

RESEARCH ARTICLE

Data-driven quality management applying sentiment analysis and statistical process control: A Malaysian restaurant case study

Jon Vince Lim¹, Chai Jian Tay^{1*} and Zhen Dong Hee²

¹Centre for Mathematical Sciences, Universiti Malaysia Pahang Al-Sultan Abdullah, Lebu Persiaran Tun Khalil Yaakob, 26300 Kuantan, Pahang, Malaysia

²UCAN Technology Sdn Bhd, Persiaran Bandar Utama, Bandar Utama, 47800 Petaling Jaya, Selangor, Malaysia

Abstract - Despite the growing availability of customer feedback data, restaurant chains often face challenges in translating customer feedback into actionable quality improvements. To address this problem, this research aims to investigate operational process inefficiencies and food quality-related aspects of customer satisfaction. For this purpose, customer review data and operational activity data are collected from the central kitchen of a Malaysian restaurant chain. The data are first pre-processed through cleaning, transformation, and dimensionality reduction. Tokenization, stemming and part of speech tagging are applied to the customer review data. Subsequently, we utilize the BERT model, PyABSA model and the exponentially weighted moving average (EWMA) control chart to achieve the objectives. The performance of the BERT and PyABSA models is then compared. The sentiment analysis results are visualized in bar chart, Pareto chart and cause-and-effect diagram. The evaluation results indicate that the PyABSA model achieves higher accuracy and weighted F1 score compared to the BERT model. The findings suggest that most of the food quality complaints are related to duck, meat, and sauce products. Next, the EWMA control chart is analysed using metrics including the out-of-control percentage, in-control percentage, average run length, and identified out-of-control points. Based on the EWMA analysis, specific out-of-control dates are identified, highlighting the need for targeted strategies to improve operational processes during these periods. The results of this study provide valuable insight for the restaurant chain to improve operational processes and effectively address customer complaints related to food quality.

Article History

Received : 23 February 2026
Revised : 7 April 2026
Accepted : 18 June 2026
Published : 30 September 2026

Keywords

Food quality control
BERT model
PyABSA model
Pareto chart
Cause-and-Effect diagram
EWMA control chart

1. Introduction

A central kitchen is a facility in which food production companies consolidate food preparation and storage processes. Centralizing storage, preparation, and processing activities in a central kitchen helps improve product quality and consistency [1]. Consequently, the central kitchen production concept enables large-scale food production that is aligned with contemporary purchasing behaviors in Malaysia [2]. In restaurant operations, the central kitchen handles key food production processes, including raw material preparation, cooking, portioning, and packaging [1]. To ensure procedural compliance and consistent food quality, statistical process control (SPC) can be used to visually assess and monitor operational performance [3]. SPC is a statistical quality control technique used to monitor process performance over time, supporting continuous quality improvement. In the context of the restaurant industry, SPC offers significant benefits by maintaining high standards of food quality through improved resource management and process optimization [3]. The most widely used SPC tool is the control chart. These charts help to reveal the stability and variability of operational processes, identifying in-control and out-of-control variations by tracking process performance over time. Previous research has demonstrated that control chart can be applied beyond manufacturing to evaluate customer satisfaction, monitor quality, and support continuous improvement of operational procedures [4-7]. For example, [5] utilized individual and moving range charts for analyzing continuous data, given the small sample size and data collection constraints. More recent studies have employed exponentially weighted moving average (EWMA) control charts to effectively detect anomalies in production processes [8].

Sentiment analysis is commonly used to analyse customer reviews to improve customer satisfaction. The techniques used in sentiment analysis include bidirectional encoder representations from transformers (BERT) [9-10], graph convolutional network (GCN) [9], valence aware dictionary for emotion reasoning (VADER) [11] and decoding-enhanced BERT with disentangled attention (DeBERTa)-based model [12-13]. For example, [9] implemented both a BERT-based and a GCN-based aspect-based sentiment analysis model to extract sentiments from customer reviews of restaurants. The BERT-based sentiment analysis is used to identify key restaurant aspects, including tastiness, service, location, price, and atmosphere. On the other hand, the GCN-based sentiment analysis is employed to reveal aspect-level sentiments and to verify the accuracy of the BERT-based model. Additionally, the algorithm is utilized to perform robustness checks on prediction performance [9]. In addition, some studies incorporated machine learning [14], deep learning [15] and artificial intelligence [16, 17] in sentiment analysis to improve the precision of the model.

Pareto analysis is a statistical technique that identifies the most frequent problems or causes in a process. In the restaurant industry, it assists in identifying the key factors contributing to operational inefficiencies in the central kitchen by ranking them in descending order of occurrence. Several researchers have applied Pareto analysis to investigate issues affecting customer satisfaction [18, 19]. For example, [19] used a Pareto chart to analyze the frequency of problems affecting the quality of the KsTU restaurant, enabling researchers to focus on the most significant causes. Customer complaints can be classified and compiled into an issue list for Pareto analysis using sentiment analysis results. The analysis results are visualized as a descending frequency distribution, with problem occurrence percentages presented on the Pareto chart. Cause-and-effect analysis helps manage operational processes by systematically identifying all potential causes and their effects across different categories. This analysis is often visualized as a “fishbone” diagram, with the problem as the “head,” cause categories as the main “bones,” and detailed causes as branches. Multiple contributing factors should be considered for each problem, rather than relying solely on the primary causes identified through Pareto analysis. Several researchers have applied cause-and-effect analysis to examine problems affecting customer satisfaction [5, 19]. For example, [5] used cause-and-effect analysis to investigate restaurant processes and determine the root causes of problems previously detected in their study. They collected and organized potential causes based on data and findings from earlier stages of the study. Their study demonstrated the importance of systematic data evaluation to identify the main contributors to operational issues.

This study seeks to improve food quality in central kitchen operations while enhancing customer satisfaction in a Malaysian restaurant chain. To the best of the authors’ knowledge, the adoption of sentiment analysis and SPC tools within Malaysian restaurant chain remains limited. This study aims to address this gap by providing insights into improving restaurant operational performance through the integration of SPC tools and sentiment analysis. This paper starts with the BERT and PyABSA models to analyse sentiments in customer reviews. The performance of these two models is then evaluated. Subsequently, bar chart, Pareto chart and cause-and-effect diagram are illustrated to explore customer satisfaction based on sentiment-derived keywords. Finally, after determining the food categories to examine, an EWMA control chart is used to evaluate the operational processes influencing food quality in the central kitchen. The insights provided in this study could assist the restaurant enhance both operational performance and customer satisfaction. This paper is organized as follows: Section 2 presents the methodology, Section 3 reports the results and discussion, and Section 4 concludes the study.

2. Methodology

2.1 Data Collection

The dataset used in this study is provided by a Malaysian restaurant chain operating 18 outlets, whose identity is kept confidential. The data comprises time-spent records for central kitchen packing and delivery activities across all 18 outlets from 1 August 2024 to 31 March 2025. In addition, customer reviews recorded from 1 August 2024 until 1 March 2025 are used for the sentiment analysis. Table 1 shows the variables in central kitchen production dataset, while Table 2 list the variables in customer reviews dataset.

Table 1. Variables in central kitchen production dataset

Variables	Descriptions
order_date	The date and time for receiving order from the outlets.
outlet	The masked name of outlet kitchen.
item_name	The name of the food items.
item_category_name	The category name of the food items.
order_qty	The quantity in each time of order in integer.
item_uom	The unit of measure for the food items.
packing_datetime	The date and time for starting packing process.
packing_complete_datetime	The date and time for completing packing process.
deliver_datetime	The date and time for starting delivery process.
received_datetime	The date and time when the outlet kitchen received the food items.

Table 2. Variables in customer reviews dataset

Variables	Descriptions
id	The unique values for the restaurant review data in string.
user_id	The unique values for the registered users in string.
created_date	The date and time for creating the review data.
service_type	The types of received food service with “DineIn” and “Others” (takeout orders placed through the app) categories.
food_compliment_code	The text key compliment codes obtained from the comments related to dishes section.

comment_for_food_compliment	The text comments from customers for reviewing the dishes section.
food_improvement_code	The text key improvement codes obtained from the comments related to dishes section.
comment_for_food_improvement	The text comments from customers for suggesting improvement for the dishes section.
service_compliment_code	The text key compliment codes obtained from the comments related to service section.
comment_for_service_compliment	The text comments from customers for reviewing the service section.
service_improvement_code	The text key improvement codes obtained from the comments related to service section.
comment_for_service_improvement	The text comments from customers for suggesting improvement for the service section.
title_for_overall	The rating category to overall food service with 5 scales, from “Excellent” to “Very Poor” categories.
comment_for_overall	The text comments from customers for reviewing the overall food service.

2.2 Data Preparation

In data preparation stage, we perform data cleaning, data transformation, and data reduction by using Python. First, duplicate data is removed to reduce redundancy and to avoid the risk of overfitting in the trained models. To handle missing values, numerical attributes are imputed using the mean, median, or most probable values, as appropriate. For missing values in string-type attributes, such as those in the customer reviews dataset, records with missing data are removed, as imputation techniques are not applicable, except in cases where customer reviews are available in other columns. In data transformation stage, feature scaling and feature encoding are employed. To enable data manipulation and date arithmetic, date and time data types must be converted into appropriate date-time formats. The Python *datetime* package is used to manipulate date and time objects, facilitating date arithmetic, date formatting, and the parsing of date strings within the dataset. Feature scaling is performed using the *MinMaxScaler* package in Python to normalize the time spent on operations, which is calculated as the difference between the completion time and the start time. Normalization rescales the attributes to a range of 0 to 1, making it suitable for variables used in the EWMA control chart. To perform sentiment analysis, the customer reviews and complaints data are transformed using text preprocessing techniques, including tokenization, stemming, and lemmatization. Local sentiment aggregation technique is used to compute the weights of token-aspect pairs while distinguishing sentiment-related features from positional information, thereby reducing the influence of domain-specific noise. As the review data contains multiple languages, a translation process is also applied to standardize the customer reviews prior to analysis. For this purpose, a neural machine translation (NMT) approach, specifically the Google Translate API is employed. Although NMT may introduce potential translation artifacts, several studies have demonstrated its reliability in cross-lingual sentiment analysis [20, 21]. One advantage of Google Translate NMT is its ability to process complex sentence structures and capture subtle linguistic nuances. Furthermore, the experimental results reported in [21] indicate that sentiment analysis on translated text can achieve accuracy exceeding 86% when using a combination of the proposed ensemble model and the Google Translate API.

Negation and sarcasm are common in customer reviews. In this study, negation is addressed at the token level through the model's self-attention mechanism. The PyABSA model calculates the contextual relationship between a negation cue and the sentiment descriptor. This allows the model to correctly identify the aspect-specific sentiment. Once the sentiment for each individual aspect is determined, the model aggregates these values to calculate the dominant sentiment of the overall review, ensuring that negated expressions are correctly reflected in the final classification. In contrast, sarcasm is inherently more complex and is typically handled more effectively by fine-tuned models. However, the fine-tuning phase is not included in this study due to the presence of text-rating review discrepancy (as discussed in Section 2.4.2). Therefore, sarcasm detection is not a primary focus; instead, this study emphasizes identifying key aspects of customer complaints.

2.3 Data Reduction

In data reduction stage, we reduce the dimensionality of the data by removing weakly relevant or unnecessary columns from the datasets. The variables selected for developing statistical process control model comprise 8 variables, namely outlet, item_name, item_category_name, total_time_minutes, order_date_only, packing_date_only, packing_time_minutes_norm and delivery_time_minutes_norm. On the other hand, 4 variables—service_type, overall_rating, sentiment, and text are chosen for developing the sentiment analysis model.

2.4 Modelling

2.4.1 Statistical process control model

We employ an EWMA control chart model to monitor the operational processes in the central kitchen for food quality control. For this purpose, daily average time spent on packing and delivery is calculated from the collected data. In the EWMA control chart, upper control limit (UCL) and lower control limit (LCL) are calculated by using Equations (1) and

(2). In Equation (1), $\bar{x}$ denotes the mean, L is control limit multiplier, SE is standard error (defined in Equation (3)), σ is standard deviation, FA is factor adjustment (given in Equation (4)), i is index and λ is weighting constant.

$$UCL = \bar{x} + (L * SE) \quad (1)$$

$$LCL = \bar{x} - (L * SE) \quad (2)$$

$$SE = \sigma \sqrt{\frac{\lambda}{2 - \lambda}} * FA \quad (3)$$

$$FA = \frac{1 - (1 - \lambda)^{2(i+1)}}{2 - \lambda} \quad (4)$$

When the control chart is in stable period after a certain point, asymptotic control limits are applied to avoid having different limits at each point. The asymptotic standard error (ASE) is calculated using Equation (5). It is then used to compute the asymptotic control limits—namely, the asymptotic upper control limit and asymptotic lower control limit as shown in Equations (6) and (7). The formulas for the EWMA plotting statistics used to construct the control chart are given in Equations (8) and (9). In Equations (8) and (9), Z_0 represents initial EWMA point value, Z_i denotes EWMA point value at day i and X_i is the time spent (in packing or delivery process) at day i .

$$ASE = \sigma \sqrt{\frac{\lambda}{2 - \lambda}} \quad (5)$$

$$\text{Asymptotic } UCL = \bar{x} + (L * ASE) \quad (6)$$

$$\text{Asymptotic } LCL = \bar{x} - (L * ASE) \quad (7)$$

$$Z_0 = \bar{x} \quad (8)$$

$$Z_i = \lambda X_i + (1 - \lambda)Z_{i-1} \quad (9)$$

2.4.2 Sentiment analysis

In this study, we implemented both BERT and PyABSA models for sentiment analysis. The comparison between the BERT and PyABSA models serves as a direct ablation study of the attention mechanism. The BERT model serves as the baseline for sentiment analysis, and it utilizes a masked language modelling technique to learn bidirectional representations from unlabelled text. In contrast, the PyABSA model, which is based on the DeBERTa architecture, can handle string datasets without removing stop words or performing lemmatization. The model classifies sentiment into three labels—positive, neutral, and negative—based on aspect words and the overall sentiment score derived from customer reviews. After extracting and classifying aspect words using the developed pipeline, the frequency of customer complaint types is visualized in a Pareto chart. These results are then analysed with a cause-and-effect diagram to facilitate drawing conclusions and formulating suggestions for addressing the identified issues. The fine-tuning phase is excluded in the implementation of both BERT and PyABSA models primarily due to the presence of text–rating review discrepancy (TRRD) observed in the dataset [22]. In TRRD, customer ratings do not consistently reflect the sentiment expressed in the corresponding textual reviews, as illustrated in Figure 1. This inconsistency introduces label noise, which may adversely affect model learning and increase the risk of overfitting. Therefore, fine-tuning is omitted for both BERT and PyABSA models to mitigate the impact of noisy labels and ensure more reliable evaluation.

Review ID: 25
Original Text: The taste not so tasty and the food is cold.
Overall Rating: 4

Review ID: 282
Original Text: The noodles are burnt when fried
Overall Rating: 5

Figure 1. TRRD observed in the dataset.

2.5 Evaluation

The performance of the BERT and PyABSA models is assessed using accuracy and weighted F1-Score. Accuracy measures the overall proportion of correctly classified text across all sentiment classes (positive, negative and neutral). In contrast, the weighted F1-score calculates the average F1 score by assigning weights to each class based on its support, thereby considering the imbalance class. This metric is particularly suitable for imbalanced datasets, as it reflects model performance more prominently on majority classes through proportionate weighting. Equation (10) presents the formula used to calculate the accuracy. Equations (11) to (14) show the formulae used to calculate the weighted F1-Score.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (10)$$

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \quad (11)$$

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \quad (12)$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (13)$$

$$F1\ score_{weighted} = \sum_{i=1}^n F1\ score \times \frac{Number\ of\ sample\ class}{Total\ sample} \quad (14)$$

To evaluate the EWMA control chart, the trends and patterns of the data points and control limits must be examined. The pattern of data points in the EWMA chart reflects variations in the operation. Operational performance is quantified using performance metrics, including the out-of-control percentage (OCP, Equation (8)), the in-control percentage (ICP, Equation (9)), and the average run length (ARL), where runs represent the distances between violations. When a data point falls outside the range defined by the UCL and LCL, it indicates that the process performance represented by that point is out of control on that specific day, and the corresponding out-of-control dates are highlighted.

$$OCP = \frac{num_of_violation}{N} * 100 \quad (15)$$

$$ICP = \frac{N - num_of_violation}{N} * 100 \quad (16)$$

3. Results and Discussion

3.1 Evaluation of Performance of the BERT and PyABSA Models

We compared the performance of the BERT and PyABSA models in terms of accuracy and weighted F1-Score, as presented in Table 3. The results indicate that the BERT model achieved an accuracy of 0.1780 and a weighted F1-Score of 0.0837, whereas the PyABSA model attained an accuracy of 0.4735 and a weighted F1-Score of 0.5253. The use of disentangled attention yielded a 0.2955 increase in accuracy and 0.4416 increase in weighted F1-Score. This finding suggests that the BERT model performs poorly in classifying the three sentiment classes compared to the PyABSA model. In contrast, PyABSA model is capable of effectively assigning negative, positive, and neutral sentiments at the aspect level. This also highlights the advantage of incorporating aspect-based sentiment analysis techniques over general-purpose language models for fine-grained sentiment classification. Based on these performance metrics, the PyABSA model demonstrated a more robust inherent understanding of aspect-level sentiment, making it a more suitable tool for sentiment analysis. Hence, PyABSA model is selected for the sentiment analysis in Section 3.2.

Table 3. Performance comparison of BERT and PyABSA models

Model	Accuracy	Weighted F1-Score
BERT	0.1780	0.0837
PyABSA	0.4735	0.5253

3.2 Sentiment Analysis Results

Figure 2 shows the horizontal dominant sentiment bar chart. Based on Figure 2, the restaurant received 138 negative sentiments, which is higher than the number of positive (98) and neutral (28) sentiments. This indicates that most customers expressed negative sentiment in their reviews of the restaurant. This trend may suggest that the restaurant failed to meet customers' expected levels of satisfaction. Based on the distribution of sentiment scores shown in Figure 3, the most frequent sentiment scores fall within the -1.0 and 1.0 groups, indicating that most customers leave either strongly negative or strongly positive reviews.

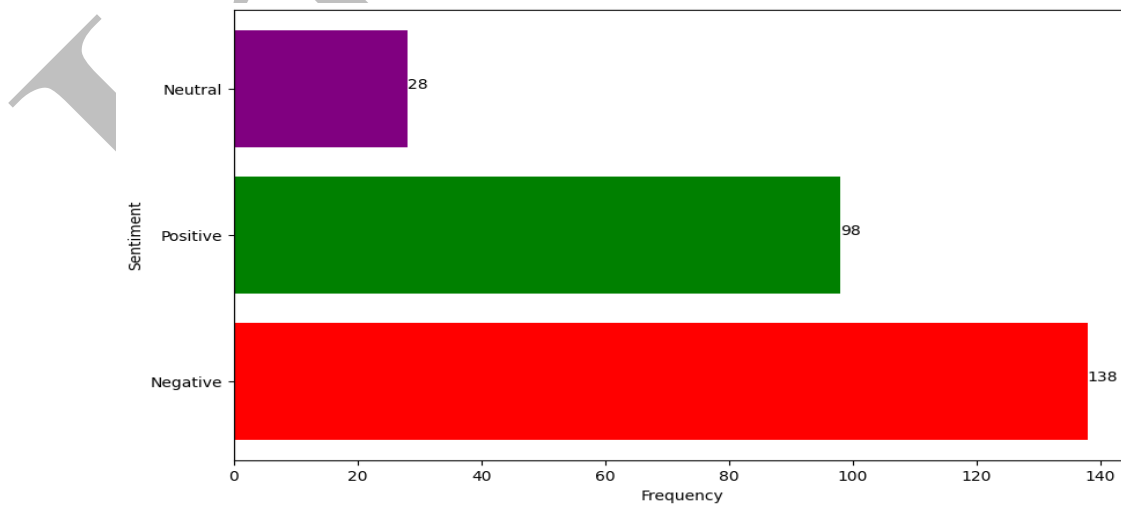


Figure 2. Horizontal dominant sentiment bar chart

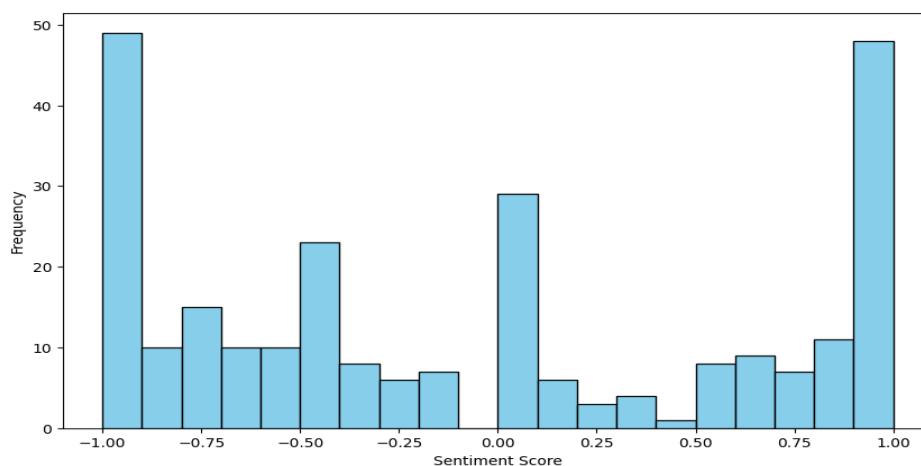


Figure 3. Distribution of sentiment scores

The Pareto chart in Figure 4 illustrates the frequency of keywords occurring in customer complaints, arranged in descending order and categorized into majority and minority groups. Based on the chart, the top three complaint keywords are “staff”, “time” and “duck”, indicating that many customers were dissatisfied with staff service, long waiting time, and the taste of the duck. According to the Pareto principle and the 80% cumulative frequency threshold, most complaint keywords extend up to “system”. The keyword “system” may suggest that customers experienced issues with the restaurant’s application system.

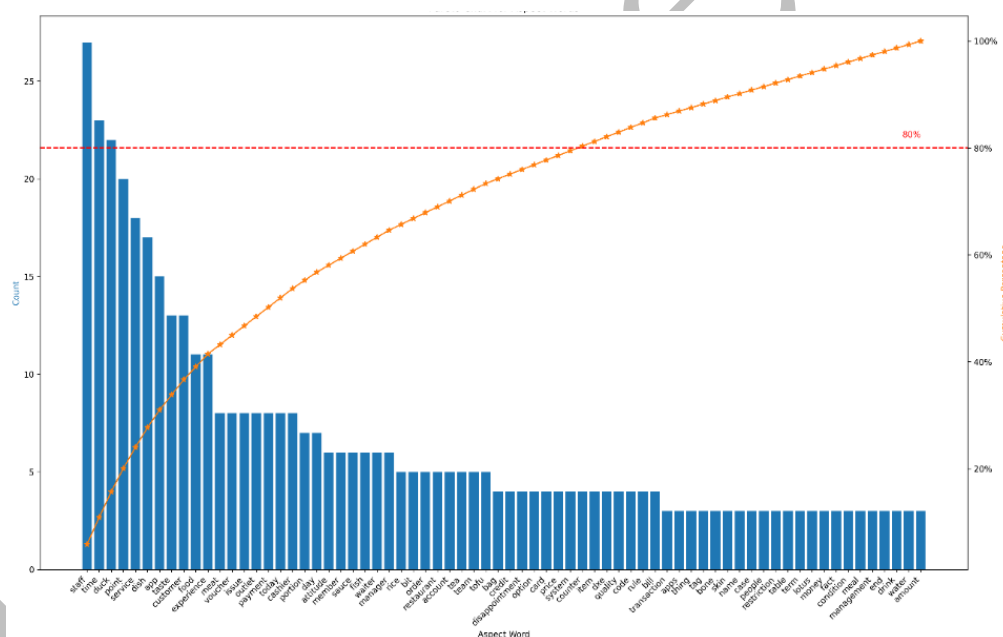


Figure 4. Frequency of keywords occurring in customer complaints

Figure 5 demonstrates the cause-and-effect diagram for major customer complaint factors. From Figure 5, several common keywords across the categories “People”, “Process & Service”, “Payment & Technology”, “Food & Menu” and “Place & Atmosphere” are identified as key factors contributing to customer complaint issues. Firstly, attention should be focused on the “Food & Menu” category, as it is most directly related to food quality and aligns with the research objectives. Based on the cause-and-effect diagram, the top three keywords in customer complaints concerning food quality are “duck”, “meat”, and “sauce”, indicating that some customers are dissatisfied with these food items. Therefore, it is necessary to examine the production processes of these items using control charts and to further investigate the underlying food quality issues. Next, the “Process & Service” category is examined to identify factors contributing to customer satisfaction issues. Among customer complaints related to processes and services, the top three keywords were “time”, “service”, and “issue”, indicating dissatisfaction with long waiting time and poor service quality. In addition, the most frequent complaints in the “People” category are “staff”, “cashier” and “attitude”, suggesting that customers are dissatisfied with the attitudes and behaviour of the restaurant’s staff and cashiers. For the remaining factors affecting customer satisfaction, we examined the “Payment & Technology” and “Place & Atmosphere” categories in the cause-and-effect diagram. In the “Payment & Technology” category, the top three complaint keywords are “point”, “app”, and “voucher”, suggesting that customers experienced issues with the points system in the restaurant’s app or are unable to

use vouchers when purchasing food. In the “Place & Atmosphere” category, the top three keywords were “experience”, “outlet” and “restaurant”, indicating that customers may have had a poor dining experience at the restaurant.

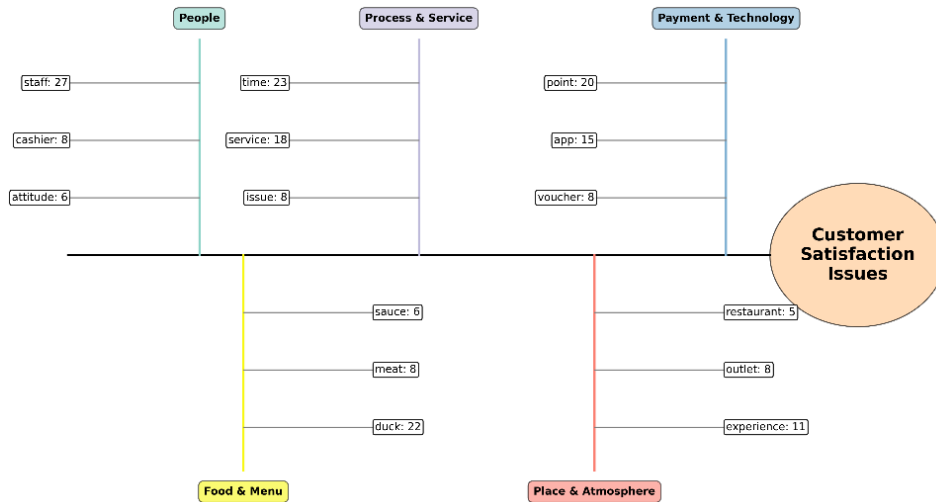


Figure 5. Cause-and-effect diagram for major customer complaint factors

3.3 EWMA Control Chart Results

In this section, we analysed the packing and delivery time spent across food categories in the operation using EWMA control charts. To detect small to moderate shifts occurring in the process average, we select a weighting constant of $\lambda = 0.20$ [6]. Figure 6 shows the EWMA control chart for packing marinated ducks, one of the primary products of the central kitchen. From Figure 6, the OCP and ICP for packing marinated ducks are 5.42% and 94.58%, respectively. This indicates that the packing process operates as expected most of the time. The ARL for packing marinated ducks is 16.83, suggesting that, on average, a process violation occurs approximately every 17 operating days. The violations in the packing process occur between 25 January 2025 and 6 February 2025, possibly due to malfunctions in the packing machine or manpower scheduling issues. It is recommended to review the incidents that occurred during this period. Figure 7 demonstrates the EWMA control chart for packing marinated meat and poultry, one of the central kitchen's primary products. Based on Figure 7, OCP and ICP for packing marinated meat and poultry are 5.83% and 94.17%, respectively. This indicates that the packing process operates as expected most of the time. The ARL for packing marinated meat and poultry is 15.77, meaning that, on average, a process violation occurs approximately every 16 operating days. The violations occurred on the following dates: 1st, 3rd, 15th, 17th, 18th, 20th, and 21st February 2025, and 1st, 4th, 5th, 13th, and 14th March 2025.

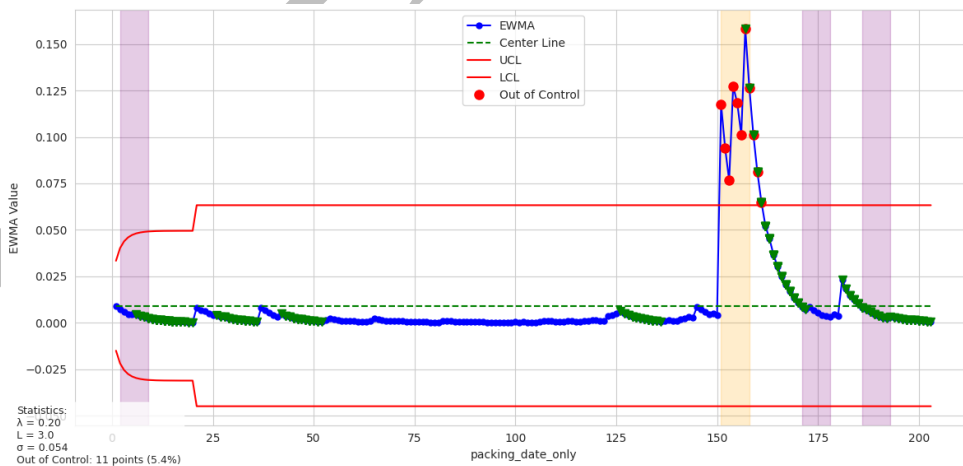


Figure 6. EWMA control chart for packing marinated ducks

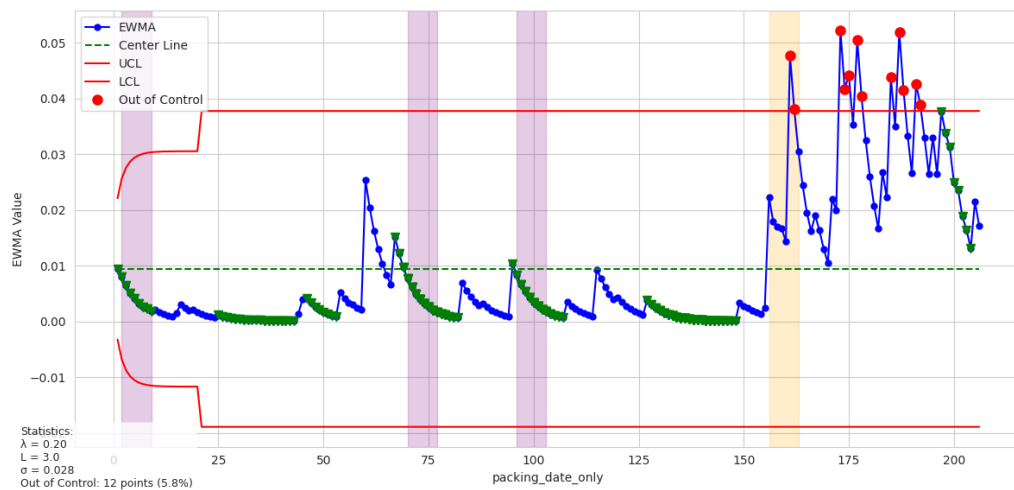


Figure 7. EWMA control chart for packing marinated meat and poultry

We also plotted the EWMA control chart for packing sauce products as shown in Figure 8. Based on Figure 8, the OCP and ICP for packing sauce products are 3.29% and 96.71%, respectively. This indicates that the packing process is generally stable, with few violations. The ARL for packing sauce products is 26.5, suggesting that, on average, a violation occurs approximately every 27 operating days. The violations occurred on the following dates: 1st, 15th, 18th, 20th, and 22nd February 2025, and 1st and 4th March 2025. This stability may be attributed to well-established packing techniques and equipment for sauce products.

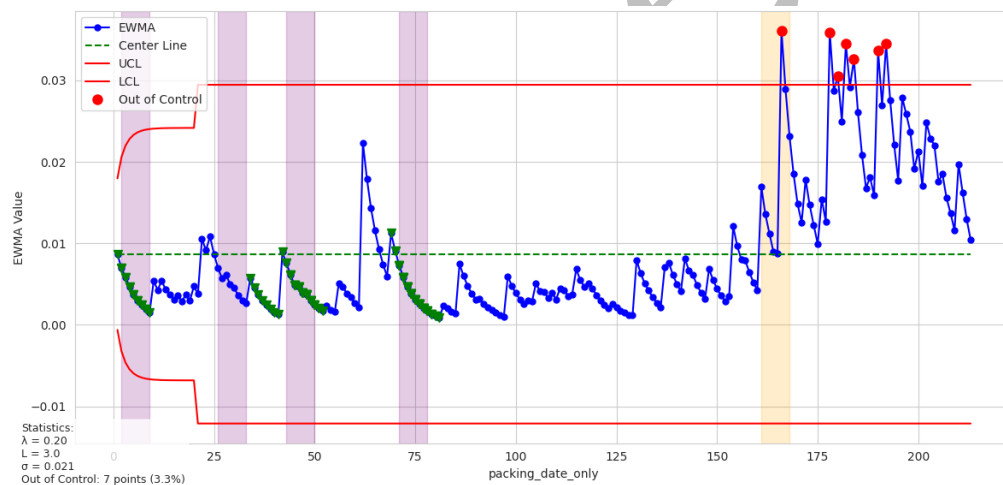


Figure 8. EWMA control chart for sauce products

After analysing the packing times for the target item categories, we also examined the delivery times from the central kitchen with the highest order volume to the outlet kitchen—specifically, Outlet 15—to investigate potential issues in the delivery process for the most popular outlet. The results of which is illustrated in Figure 9. Based on the EWMA control chart results in Figure 9, the OCP and ICP for delivery time to Outlet 15 are 4.87% and 95.13%, respectively. This indicates that the delivery process is generally stable, with a few violations. The ARL for delivery to Outlet 15 is 18.75, meaning that, on average, a violation occurs approximately every 19 operating days. These violations may be caused by factors such as traffic congestion, route construction, or driver scheduling issues. The violations occur between 10th January and 19th January 2025, with additional occurrences on 7th and 8th February, and 21st March 2025.

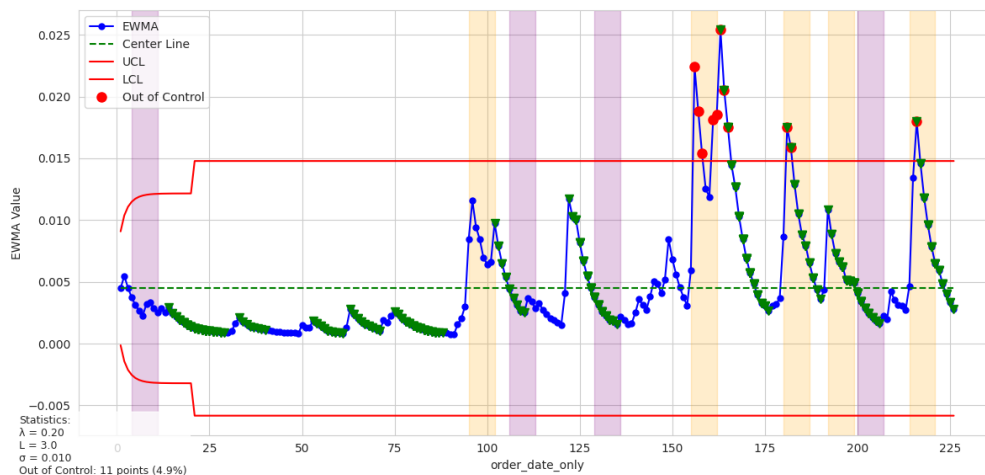


Figure 9. EWMA control chart for delivery time in outlet 15

3.4 Discussion

In this study, we applied the sentiment analysis and statistical process control to investigate the causes of customer complaints and to identify dates on which packing and delivery operations are out of control or in violation of established standards. The sentiment analysis suggests that the most common food quality complaints are related to duck, meat, and sauce products. The EWMA control chart enables restaurant managers to pinpoint specific dates for further investigation, helping them better understand the underlying causes of process deviations and implement improvements in food quality by strengthening central kitchen operating procedures. Based on the analysis results, we recommend that the restaurant first reexamine the cooking, packing, and delivery processes for duck, meat, and sauce products to ensure the freshness, taste, and overall quality of these food categories. In addition, the restaurant should address other contributing factors to customer dissatisfaction, such as long waiting time, poor service quality, and negative interactions with staff. Providing regular staff training can help ensure that employees consistently follow standard operating procedures and reduce the recurrence of complaints. Finally, improvements should also be made to the restaurant's mobile app, particularly regarding the points system and the use of vouchers during food purchases, as these issues were also highlighted in customer feedback.

Besides, the restaurant should closely investigate operations between January and March 2025, as most out-of-control dates fall within this period. For a more targeted investigation, management should review the specific events that occurred on each out-of-control date to better understand the underlying causes. This understanding will enable the restaurant to develop effective strategies and procedures to improve operational processes, ensure consistent food quality, and support sustainable business growth. One limitation of this study is that some customer reviews were translated from Mandarin into English using the Google Translate API, and the aspect word extraction and classification tasks were performed with the English-based Natural Language Toolkit (NLTK). Consequently, reviews in languages that cannot be accurately translated by Google Translate may be unprocessable. Furthermore, using automated translation may not fully capture the original sentiment of the reviews due to potential inaccuracies in the translation. For future work, customer review data should be collected following the implementation of service improvements. This would enable pre- and post-intervention comparisons of customer sentiment using appropriate statistical significance tests. Such analysis would allow for the assessment of whether observed changes in sentiment are statistically meaningful rather than due to random variation, thereby providing stronger validation of the effectiveness of the proposed PyABSA model in enhancing customer satisfaction.

4. Conclusions

In conclusion, this study employs sentiment analysis and statistical process control model to address operational process issues and food quality-related customer satisfaction for a Malaysian restaurant chain with 18 outlets. For this purpose, we implemented BERT and PyABSA models and compared the performance of these two models. The evaluation results demonstrate that the PyABSA model outperforms the BERT model, achieving higher accuracy and weighted F1 scores, thereby highlighting the advantages of aspect-based sentiment analysis. Subsequently, the PyABSA model is used to generate the sentiment analysis results. These results are visualized using bar chart, Pareto chart and cause-and-effect diagrams, facilitating the identification and investigation of food quality issues. The sentiment analysis findings suggest that the most frequent food quality complaints are related to duck, meat, and sauce products. We also applied a statistical process control model using EWMA control charts to monitor operational processes. This model allows the analysis of average packing and delivery times, and the identification of dates on which operations are out of control or non-compliant. The EWMA control charts reveals that most out-of-control dates occurred between January and March 2025, highlighting the need for targeted strategies to improve operational processes during this period.

Acknowledgement

The authors thank the Centre for Mathematical Sciences, Universiti Malaysia Pahang Al-Sultan Abdullah for providing the support. The authors also would like to thank UCAN Technology Sdn Bhd for providing the data for this study.

Funding

This research received no specific grant from any funding agency in the public, private, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare no conflict of interest.

CRedit Authorship Contribution Statement

Jon Vince Lim (Conceptualization; Methodology; Data curation; Investigation; Software; Formal analysis; Visualization; Writing – original draft; Writing – review and editing; Project administration)

Chai Jian Tay (Conceptualization; Methodology; Supervision; Writing – review and editing)

Zhen Dong Hee (Conceptualization; Resources; Data curation; Writing – review and editing)

Availability of the Data and Materials

The data used to support the findings of this study are included within the article.

Ethical Declarations

No artificial intelligence tools were used in the preparation of this manuscript. All content was developed manually by the authors. This study did not involve human participants or animals. Ethical approval was therefore not required.

Generative Artificial Intelligence Declarations

The authors claim that artificially intelligent-assisted technologies in the form of generative AI were not used to generate content, ideas, or theories. We have just utilised AI to enhance readability and refine the language. This was used with extreme human control and oversight. The authors take full responsibility for reviewing and approving the content.

References

- [1] Jia Y, Hu L, Liu R et al. Innovations and challenges in the production of prepared dishes based on central kitchen engineering: A review and future perspectives. *Innovative Food Science & Emerging Technologies*. 2024 Jan 1;91:103521.
- [2] Hasnan NZ, Ramli SH. Modernizing the preparation of the Malaysian mixed rice dish (MRD) with Cook-Chill Central Kitchen and implementation of HACCP. *International Journal of Gastronomy and Food Science*. 2020 Apr 1;19:100193.
- [3] Lim SA, Antony J, Albliwi S. Statistical Process Control (SPC) in the food industry—A systematic review and future research agenda. *Trends in Food Science & Technology*. 2014 Jun 1;37(2):137-51.
- [4] Aslam MU, Xu S, Noor-ul-Amin M, Hussain S, Waqas M. Fuzzy control charts for individual observations to analyze variability in health monitoring processes. *Applied Soft Computing*. 2024 Oct 1;164:111961.
- [5] Mirzaei N, Niroomand S, Zare R. Application of statistical process control in service industry: A case study of the restaurant sector. *Journal of Modelling in Management*. 2016 Aug 8;11(3):763-82.
- [6] Sanaullah A, Chaudhary AM, Hanif M, Sharma P. EWMA control charts based on robust estimators: A powerful tool for monitoring a process with a non-normal distribution. *Journal of Computational and Applied Mathematics*. 2024 Oct 15;449:115919.
- [7] Xue L, Wang Q, An L, He Z, Feng S, Zhu J. A nonparametric adaptive EWMA control chart for monitoring mixed continuous and categorical data using self-starting strategy. *Computers & Industrial Engineering*. 2024 Feb 1;188:109930.
- [8] Mishra AK, Das B. A scheme for electricity theft detection based on EWMA control chart. *Electric Power Systems Research*. 2024 May 1;230:110277.
- [9] Li H, Bruce XB, Li G, Gao H. Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews. *Tourism Management*. 2023 Jun 1;96:104707.
- [10] Almalki AJ. Enhanced sentiment analysis framework: Ensemble attention enhanced bidirectional long-short-term encoder for accurate classification of consumer reviews. *Alexandria Engineering Journal*. 2025 Aug 1;127:265-83.
- [11] Pleerux N, Nardkulpat A. Sentiment analysis of restaurant customer satisfaction during COVID-19 pandemic in Pattaya, Thailand. *Heliyon* 9 (11): e22193 [Internet]. 2023.
- [12] Yang H, Zhang C, Li K. Pyabsa: A modularized framework for reproducible aspect-based sentiment analysis. In *Proceedings of the 32nd ACM international conference on information and knowledge management* 2023 Oct 21 (pp. 5117-5122).
- [13] Vybhavi B, Karthika S, Vadivel R. Sentiment Analysis and Opinion Mining using Decoding Enhanced BERT with Disentangled Attention and Light Gradient Boosting Machine. In *2025 3rd International Conference on Data Science and Information System (ICDSIS)* 2025 May 16 (pp. 1-6). IEEE.
- [14] Wang X, Feng Y, Wang Y, Zhu H, Song D, Shen C, Luo Y. Enhancing optical non-destructive methods for food quality and safety assessments with machine learning techniques: A survey. *Journal of Agriculture and Food Research*. 2025 Mar 1;19:101734.

- [15] Khadivar A. Utilizing Deep Learning for Aspect-Based Sentiment Analysis in Restaurant Reviews. *Journal of Information Technology Management*. 2025 Jun 1;17(2):50-68.
- [16] Ziani I, Bouakline H, El Guerraf A, El Bachiri A, Fauconnier ML, Sher F. Integrating AI and advanced spectroscopic techniques for precision food safety and quality control. *Trends in Food Science & Technology*. 2025 Feb 1;156:104850.
- [17] Naseem S, Rizwan M. The role of artificial intelligence in advancing food safety: A strategic path to zero contamination. *Food Control*. 2025 Sep 1;175:111292.
- [18] Chirinza N, Muguirrima P, Zerpa FA, Martin AR, Pino CA. Resources optimization using Pareto analysis for sea water desalination plants. *Cleaner Water*. 2024 Dec 1;2:100054.
- [19] Yeboah MA, Anyan FY, Abayase R. Measuring and Evaluating the Services of Kumasi Technical University Restaurant by Employing Statistical Process Control Tools. *Asian Journal of Probability and Statistics*. 2023 Oct 6;25(2):1-6.
- [20] Saadany H, Orasan C. Is it great or terrible? preserving sentiment in neural machine translation of arabic reviews. *InProceedings of the fifth Arabic natural language processing workshop 2020 Dec* (pp. 24-37).
- [21] Miah MS, Kabir MM, Sarwar TB, Safran M, Alfarhood S, Mridha M. A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and LLM. *Sci. Rep*.
- [22] Almansour A, Alotaibi R, Alharbi H. Text-rating review discrepancy (TRRD): an integrative review and implications for research. *Future Business Journal*. 2022 Feb 22;8(1):3.

Im-Press