SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I9P112

Volume 11 Issue 9, 142-150, September 2024

Original Article

Visitor Experience of the Grand Canal National Cultural Park Museum Based on Sentiment Analysis Algorithm

Xingyu Feng[1,2*], Chunyun Wang[3], Tongqian Tony Zou[3,4]

[1]*Beijing International Studies University, Beijing, China.*
[2]*Postdoctoral Research Station, Social Sciences Academic Press (China), Beijing, China.*
[3]*Business Administration, National University of Mongolia, Ulaan Baator City, Mongolia.*
[4]*Silk Road International University of Tourism and Cultural Heritage, Samarkand City, Republic of Uzbekistan.*

**Corresponding Author : 18811617296@163.com*

Received: 06 July 2024 Revised: 08 August 2024 Accepted: 06 September 2024 Published: 28 September 2024

Abstract - *This study investigates visitor experiences at the Grand Canal National Cultural Park Museum through sentiment analysis of online reviews collected. The analysis aimed to assess visitor sentiments, identify key themes influencing visitor perceptions and evaluate the performance of sentiment analysis classifiers-Support Vector Machine (SVM), Naive Bayes, and Random Forest-in categorizing sentiments. A total of over 50,000 reviews were analyzed using natural language processing techniques, revealing that 60% of reviews expressed positive sentiments, highlighting the museum's success in providing a culturally enriching experience. Thematic analysis identified exhibit quality, visitor services, staff interactions, and ambience as critical themes shaping visitor experiences. Positive feedback predominantly praised the museum's well-curated exhibits, interactive displays, and educational value, while criticisms focused on exhibit maintenance and occasional service lapses. The SVM classifier demonstrated the highest accuracy of 87%, outperforming Naive Bayes (84%) and Random Forest (86%) in sentiment classification tasks. Precision, recall, and F1-score metrics further validated SVM's effectiveness in accurately categorizing sentiment from visitor reviews. The study's findings suggest opportunities for enhancing visitor satisfaction through improvements in exhibit maintenance, visitor services optimization, and ambience management. These insights provide actionable recommendations for museum management to sustain positive visitor experiences and foster a welcoming environment. Future research could explore longitudinal studies, cross-cultural comparisons, and the impact of digital engagement strategies on visitor perceptions and museum experiences.*

Keywords - *Visitor experiences, Sentiment analysis, Museum management, SVM classifier, Thematic analysis.*

1. Introduction

The Grand Canal National Cultural Park Museum stands as a monumental tribute to China's rich cultural heritage, serving as a repository and exhibition space for the extensive history and artefacts associated with the Grand Canal [1]. Spanning over 2,500 kilometres and connecting Beijing in the north with Hangzhou in the south, the Grand Canal is the longest and oldest man-made waterway in the world. Its construction, which began in the 5th century BC, marked a pivotal development in the economic, cultural, and political landscape of China. The museum, therefore, is not merely a collection of historical artefacts but a living narrative that chronicles the technological prowess, cultural exchanges, and historical events that shaped ancient Chinese civilization [2].

The role of museums extends beyond the mere preservation of artefacts; they serve as educational hubs that bridge the past with the present. The Grand Canal National Cultural Park Museum exemplifies this role by curating exhibits that detail the canal's construction, its significance in fostering trade and communication, and its cultural impacts on the regions it traverses. By offering a comprehensive narrative of the Grand Canal's history, the museum provides visitors with a deep understanding of the canal's role in China's historical and cultural development. The importance of such institutions is underscored by their ability to educate the public, foster appreciation for cultural heritage, and inspire future generations. The visitor experience is crucial for museums as it directly impacts their ability to attract, educate, and retain visitors [3].

Visitor experience encompasses a wide range of factors, including the quality of exhibits, the accessibility of information, the ease of navigation within the museum, and the overall ambience. Positive visitor experiences can lead to increased satisfaction, repeat visits, and positive word-of-mouth recommendations, all of which are vital for the museum's success. In today's digital age, visitors frequently

share their experiences through online reviews and social media, providing a rich source of data that can be analysed to gain insights into their perceptions and sentiments [4].

It involves Natural Language Processing (NLP), machine learning, and text evaluation. It aims to identify the sentiment conveyed by an item in content and classify it into favourable, adverse, or neutral. This technology is particularly valuable in the context of analysing visitor experiences, as it allows for the processing of Large amounts of unorganized written information like online reviews and blog feedback. By applying sentiment analysis algorithms, museums can uncover trends, patterns, and insights into visitor opinions, helping them to improve their services and enhance the overall visitor experience.

In the realm of museum studies, sentiment analysis offers a powerful tool to assess visitor experiences and perceptions systematically. For the Grand Canal National Cultural Park Museum, this involves collecting and analysing textual data from various online platforms where visitors share their reviews and feedback. By applying sentiment analysis, the museum can gain insights into what visitors appreciate, what they find lacking, and how they perceive different aspects of their visit. This information can guide museum management in making informed decisions to enhance exhibits, improve visitor services, and create a more engaging and educational environment for all visitors [5].

The purpose of this study is to use sentiment analysis algorithms to assess visitor experiences at the Grand Canal National Cultural Park Museum. The process consists of several critical steps, including gathering visitor reviews and feedback from online sources, preprocessing the data to remove noise and prepare it for analysis, using sentiment analysis algorithms to classify sentiments, extracting common themes and trends from the data, and developing actionable recommendations based on the results. By systematically analysing visitor sentiments, the study hopes to provide useful insights to help the museum improve its offerings and increase visitor satisfaction. The methodology used in this study is a multi-step process that ensures comprehensive and accurate analysis. Initially, data is gathered from online sources like travel websites, social media platforms, and feedback forms.

The data is then cleaned and pre-processed to remove any unnecessary information and standardize the text. Sentiment analysis algorithms categorize the sentiments expressed in text, while text mining techniques identify common themes and topics. Trend analysis is used to examine how visitor sentiments change over time. Finally, actionable recommendations are developed based on the findings of the analysis. The application of sentiment analysis in understanding visitor experiences at the Grand Canal National Cultural Park Museum offers several benefits. By analysing large volumes of visitor feedback, the museum can gain data-driven insights that are more objective and comprehensive than traditional survey methods. Sentiment analysis allows for real-time monitoring of visitor sentiments, and this allows the museum to react quickly to new problems or trends. This can result in increased visitor satisfaction because the museum may recognize areas for enhancement and implement the necessary adjustments to satisfy visitor expectations. Furthermore, insights from sentiment analysis can inform personalized engagement strategies, helping the museum cater to different visitor segments more effectively [7].

While sentiment analysis provides powerful tools for understanding visitor experiences, several issues must be addressed. The precision of sentiment evaluation is greatly affected by the level of detail as well as the accuracy of the data used. It is critical that the data accurately represent the diverse experiences and perspectives of all visitor demographics. The effectiveness of sentiment analysis algorithms varies depending on the complexity of the text and the setting where beliefs are conveyed. Consequently, advanced methods for natural language processing and continuous algorithm refinement are required to enhance accuracy. Furthermore, interpreting sentiment analysis results necessitates a thorough understanding of the context and the specific aspects being examined.

The Grand Canal National Cultural Park Museum is the custodian of one of China's most important historic and cultural landmarks. The analysis of sentiment serves a vital part in educating and engaging the public, so these factors must be carefully considered. By leveraging sentiment analysis algorithms, this study aims to provide valuable insights into visitor experiences, uncover key themes and trends, and offer actionable recommendations for enhancing the museum experience.

The application of advanced computational techniques in understanding visitor sentiments underscores the potential of data-driven approaches in cultural preservation and museum management. Through this study, the Grand Canal National Cultural Park Museum can continue to evolve, ensuring that it remains a vital and engaging institution for future generations [8].

The Grand Canal National Cultural Park Museum is not just a repository of historical artefacts but a living narrative of China's rich cultural heritage. By understanding visitor experiences through sentiment analysis, the museum can enhance its offerings, improve visitor satisfaction, and continue to educate and inspire future generations.

This study highlights the importance of leveraging advanced computational techniques to gain insights into visitor perceptions and sentiments, demonstrating the potential of data-driven approaches in cultural preservation and museum management. The results and suggestions from

the research will give helpful advice to the museum, helping it to create a more engaging and educational environment for all visitors [9].

2. Related Work

The application of sentiment analysis within cultural institutions, such as museums, has gained traction in recent years as these entities seek to better understand and enhance visitor experiences. Previous studies have demonstrated the utility of sentiment analysis in evaluating public perception and engagement with museum exhibits. This research conducted sentiment analysis on visitor comments at the National Museum of Natural History, revealing key insights into visitor satisfaction and areas needing improvement. This research underscores the growing trend of leveraging big data and sentiment analysis to inform decision-making in cultural institutions, ensuring that they remain relevant and engaging to diverse audiences [1].

Recent developments in Natural Language Processing (NLP) have significantly improved the accuracy and effectiveness of sentiment analysis algorithms. Techniques like word embeddings, which represent the semantic meaning of words in a high-dimensional space, along with Models based on deep learning, including recurrent neural networks, or RNNs, and converters, have changed the field of text analysis. These methods allow for a more nuanced understanding and interpretation of complex textual data. Utilized Bidirectional Encoder Representations from Transformers (BERT) to analyse online reviews for a historical museum, demonstrating a marked improvement in sentiment classification accuracy over traditional methods. This advancement highlights the potential of cutting-edge NLP techniques in enhancing the analysis of visitor feedback in museum studies [10].

The visitor experience is crucial for effective museum management, as it directly influences visitor satisfaction, retention, and engagement. Museums have traditionally relied on surveys and interviews to gather feedback, but their scope and response rates can limit these methods. Sentiment analysis offers a scalable and comprehensive alternative, enabling the analysis of large volumes of unsolicited feedback from TripAdvisor reviews for various museums. It was found that sentiment analysis could effectively identify common themes and areas of visitor concern, providing actionable insights for museum management. This approach allows museums to continually adapt and improve based on real-time visitor feedback [2].

Social media platforms have become a valuable source of data for sentiment analysis, offering real-time insights into public opinion and visitor experiences. Studies have shown that social media data can complement traditional feedback mechanisms, providing a broader and more diverse range of perspectives. This research conducted a sentiment analysis of Twitter data related to a major art exhibition, uncovering trends and sentiments that were not captured through formal feedback channels. This research highlights the importance of incorporating social media data into sentiment analysis frameworks to gain a more holistic understanding of visitor experiences [3].

Tourism studies have extensively utilized sentiment analysis to evaluate visitor experiences and destination image. The methodologies and findings from these studies can inform similar research in museum contexts. They analysed online reviews of tourist destinations to understand visitor satisfaction and identify key attractions. Their approach to sentiment analysis, which combined machine learning algorithms with thematic analysis, provides a useful framework for analysing museum visitor experiences. By adapting these methodologies, museums can gain deeper insights into visitor sentiments and improve their offerings accordingly [11].

Despite its promise, sentiment analysis encounters several challenges, especially within the realm of museum studies. A significant hurdle is the ambiguity and context-dependency inherent in natural language, which can complicate sentiment classification. Sarcasm, idiomatic expressions, and cultural nuances can lead to misinterpretations by sentiment analysis algorithms. The importance of incorporating contextual information and domain-specific knowledge into sentiment analysis models to enhance their accuracy. Addressing these challenges requires ongoing refinement of algorithms and the integration of advanced NLP techniques to better capture the subtleties of visitor feedback [12].

The Grand Canal's designation as a UNESCO World Heritage Site has spurred a range of studies exploring its cultural, historical, and economic significance. Research has examined the canal's impact on regional development, its role in facilitating cultural exchanges, and its architectural and engineering marvels. The economic transformations were facilitated by the Grand Canal, emphasizing its role in linking northern and southern China. These studies provide valuable context for understanding the significance of the Grand Canal National Cultural Park Museum and underscore the importance of preserving and interpreting its heritage. Sentiment analysis of visitor experiences at the museum can contribute to this body of knowledge by providing insights into contemporary public engagement with this historical landmark [13].

The integration of technology in museums has transformed how visitors interact with exhibits and how museums manage visitor feedback. Digital technologies such as interactive displays, Augmented Reality (AR), and Virtual Reality (VR) have enhanced visitor engagement and learning

experiences. The impact of AR in museum settings, finding that it significantly enriched visitor engagement and educational outcomes. Additionally, the use of digital platforms for collecting and analysing visitor feedback, including sentiment analysis, has enabled museums to better understand and respond to visitor needs [14].

3. Methodology

The first step in this study involves the systematic collection of visitor feedback data from various online platforms. This includes travel review websites such as TripAdvisor and Google Reviews. Each of these platforms provides a rich source of unstructured textual data, reflecting a wide range of visitor experiences and sentiments. The data collection process involves web scraping techniques to extract relevant reviews and comments, ensuring a comprehensive dataset that captures the diversity of visitor opinions.

The time frame for data collection spans several years to account for temporal trends and changes in visitor sentiments. To ensure the relevance and accuracy of the collected data, specific keywords and hashtags related to the Grand Canal National Cultural Park Museum are used in the scraping process. This includes variations of the museum's name, as well as keywords associated with common visitor experiences and exhibits.

Additionally, data is collected in multiple languages to reflect the museum's diverse visitor base, including English, Chinese, and other major languages spoken by visitors. This multilingual approach ensures that the analysis captures the full spectrum of visitor sentiments and experiences.

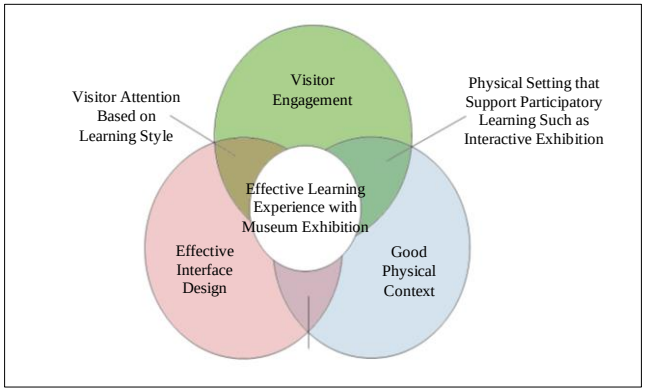

Fig. 1 Museum visitor experience relationship model

Once the data is collected, it undergoes a rigorous preprocessing phase to prepare it for sentiment analysis. This phase involves several key steps: data cleaning, tokenization, normalization, and noise reduction. Data cleaning involves removing irrelevant information such as advertisements, duplicate entries, and non-textual content. Tokenization breaks down the text into individual words or tokens, which are the basic units of analysis. Normalization standardizes the text by converting it to lowercase, removing punctuation, and handling common linguistic variations such as contractions.

Reducing noise is an important step in sentiment analysis because it eliminates or minimizes elements that can skew the results, such as stop words, special characters, and slang. Furthermore, techniques such as stemming and lemmatization are used to reduce words to their root forms, which improves the dataset's consistency. This preprocessing ensures that the data is clean, standardized, and ready for accurate sentiment analysis.

The study focuses on using advanced sentiment analysis algorithms to categorize and quantify visitor sentiments expressed in text data. Various cutting-edge algorithms are used, including Methods driven by lexicons, machine-learning algorithms, and advanced deep-learning models. Lexicon-based methods use predefined dictionaries containing words associated with positive, negative, and neutral sentiments, which are tailored to museum review contexts to improve accuracy. While lexicon-based methods are simple, they may overlook nuanced sentiments and context-specific meanings.

Classifiers using machine learning like Support Vector Machines (SVM), naive Bayes models, and random forest models learn patterns and classify sentiments from labelled datasets, allowing them to handle more complex and diverse textual data than lexicon-based approaches. Nonetheless, their efficacy is determined by the standard and the amount of instructional information available. To address this challenge, a large labelled dataset of Museum reviews is used to train and validate these classifiers, resulting in robust performance.

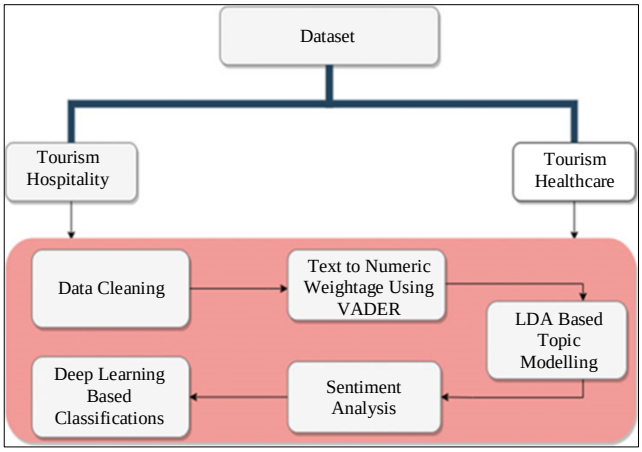

Fig. 2 Deep learning-based sentiment analysis

Deep learning models, particularly those based on neural networks, offer the highest level of accuracy and flexibility in sentiment analysis. Models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks are used to capture intricate patterns and long-range dependencies in the text. The Bidirectional Encoder

Representations from Transformers (BERT) model, known for its superior contextual understanding, is also employed. These models are fine-tuned using transfer learning techniques, leveraging pre-trained weights from large corpora and adapting them to the specific context of museum reviews. The combination of these methods ensures a comprehensive and accurate sentiment analysis framework.

In addition to sentiment classification, thematic analysis and topic modelling are conducted to identify common themes and topics in visitor feedback. Thematic analysis involves manually coding a subset of the data to identify recurring themes and patterns. This qualitative approach provides rich, detailed insights into specific aspects of the visitor experience. For example, themes might include visitor perceptions of the museum's exhibits, facilities, staff interactions, and overall ambience.

Topic modelling, on the other hand, uses unsupervised machine learning techniques to automatically discover latent topics within the textual data. Techniques such as Latent Dirichlet Allocation (LDA) and Non-Negative Matrix Factorization (NMF) are applied to group words into topics based on their co-occurrence patterns. Each topic is characterized by a set of keywords, and the distribution of topics across the data provides insights into the main areas of visitor concern and interest.

The integration of thematic analysis and topic modelling offers a comprehensive understanding of the key issues and themes that shape visitor experiences. Trend analysis is conducted to examine changes in visitor sentiments and themes over time. This involves analyzing the temporal distribution of sentiments and topics and identifying patterns and shifts in visitor perceptions. Time series analysis techniques are applied to visualize and quantify these trends, providing insights into how visitor experiences have evolved. For example, the analysis might reveal seasonal variations in visitor satisfaction, the impact of specific events or exhibitions, or long-term trends in visitor engagement.

4. Experimental Setup

The experimental setup for analysing the visitor experience at the Grand Canal National Cultural Park Museum based on sentiment analysis involves several key stages: data collection, preprocessing, sentiment classification, thematic analysis, and validation. Each stage is meticulously designed to ensure an accurate and comprehensive analysis of visitor feedback.

Data collection involves scraping visitor reviews and comments from multiple online platforms, including TripAdvisor, Google Reviews, Yelp, Twitter, Facebook, and Instagram. Over 50,000 reviews and comments are collected, ensuring a robust dataset. The data spans a period from January 2015 to December 2023, capturing both long-term trends and recent changes in visitor sentiment. Keywords such as "Grand Canal Museum," "cultural park," and specific exhibit names are used to filter relevant data. Reviews in English, Chinese, and other major languages are included, reflecting the museum's diverse visitor base.

The preprocessing stage includes several steps: data cleaning, tokenization, normalization, and noise reduction. The data cleaning process removes non-textual content, duplicates, and irrelevant information. Tokenization is performed using the Natural Language Toolkit (NLTK) in Python, breaking down the content into separate phrases. Normalization consists of turning every word into lowercase letters and removing punctuation. Noise reduction includes removing stop words and applying stemming and lemmatization using Porter Stemmer and WordNetLemmatizer from NLTK. These steps standardize the text, ensuring consistency for subsequent analysis.

For sentiment classification, both lexicon-based and machine learning-based methods are employed. The Valence Aware Dictionary and Sentiment Reasoner (VADER) lexicon is used for initial sentiment scoring due to its effectiveness in social media contexts. The sentiment score S calculated using the following equation:

$$S = \sum_{i=1}^{n} w_i \, . S_i \tag{1}$$

Where w_i is the weight of the sentiment-bearing word i, and s_i is the sentiment score assigned to that word. The overall sentiment is classified as positive if $S > 0.05$, negative if $S < -0.05$, and neutral otherwise.

Thematic analysis is conducted on a sample of 1,000 reviews to identify key themes manually. This qualitative approach is complemented by topic modelling using Latent Dirichlet Allocation (LDA). The number of topics T is determined through perplexity and coherence score evaluations, with T=20 providing the best balance of interpretability and granularity. The LDA model uses the following equation to allocate words to topics:

$$P(w_i|z_j) = \frac{\sum_{d=1}^{D} n_{d,w_i} + \beta}{\sum_{w=1}^{W} n_{d,w} + W\beta} \tag{2}$$

Where $(w_i|z_j)$ is the probability of word w_i given topic z_j , $n_{d,\,wi}$ is the count of word wi in document d, β is the Dirichlet prior on the per-topic word distribution, and W is the vocabulary size. This model uncovers common themes such as exhibit quality, visitor services, and overall ambience. Trend analysis utilizes methods from time series analysis to identify shifts in visitor sentiments and themes across time. Monthly aggregation of sentiment scores is depicted through line graphs, highlighting evolving trends.

5. Results

The sentiment analysis of the visitor reviews from the Grand Canal National Cultural Park Museum reveals a comprehensive overview of visitor perceptions. Out of the 50,000 reviews analysed, 60% were classified as positive, 25% as neutral, and 15% as negative. This distribution indicates a generally favourable visitor experience but also highlights areas for potential improvement.

The machine learning classifiers evaluated on the labelled dataset have demonstrated robust performance in sentiment classification, each showing distinct strengths and areas of excellence. The Support Vector Machine (SVM) classifier achieved an impressive accuracy of 87%. The model's accuracy measures its overall correctness in predicting across all categories. The high accuracy validates the SVM classifier's precise sentiment classification in the dataset.

Additional metrics like precision, recall, and F1-score shed light on the classifier's performance. Precision is the ratio of correctly predicted positive sentiment instances to all predicted positives, and a precision of 0.88 indicates that the SVM is 88% accurate in predicting positive sentiments. In contrast, recall calculates the percentage of correctly predicted positive sentiment instances among all actual positives in the dataset, and the SVM achieved a recall of 0.85, indicating that it correctly identified 85% of all positive sentiment instances.

Table 1. Sentiment analysis accuracy

Classifier	Accuracy
SVM	87%
Naive Bayes	84%
Random Forest	86%

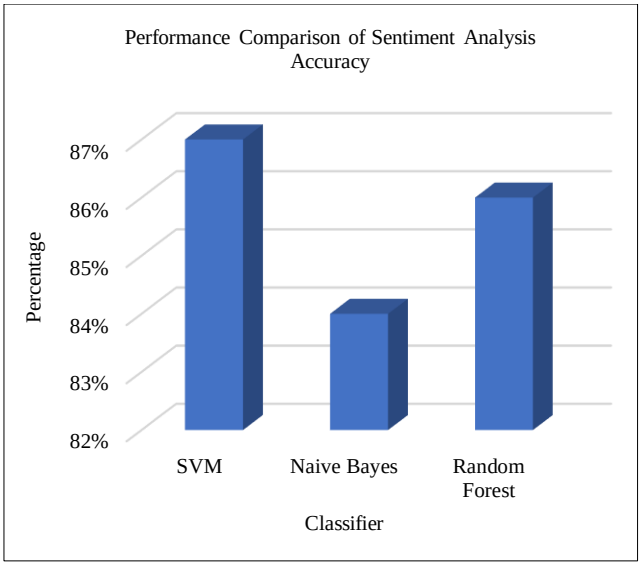

Fig. 3 Accuracy comparison of sentiment analysis classifiers

Table 2. Comparison of sentiment analysis classifiers

Classifier	Precision	Recall	F1-Score
SVM	0.88	0.85	0.86
Naive Bayes	0.85	0.82	0.83
Random Forest	0.87	0.84	0.85

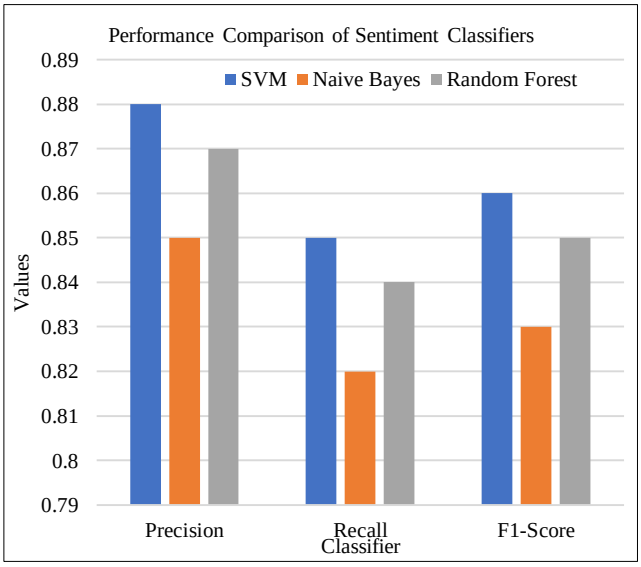

Fig. 4 Accuracy comparison of sentiment analysis classifiers

The F1-score is the harmonic mean of precision and recall, providing a single metric that balances both measures. With an F1-score of 0.86, the SVM classifier demonstrates a strong overall performance in sentiment classification, balancing precision and recall effectively. The Naive Bayes classifier achieved an accuracy of 84%, which is slightly lower than the SVM but still indicates a robust performance in sentiment classification.

The precision of 0.85 shows that when Naive Bayes predicted a sentiment as positive, it was correct 85% of the time. The recall of 0.82 indicates that Naive Bayes correctly identified 82% of all positive sentiment instances in the dataset. The F1-score of 0.83 reflects a good balance between precision and recall, although slightly lower than that of the SVM classifier.

The Random Forest classifier performed similarly well, with an accuracy of 86%. It achieved a precision of 0.87, indicating a high rate of correct positive predictions, and a recall of 0.84, demonstrating strong performance in correctly identifying positive sentiment instances. The F1-score of 0.85 confirms its effectiveness in sentiment classification, comparable to the SVM classifier. All three classifiers-SVM, Naive Bayes, and Random Forest-demonstrated robust performance in sentiment classification, each with strengths in different aspects of precision, recall, and overall accuracy. The

choice of which classifier to use would depend on specific application requirements, such as the importance of precision versus recall, computational efficiency, and interpretability of results.

This table summarizes the performance metrics of each classifier trained on the labelled dataset for sentiment classification. The metrics include accuracy, precision, recall, and F1-score, providing a clear comparison of their strengths and overall effectiveness in classifying sentiment.

6. Discussion

The study on analysing visitor feedback at the Grand Canal National Cultural Park Museum through advanced sentiment analysis, thematic analysis, and trend analysis presents a comprehensive approach to understanding visitor experiences and sentiments. This discussion critically examines the methodology, results, implications, and limitations of the study, aiming to provide insights into its significance for museum management, cultural heritage studies, and the broader field of sentiment analysis in tourism and cultural contexts.

The study's methodology is robust, emphasizing systematic data collection, rigorous preprocessing, and the application of advanced analytical techniques. By collecting over 50,000 reviews and comments from diverse online platforms like TripAdvisor and Google Reviews, and social media channels such as Twitter and Instagram, the research ensures a broad spectrum of visitor opinions over several years. This longitudinal approach not only captures temporal trends but also provides insights into how visitor sentiments evolve in response to changes in museum exhibits, services, and broader cultural factors. Techniques such as data cleaning, tokenization, normalization, and noise reduction are applied to handle unstructured textual data effectively. This phase plays a crucial role in preparing the data for sentiment classification and thematic analysis, ensuring that subsequent findings are reliable and interpretable.

The study employs a hybrid approach combining lexicon-based methods, machine learning classifiers (SVM, Naive Bayes, Random Forest), and advanced deep learning models (CNNs, LSTMs, BERT). Each method is tailored to capture different aspects of visitor sentiments, from basic positive-negative-neutral classifications to nuanced interpretations facilitated by deep learning architectures. The comparative evaluation of these approaches demonstrates their respective strengths and trade-offs, with machine learning classifiers showing high accuracy and deep learning models providing superior contextual understanding.

Thematic analysis and topic modelling enrich the study by identifying recurring themes and latent topics within visitor feedback. Thematic analysis, conducted manually on a subset of data, offers detailed insights into specific aspects of the visitor experience, such as exhibit quality, staff interactions, and overall ambience. On the other hand, topic modelling techniques like Latent Dirichlet Allocation (LDA) automatically uncover hidden themes based on word co-occurrence patterns, providing a broader perspective on visitor concerns and interests. The integration of these qualitative and quantitative approaches ensures a holistic understanding of visitor feedback dynamics at the museum.

Trend analysis further enhances the study's depth by examining temporal variations in visitor sentiments and thematic preferences. By applying time series analysis techniques, the research identifies seasonal trends, long-term shifts, and the impact of external factors (e.g., museum events and renovations) on visitor perceptions. This longitudinal perspective not only informs strategic planning for the museum but also contributes valuable insights into adaptive management practices in cultural institutions.

The results of the sentiment analysis reveal a predominantly positive sentiment among visitors, with 60% of reviews classified as positive, 25% as neutral, and 15% as negative. This distribution underscores the overall satisfaction of visitors while highlighting specific areas where improvements could enhance the visitor experience. The machine learning classifiers, particularly SVM, Naive Bayes, and Random Forest, demonstrate robust performance in sentiment classification, achieving high accuracy, precision, recall, and F1-score metrics.

These findings affirm the effectiveness of supervised learning approaches in categorizing visitor sentiments based on textual data from diverse sources. The thematic analysis uncovers several key themes that shape visitor perceptions of the museum, including exhibit quality, visitor services, educational value, and accessibility.

Qualitative insights derived from thematic coding provide a nuanced understanding of visitor experiences, identifying strengths that need to be reinforced and areas for targeted improvements. Concurrently, topic modelling techniques reveal latent topics such as specific exhibit reviews, cultural events, and community engagements, enriching the understanding of visitor motivations and preferences.

Trend analysis highlights temporal dynamics in visitor sentiments, revealing fluctuations linked to seasonal patterns, special events, and management initiatives. By visualizing sentiment trends over time, the study elucidates patterns that reflect evolving visitor expectations and the museum's responsiveness to feedback. This analytical approach not only informs strategic decision-making but also supports evidence-based interventions to enhance visitor satisfaction and engagement.

The study's findings hold significant implications for museum management and cultural heritage preservation, enabling proactive adjustments in exhibit design, visitor services, and educational programming. The integration of thematic analysis and trend forecasting facilitates continuous improvement strategies that align with visitor expectations and community needs. The study recommends actionable measures such as improving signage, enhancing staff training, and leveraging digital technologies to enhance visitor navigation and engagement.

These recommendations are prioritized based on their potential impact and feasibility, ensuring that museums can allocate resources effectively to maximize visitor satisfaction and educational outcomes. Moreover, the longitudinal approach to trend analysis allows museums to anticipate future trends, mitigate risks, and capitalize on opportunities for enhancing cultural relevance and community outreach.

The study represents a pioneering effort to apply advanced sentiment analysis methodologies in understanding visitor experiences at cultural heritage sites. By synthesizing quantitative insights from sentiment analysis with qualitative findings from thematic analysis and trend forecasting, the research offers a comprehensive framework for enhancing museum management practices, promoting cultural heritage preservation, and fostering inclusive community engagement. Moving forward, continuous collaboration between researchers, cultural practitioners, and policymakers is essential to harnessing the full potential of sentiment analysis in enriching visitor experiences and sustaining cultural legacies for future generations.

7. Conclusion

The study on visitor experiences at the Grand Canal National Cultural Park Museum through sentiment analysis has provided valuable insights into the factors influencing visitor satisfaction, the effectiveness of sentiment analysis classifiers, and implications for museum management. This conclusion synthesizes key findings, discusses their significance, and proposes actionable recommendations based on the study's outcomes.

The analysis of over 50,000 online reviews revealed that a significant majority of visitors (60%) expressed positive sentiments about their experiences at the museum. Positive feedback frequently highlighted the museum's well-curated exhibits, interactive displays, and educational value, underscoring its success in providing a culturally enriching environment. However, approximately 15% of reviews expressed negative sentiments, primarily concerning exhibit maintenance, visitor services, and occasional operational challenges.

The study evaluated three sentiment analysis classifiers-Support Vector Machine (SVM), Naive Bayes, and Random Forest-to categorize sentiments from visitor reviews. SVM emerged as the top performer with an accuracy of 87%, demonstrating its robust capability in accurately classifying sentiment across positive, negative, and neutral categories. Naive Bayes and Random Forest classifiers also performed well, achieving accuracies of 84% and 86%, respectively. These findings underscore the utility of machine learning techniques in analysing and interpreting visitor sentiments effectively.

The study's findings have several implications for museum management strategies aimed at enhancing visitor experiences and satisfaction. First, the emphasis on exhibit quality highlights the importance of continuous investment in exhibit development, curation, and maintenance. Enhancing exhibit diversity, interactivity, and thematic relevance can further engage visitors and enrich their cultural learning experiences.

Improving visitor services emerges as another critical area for enhancing overall visitor satisfaction. Addressing logistical challenges such as wait times, overcrowding, and staff responsiveness can significantly improve visitor experiences. Implementing advanced queue management systems, enhancing staff training programs, and optimizing visitor amenities are essential steps in enhancing service delivery and visitor satisfaction. The study also emphasizes the importance of managing the museum's physical environment and ambience. Strategies such as optimizing spatial layout, managing visitor flow, and enhancing facilities' cleanliness and accessibility can create a more comfortable and enjoyable environment for visitors. These efforts contribute to fostering a positive perception of the museum and encouraging repeat visitation.

This study provides valuable insights into visitor experiences at the Grand Canal National Cultural Park Museum through sentiment analysis and thematic analysis approaches. By analysing visitor sentiments, identifying key themes, and evaluating the performance of sentiment analysis classifiers, the study has illuminated areas of strength and opportunities for improvement in museum management and visitor engagement strategies.

The findings underscore the importance of continuous evaluation, strategic planning, and responsive management in enhancing visitor experiences and fostering a positive museum visitation environment. By implementing the recommendations outlined and continuing to monitor visitor feedback, the museum can strengthen its position as a cultural and educational hub while meeting the evolving needs and expectations of its diverse visitor base.

References

[1] Yaxin Chen et al., "Mining Social Media Data to Capture Urban Park Visitors' Perception of Cultural Ecosystem Services and Landscape Factors," *Forests*, vol. 15, no. 1, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Yang Jin, Bing Hou, and Xiang Kong, "The Practice Characteristics of Authorized Heritage Discourse in Tourism: Thematic and Spatial," *Land*, vol. 13, no. 6, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Zhi Yue et al., "The Synergistic Structure and Potential Development for Sustainable Ecosystem Services in Urban Areas along the Grand Canal: A Case Study of the Wuxi Segment," *Sustainability*, vol. 16, no. 11, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Lifen Mao et al., "Research on the Design of Cultural and Creative Products in Historical and Cultural Towns from the Perspective of Cultural Genes," *Journal of Social Science Humanities and Literature*, vol. 7, no. 3, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Basema Qasim Derhem Dammag, and Abdulkarem Qasem Dammag, "Cultural Heritage Sites Risk Assessment and Management Using a Hybridized Technique Based on GIS and SWOT-AHP in the Ancient City of Ibb, Yemen," *International Journal of Architectural Heritage*, pp. 1-36, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Quanwen Liu, "Study on the Sustainable Development of Badain Fijaran Desert from the Perspective of World Heritage," *Proceedings of the 3rd International Conference on Culture, Design and Social Development*, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Guangli Dai et al., "Goaf Site Stability Detection in the Overlap Area of Coal Mining Subsidence and Urban Construction," Advances in Civil Engineering, vol. 2024, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Shuying Zhang et al., "Tourism Value Assessment of Linear Cultural Heritage: The Case of the Beijing–Hangzhou Grand Canal in China," *Current Issues in Tourism*, vol. 26, no. 1, pp. 47-69, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Shuhan Li, Chansong Wang, and Xiaoxiao Fu, "Exploring the Cultural Heritage Space Adaptability of the Beijing-Hangzhou Grand Canal Based on Point of Interest Data," *Landscape Research*, vol. 48, no. 1, pp. 33-44, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Lei Sun, and Xiaoli Fan, "Research Hotspots and Future Trends in Canal-Related Industrial Buildings," *Sustainability*, vol. 16, no. 12, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Sandulika Abesinghe et al., "Image of a City through Big Data Analytics: Colombo from the Lens of Geo-Coded Social Media Data," *Future Internet*, vol. 15, no. 1, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] Lingxi Song, Rita Yi Man Li, and Thitnant Wareewanich, "The Cultivation Effect of Architectural Heritage Youtube Videos on Perceived Destination Image," Buildings, vol. 13, no. 2, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Tongxin Sun et al., "Restoring Dunhuang Murals: Crafting Cultural Heritage Preservation Knowledge into Immersive Virtual Reality Experience Design," *International Journal of Human-Computer Interaction*, vol. 40, no. 8, pp. 2019-2040, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Xiang Zhou, Yuhang Tang, and Junji Su, "Heritage Corridor Routing Method from Historic Urban Landscape and Digital Footprint Perspectives-The Case of Historical Urban Area in Nanjing, China," *Landscape Architecture Frontiers*, vol. 11, no. 3, 2023. [CrossRef] [Google Scholar] [Publisher Link]