Available Online at www.e-iph.co.uk
Indexed in Clarivate Analytics WoS, and ScienceOPEN

E-B
Environment - Behaviour
Proceedings Journal

KICSS2024
Kedah International Conference on Social Sciences:
https://kicss2024.wixsite.com/kicss2024

e-IPH
e-International Publishing House Ltd., United Kingdom

InforMaTIC2024: Information Management and Technology International Conference
Virtual Conference, UiTM Kedah, Malaysia
Organised by: Universiti Teknologi MARA, Kedah, Malaysia

InsightVista: Unveiling visitor sentiments and trends for Terengganu State Museum using text analytics

Nurul Husna Tarmizi, Nur Iman Husna A'riff, Norulhidayah Isa*, Sarah Yusoff, Nor Hasnul Azirah Abdul Hamid
**Corresponding Author*

College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Cawangan Terenggnu, Kampus Kuala Terengganu, Terengganu, Malaysia.

husnatarmizi296@gmail.com, husnariffy29@gmail.com, norul955@uitm.edu.my, sarahyusoff@uitm.edu.my,hasnulazirah@uitm.edu.my
Tel: +601112381983

Abstract
This project analyzes 1,973 visitor reviews from Google Maps and TripAdvisor using sentiment analysis and K-means clustering. With Support Vector Machine (SVM), the model achieved 87.66% accuracy and identified three main clusters. Visitors often praised the museum's architecture, while accessibility issues were a common concern. These findings suggest that improving facilities could boost inclusivity and satisfaction. An interactive dashboard was developed to present the insights, helping museum management make data-driven decisions. Despite some data limitations, the project provides actionable recommendations to enhance the overall visitor experience at the Terengganu State Museum.

Keywords: sentiment analysis, clustering, visitor experiences, CRISP-DM

DOI: https://doi.org/10.21834/e-bpj.v10iSI31.6938

1.0 Introduction

In today's digital era, museums increasingly rely on visitor feedback to enhance their services. As the largest museum in Malaysia, the Terengganu State Museum plays a vital role in preserving cultural heritage and aims to be a national leader. A successful museum visit is not just about exhibits, but about engaging experiences, staff interaction, and the surrounding environment (Falk & Dierking, 2018).

To meet evolving expectations, museums are embracing technologies such as augmented reality (AR) and virtual reality (VR) to enrich exhibits and attract tech-savvy audiences (Liao & Huang, 2019). Equally important is the role of museum staff. Friendly, well-informed interactions can significantly enhance a visit, while negative experiences can detract from even the best exhibits (Smith & McCartney, 2019). As Xu and Li (2020) emphasise, ongoing staff training—covering both content delivery and interpersonal skills—is essential for maintaining high visitor satisfaction.

Yet, to make meaningful improvements, museums must deeply understand what visitors are expressing in their reviews. Online platforms like Google Maps and TripAdvisor now serve as rich sources of feedback, offering insights into visitor preferences and expectations (Liu, Huang, & Chen, 2020). However, manually analysing thousands of reviews is time-consuming and error-prone (Li & He, 2019). To overcome this, many institutions are turning to advanced text analytics, particularly sentiment analysis and clustering, which enable the effective processing of large volumes of unstructured review data (Yang & Zhang, 2021). These techniques uncover patterns and trends that can guide decision-making.

This study aims to apply sentiment analysis and K-Means clustering to visitor reviews of the Terengganu State Museum. The objective is to uncover sentiments, key themes, and group feedback into clusters that reflect visitor experiences. These insights will support management in addressing issues, enhancing inclusivity, and improving engagement (Xu & Li, 2020). This project involves collecting and preprocessing review data, applying machine learning models, and evaluating the model's ability to interpret review content. By integrating sentiment analysis with clustering, a comprehensive view of visitor experiences can be achieved.

2.0 Literature Review

The application of text analytics to understand visitor sentiments and emerging trends in cultural institutions—particularly museums—has received growing attention in recent years. As museums adapt to the needs of a digitally engaged audience, data-driven analysis of visitor feedback has become a crucial tool for enhancing engagement and satisfaction. This section explores key aspects of text analytics in the context of visitor review analysis. Additionally, it outlines the core criteria that contribute to a positive and meaningful museum experience.

Text analytics, a branch of natural language processing (NLP), involves deriving meaningful insights from unstructured text data. In the museum context, it is beneficial to analyze visitor feedback from online reviews and surveys. As noted by Li and He (2019), text analytics allows institutions to systematically process large volumes of feedback, uncovering sentiments and trends that might otherwise go unnoticed. This approach supports a deeper understanding of visitor preferences and aids in enhancing service delivery.

Recent studies have demonstrated the value of sentiment analysis and clustering in this domain. For example, Liu et al. (2020) effectively used sentiment analysis to classify visitor reviews into positive, negative, and neutral categories, providing a clearer overview of visitor satisfaction. Similarly, Yang and Zhang (2021) applied K-Means clustering to identify key thematic areas within the feedback, enabling museums to prioritize improvements based on visitor concerns and interests.s.

3.0 Methodology

This section outlines the methodology used to analyze visitor sentiments and trends at the Terengganu State Museum, employing text analytics within the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework. This structured approach ensures a systematic workflow, encompassing phases from business understanding to model evaluation. Fig. 1 presents the conceptual framework of the project.

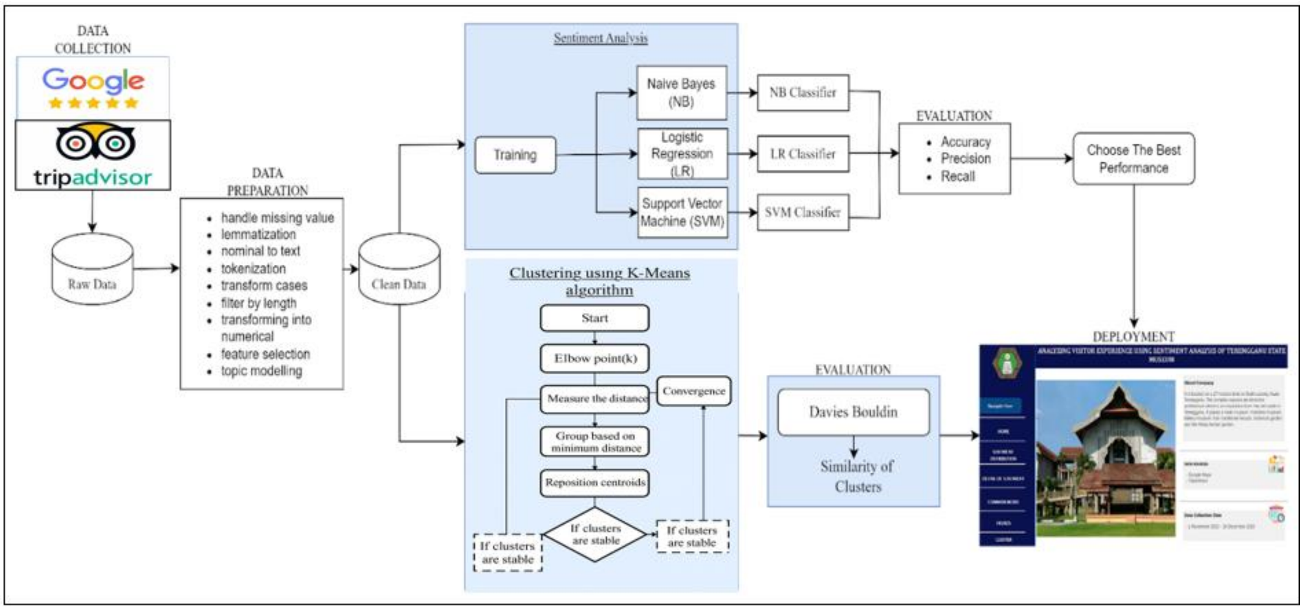

Fig. 1: Conceptual framework for InsightVista

3.1 Business Understanding

The project begins by aligning with the Terengganu State Museum's goals of enhancing visitor satisfaction and becoming a leading conservation and reference centre. It aims to extract actionable insights from visitor reviews to guide service improvements. Initial interviews with museum staff helped identify feedback management challenges and align the analysis with institutional priorities.

3.2 Data Understanding

To understand the visitor experience at the Terengganu State Museum, 1,985 reviews were collected from Google Maps and TripAdvisor over a one-month period (2 November – 26 December 2023). The dataset includes variables such as author, date, content, rating, and likes, providing a solid basis for analysis. All reviews were translated into English for consistency. Where Google Translate proved insufficient due to informal language or abbreviations, manual preprocessing was applied (e.g., "drpd" expanded to "daripada") to ensure accurate sentiment interpretation.

3.3 Data Preparation

Data preparation is essential for converting raw, unstructured text into a format suitable for analysis. The process begins with text cleaning, including the removal of special characters, numbers, and stopwords. This is followed by tokenisation and lemmatisation to standardise word forms. The text is then vectorised using TF-IDF, enabling numerical analysis. Additional steps include handling missing values and translating non-English reviews to ensure dataset consistency. Fig. 2 illustrates the preprocessing workflow.

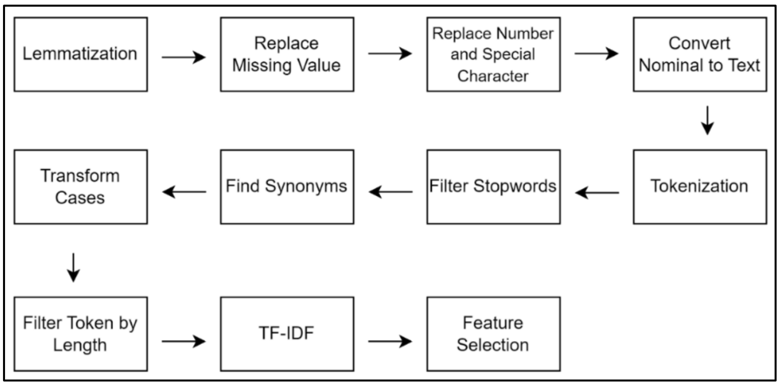

Fig. 2: Detail data preprocessing steps.

3.4 Modeling

The K-Means clustering algorithm was used to group similar reviews based on textual content. The optimal number of clusters (k) was determined using the Elbow Method, which evaluates the within-cluster sum of squares (WCSS) across different k values. As shown in Fig. 3, the elbow point appears at k = 3, indicating the most appropriate number of clusters (Singh et al., 2020). In this approach, WCSS is calculated iteratively from k = 1 to k = n, measuring the squared distance between each point and its assigned centroid. The optimal k is identified where the WCSS curve begins to flatten, suggesting minimal gain in cohesion with additional clusters. Each review is then assigned to the nearest cluster, with the algorithm updating centroids until convergence. The resulting clusters reflect distinct themes or topics within the reviews, enabling the museum to identify key areas of interest or concern among visitors.

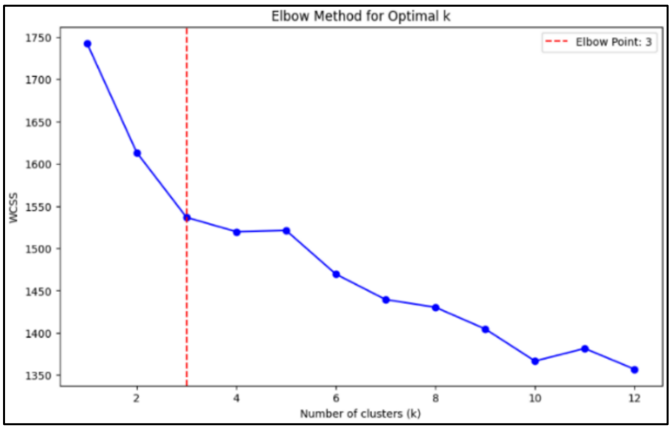

Fig. 3: Elbow Method Chart

Sentiment analysis was conducted to classify the tone of visitor reviews within each cluster as either positive or negative. This process utilized natural language processing (NLP) techniques to evaluate emotional content, offering more profound insight into visitor sentiment patterns. Three supervised machine learning algorithms were applied for training and classification: Support Vector Machine (SVM), Naïve Bayes (NB), and Logistic Regression.

SVM is well-regarded for its effectiveness in classification tasks, operating by identifying the optimal hyperplane that separates sentiment classes with maximum margin (Reddy, 2023). Naïve Bayes, grounded in Bayes' Theorem, estimates the probability of sentiment categories based on word frequency, offering a fast and efficient approach to text classification (Ashfaque, 2023). Logistic Regression models the likelihood of binary outcomes and is suitable for predicting sentiment polarity based on input features.

3.5 Evaluation

The quality of clustering was assessed using the Davies-Bouldin Index (DBI), a widely used metric for evaluating cluster validity (Roni et al., 2023). DBI measures the average similarity between each cluster and its most similar counterpart, where lower values indicate better-defined and more distinct clusters. The DBI formula is shown in Fig. 4. In addition to this quantitative evaluation, the clustering results were manually reviewed to ensure logical and meaningful groupings, thereby validating the reliability of the automated process.

$$R_{i,j,\dots,n} = \frac{SSW_i + SSW_j + \dots + SSW_n}{SSB_{i,j} + \dots + SSB_{ni,nj}}$$

$$DBI = \frac{1}{k}\sum_{i=1}^{k} max_{i \neq j}(R_{i,j,\dots,k})$$

Fig. 4 Davies-Bouldin Index Formula (Source: Wati et al., 2022)

The sentiment analysis component was validated by comparing predicted sentiments against a manually labelled subset of reviews, ensuring alignment with the actual emotional tone. To improve reliability, k-fold cross-validation was applied, where the dataset was divided into multiple folds, and the model was iteratively trained and tested using different fold combinations. This method provides a more robust and unbiased performance estimate. Model performance was measured using the accuracy metric, which reflects the proportion of correct predictions out of the total. A higher accuracy indicates better classification performance. The calculation is presented in Equation (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

Where:

TP (True Positives): The number of positive cases correctly identified by the model.
TN (True Negatives): The number of negative cases correctly identified by the model.
FP (False Positives): The number of negative cases incorrectly identified as positive by the model.
FN (False Negatives): The number of positive cases incorrectly identified as negative by the model.

3.6 Deployment

The outcomes from the modelling and evaluation phases were visualized using a dashboard developed in Microsoft Power BI, as shown in Fig. 5. This dashboard presents key insights from the clustering and sentiment analysis, enabling the museum's management to interpret results easily and make informed, data-driven decisions. By consolidating complex findings into an interactive format, the dashboard supports quicker response times and more targeted service improvements.

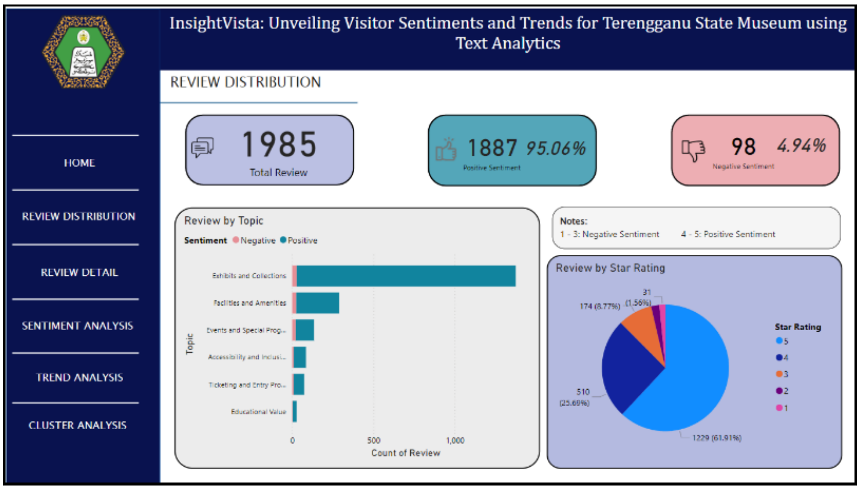

Fig. 5: InsightVista Dashboard

4.0 Findings and Discussions

This chapter presents the results and findings from the sentiment analysis conducted on visitor reviews of the Terengganu State Museum. The data was carefully evaluated to ensure consistency with the established methodology. Key patterns and themes emerging from visitor feedback are discussed, offering insights into overall experiences and perceptions. In addition, the chapter assesses the performance of the machine learning models used for sentiment classification, focusing on their effectiveness and accuracy. The analysis aims to both validate the approach and deepen understanding of visitor sentiment.

4.1 Topic Modelling

Latent Dirichlet Allocation (LDA) was used to model the topic of the visitor reviews to uncover deeper themes. This unsupervised machine learning method identifies groups of frequently co-occurring words, revealing hidden topics within large text datasets. The analysis produced six distinct topics, each representing a key aspect of the museum experience. These are summarized in Table 1, based on the dominant keywords associated with each topic cluster.

Table 1. Topics and Descriptions

Topic	Descriptions
Exhibits and Collections	Most of the reviews focused on the quality and variety exhibits and collections at the museum.
Facilities and Access	Visitors commented on the museum facilities, including parking, stairs, and accessibility for disabled visitors.
Staff and Services	Feedback on the helpfulness and friendliness of the museum staff.
Educational Value	Special events and activities offered by the museum were highlighted in the reviews.
Visitor Experience	General comments on the overall visitor experience, including the atmosphere of the museum.

To demonstrate the model's output, Fig. 6 displays selected review excerpts matched with key terms and their corresponding thematic categories. Reviews were assigned based on dominant terms identified by the LDA model and later validated manually. For instance, the term "damaged" is linked to the Facilities and Accessibility theme, reflecting concerns about maintenance. On the other hand, keywords like "program", "environment", and "informative" correspond to positive experiences involving exhibitions and events.

Category	Word	Review	Author	Sentiment	Rating
Facilities and Accessibility	damaged	many exterior part damaged look like awake yet	Hairi Jaafa	Negative	2
Events and Activities	program	lot upgrade lot new program	Nizam Mus	Positive	5
Visitor Experience	environment	learn history terengganu state place environment nice	Zan Zan	Positive	4
Exhibits and Collections	informative	beautiful architecture old style palace many historical informative item cover many aspe	IzarsZ	Positive	5
Educational Value	primary	super educational	Izzati Nuru	Positive	5
Staff and Service	friendly	beautiful place friendly staff highly recommend come lot learned	Ahmad Ru	Positive	5

Fig. 6: Example Word Based on Topic

4.2 Exploratory Data Analysis

The word clouds Fig. 7 and 8 provide a clear contrast in thematic focus between positive and negative sentiments. From a sentiment polarity perspective, terms like "beautiful" and "interactive" indicate emotionally rich, positive experiences, while words such as "maintenance" and "closed" highlight dissatisfaction. This aligns with the Expectation-Confirmation Theory, as positive feedback reflects fulfilled expectations, and negative terms indicate unmet needs, particularly in accessibility and facility maintenance.

Additionally, the recurrent appearance of terms related to architecture, culture, and exhibitions suggests that visual and educational stimuli play a key role in shaping visitor impressions, supporting insights from Dual Coding Theory. When these aspects align with visitor expectations, satisfaction is reinforced; when they fail (e.g., broken displays or poor accessibility), the same attributes become sources of frustration. Finally, drawing from the Contextual Model of Learning, the positive terms emphasise engagement within personal and sociocultural contexts. At the same time, negative experiences predominantly stem from physical context issues—signalling areas for operational improvement.

Fig. 7: Word cloud for positive sentiments

Fig. 8: Word cloud for negative sentiments

4.3 Clustering Analysis

For the clustering analysis, results were visualised using tables, word clouds, Venn diagrams, and a dashboard to enhance interpretability. Both automated and manual techniques were employed to understand the data structure better and derive new features from the clusters (Zhang et al., 2019). As shown in Table 2, each cluster includes reviews across all rating levels (1–5), though rating 5 dominates, reflecting the overall review distribution. Cluster 3 contained the most reviews, while Cluster 2 had the fewest. This imbalanced distribution made it difficult to characterise Cluster 3 specifically.

Table 2. Clustering Result

Cluster/Rating	1	2	3	4	5	Total
1	2	2	19	89	211	323
2	1	4	12	68	207	292
3	27	35	143	348	805	1358
Total	30	41	174	505	1223	1973

Fig. 9 shows a Venn diagram illustrating the distribution of shared and unique terms across the three clusters. Cluster 3 contains the highest number of unique terms (1,385), while 495 terms are common to all clusters.

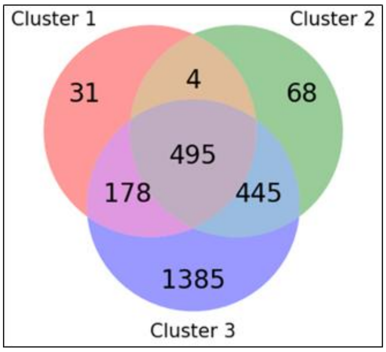

Fig. 9: Venn Diagram

Among the clusters, Cluster 3 had the highest number of reviews, while Cluster 2 had the fewest, resulting in an imbalanced distribution that made it challenging to interpret Cluster 3. To enhance understanding, word clouds (Fig. 10–12) were generated using TF-IDF values, with word size indicating importance. The thematic focus of each cluster is summarized as follows:

Fig. 10: Cluster 1

- Cluster 1 (Fig. 11) emphasizes emotion, safety, and personal comfort, with keywords like *"grateful"*, *"safe"*, and *"cheerful"*. Terms such as *"woman"* and *"heaven"* suggest a peaceful environment, possibly resonating with elderly or family-oriented visitors.
- Cluster 2 (Fig. 12) reflects event-specific and historical references, with terms like *"parrot"*, *"ruler"*, *"British"*, and *"February"*. These reviews appear linked to temporary exhibitions or seasonal events. The term *"boil"*, although unusual, may reflect a specific artefact or metaphorical description.
- Cluster 3 (Fig. 13) showcases interactivity and educational engagement, featuring terms such as *"fantastic"*, *"interactive"*, *"invaluable"*, and *"fact"*. Words like *"fruit"*, *"wood"*, and *"clean"* suggest nature-themed displays and hands-on learning activities.

These clusters offer granular insights into visitor sentiment and experiences, supporting strategic improvements in exhibit design, event planning, and audience engagement. Due to the size and diversity of Cluster 3, sub-clustering is recommended for further analysis, aligning with recommendations by Cambria et al. (2018).

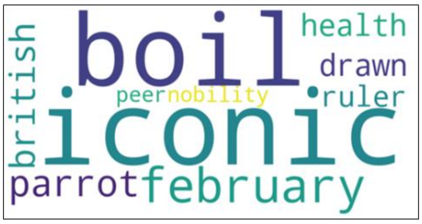

Fig. 11: Cluster 2

Fig. 12: Cluster 3

4.4 Sentiment Analysis Result

Sentiment analysis was performed on visitor reviews of the Terengganu State Museum using three widely adopted classification algorithms: Support Vector Machine (SVM), Naïve Bayes, and Logistic Regression. Model evaluation was conducted using both 5-fold and 10-fold cross-validation, with performance assessed based on accuracy. As shown in Fig. 13, SVM achieved the highest accuracy, recording 87.05% with 10-fold and 86.95% with 5-fold cross-validation, indicating strong predictive performance. In contrast, Naïve Bayes yielded the lowest accuracy, with 38.39% (10-fold) and 39.90% (5-fold). Logistic Regression performed moderately, achieving 56.37% (10-fold) and 57.53% (5-fold). These results highlight SVM as the most effective algorithm for sentiment classification in this context, particularly when evaluated using 10-fold cross-validation.

Cross Validation Method						
Model	**Folds**	**Accuracy**	**Class Precision (Positive)**	**Class Precision (Negative)**	**Class Recall (Positive)**	**Class Recall (Negative)**
Support Vector Machine (SVM)	10	87.05%	87.81%	28.00%	98.96%	2.85%
	5	86.95%	87.68%	19.05%	99.02%	1.63%
Naïve Bayes	10	38.39%	84.86%	10.76%	36.11%	54.47%
	5	39.90%	85.09%	10.77%	38.07%	52.85%
Logistic Regression	10	56.37%	88.26%	13.27%	57.91%	45.53%
	5	57.53%	88.75%	13.99%	59.00%	47.15%

Fig. 13: Performance algorithms

Further analysis was conducted on the Support Vector Machine (SVM) classification output to interpret the influence of individual features. As illustrated in Fig. 14, each bar represents an attribute, with the bar's length corresponding to the magnitude of its weight in the model. The horizontal axis displays the attributes, while the vertical axis indicates the weight values. Positive weights suggest that a feature contributes toward predicting a positive sentiment, whereas negative weights influence the prediction of a negative sentiment.

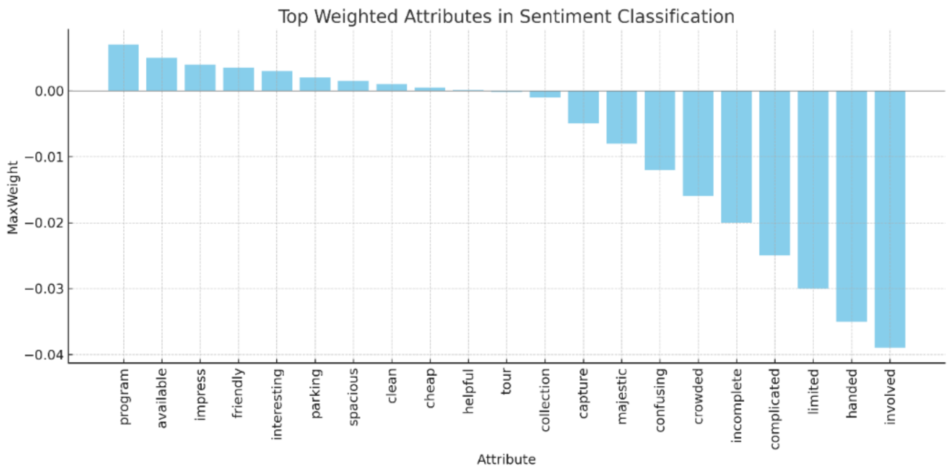

Fig. 14 Feature Importance Based on Model Weights

5.0 Conclusion and Recommendations

This project applied K-Means clustering and sentiment analysis to 1,985 visitor reviews of the Terengganu State Museum, uncovering key themes and sentiment trends. The study identified three main clusters and six topics related to various aspects of the visitor experience. Sentiment results showed that 87% of reviews were positive and 13% negative, indicating overall satisfaction but also presenting a limitation in capturing detailed negative sentiment. Another notable limitation involved the manual translation of reviews from Malay to English, where informal expressions, abbreviations, or cultural nuances may have led to misinterpretation despite efforts to standardize the data.

The findings indicate that while visitors praised the museum's exhibits and environment, there is room for improvement in interactivity, staff engagement, and accessibility. Recurring issues in the negative feedback included unclear exhibit descriptions, long waiting times, and limited seating for elderly visitors. However, such concerns were underrepresented due to the imbalance in sentiment distribution. This highlights the need for a deeper exploration of minority sentiments. Overall, the insights align with the museum's strategic objective of enhancing visitor satisfaction.

To address the findings, it is recommended that the museum introduce more interactive features, strengthen staff training, and improve facility accessibility. Future research should consider refining sentiment models, expanding the range of data sources, and integrating real-time feedback systems to support continuous improvement. These steps will help the Terengganu State Museum strengthen its role as a leading cultural and educational institution in Malaysia.

Acknowledgement

The authors gratefully acknowledge the Terengganu State Museum for their support and collaboration. Special thanks to the museum's management and staff for their valuable insights and assistance, which were instrumental to the success of this project.

Paper Contribution to Related Field of Study

This research contributes to information management and text analytics by introducing a method to analyze review data in cultural institutions. Through sentiment analysis and clustering, it enhances understanding of visitor perceptions and offers a practical approach to managing large-scale feedback.

Information Management

This paper introduces a method for analyzing large-scale visitor reviews using sentiment analysis and clustering. By processing unstructured data from platforms like Google Maps and TripAdvisor, the approach extracts key insights to support data-driven decisions. It aligns with the museum's goals by enhancing visitor experience and guiding strategic planning..

Text Analytics in Cultural Institutions

This research contributes to the field of text analytics by integrating sentiment analysis and clustering to better understand user perceptions of cultural institutions, particularly museums. The combined approach offers a comprehensive view of visitor sentiment and thematic patterns, uncovering key insights to guide museum management. The study demonstrates how text analytics can inform strategic improvements and enhance overall visitor satisfaction and engagement.

References

Ashfaque, Z. (2023). Sentiment Analysis with Naive Bayes Algorithm - Zubair Ashfaque - Medium. Medium. https://medium.com/@zubairashfaque/sentiment-analysis-with-naive-bayes-algorithm-a31021764fb4

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2018). New Avenues in Knowledge Bases for Natural Language Processing. Journal of Artificial Intelligence Research, 60, 239-270. https://doi.org/10.1613/jair.5551

Falk, J. H., & Dierking, L. D. (2018). "Enhancing Visitor Engagement in Museums: The Role of Personal, Social, and Physical Contexts." Museum Management and Curatorship, 33(1), 55-70. https://doi.org/10.1080/09647775.2018.1388832

Li, X., & He, Z. (2019). "Text Analytics for Enhancing Visitor Experience in Museums: A Case Study." Journal of Cultural Heritage Management and Sustainable Development, 9(2), 167-180. https://doi.org/10.1108/JCHMSD-01-2019-0003

Liao, L., & Huang, Y. (2019). "Interactive Technologies in Museums: Enhancing Visitor Engagement Through Digital Experiences." Journal of Cultural Heritage, 36, 170-179. https://doi.org/10.1016/j.culher.2018.09.016

Liu, Y., Huang, Q., & Chen, J. (2020). "Sentiment Analysis of Online Reviews for Museums: A Study of Visitor Experiences." Journal of Hospitality and Tourism Technology, 11(3), 315-330. https://doi.org/10.1108/JHTT-11-2019-0120

Reddy, V. (2023). Sentiment Analysis using SVM - ScrapeHero - Medium. Medium. https://medium.com/scrapehero/sentiment-analysis-using-svm-338d418e3ff1

Roni, E. F., Haryani, C. A., Mitra, A. R., & Aribowo*, A. (2023). Sentiment Analysis of Crypto Coin on Twitter Data Using Text Mining Method with K-Means Clustering Case Study: Bitcoin, Ethereum, and Binance. 2023 IEEE 9th Information Technology International Seminar (ITIS). https://doi.org/10.1109/ITIS59651.2023.10420085

Singh, S., & Srivastava, S. (2020). Review of clustering techniques in control system: review of clustering techniques in control system. Procedia Computer Science, 173, 272-280.

Smith, P., & McCartney, G. (2019). "The Impact of Staff Behavior on Visitor Satisfaction in Museums." Journal of Service Management, 30(3), 335-351. https://doi.org/10.1108/JOSM-10-2018-0334

Wati, M., Noorlah, M. S., Tejawati, A., Septiarini, A., Jamil, M., & Puspitasari, N. (2022). Implementation of The K-means Clustering for The Public Health Center Data. 2022 International Conference on Electrical Engineering, Computer and Information Technology. https://doi.org/10.1109/ICEECIT55908.2022.10030452

Xu, L., & Li, Y. (2020). "Enhancing Customer Service in Museums Through Staff Training: A Case Study Approach." Journal of Hospitality and Tourism Insights, 3(2), 155-172. https://doi.org/10.1108/JHTI-09-2019-0092

Yang, S., & Zhang, W. (2021). "Clustering Visitor Feedback Using K-Means: Insights for Cultural Institutions." Information & Management, 58(6), 103-116. https://doi.org/10.1016/j.im.2021.103116

Zhang, X., Xu, Z., Wang, X., & Lin, X. (2019). An Improved Text Clustering Method Based on Sentiment Analysis. IEEE Access, 7, 87074-87084. https://doi.org/10.1109/ACCESS.2019.2925862