying in Bangla and Romanized Bangla Text: A Comparative Study," 2021, doi: 10.1109/icaect49130.2021.9392608.
- [17] M. Rahman, S. Haque, and Z. Rahman, "Identifying and Categorizing Opinions Expressed in Bangla Sentences Using Deep Learning Technique," *International Journal of Computer Applications*, vol. 176, no. 17, pp. 13-17, 2020, doi: 10.5120/ijca2020920119.
- [18] M. Dam, S. Roy, D. Sarkar, and S. K. Mahata, "Sentiment Analysis Using LSTM Model for Emotion Detection," Kolkata, India, 2025, pp. 1-5, doi: 10.1109/IEMENTech65115.2025.10959455.
- [19] F. Ricchiuti and G. Sperli, "An Advisor Neural Network Framework Using LSTM-Based Informative Stock Analysis," *Expert Systems with Applications*, vol. 259, p. 125299, 2025.
- [20] F. Z. Xing, E. Cambria, and R. E. Welsch, "Natural Language-Based Financial Forecasting: A Survey," *Artificial Intelligence Review*, vol. 50, pp. 49-73, 2018.
- [21] A. Mirzaee, M. Zeynali, A. Ghorbanzadeh, and P. Ghorbanzadeh, "Personal Recommender Model and Predicting Consumer Behavior in Digital Marketing Based on Deep Learning," *Transactions on Machine Intelligence*, vol. 7, no. 3, pp. 179-193, 2024, doi: 10.47176/TMI.2024.179.
- [22] M. S. Hossain, "Textual Feature Engineering for Purchase Intent and Customer Satisfaction: Insights from Marketing 4.0 and Sentiment," *Sustainable Futures*, vol. 8, p. 100385, 2024.
- [23] I. H. Sarker, M. H. Furhad, and R. Nowrozy, "AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling, and Research Directions," *SN Computer Science*, vol. 2, no. 3, 2021, doi: 10.1007/s42979-021-00557-0.
- [24] T. Susnjak, "Applying BERT and ChatGPT for Sentiment Analysis of Lyme Disease in Scientific Literature," 2024, pp. 173-183.
- [25] B. Huang *et al.*, "Aspect-Level Sentiment Analysis with Aspect-Specific Context Position Information," *Knowledge-Based Systems*, vol. 243, p. 108473, 2022.
- [26] H. T. Phan, N. T. Nguyen, and D. Hwang, "Aspect-Level Sentiment Analysis: A Survey of Graph Convolutional Network Methods," *Information Fusion*, vol. 91, pp. 149-172, 2023.
- [27] A. Kazance, R. Ventura, and J. Smith, "Deep Learning-Based Sentiment Analysis for Persian User Reviews," *Expert Systems with Applications*, vol. 225, p. 119862, 2023, doi: 10.1016/j.eswa.2023.119862.
- [28] A. Achour, M. A. Kammoun, and Z. Hajej, "Towards Optimizing Multi-Level Selective Maintenance via Machine Learning Predictive Models," *Applied Sciences*, vol. 14, p. 313, 2023.
- [29] P. Amjadian, "Remaining Useful Life Estimation of Bearings Using Vibration Signal Processing Based on Continuous Wavelet Transform and LSTM Deep Learning Network," *Transactions on Machine Intelligence*, vol. 7, no. 1, pp. 61-69, 2024, doi: 10.47176/TMI.2024.61.

- [30] L. Rognone, S. Hyde, and S. S. Zhang, "News Sentiment in the Cryptocurrency Market: An Empirical Comparison with Forex," *International Review of Financial Analysis*, vol. 69, p. 101462, 2020.
- [31] Q. Zhang, K. Yang, Y. Hu, J. Jiao, and S. Wang, "Unveiling the Impact of Geopolitical Conflict on Oil Prices: A Case Study of the Russia-Ukraine War and Its Channels," *Energy Economics*, vol. 126, p. 106956, 2023.
- [32] M. R. Baker, Y. N. Taher, and K. H. Jihad, "Prediction of People Sentiments on Twitter Using Machine Learning Classifiers during Russian Aggression in Ukraine," *Jordanian Journal of Computers & Information Technology*, vol. 9, 2023.
- [33] K. H. Jihad, M. R. Baker, M. Farhat, and M. Frikha, "Machine Learning-Based Social Media Text Analysis: Impact of Rising Fuel Prices on Electric Vehicles," 2022, pp. 625-635.