

International Journal on Artificial Intelligence Tools

Sentiment Analysis of Water Metro Passengers Review:A Comprehensive Review

--Manuscript Draft--

Manuscript Number:	
Full Title:	Sentiment Analysis of Water Metro Passengers Review:A Comprehensive Review
Article Type:	Review
Keywords:	Sentiment Analysis; Water Metro; Passenger Reviews; Public Transportation; Natural Language Processing; Machine Learning
Corresponding Author:	ANMIGHA N M Amal Jyothi College of Engineering INDIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Amal Jyothi College of Engineering
Corresponding Author's Secondary Institution:	
First Author:	ANMIGHA N M
First Author Secondary Information:	
Order of Authors:	ANMIGHA N M Mr. Jinson Devis
Order of Authors Secondary Information:	
Abstract:	In this review paper, we examine the various techniques that have been developed for conducting sentiment analysis of passenger reviews, with the particular emphasis on Water Metro services. As Water Metro becomes more and more popular as a transportation alternative, understanding the passengers' sentiment becomes important to enhancing service quality. The paper surveys existing trends, challenges, and plausible receptacles of organizational transformation through concrete evidence of sentiment analysis in public transport systems. In the paper, this analysis is augmented by the analysis of the sentiment of Water Metro passenger's reviews, in terms of service improvement.
Suggested Reviewers:	

Sentiment Analysis of Water Metro Passengers Review: A Comprehensive Review

Anmigha N M¹, Jinson Devis²

¹PG Scholar, Dept. of Computer Applications, Amal Jyothi College of Engineering,
Kanjirappally, Kottayam, Kerala (India), anmighanm2025@mca.ajce.in

² Assistant Professor, Dept. of Computer Applications, Amal Jyothi College of Engineering,
Kanjirappally, Kottayam, Kerala (India), jinsondevis@amaljyothi.ac.in

Abstract

In this review paper, we examine the various techniques that have been developed for conducting sentiment analysis of passenger reviews, with the particular emphasis on Water Metro services. As Water Metro becomes more and more popular as a transportation alternative, understanding the passengers' sentiment becomes important to enhancing service quality. The paper surveys existing trends, challenges, and plausible receptacles of organizational transformation through concrete evidence of sentiment analysis in public transport systems. In the paper, this analysis is augmented by the analysis of the sentiment of Water Metro passenger's reviews, in terms of service improvement.

Keywords:

Sentiment Analysis, Water Metro, Passenger Reviews, Public Transportation, Natural Language Processing, Machine Learning

1. Introduction

1.1 Background

The water metro is a subtle deviation from the conventional road and rail network. Urban water transport such as Water Metro thus forms an environment-friendly and efficient means of transport. Passengers' satisfaction is one of the key indicators of success regarding the services rendered, and their feedback analyzed through sentiment analysis gives better insights into the service performance.

1.2 Importance of Sentiment Analysis in Transport

The extraction and examination of subjective information from textual data is the focus of sentiment analysis, a subfield of natural language processing. It would be helpful in transportation to understand the perceptions of passengers, give insights for areas of improvement, and allow enhancements in the overall service quality.

1.3 Scope and Objective

The paper reviews existing research in the application of sentiment analysis in transport services, putting more emphasis on Water Metro passenger reviews. In this context, a review has been given to methodologies

employed for extracting and interpreting passenger sentiments, with the unique challenges presented by aquatic modes of transportation. Further, the review aims to find out lacunae in the exiting research on real-time data analysis integrated with multi-lingual sentiment interpretation. While discussing these challenges, the paper also charts future directions of research which will further enhance accuracy and applicability of Sentiment Analysis in passenger experience and service quality improvement within the Water Metro system.

2. Literature Review

Sentiment analysis has become a vital tool across various sectors, particularly in transportation, where it helps organizations understand customer emotions and opinions through reviews on travel-related apps like Red Bus, Make My Trip, and Yatra.com [1]. Research from the last few years has pointed out the use of Sentiment Analysis to evaluate and improve service quality in the airline industry through categorizing customer comments as positive, negative, or neutral. This process includes data collection, preprocessing, sentiment analysis techniques, and visualization, so that the airlines can assess the quality of their service and change their strategies to better serve the passengers [2]. Additionally, sentiment analysis in public transportation systems offers insights into commuter preferences, aiding in the planning of multimodal journey options and improving overall user experience. This can be achieved by combining sentiment analysis with current innovations such as car-sharing plans and

autonomous vehicles, enabling transportation companies to offer more individualized and user-friendly services, which in turn will lead to increased customer satisfaction and ultimately the use of public and shared transportation [3].

With the explosion of social networking sites, there has been an outpouring of unstructured textual data which leads to the necessity for more sophisticated sentiment and emotion analysis techniques [8]. The present techniques, lexicon based, and corpus based, are good in some ways and not so good in others, but recurrent neural networks (RNNs) such as LSTM can capture long term dependencies very well from a large amount of data. Preprocessing and feature extraction problems remain, especially with implicit features in text. Sentiment analysis has become an integral part of understanding user opinions on many different forums, and new methods using part-of-speech tagging seem to be very promising in achieving greater accuracy [9]. A review of sentiment analysis techniques highlights the widespread use of lexicon-based methods and machine learning approaches, such as Naïve Bayes and SVM, particularly for analysing data from platforms like Twitter [10]. This analysis implies that although lexicon-based methods are accurate for smaller corpus, machine learning methods are more appropriate for larger more complicated data, and it calls for the creation of general models for different corpora and for the exploration of other wikis.

Recent developments in natural language processing (NLP) and artificial intelligence (AI), especially with models like BERT and

RoBERTa, have greatly enhanced the ability to analyze feelings and opinions in many fields, including aviation [4]. These models have performed very well in categorizing sentiments, especially when looking at reviews from airline customers, with RoBERTa being the most accurate. Techniques like Random Forest, a type of machine learning, have also been successful in sorting out sentiments from social media, giving airlines useful information to improve their services [5]. Moreover, methods like text mining and topic modeling help analyze customer reviews, pinpointing areas where airlines can improve their services and ensuring they meet customer needs better [6]. Using sentiment analysis across various languages and platforms highlights its importance in handling large amounts of unstructured data [7].

Due to the proliferation of digital platforms like blogs and social media, which produce enormous volumes of data that are rich in opinions, sentiment analysis has grown in importance. To determine public opinion on a range of subjects, goods, and services, this procedure entails removing and categorizing subjective information from text. Because of their ease of use and precision, supervised machine learning algorithms such as Naïve Bayes (NB) and Support Vector Machines (SVM) are commonly used in text mining and natural language processing (NLP). These techniques have shown their worth in drawing useful conclusions from huge datasets in applications like social media monitoring and consumer feedback analysis. [11].

Despite its advancements, sentiment analysis faces challenges related to the complexity of human language, including sarcasm, domain-specific terminology, and varying contextual meanings. Issues like domain dependence and sentiment polarity interpretation difficulties persist. Addressing these challenges requires expanding datasets, improving algorithms, and exploring new methodologies. Sentiment analysis's ability to provide actionable insights for businesses—by automating sentiment classification and obtaining real-time feedback—enhances decision-making and reduces manual analysis costs. As technology evolves, sentiment analysis is expected to offer increasingly nuanced understandings of public opinion and sentiment across various sectors [12].

Sentiment analysis (SA), which examines textual data from social media and other online platforms, has emerged as a crucial method for comprehending public sentiment. Text may now be effectively classified as positive, negative, or neutral thanks to the development of SA approaches, which include the application of machine learning algorithms like Naïve Bayes and Support Vector Machines (SVM). These techniques are well known for their effectiveness in processing vast amounts of data and deriving insightful conclusions. By handling the enormous volume of data produced on social media platforms, the incorporation of big data technologies, such as Hadoop, has further improved sentiment analysis's efficacy. [13].

Additionally, sentiment analysis of social media data, especially from Twitter, has

shown significant value in understanding consumer sentiment. Various machine learning approaches, combined with semantic analysis, have been explored to improve sentiment classification accuracy. Research indicates that Naïve Bayes outperforms other methods like Maximum Entropy and SVM when using a unigram model, achieving higher accuracy. Combining semantic analysis with machine learning techniques and expanding the training dataset can further refine sentiment classification. These advancements offer a practical approach for analysing unstructured data from social media platforms [14].

2.1 Sentiment Analysis Techniques

Lexicon-Based Methods:

Lexicon-based sentiment analysis methods rely on predefined dictionaries that categorize words into positive or negative sentiment categories. Tools such as Sent WordNet or AFINN provide extensive lists of words along with sentiment scores, allowing for straightforward sentiment classification. This technique calculates the sentiment score of a text by summing the scores of individual words. While effective for smaller and less complex datasets, lexicon-based methods may face limitations in dealing with context-specific meanings, subtle nuances of language, and sarcasm. For instance, a phrase like "I love waiting" could be misclassified as positive if the sentiment lexicon does not recognize the sarcasm.

Machine Learning Approaches:

Machine learning techniques involve training models on labelled datasets to classify sentiments. Common algorithms include:

- **Naive Bayes:** a probabilistic classifier that assumes feature independence and is based on the Bayes theorem. It is a well-liked option for sentiment analysis because of its reputation for being straightforward and successful in text classification. On complicated datasets, nevertheless, its efficiency may occasionally be constrained by its feature independence assumption.
- **Support Vector Machines (SVM):** SVMs are powerful classifiers that work well with high-dimensional data and are effective in finding the optimal hyperplane to separate different classes. SVMs can handle complex patterns in data but might require extensive computational resources and careful parameter tuning.
- **Decision Trees:** These provide a clear model structure and are interpretable, showing how decisions are made based on feature values. While Decision Trees can be prone to overfitting, techniques like pruning can mitigate this issue. They are useful for understanding how different features impact sentiment classification.

Machine learning methods are well-suited for larger datasets where nuanced language patterns need to be captured, but they require a substantial amount of labelled data for training and validation.

Deep Learning Models:

Deep learning techniques, including Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs), offer advanced capabilities for sentiment analysis:

- **LSTM Networks:** Long-term dependencies can be captured and sequential data can be handled by LSTMs. They are appropriate for sentiment analysis of large and complex materials because they are especially good at handling long-term dependencies in text and comprehending context.
- **Convolutional Neural Networks (CNNs):** CNNs were initially created for image processing, but by identifying local patterns and correlations in text, they have demonstrated potential in text classification. They can be effective for identifying sentiment-related features in text, especially when combined with other deep learning methods.

Deep learning models offer superior accuracy and can handle complex language patterns but require significant computational resources and large annotated datasets.

2.2 Data Sources and Preprocessing

Data Sources:

Sentiment analysis for transportation reviews often utilizes various data sources:

- **Social Media Platforms:** Social media platforms like Twitter and Facebook are rich sources of real-time feedback and personal experiences. They provide a

large volume of user-generated content, which can be useful for sentiment analysis but may also include informal language and noise.

- **Online Review Sites:** Sites like TripAdvisor and Yelp offer structured feedback from users regarding their travel experiences. Reviews on these platforms often include detailed comments on service quality, which can be valuable for sentiment analysis.
- **Passenger Surveys:** Structured surveys conducted with passengers provide direct feedback on specific aspects of transportation services. These surveys are often designed to capture detailed sentiments and can be analyzed to understand passenger satisfaction.
- **Mobile App Feedback:** Feedback and ratings from transportation apps provide another data source. This feedback is usually concise but can offer insights into user experiences and preferences.

Preprocessing Techniques:

To prepare data for sentiment analysis, several preprocessing steps are essential:

- **Tokenization:** This process involves splitting text into individual words or tokens, which is crucial for analyzing text data.
- **Stemming and Lemmatization:** These techniques reduce words to their root forms to standardize variations and improve analysis consistency. For example, "running" and "runner" are reduced to "run."
- **Handling Stop Words:** Common words like "the," "is," and "and" are removed as

they do not contribute significant meaning to sentiment analysis.

- **Noise Removal:** Filtering out irrelevant information such as advertisements, spam, or unrelated content ensures that the analysis focuses on genuine feedback.

2.3 Case Studies in Public Transportation

Sentiment Analysis in Bus and Rail Systems:

Studies have shown that sentiment analysis can effectively identify factors impacting passenger satisfaction in bus and rail systems. For example, research might focus on aspects such as:

- **Punctuality:** Analysis of reviews can reveal complaints about delays and their impact on overall satisfaction.
- **Cleanliness:** Feedback related to cleanliness and hygiene can help identify areas requiring improvement.
- **Customer Service:** Sentiment analysis can highlight issues related to staff behavior and service quality.

Such insights can inform operational improvements and enhance passenger experience.

Application in Water Transport:

Sentiment analysis in water transport systems, such as ferries and Water Metro, is less common but provides valuable insights:

- **Service Quality:** Analysis of passenger feedback can assess the quality of service, including aspects like comfort and safety.

- **Weather Conditions:** Water transport services are affected by weather conditions, which can be reflected in passenger feedback. Sentiment analysis can help understand how weather impacts service satisfaction.
- **Unique Aspects:** Factors specific to water transport, such as scheduling and route changes, add complexity to the analysis but also offer opportunities to address unique challenges.

2.4 Challenges in Sentiment Analysis

Language and Context:

Understanding language nuances, such as sarcasm and idioms, poses a significant challenge. For example, "The bus service is just great" could be interpreted differently based on context. Ensuring that sentiment analysis algorithms are trained on diverse datasets can help address these challenges.

Data Imbalance:

An imbalance in the distribution of sentiment classes (positive, negative, neutral) can affect the accuracy of sentiment classification. Techniques like oversampling minority classes, under sampling majority classes, or using balanced datasets can help mitigate this issue.

Real-Time Analysis:

Implementing real-time sentiment analysis requires robust systems capable of handling large volumes of data quickly. This can be challenging in dynamic environments where

feedback is continuously generated. Solutions might include leveraging scalable cloud infrastructure and efficient data processing pipelines.

3. Methodology

The review examined various techniques for conducting sentiment analysis of passenger reviews, with a particular emphasis on Water Metro services. It surveyed existing trends, challenges, and potential applications of sentiment analysis in public transportation systems. The analysis was augmented by an examination of sentiment in Water Metro passenger reviews specifically, focusing on service improvement.

The evaluation covered a variety of sentiment analysis approaches, such as lexicon-based techniques, deep learning models like CNN and LSTM, and machine learning techniques like Naive Bayes and Support Vector Machines. It also touched on data sources for sentiment analysis, including social media platforms, online review sites, passenger surveys, and mobile app feedback. Preprocessing techniques like tokenization, stemming/lemmatization, and stop word removal were mentioned as essential steps in preparing data for analysis.

The review highlighted challenges in sentiment analysis, such as understanding language nuances, dealing with data imbalance, and implementing real-time analysis. It noted that machine learning techniques like Random Forest had been successful in sorting sentiments from social media, providing insights for service improvement. Additionally, methods like

text mining and topic modeling were mentioned as helpful tools for analyzing customer reviews and identifying areas for service enhancement.

The review aimed to find gaps in existing research on real-time data analysis integrated with multilingual sentiment interpretation for Water Metro systems. It also charted future directions for research to enhance accuracy and applicability of sentiment analysis in passenger experience and service quality improvement within the Water Metro system.

4. Conclusion

Sentiment analysis is a crucial technique for comprehending and enhancing the opinions of users of public transit systems, such as the Water Metro. By using different methods, such as looking at word meanings, teaching computers to learn, and using complex learning techniques, sentiment analysis gives us useful information from passenger comments. This information helps us find where we can make things better and guide improvements in services. However, there are still challenges, such as preparing data correctly, understanding the context of comments, and analyzing comments in different languages, which can affect how well sentiment analysis works.

To solve these problems, future studies need to work on creating advanced models and methods. It's important to make data preparation better, understand context more deeply, and develop models that can analyze feelings in many languages. By working on these areas, the study of sentiment analysis can help transportation systems provide

better, more responsive services that meet people's needs and increase their satisfaction.

References

1. Gupta, A., Dwivedi, D., & Singh, J. (2022). Sentiment analysis of travelling passengers using machine learning. *International Journal of Advance Research and Innovative Ideas in Education*, 8(3), 16776-1685. ISSN(O)-2395-4396. Available at: www.ijariie.com
2. Idris, S. L., & Mohamad, M. (2022). A study on sentiment analysis on airline quality services: A conceptual paper. *International Journal of Advanced Research in Computer Science and Software Engineering*, 12(6), 345-355.
3. Kumar, G. L., Reddy, I. K., & Koduru, S. (2021). *Sentiment Analysis in Transportation System*. *International Journal of Creative Research Thoughts (IJCRT)*, 9(11), IJCRT2111075. ISSN: 2320-2882. Retrieved from www.ijcrt.org.
4. Li, Z., Yang, C., & Huang, C. (2024). A Comparative Sentiment Analysis of Airline Customer Reviews Using Bidirectional Encoder Representations from Transformers (BERT) and Its Variants. *Mathematics*, 12(1), 53. <https://doi.org/10.3390/math12010053>
5. Raihen, M. N., & Akter, S. (2024). Sentiment analysis of passenger feedback on U.S. airlines using machine learning classification methods. *World Journal of Advanced Research and Reviews*, 23(01), 2260–2273.
6. Farzadnia, S., & Vanani, I. R. (2022). Identification of opinion trends using sentiment analysis of airlines passengers' reviews. *Journal of Air Transport Management*, 102, 102232. <https://doi.org/10.1016/j.jairtraman.2022.102232>
7. Redhu, S., Srivastava, S., Bansal, B., & Gupta, G. (2018). Sentiment Analysis Using Text Mining: A Review. *International Journal on Data Science and Technology*, 4(2), 49-53. <https://doi.org/10.11648/j.ijdst.20180402.12>
8. Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Springer-Verlag GmbH Austria, part of Springer Nature*. <https://doi.org/10.1007/s00542-021-06384-z>
9. Sultana, N., Kumar, P., Patra, M. R., Chandra, S., & Alam, S. K. (2024). Sentiment analysis for product review. *Department of Computer Science and Engineering, Calcutta Institute of Technology, India*.
10. Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. In *The Fifth Information Systems International Conference 2019*. Azman Hashim

<https://doi.org/10.30574/wjarr.2024.23.1.2183>

International Business School, Kuala Lumpur, Malaysia. Available at <https://doi.org/10.1016/j.jairtraman.2022.102232>

11. Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55, 5731–5780.
<https://doi.org/10.1007/s10462-021-09983-0>
12. Grana, P. A. (2022). Sentiment analysis of text using machine learning models. *International Research Journal of Modernization in Engineering Technology and Science*, 4(5), 2952. e-ISSN: 2582-5208. Retrieved from www.irjmets.com
13. Aqlan, A. A. Q., Manjula, B., & Naik, R. L. (2019). A study of sentiment analysis: Concepts, techniques, and challenges. In *Proceedings of International Conference on Computational Intelligence and Data Engineering* (Vol. 28, pp. 147-156). Springer Nature Singapore Pte Ltd.
https://doi.org/10.1007/978-981-13-6459-4_16
14. Ashique, M., Kumar, S., Vij, A., & Panwar, S. (2021). Sentiment analysis using machine learning approaches of Twitter data and semantic analysis. *Turkish Journal of Computer and Mathematics Education*, 12(6), 5181-5192.