

Unveiling the impact of service attributes and review scores on sentiment: A deep learning and feature engineering approach to UberEats reviews

*International Journal of Engineering**Business Management*

Volume 17: 1–14

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/18479790251341980

journals.sagepub.com/home/enb**Md Kamal Hossain¹, Le Li^{2,3}, Usman Ghani¹ and Md Shamim Hossain⁴**

Abstract

This study investigates the impact of SERVQUAL dimensions (Assurance, Reliability, Tangibles, Empathy, and Responsiveness) and review scores on customer sentiment. We analyze a large dataset of 920,407 UberEats reviews from the Google Play Store, classifying sentiment based on star ratings and using a Long Short-Term Memory (LSTM) model to predict sentiment from review content. Using text mining and sentiment analysis, the study employs robust feature engineering techniques to extract and quantify SERVQUAL components from customer reviews. The LSTM model demonstrated high accuracy (89.64%) in predicting sentiment, validating the alignment between predicted and assigned sentiments. Our analysis reveals that all SERVQUAL dimensions and review scores have a positive and significant impact on overall sentiment. Specifically, the Ordinary Least Squares (OLS) regression results highlight Empathy as the most influential SERVQUAL component, followed by Responsiveness, Reliability, Tangibles, and Assurance. Furthermore, review score emerged as the strongest predictor of customer sentiment. These findings provide actionable insights for service providers aiming to enhance customer satisfaction by optimizing key SERVQUAL dimensions and addressing review score trends.

Keywords

SERVQUAL, customer sentiment, feature engineering, LSTM model, UberEats reviews

Date received: 28 January 2025; accepted: 28 April 2025

Introduction

The food delivery service (FDS) industry has experienced rapid growth in recent years, driven by shifting consumer behaviors and an increasing demand for convenience, safety, and efficiency. The COVID-19 pandemic accelerated this trend, as consumers increasingly turned to digital platforms for contactless delivery, reinforcing the centrality of convenience in modern dining experiences.¹ Platforms like UberEats, Zomato, and Swiggy have become integral parts of the food service industry, providing customers with quick access to a wide variety of meals and a seamless ordering experience from the comfort of their homes. In this growing industry, customer satisfaction plays a critical role

¹Doctoral School of Regional and Business Administration Sciences, Széchenyi István Egyetem, Győr, Hungary

²School of Business and Economics, Universiti Putra Malaysia, Serdang, Malaysia

³School of Management Engineering, Anhui Institute of Information Technology, Wuhu, China

⁴Department of Marketing, Hajee Mohammad Danesh Science and Technology University (HSTU), Dinajpur, Bangladesh

Corresponding author:

Md Shamim Hossain, Department of Marketing, Hajee Mohammad Danesh Science and Technology University (HSTU), Dinajpur 5200, Bangladesh.

Email: shamim.mkt@hstu.ac.bd



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without

further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

in success, with businesses continually seeking ways to improve service quality and foster loyalty. Customer reviews and ratings on platforms like UberEats serve as key tools for assessing service quality and understanding customer sentiment, offering valuable insights into customer experiences and perceptions.²

Feature engineering plays a crucial role in extracting valuable insights from customer reviews. By applying various techniques such as sentiment analysis, keyword extraction, and clustering, businesses can understand customer satisfaction, purchasing intent, and emotional responses. Studies like Hossain³ demonstrate how binary features derived from review keywords can correlate strongly with sentiment and customer satisfaction. In this study, feature engineering was implemented to analyze customer sentiment in relation to the five SERVQUAL dimensions: Assurance, Reliability, Tangibles, Empathy, and Responsiveness. The SERVQUAL model, widely applied across various industries including hospitality, banking, and retail, provides a comprehensive framework for evaluating service quality. It encompasses five essential dimensions: Assurance, Reliability, Tangibles, Empathy, and Responsiveness.⁴ These dimensions assess various facets of the service encounter, from the tangible aspects of the service (such as the app interface and food presentation) to the intangible elements (such as customer trust, responsiveness to issues, and personalized service). In the context of food delivery services, these dimensions are particularly relevant, as they capture both the physical and emotional aspects of the customer experience. Although the SERVQUAL model has been widely used in traditional industries, its application in the context of modern, data-driven approaches like sentiment analysis is still underexplored within the FDS sector. Sentiment analysis, a subfield of natural language processing (NLP), has become a powerful tool for extracting insights from large volumes of customer feedback. Machine learning (ML) algorithms, such as support vector machines and deep learning models, have demonstrated effectiveness in analyzing customer reviews and predicting sentiment, offering businesses valuable insights into service quality.⁵ The integration of explainable artificial intelligence (XAI) techniques further enhances the interpretability and trustworthiness of deep learning models.¹ In the FDS industry, sentiment analysis uncovers patterns in customer sentiment related to specific service attributes, helping businesses identify areas of strength and opportunities for improvement. By combining sentiment analysis with the SERVQUAL model, it becomes possible to gain a deeper understanding of how each service dimension influences customer satisfaction.

This study investigates the relationship between the five SERVQUAL dimensions—Assurance, Reliability, Tangibles, Empathy, and Responsiveness—and customer sentiment in the UberEats platform. UberEats was selected as the

case study due to its global presence and extensive customer base, making it a representative platform for analyzing food delivery service (FDS) quality. With a large volume of user-generated reviews, UberEats provides a rich dataset for sentiment analysis and SERVQUAL-based evaluation. Additionally, its well-structured review system, including textual feedback and review scores, allows for a comprehensive, data-driven assessment of service quality. This enables a more generalizable analysis compared to smaller or region-specific platforms. Using a dataset of over 920,000 customer reviews, it identifies the key service attributes that influence customer sentiment and satisfaction. Feature engineering techniques are applied to extract keywords associated with each SERVQUAL dimension from the reviews, and the relationship between these attributes and sentiment is analyzed. The study also examines how review scores (numerical ratings) impact customer perceptions. By exploring the interplay between service quality dimensions and review scores, this research provides insights into the factors that shape customer sentiment, offering valuable feedback for businesses seeking to improve service delivery. Additionally, the study assesses different sentiment classification techniques, including both traditional methods based on review scores and more advanced machine learning models. Specifically, deep learning models, such as Long Short-Term Memory (LSTM) networks, are employed to predict sentiment based on the content of customer reviews. These models help validate the relationship between review scores and sentiment, providing a robust framework for analyzing customer feedback in the FDS industry. The insights from this study are valuable for both researchers and practitioners in the food delivery sector. Understanding the key drivers of customer sentiment enables businesses like UberEats to improve service quality, address customer concerns, and foster customer loyalty. By leveraging the SERVQUAL model and sentiment analysis, this research offers a deeper understanding of the factors that shape customer experiences in the food delivery industry, providing actionable recommendations for enhancing customer satisfaction and business performance.

Literature review

Data mining in food delivery

The rapid growth of food delivery services (FDS) has been fueled by changing consumer preferences, particularly during the COVID-19 pandemic, where convenience and safety became top priorities.¹ Customer reviews and ratings on platforms like Zomato, UberEats, and Swiggy have emerged as critical tools for understanding service quality and customer satisfaction.² Adak et al.¹ highlighted the prevalence of machine learning (ML) techniques, noting the

underutilization of deep learning (DL) models due to challenges in model interpretability and explainability. The study suggested integrating explainable artificial intelligence (XAI) into DL to build trust and enhance adoption. Hossain et al.⁵ further demonstrated the effectiveness of ML models in sentiment analysis, with support vector machines achieving the highest accuracy and logistic regression excelling in recall and F1 scores. Additionally, Bannor and Amponsah⁶ identified factors such as convenience, hygiene, pricing, and app reviews as significant influencers of FDS adoption in Africa. Khan et al.² underscored the importance of hygiene and pricing in shaping customer sentiment and ratings. Furthermore, Dogra et al.⁷ found that app quality and customer reviews play a vital role in fostering trust and repurchase intention. Collectively, these studies highlight the transformative role of sentiment analysis in improving FDS performance and customer satisfaction. In this study, we employed feature engineering techniques, specifically keyword search methods, to extract the five key service quality dimensions—Assurance, Reliability, Tangibles, Empathy, and Responsiveness—from customer reviews. By analyzing the frequency and context of keywords associated with each dimension, we were able to quantify their impact on customer sentiment. Additionally, we measured the influence of review scores on sentiment, considering how both the review content and the numerical ratings interact to shape overall customer satisfaction. This approach enabled us to understand the nuanced ways in which service quality dimensions and review scores contribute to sentiment, providing deeper insights into the factors that drive customer perceptions and experiences in the food delivery industry.

Feature engineering in sentiment analysis

Feature engineering plays a critical role in sentiment analysis, particularly in leveraging textual data from online reviews to uncover valuable insights about customer behavior and service quality. Numerous studies have emphasized the importance of extracting relevant features to enhance the performance of machine learning (ML) models and gain a deeper understanding of consumer preferences, satisfaction, and behavior. Hossain³ explores the role of textual features in shaping consumer behavior metrics like Purchase Intention and Customer Satisfaction. By analyzing 329,658 Yelp reviews, the study highlights how sentiment, along with features such as Brand Identity, Brand Image, and Customer Interaction, influences these behavior metrics. The study employs feature engineering methods like keyword extraction and the use of ChatGPT for generating binary features, ultimately demonstrating a strong correlation between sentiment and customer satisfaction. This approach emphasizes how nuanced features derived from text data can drive more accurate predictions and provide

actionable insights for businesses. Similarly, Bilal et al.⁸ tackle the challenge of information overload caused by the rapid growth of online customer reviews. By examining the “helpfulness” of reviews, the study introduces the concept of Social Network Strength (SNS), a key feature in determining the quality and credibility of online feedback. Their feature engineering approach combines SNS attributes with review content, leading to improved performance in both regression and classification models. This study underscores the need to consider the broader context of customer reviews, including social influence, when extracting features that predict review helpfulness. Chang et al.⁹ focus on the hospitality industry, particularly luxury hotel reviews, and propose an integrated sentiment analysis model that combines feature engineering techniques like Term Frequency-Inverse Document Frequency (TF-IDF) and Doc2Vec. Their work emphasizes the significance of using both traditional and modern methods to generate features that capture the complexities of customer sentiment. The model’s ability to integrate multiple data types—ranging from textual reviews to geospatial and multimedia data—illustrates the growing demand for comprehensive feature engineering strategies that can handle diverse datasets in sentiment analysis.

In the realm of e-commerce and personalized marketing, Alves Gomes and Meisen¹⁰ examine customer segmentation methods, highlighting how feature engineering plays a central role in accurately representing customer behavior. They present a four-phase process for customer analysis, beginning with data collection and moving through to segmentation and targeting. The paper stresses the importance of manual feature selection and the application of segmentation techniques such as k-means clustering to improve customer-targeting strategies in personalized marketing. A different approach to feature engineering is demonstrated by Wang et al.,¹¹ who explore affective attributes in online reviews using a deep learning method based on Kansei engineering. This method extracts customer emotions and classifies them into various affective pairs like “useful-useless” and “reliable-unreliable,” showcasing the potential of feature engineering to enhance sentiment analysis by considering not only the sentiment polarity but also the emotional nuances present in reviews. Razali et al.¹² also tackle the challenges of sentiment classification in niche sectors, such as gastronomy tourism. Their study combines feature engineering with data augmentation to improve the recognition of minority sentiment classes, which often remain underrepresented in sentiment datasets. By optimizing feature extraction through n-gram and synonym augmentation strategies, they demonstrate how feature engineering can address imbalances in sentiment data and enhance the performance of machine learning models. Kumar et al.¹³ introduce a novel feature engineering approach to detect fraudulent reviews, focusing on

the behavior of fake reviewers. By combining review-centric and reviewer-centric features, the study develops a probabilistic model that improves the accuracy of fraudulent review detection. This work exemplifies how feature engineering can be utilized not just for analyzing genuine customer sentiment but also for identifying and mitigating fraudulent activities in online reviews.

Additionally, Kuppusamy and Thangavel¹⁴ present a deep non-linear learning method for opinion mining, proposing a new feature extraction process that focuses on computational efficiency and accuracy. Their approach, which utilizes a skip-gram input layer and non-linear ReLU activation functions, highlights the increasing sophistication of feature engineering in deep learning-based sentiment analysis. By enhancing feature extraction, they achieve higher classification accuracy compared to traditional methods. Also, Kim et al.¹⁵ investigate the determinants of review helpfulness using feature engineering and explainable AI (XAI) techniques. By extracting features such as sentiment and reviewer reputation, they improve the prediction of review helpfulness, further demonstrating how feature engineering can enhance the interpretability and performance of machine learning models in real-world applications. Together, these studies illustrate the diverse applications of feature engineering in sentiment analysis across various domains. Whether in predicting customer satisfaction, detecting fraudulent reviews, or enhancing marketing strategies, feature engineering remains a pivotal component in unlocking the full potential of online customer reviews. These efforts underscore the growing importance of tailored feature extraction methods to address the complexities of modern sentiment analysis tasks. In our study, feature engineering plays a pivotal role in extracting meaningful insights from customer reviews on UberEats. The goal was to effectively quantify the SERVQUAL dimensions (Assurance, Reliability, Tangibles, Empathy, and Responsiveness) and evaluate their impact on customer sentiment. We implemented several feature engineering techniques to preprocess the textual data, allowing us to capture the key attributes that influence customer perceptions and satisfaction.

Applying the SERVQUAL model to evaluate service quality in food delivery apps

Applying the SERVQUAL model to analyze food delivery apps' reviews allows for a nuanced understanding of service quality and customer satisfaction. SERVQUAL, which includes five key dimensions—tangibility, reliability, responsiveness, assurance, and empathy—can be adapted to the food delivery industry to assess consumer experiences. Tangibility pertains to the physical aspects of the service, including the app's interface and food presentation. In the

context of food delivery apps, reviews often highlight the design and functionality of the app, as well as the quality of packaging and food's appearance upon arrival.¹⁶ Reliability is another vital dimension, referring to the accuracy and consistency of service delivery. Customers frequently comment on the timeliness and correctness of their orders, with any discrepancies often leading to dissatisfaction.¹⁷ Responsiveness relates to how quickly and effectively the service addresses issues, such as delivery delays or incorrect orders. Positive reviews typically highlight helpful and prompt customer service, which is crucial for customer loyalty.¹⁸ Assurance, a key aspect of customer trust, is increasingly significant in the context of food delivery, especially in terms of hygiene and safety. Customers have noted the importance of hygiene practices, particularly in light of the COVID-19 pandemic, as part of their assurance when ordering food.¹⁸ Finally, empathy involves personalized service and attentiveness to individual customer needs. Reviews reflecting a high level of empathy often mention how food delivery apps accommodate special requests or make efforts to enhance the overall customer experience.¹⁶ By leveraging these SERVQUAL dimensions, food delivery apps can gain valuable insights into areas for service improvement and enhance customer satisfaction. In the current study, we measured the impact of various components of the SERVQUAL service quality model, alongside review scores, on the overall sentiment of users of the UberEats app. By analyzing the relationship between the five SERVQUAL dimensions—Assurance, Reliability, Tangibles, Empathy, and Responsiveness—and the numerical ratings provided in customer reviews, we assessed how these factors influence user sentiment. This approach provided a deeper understanding of the specific service attributes that contribute to user satisfaction and dissatisfaction, offering valuable insights for improving the UberEats customer experience.

This study enhances food delivery service (FDS) quality assessment by analyzing text reviews through the integration of SERVQUAL components and review scores to evaluate sentiment, addressing key limitations in existing research. Previous studies primarily rely on surveys or qualitative methods, which often suffer from small sample sizes and subjective biases. While Long Short-Term Memory (LSTM) models have been widely applied in sentiment analysis of customer reviews, their integration with the SERVQUAL model and review scores for FDS assessment remains largely unexplored.

To bridge this gap, we utilize a large-scale dataset of 920,407 UberEats reviews from the Google Play Store, enabling a more comprehensive, data-driven evaluation of customer sentiment and service attributes. Unlike prior research, our approach combines LSTM-based sentiment classification with SERVQUAL dimensions—Assurance, Reliability, Tangibles, Empathy, and Responsiveness—

alongside review scores, offering a dual-layered analysis of service quality. This integration improves sentiment classification accuracy and provides deeper insights into the relationship between service attributes, sentiment, and overall customer satisfaction. By leveraging deep learning on large-scale real-world data, this study introduces a novel and scalable approach to evaluating service quality in the food delivery industry. By addressing these limitations, our research provides valuable empirical insights for both academics and industry practitioners aiming to enhance service quality and customer satisfaction in food delivery platforms.

Hypotheses development

Based on the findings from prior studies, the SERVQUAL model's components—Assurance, Reliability, Tangibles, Empathy, and Responsiveness—are essential in shaping the overall sentiment of UberEats app users. Each of these components influences how customers perceive and interact with the service, ultimately impacting their satisfaction and loyalty. Assurance, as highlighted by Zygiaris et al.¹⁹ and Meesala and Paul,²⁰ plays a crucial role in establishing customer confidence in the service. For UberEats, this involves ensuring that users feel assured about the reliability of order deliveries, the transparency of service, and the security of their transactions. When these elements are assured, users are more likely to have a positive sentiment toward the app. Reliability is another important component, emphasized by both Zygiaris et al.¹⁹ and Jain et al.²¹ Reliability in the context of UberEats pertains to the consistency and timeliness of order fulfillment. When users can depend on UberEats to deliver their food as promised, their satisfaction increases, which, in turn, enhances their overall sentiment toward the app. While Tangibles are not directly discussed in all the studies, de Kervenoael et al.²² underline the importance of tangible elements like service presentation, which can impact customer experience. For UberEats, the design and user interface of the app are crucial. A visually appealing and intuitive app increases user satisfaction and contributes positively to the overall sentiment.

Empathy, discussed by Hossain and Rahman,²³ Zhang et al.,²⁴ and Lv et al.,²⁵ involves understanding and responding to customers' emotional needs. In the UberEats context, empathetic customer service—whether it involves addressing concerns with patience or resolving complaints thoughtfully—plays a vital role in fostering positive sentiment. When users feel understood and valued, they are more likely to have favorable attitudes toward the service. Lastly, Responsiveness is integral to customer satisfaction, as shown by Zygiaris et al.¹⁹ and Jain et al.²¹ In the case of UberEats, responsiveness refers to the ability to quickly address issues such as delayed deliveries, order inaccuracies, or technical problems. A timely and effective

response to customer queries and complaints can significantly enhance the overall user experience, contributing to a more positive sentiment. Together, these SERVQUAL components are hypothesized to have a positive and significant impact on the overall sentiment of UberEats app users, as they directly influence the users' perceptions of service quality, their emotional reactions, and their continued engagement with the app. Thus

H1: The components of the SERVQUAL model, including Assurance (H1a), Reliability (H1b), Tangibles (H1c), Empathy (H1d), and Responsiveness (H1e), have a positive and significant impact on the overall sentiment of UberEats app users.

Additionally, prior studies have consistently shown that user ratings or scores significantly affect customer sentiment and satisfaction. For example, Hossain & Rahman²³ and Lv et al.²⁵ explored the role of customer reviews and ratings in shaping emotional responses and behaviors, demonstrating that positive ratings foster favorable sentiments, while negative ones can lead to dissatisfaction. These findings suggest that a user-provided score, reflecting the level of satisfaction with services like UberEats, directly influences overall sentiment. Similarly, studies such as Zhang et al.²⁴ and Meesala & Paul²⁰ reinforce the idea that positive user ratings contribute to increased customer loyalty and satisfaction, further enhancing sentiment. Moreover, Zygiaris et al.¹⁹ and Jain et al.²¹ highlighted the impact of feedback on service quality, emphasizing that the scores users provide not only reflect their immediate experiences but also influence the app's future service improvements. In the case of UberEats, positive scores indicate user satisfaction, which directly correlates with more favorable overall sentiment, while negative scores can suggest dissatisfaction and potentially alter sentiment in the opposite direction. Furthermore, studies like de Kervenoael et al.²² and Zhou et al.²⁶ have shown that customer ratings affect future behavior, emotional reactions, and customer engagement, which supports the hypothesis that user-provided scores significantly influence the sentiment of UberEats app users. Thus, the relationship between user ratings and overall sentiment is well-established in prior research, making H2 a plausible assertion.

H2: The user-provided score has a positive and significant impact on the overall sentiment of UberEats app users.

Method

Using Google Colab, we collected a dataset of 929,470 customer reviews for the UberEats app from the Google Play Store. After downloading the data to a local

computer, we conducted the analysis in Jupyter Notebook. The analysis involved importing essential Python libraries, including pandas, statsmodels, numpy, sklearn, tensorflow, and others. The dataset was preprocessed by cleaning the content column. Missing values were replaced with empty strings, and all text was converted to lowercase. Non-alphabetic characters and emojis were removed using regular expressions, and extra spaces were stripped. Blank entries were excluded, reducing the dataset to 920,407 reviews, which was then used for analysis.

A new column, sentiment, was added to the dataset to classify customer sentiments based on their review scores. Following a simple mapping approach, reviews with scores of 0 or 1 were categorized as negative sentiment (−1), scores of 3 as neutral sentiment (0), and scores of 4 or 5 as positive sentiment (1). This approach to assigning sentiment based on review scores is supported by the methodologies of Pashchenko et al.²⁷ and Hossain.³ To validate the sentiment assignment process based on review star, we predicted sentiment based on the content of the reviews. Initially, sentiment labels were assigned based on review star ratings: scores of 0 or 1 were classified as negative (−1), 3 as neutral (0), and 4 or 5 as positive (1). To further validate this approach, a deep learning model was trained to predict sentiment directly from the text. The dataset, including both text reviews and sentiment labels, was preprocessed by tokenizing the reviews and encoding the sentiment labels into one-hot vectors. Text data was padded to ensure uniform sequence lengths for model input. An LSTM (Long Short-Term Memory) model was used for sentiment classification, which is well-suited for sequential data like text. The model classified sentiments into three categories: negative (−1), neutral (0), and positive (1). The Long Short-Term Memory (LSTM) model is a type of recurrent neural network (RNN) designed to capture long-term dependencies in sequential data, making it highly effective for sentiment analysis of customer reviews. The LSTM model processes text data by passing it through multiple layers of memory cells, each governed by a set of gating mechanisms. Each LSTM unit consists of three gates—input, forget, and output—which regulate the flow of information through the network. The mathematical formulation of an LSTM cell is as follows:

- (i) Forget Gate-Determines how much past information is retained:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

where f_t is the forget gate activation, W_f represents learned weights, h_{t-1} is the hidden state from the previous time step, x_t is the current input, and b_f is the bias term.

- (ii) Input Gate – Regulates new information added to memory:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\check{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

where i_t determines the importance of new information, and $\check{C}_t$ is the candidate memory update.

- (iii) Cell State Update – Updates memory with new and retained information:

$$c_t = f_t * c_{t-1} + i_t * \check{C}_t$$

where C_t is the updated memory cell state.

- iv. Output Gate – Determines the final hidden state for sentiment classification:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

where h_t is the final hidden state, used to classify sentiment.

The model was implemented using TensorFlow with an embedding layer, two LSTM layers (128 and 64 units), dropout (0.5), and a dense softmax layer. It was trained using categorical cross-entropy loss and the Adam optimizer.

Additionally, prior studies have leveraged sentiment analysis techniques to extract customer insights from social media and online reviews. Waluyo and Juwono²⁸ applied an LSTM model with TensorFlow for classifying negative social media comments, achieving high validation accuracy. Similarly, Kaur²⁹ demonstrated the effectiveness of machine learning-based sentiment analysis combined with web scraping for analyzing live news data. These studies highlight the significance of deep learning models in text-based sentiment classification, reinforcing our choice of LSTM for evaluating UberEats reviews. Furthermore, Lee et al.³⁰ explored customer sentiment and service quality in accommodation-sharing services, integrating social media analytics. Their findings emphasize the impact of sentiment on customer loyalty, which aligns with our research objective of assessing food delivery service quality through sentiment analysis and SERVQUAL dimensions.

The model's performance was evaluated on a test set, and predictions were made to assess the accuracy of sentiment prediction. The sentiment assignment based on review scores acts as the initial label generation process, while the deep learning model and evaluation (with confusion matrix, classification report, and accuracy score) validate the effectiveness and alignment of those sentiment labels with the predicted sentiments from the review content.

The inclusion of this feature provided a solid foundation for analyzing and predicting customer sentiment in subsequent stages of the study. Using feature engineering and

keyword search techniques, we can gain valuable customer insights from text reviews.³ Keyword lists were defined for five key service attributes—Assurance, Reliability, Tangibles, Empathy, and Responsiveness—each containing 50 distinct keywords (presented in [appendix 1](#)). These keywords were carefully selected to represent terms associated with each attribute. To ensure the quality of the data, duplicates were identified and removed both within each attribute and across all dimensions. This process revealed redundancies, which were addressed accordingly. The refined keyword lists formed a strong foundation for feature extraction, enabling effective analysis of customer sentiment. Binary features were created for each service attribute by checking if keywords from predefined lists appeared in the review content. If a keyword was found, a feature value of 1 was assigned; otherwise, it remained 0. Sentiment adjustments were made, converting feature values from 1 (positive) to -1 (negative) for negative reviews. This process enhanced the dataset by aligning features with customer sentiment for further analysis.

To analyze the relationships between sentiment and key service attributes, we first calculated the correlation between sentiment, service attributes (Assurance, Reliability, Tangibles, Empathy, and Responsiveness), and review scores. Next, we performed an Ordinary Least Squares (OLS) regression with sentiment as the dependent variable and the service attributes and score as independent variables. This regression model helped assess the influence of each attribute on sentiment. Additionally, to check for multicollinearity among the features, we calculated the Variance Inflation Factor (VIF) for each variable, ensuring the robustness of our model.

Finding and discussion

[Table 1](#) highlights the sentiment around customer satisfaction. A significant majority of reviews (52.61%) were rated 5 stars, indicating strong positive sentiment and high customer satisfaction. However, the 30.49% of 1-star reviews reveal a considerable amount of negative sentiment, suggesting areas where the service is falling short and customers are dissatisfied. The 4-star reviews (9.29%) reflect somewhat positive sentiment with minor concerns, while 2-star reviews (3.36%) show mild dissatisfaction. The low proportion of 3-star reviews (4.25%) indicates a polarized sentiment, where customers are either highly satisfied or highly dissatisfied, further emphasizing the need for improvements in areas causing negative sentiment. Overall, while UberEats enjoys positive sentiment from the majority of its users, the considerable negative feedback highlights the importance of addressing areas that are impacting customer satisfaction.

To validate the sentiment assignment process based on review stars, we predicted sentiment based on the content of

Table 1. Distribution of UberEats reviews by score with corresponding percentages.

Score (Review star)	Number of reviews	Percentage
5	484247	52.61
4	85507	9.29
3	39088	4.25
2	30887	3.36
1	280678	30.49
Total	920407	100.00

the reviews using a Long Short-Term Memory (LSTM) model. The results, shown in [Table 2](#), demonstrate the model's performance. After 10 epochs, the model achieved a test accuracy of 89.64%, with the training accuracy increasing from 88.92% in Epoch 1 to 91.09% in Epoch 10. The validation accuracy reached 89.64%, showing strong performance, though slightly lower than the training accuracy due to some overfitting. The loss decreased steadily from 0.3164 to 0.2519 on the training set, while the validation loss plateaued at 0.3021 by the final epoch. These results confirm that the sentiment prediction based on review content aligns well with the sentiment assigned based on review stars.

The Classification Report ([Figure 1](#)) further highlights the model's overall performance, with an accuracy of 89.64%. The Cohen's Kappa and Matthews Correlation Coefficient both reached 0.78, indicating a strong agreement between predicted and actual sentiments. The model showed a balanced performance across the three sentiment categories, confirming that sentiment prediction based on review content aligns well with the sentiment assigned based on review stars. These results validate the effectiveness of the model in capturing the overall sentiment from UberEats reviews.

Additionally, [Table 3](#) provides the average values of service attributes and scores across different sentiment categories. The data reveals notable differences in how each sentiment category perceives key service attributes—Assurance, Reliability, Tangibles, Empathy, and Responsiveness—and their corresponding average review scores. For negative sentiment (-1), the attribute values are predominantly negative, with the lowest average value observed for Responsiveness (-0.09445), suggesting dissatisfaction in this area. The average score for this category is 1, reflecting strong negative sentiment. For neutral sentiment (0), the attribute values are generally positive, albeit lower compared to the positive sentiment category. Reliability (0.08509) and Responsiveness (0.0877) are slightly more prominent, with an average review score of 3, indicating mixed or moderate customer experiences. For positive sentiment (1), the attributes are most highly rated, with Responsiveness achieving the highest average value

Table 2. Model training and validation results for sentiment prediction.

Epoch 1/10
11506/11506 - 1516 s - loss: 0.3164 - accuracy: 0.8892 - val_loss: 0.2989 - val_accuracy: 0.8950 - 1516 s/epoch - 132 ms/step
Epoch 2/10
11506/11506 - 2031 s - loss: 0.2887 - accuracy: 0.8985 - val_loss: 0.2935 - val_accuracy: 0.8972 - 2031 s/epoch - 177 ms/step
Epoch 3/10
11506/11506 - 2119 s - loss: 0.2792 - accuracy: 0.9015 - val_loss: 0.2899 - val_accuracy: 0.8977 - 2119 s/epoch - 184 ms/step
Epoch 4/10
11506/11506 - 2139 s - loss: 0.2724 - accuracy: 0.9037 - val_loss: 0.2906 - val_accuracy: 0.8979 - 2139 s/epoch - 186 ms/step
Epoch 5/10
11506/11506 - 2110 s - loss: 0.2672 - accuracy: 0.9055 - val_loss: 0.2914 - val_accuracy: 0.8981 - 2110 s/epoch - 183 ms/step
Epoch 6/10
11506/11506 - 2137 s - loss: 0.2630 - accuracy: 0.9070 - val_loss: 0.2946 - val_accuracy: 0.8975 - 2137 s/epoch - 186 ms/step
Epoch 7/10
11506/11506 - 2184 s - loss: 0.2594 - accuracy: 0.9083 - val_loss: 0.2931 - val_accuracy: 0.8965 - 2184 s/epoch - 190 ms/step
Epoch 8/10
11506/11506 - 2151 s - loss: 0.2562 - accuracy: 0.9094 - val_loss: 0.3002 - val_accuracy: 0.8965 - 2151 s/epoch - 187 ms/step
Epoch 9/10
11506/11506 - 2084 s - loss: 0.2539 - accuracy: 0.9103 - val_loss: 0.3013 - val_accuracy: 0.8959 - 2084s/epoch - 181 ms/step
Epoch 10/10
11506/11506 - 2184 s - loss: 0.2519 - accuracy: 0.9109 - val_loss: 0.3021 - val_accuracy: 0.8964 - 2184 s/epoch - 190 ms/step
5753/5753 [= = = = =] - 79 s 14 ms/step - loss: 0.3021 - accuracy: 0.8964
Test Loss: 0.3021138906478882, Test Accuracy: 0.8963614106178284

(0.101928). Assurance also stands out with the highest average value (0.119651) across all sentiment categories. The average score for this category is 4.70337, consistent with strong positive sentiment. These findings highlight

how customer perceptions of specific service attributes correlate with overall sentiment and scores, emphasizing the critical role of Responsiveness and Assurance in driving positive experiences.

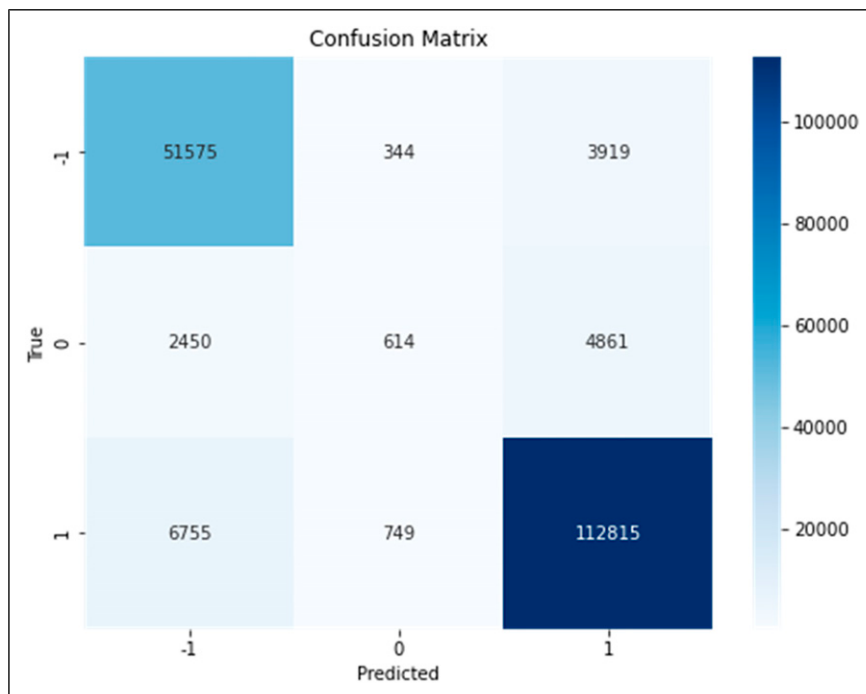
**Figure 1.** Confusion matrix. Note. Accuracy: 89.64%; Cohen's Kappa: 0.78; Matthews Correlation Coefficient: 0.78.

Table 3. Average service attributes and scores by sentiment categories.

Sentiment	Assurance	Reliability	Tangibles	Empathy	Responsiveness	Score
–1	–0.05967	–0.06895	–0.03743	–0.04178	–0.09445	1
0	0.069075	0.08509	0.035279	0.038861	0.0877	3
1	0.119651	0.084806	0.050756	0.024572	0.101928	4.70337

Then we have calculated correlations among variables, as shown in Table 4, to examine the relationships between sentiment, service attributes (Assurance, Reliability, Tangibles, Empathy, and Responsiveness), and review scores. The analysis shows that all service attributes have positive correlations with sentiment, highlighting their contributions to customer satisfaction. Notably, score has the strongest positive correlation with sentiment (0.945633), reaffirming that review scores are a reliable indicator of customer sentiment. Among the service attributes, Responsiveness (0.283544) and Assurance (0.263673) exhibit the highest positive correlations with sentiment, underscoring their critical role in shaping positive customer experiences. Reliability also shows a moderately positive correlation with sentiment (0.246476), emphasizing its influence. Additionally, while Tangibles (0.186928) and Empathy (0.16965) have relatively smaller positive correlations, they still contribute positively to customer sentiment. The correlations among the service attributes themselves also highlight significant interrelationships. For instance, Reliability is strongly positively correlated with Responsiveness (0.681794) and moderately with Assurance (0.337462), indicating these attributes often work together to enhance customer perceptions. Overall, the analysis reveals that all service attributes positively impact customer sentiment, with Responsiveness and Assurance standing out as key drivers of positive sentiment and higher review scores. This emphasizes the importance of focusing on these attributes to improve customer satisfaction further.

To understand the impact of service attributes and review scores on sentiment, we conducted an Ordinary Least Squares (OLS) regression analysis, with results presented in

Table 5. The regression model achieved a high R-squared value of 0.896, indicating that 89.6% of the variance in sentiment can be explained by the independent variables: Assurance, Reliability, Tangibles, Empathy, Responsiveness, and score. The results demonstrate that all variables have a positive and statistically significant impact on sentiment ($p < 0.001$). Among the service attributes, Empathy shows the strongest positive influence on sentiment (coef = 0.1178), followed by Responsiveness (coef = 0.0533) and Reliability (coef = 0.0438). Tangibles (coef = 0.0385) and Assurance (coef = 0.0193) also contribute positively to sentiment, though their impact is relatively smaller compared to other variables. Notably, review score has the highest coefficient (coef = 0.4761) and the most significant t-statistic, confirming its critical role in determining sentiment. This finding aligns with the earlier observed strong correlation between sentiment and review scores.

The findings from the OLS regression analysis strongly support all hypotheses (H1a to H1e and H2). As hypothesized, the components of the SERVQUAL model—Assurance, Reliability, Tangibles, Empathy, and Responsiveness—each have a positive and statistically significant impact on the overall sentiment of UberEats app users. Among these, Empathy emerges as the most influential service attribute, reinforcing the importance of understanding and addressing customers' emotional needs in the context of food delivery services. The significant coefficients for Responsiveness and Reliability further emphasize the need for timely and dependable service, while Tangibles and Assurance, although having smaller coefficients, still contribute positively to user sentiment. Additionally, the review score shows a remarkably strong positive effect on sentiment, with the highest coefficient,

Table 4. Correlations.

	Sentiment	Assurance	Reliability	Tangibles	Empathy	Responsiveness	Score
sentiment	1	0.263673	0.246476	0.186928	0.16965	0.283544	0.945633
Assurance	0.263673	1	0.337462	0.075076	0.256069	0.269312	0.259204
Reliability	0.246476	0.337462	1	0.164551	0.109019	0.681794	0.230452
Tangibles	0.186928	0.075076	0.164551	1	0.06834	0.164513	0.183397
Empathy	0.16965	0.256069	0.109019	0.06834	1	0.132193	0.151514
Responsiveness	0.283544	0.269312	0.681794	0.164513	0.132193	1	0.268563
score	0.945633	0.259204	0.230452	0.183397	0.151514	0.268563	1

Table 5. OLS regression results.

	coef	Std err	t	P> t	[0.025	0.975]
const	−1.3260	0.001	−1903.998	0.000	−1.327	−1.325
Assurance	0.0193	0.001	17.490	0.000	0.017	0.022
Reliability	0.0438	0.002	28.422	0.000	0.041	0.047
Tangibles	0.0385	0.001	25.984	0.000	0.036	0.041
Empathy	0.1178	0.002	64.150	0.000	0.114	0.121
Responsiveness	0.0533	0.001	38.952	0.000	0.051	0.056
score	0.4761	0.000	2578.428	0.000	0.476	0.476

confirming the crucial role of user ratings in shaping overall customer satisfaction. These results provide strong empirical evidence for the hypotheses and underscore the value of focusing on both service attributes and customer feedback to improve sentiment and satisfaction in the UberEats app.

The high F-statistic ($1.321e + 06$, $p < 0.001$) and statistically significant coefficients validate the robustness of the model, highlighting the importance of both service attributes and scores in influencing customer sentiment. These insights underscore the necessity for businesses to focus on enhancing attributes like Empathy and Responsiveness, as they have the most substantial effect on positive customer sentiment. The regression model achieved a very high R-squared value of 0.896, indicating that 89.6% of the variance in sentiment can be explained by the independent variables: Assurance, Reliability, Tangibles, Empathy, Responsiveness, and score. Additionally, we conducted a Variance Inflation Factor (VIF) analysis to check for multicollinearity among the independent variables, as shown in Table 6. The VIF values are generally low, indicating that there is no significant multicollinearity between the variables, which strengthens the reliability of our regression results. Furthermore, all VIF values are below the threshold of 5, with Assurance having the lowest VIF value of 1.232246, followed by Tangibles (1.055931) and Empathy (1.084117). Other variables, such as Reliability (1.974930), Responsiveness (1.928778), and score (1.497386), also show moderate VIF values, all of which are well within acceptable limits. The constant has a VIF of 9.185573, which is expected as it is not an independent variable. These results confirm that multicollinearity is not a concern in our model, allowing us to confidently interpret the individual effects of the service attributes and review score on sentiment. The absence of high VIF values ensures that the regression coefficients are stable and reliable. This highlights the strong explanatory power of the model in understanding customer sentiment.

The key assumptions underlying the OLS regression analysis were explicitly considered in the current study. Linearity was ensured through significant correlations

Table 6. VIF.

	Feature	VIF
0	const	9.185573
1	Assurance	1.232246
2	Reliability	1.974930
3	Tangibles	1.055931
4	Empathy	1.084117
5	Responsiveness	1.928778
6	score	1.497386

among variables, confirming a consistent relationship between sentiment and the independent variables. Multicollinearity was assessed using Variance Inflation Factor (VIF) analysis, with all values remaining below the threshold of 5, indicating no significant collinearity issues. The independence of errors was assumed, as each review represents an independent customer experience. Homoscedasticity was supported by the model's high explanatory power ($R^2 = 0.896$), while the normality of residuals was justified by the large sample size, aligning with the Central Limit Theorem. These assumptions reinforce the validity and reliability of the regression analysis.

Applications

Managerial applications

The findings of this study underscore the importance of service attributes such as Empathy and Responsiveness in shaping customer sentiment and satisfaction. Managers can use this information to enhance the quality of customer service by focusing on areas like timely responses and empathetic communication. Training programs aimed at improving these skills for customer service representatives and delivery personnel can have a significant impact on customer satisfaction, reducing negative feedback and fostering customer loyalty. Moreover, responsiveness plays a key role in ensuring positive sentiment. By streamlining processes to address customer concerns quickly—whether

through faster deliveries or more efficient customer support systems—managers can improve the overall customer experience. Ensuring that customers' needs are met in a timely manner can lead to higher satisfaction and repeat business.

Personalized service is another area where managers can apply the study's insights. Offering tailored promotions or addressing specific concerns based on past customer interactions demonstrates an understanding of individual preferences, building a stronger connection between customers and the brand. This personalized approach can significantly boost sentiment, especially when combined with quick problem resolution. The study also emphasizes the impact of Tangibles and Responsiveness on customer sentiment. Ensuring that the food quality and packaging meet customer expectations, while also maintaining an efficient and responsive service, is crucial. Managers can improve both tangible and intangible aspects of the service to enhance the overall customer experience, reducing the likelihood of negative reviews. Additionally, the study highlights the need for enhanced feedback mechanisms. By collecting real-time customer insights through surveys, social listening, and direct feedback channels, managers can identify service gaps quickly. This proactive approach enables prompt corrective action, reducing the impact of negative reviews and ensuring that customers feel valued and heard. Finally, the study's emphasis on data-driven decision-making encourages managers to leverage analytics to monitor key service attributes. Regularly assessing performance metrics related to Responsiveness, Empathy, and Assurance can help managers identify trends, track areas of improvement, and refine operational strategies. This allows for continuous enhancement of service quality, which ultimately strengthens customer satisfaction and sentiment.

Theoretical applications

This study provides significant theoretical contributions to the understanding of customer sentiment, particularly in the food delivery sector. By integrating the SERVQUAL model with sentiment analysis and machine learning techniques, the study extends the traditional service quality dimensions into the realm of modern, data-driven environments. The findings reinforce the importance of service attributes such as Responsiveness, Assurance, Tangibles, Empathy, and Reliability in shaping customer perceptions and overall sentiment. It highlights that different service dimensions have varying levels of influence on customer sentiment, with Empathy emerging as the most influential, followed by Responsiveness, Reliability, Tangibles, and Assurance. This challenges previous assumptions that all SERVQUAL dimensions have equal influence and encourages further

exploration of how these dimensions may vary across different industries.

Additionally, the study also advances the theoretical understanding of the interplay between customer sentiment and service quality. It suggests that customer sentiment, which is influenced by service attributes, has a direct impact on customer loyalty and retention. This strengthens the theoretical link between service quality and long-term business success, offering new insights into the causal relationships between customer sentiment, service quality, and overall business performance. Furthermore, by employing machine learning techniques to analyze customer sentiment from text data, this research introduces a novel approach to the integration of traditional service quality frameworks with modern AI methodologies. The use of sentiment analysis and feature engineering techniques offers new theoretical perspectives on how service attributes can be quantified and evaluated in the context of customer reviews. This approach opens new avenues for applying machine learning models to the analysis of customer sentiment in other service sectors, thus bridging the gap between service quality theory and practical AI applications in customer experience management.

This study also provides a foundation for future research in the development of more nuanced models for understanding customer sentiment in service industries. Researchers can use these insights to refine and expand the SERVQUAL framework, possibly creating a model that captures the complexities of customer sentiment more accurately. The study also lays the groundwork for future investigations into sentiment prediction models and their broader applications in understanding customer behavior across various industries, particularly in the digital and online platform contexts.

Conclusion

This study examined the impact of SERVQUAL dimensions—Assurance, Reliability, Tangibles, Empathy, and Responsiveness—and review scores on customer sentiment in UberEats reviews. Using a dataset of 920,407 reviews, we applied advanced text mining, sentiment analysis, and feature engineering techniques to analyze customer feedback. Sentiment prediction using a Long Short-Term Memory (LSTM) model achieved high accuracy (89.64%), supporting the validity of sentiment classification based on review scores. The findings indicate that all five SERVQUAL dimensions have a significant association with customer sentiment, with Empathy emerging as the most influential factor, followed by Responsiveness, Reliability, Tangibles, and Assurance. Additionally, review scores were identified as the strongest predictor of sentiment,

underscoring the role of customer ratings in shaping overall service perceptions. These insights highlight key areas where service providers can enhance customer experiences by focusing on service attributes that contribute most to positive sentiment. This study contributes to the growing literature on sentiment analysis and service quality by demonstrating how deep learning techniques can effectively analyze large-scale customer reviews.

Limitations and future research directions

While this study provides valuable insights into the relationship between SERVQUAL dimensions, review scores, and customer sentiment, there are several limitations that warrant consideration. First, the sentiment classification was based on a simple mapping approach that associates review scores with sentiment labels (negative, neutral, and positive). Although this method was validated with an LSTM model, it does not account for potential variations in sentiment that could arise from other factors, such as the context of the review or the intensity of the language used. Future research could explore more sophisticated sentiment analysis techniques, such as fine-grained sentiment classification or multi-class models, to capture these nuances. Additionally, the study focused solely on the UberEats platform, which may limit the generalizability of the findings to other food delivery services or industries. Future studies could extend this research to include reviews from other platforms like DoorDash, Grubhub, or Postmates to assess whether the findings hold true across different contexts. Another limitation is the reliance on keyword-based feature engineering to extract the SERVQUAL dimensions from the review text. While this approach is effective, it may not capture all relevant aspects of service quality, particularly those expressed indirectly or in less explicit terms. Future research could integrate more advanced natural language processing (NLP) techniques, such as transformer-based models, to more accurately capture the underlying themes of customer feedback. Furthermore, this study focuses on LSTM for sentiment analysis, achieving the targeted accuracy. However, LSTMs are computationally intensive. Future research could explore more efficient models like Transformers. Additionally, a comparative analysis with other machine learning models could provide further insights. Expanding the dataset beyond Google Play Store reviews and incorporating multimodal analysis could enhance service quality assessment. Moreover, while this study contributes to the understanding of how service attributes and review scores impact customer sentiment in the context of UberEats, there are several avenues for further exploration. By addressing these

limitations and expanding the scope of the research, future studies can provide even deeper insights into how businesses can leverage customer feedback to improve service quality and enhance customer satisfaction.

ORCID iD

Md Shamim Hossain  <https://orcid.org/0000-0003-1645-7470>

Ethical statement

This research does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

We give our consent for the publication.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by Széchenyi István Egyetem, Győr, Hungary, under Publication Grant Application as part of university publication support program regarding the coverage of APC. Furthermore, university affiliation and publication platforms shall support indexing of this article in bibliographic profiles and records.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Adak A, Pradhan B and Shukla N. Sentiment analysis of customer reviews of food delivery services using deep learning and explainable artificial intelligence: systematic review. *Foods* 2022; 11(10): 1500.
2. Khan FM, Khan SA, Shamim K, et al. Analysing customers' reviews and ratings for online food deliveries: a text mining approach. *Int J Consum Stud* 2023; 47(3): 953–976.
3. Hossain MS. Textual feature engineering for purchase intent and customer satisfaction: insights from marketing 4.0 and sentiment. *Sustain Futur.* 2024; 8: 1–19. doi:[10.1016/j.sft.2024.100385](https://doi.org/10.1016/j.sft.2024.100385).
4. Parasuraman A, Zeithaml VA and Berry LL. SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *J Retailing* 1988; 64: 12–40.
5. Hossain MS, Begum H, Rouf MA, et al. Investigation and prediction of users' sentiment toward food delivery apps applying machine learning approaches. *J Contemp Mark Sci* 2023; 6(2): 109–127.
6. Bannor RK and Amponsah J. The emergence of food delivery in Africa: a systematic review. *Sustain Technol Entrep* 2024; 3: 100062.

7. Dogra N, Adil M, Sadiq M, et al. Unraveling customer repurchase intention in OFDL context: an investigation using a hybrid technique of SEM and fsQCA. *J Retail Consum Serv* 2023; 72: 103281.
8. Bilal M, Marjani M, Hashem IAT, et al. Profiling reviewers' social network strength and predicting the "Helpfulness" of online customer reviews. *Electron Commer Res Appl* 2021; 45: 101026.
9. Chang V, Liu L, Xu Q, et al. An improved model for sentiment analysis on luxury hotel review. *Expert Syst.* 2023; 40: 1–29. doi:[10.1111/exsy.12580](https://doi.org/10.1111/exsy.12580).
10. Alves Gomes M and Meisen T. A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. *Inf Syst E-Bus Manage* 2023; 21(3): 527–570.
11. Wang WM, Wang JW, Li Z, et al. Multiple affective attribute classification of online customer product reviews: a heuristic deep learning method for supporting Kansei engineering. *Eng Appl Artif Intell* 2019; 85: 33–45.
12. Razali MN, Manaf SA, Hanapi RB, et al. Enhancing minority sentiment classification in gastronomy tourism: a hybrid sentiment analysis framework with data augmentation, feature engineering and business intelligence. *IEEE Access* 2024; 12: 49387–49407.
13. Kumar A, Gopal RD, Shankar R, et al. Fraudulent review detection model focusing on emotional expressions and explicit aspects: investigating the potential of feature engineering. *Decis Support Syst* 2022; 155: 113728.
14. Kuppusamy S and Thangavel R. Deep non-linear and unbiased deep decisive pooling learning-based opinion mining of customer review. *Cogn Comput* 2023; 15(2): 765–777.
15. Kim J, Lee H and Lee H. Mining the determinants of review helpfulness: a novel approach using intelligent feature engineering and explainable AI. *Data Technol Appl* 2023; 57(1): 108–130.
16. Kalantarzadeh Tezerjany SF. Appraise the role of novelty-seeking on consumers' satisfaction using online food delivery applications. *Int J Qual Reliab Manag* 2024; 41(4): 1142–1164.
17. Alghamdi SY, Kaur S, Qureshi KM, et al. Antecedents for online food delivery platform leading to continuance usage intention via e-word-of-mouth review adoption. *PLoS One* 2023; 18: e0290247.
18. Sangpikul A. Acquiring an in-depth understanding of assurance as a dimension of the SERVQUAL model in regard to the hotel industry in Thailand. *Curr Issues Tour* 2023; 26: 347–352.
19. Zygiaris S, Hameed Z, Ayidh Alsubaie M, et al. Service quality and customer satisfaction in the post-pandemic world: a study of Saudi auto care industry. *Front Psychol* 2022; 13: 842141.
20. Meesala A and Paul J. Service quality, consumer satisfaction and loyalty in hospitals: thinking for the future. *J Retail Consum Serv* 2018; 40: 261–269.
21. Jain NK, Singh AK and Kaushik K. Evaluating service quality in automobile maintenance and repair industry. *Asia Pac J Mark Logist* 2020; 32(1): 117–134.
22. de Kervenoael R, Hasan R, Schwob A, et al. Leveraging human-robot interaction in hospitality services: incorporating the role of perceived value, empathy, and information sharing into visitors' intentions to use social robots. *Tour Manag* 2020; 78: 104042.
23. Hossain MS and Rahman MF. Detection of potential customers' empathy behavior towards customers' reviews. *J Retail Consum Serv* 2022; 65: 102881.
24. Zhang Y, Zhang L, Zhang X, et al. Hospital service quality and patient loyalty: the mediation effect of empathy. *J Bus Ind Mark* 2018; 33(8): 1176–1186.
25. Lv X, Yang Y, Qin D, et al. Artificial intelligence service recovery: the role of empathic response in hospitality customers' continuous usage intention. *Comput Human Behav* 2022; 126: 106993.
26. Zhou W, Xiao S, Feng C, et al. Measuring the quality of mental health services from the patient perspective in China: psychometric evaluation of the Chinese version of the World Health Organization responsiveness performance questionnaire. *Glob Health Action* 2022; 15(1): 2035503.
27. Pashchenko Y, Rahman MF, Hossain MS, et al. Emotional and the normative aspects of customers' reviews. *J Retail Consum Serv* 2022; 68: 103011.
28. Waluyo IG and Juwono. Sentiment analysis of negative comments on social media using Long Short-Term Memory (LSTM) method with TensorFlow framework. *International Journal of Integrative Sciences* 2023; 2(7): 1015–1030.
29. Kaur P. Sentiment analysis using web scraping for live news data with machine learning algorithms. In: *Materials Today: Proceedings*. Elsevier Ltd, 2022, pp. 3333–3341. DOI: [10.1016/j.matpr.2022.05.409](https://doi.org/10.1016/j.matpr.2022.05.409).
30. Lee CKH, Tse YK, Leung EKH, et al. Causal recipes of customer loyalty in a sharing economy: integrating social media analytics and fsQCA. *J Bus Res* 2024; 181: 114747.

Appendix

Appendix I.

Keywords for feature engineering across service attributes in UberEats reviews.

Assurance	Reliability	Tangibles	Empathy	Responsiveness
"trustworthy", "reliable", "secure", "credible", "dependable", "professional", "confident", "guaranteed", "safe", "certified", "supportive", "authentic", "loyal", "sincere", "verified", "backed", "insured", "trusted", "consistent", "robust", "solid", "honest", "ethical", "dedicated", "caring", "accurate", "committed", "assured", "seamless", "smooth", "easy", "effortless", "positive", "skilled", "knowledgeable", "understanding", "mindful", "assuredly", "considerate", "helpful", "polite", "clear", "transparent", "upfront", "customer-first", "quality-focused", "elite", "courteous", "accountable", "attentive"	"punctual", "consistent", "steady", "efficient", "reliable", "seamless", "smooth", "flawless", "timely", "fast", "available", "responsive", "accurate", "predictable", "committed", "focused", "uninterrupted", "organized", "practical", "exact", "swift", "hassle-free", "orderly", "structured", "scheduled", "error-free", "honest", "proven", "maintained", "supported", "coordinated", "dependable", "top-notch", "trustworthy", "dedicated", "effort-free", "on-time", "unwavering", "guaranteed", "efficient-service", "predictable-service", "timely-response", "dependable-service", "reliable-delivery", "constant", "smooth-operations", "assured", "unfailing", "resilient", "responsible", "functional"	"fresh", "clean", "attractive", "appealing", "appetizing", "tasty", "delicious", "high-quality", "neat", "modern", "well-presented", "hygienic", "organized", "premium", "visible", "tangible", "touchable", "aesthetic", "beautiful", "colorful", "rich", "tempting", "stylish", "efficient", "updated", "sleek", "fancy", "alluring", "superior", "structured", "tasteful", "elegant", "amazing", "visually-pleasing", "designed", "balanced", "full", "intact", "curated", "trendy", "polished", "premium-grade", "sumptuous", "spotless", "innovative", "refined", "vibrant", "mouthwatering", "flavorful", "luxurious"	"caring", "understanding", "considerate", "compassionate", "helpful", "supportive", "thoughtful", "kind", "warm", "accommodating", "approachable", "empathetic", "humane", "respectful", "mindful", "patient", "generous", "polite", "responsive", "personal", "tailored", "flexible", "heartfelt", "appreciative", "welcoming", "dedicated", "inclusive", "courteous", "open-minded", "proactive", "personal-touch", "kindhearted", "personable", "service-oriented", "sincere", "genuine", "attentive", "close", "cooperative", "familial", "trustworthy", "gentle", "humble", "reassuring", "nurturing", "attuned", "warm-hearted", "tender", "sympathetic", "forgiving"	"prompt", "quick", "instant", "fast", "responsive", "ready", "available", "attentive", "alert", "speedy", "immediate", "swift", "diligent", "on-time", "proactive", "agile", "accessible", "resourceful", "rapid", "connected", "active", "reactive", "supportive", "flexible", "timely", "hyper-responsive", "informed", "forward-thinking", "prepared", "engaging", "prompt-response", "courteous", "reachable", "efficient", "adaptive", "thoughtful", "fast-moving", "detail-oriented", "communicative", "reliable", "competent", "mindful", "effective", "sharp", "urgent", "alert-minded", "on-demand", "immediate-reply", "hands-on", "proactive-response"