Social Media Analytic: Comparative Study of Public Transportation in Jakarta and Semarang

Hany Wijaya[1,a)], Jesslyn Hok[1, b)], Messya Carment[1, c)], Muhammad Fikri Hasani[1, d)], Dwinanda Kinanti Suci Sekarhati[1, e)]

[1]*Computer Science Department, School of Computer Science, Universitas Bina Nusantara, Jakarta, 11480, email:*

[a)] hany.wijaya@binus.ac.id
[b)] jesslyn.hok@binus.ac.id
[c)] messya.carment@binus.ac.id
[d)]Corresponding author: muhammad.fikri003@binus.ac.id
[e)] nanda.kinanti@binus.ac.id

Abstract. As a country with a population of 278.6 million, public transportation is crucial for supporting urban mobility in Indonesia. However, factors such as the low quality of services still present significant challenges to public interest in using public transportation. As a result, most Indonesians prefer private vehicles, leading to severe problems like traffic congestion and air pollution. To mitigate these issues, this study offers a solution through a dashboard designed to compare public transportation services in Jakarta and Semarang, specifically Transjakarta and Trans Semarang, through social media sentiment analysis. The developed dashboard integrates BERTopic and sentiment classification models with collected data from the X platform, focusing on public perceptions of Transjakarta and Trans Semarang in 2024. The features include sentiment distribution, tweet distribution, key topics, and popular tweets. The study's results indicate that sentiment towards Trans Semarang is relatively more favorable than Transjakarta's. Performance evaluation also showed satisfactory results, with coherence scores of 0.71 and 0.70 for topic extraction, an accuracy of 0.78 for the sentiment classification model, and an average dashboard user satisfaction score of 8.56 out of 10. This study's findings hopefully can improve Jakarta's and Semarang's quality of public transportation services by providing a foundation for developing more accurate sentiment analysis models and enhancing more innovative and informative dashboards..

Keywords: sentiment analysis, BERTopic, classification, public transportation, social media

INTRODUCTION

Indonesia has a population of 278.6 million [1], which includes urban people that are depending on public transportation in their daily mobility are depending on public transportation in their daily mobility. An affordable, convenient, and time-efficient public transport is very vital to acquire to support the growing population [2], [3]. One of the reasons why people prefer private conveyance is the fact of insufficient service quality, and with COVID-19, this has only made the matter worse, causing traffic congestion and air pollution in major Indonesian cities like Jakarta, among others [4], [5].

While Jakarta is ranked the 29th in the world for experiencing the worst traffic jam [6], the city has developed an integrated public transport system known as Jak Lingko, which integrates KRL Commuter Line, MRT, LRT, and Transjakarta [7]. The availability of such Busways in other cities—of its many states, the case with Semarang, which offers BRT Trans Semarang and Trans Jateng—has advertised the use by the local government with very affordable prices of USD 0.3-0.5 per route [8]. In contrast, both cities sort of face very little patronage of public transport due to problems in efficiency and integrated facilities that do not meet the required standards [8], [9], compounded by high urbanization and low awareness in Semarang. The appearance of these buses can be seen on **FIGURE 1**.

FIGURE 1. Tranjakarta and BRT Semarang

Government effort is therefore needed for improvement in the facilities, comfort, safety, and assimilation of the system. Proper public feedback is also necessary to identify the aspects of improvement. This research propose a dashboard analyzing public opinions on Transjakarta and Trans Semarang as one of public using topic modelling and sentiment classification, leveraging several state-of-the-art approach, aiming to serve as a reference for the government to understand system strengths and weaknesses, ultimately enhancing public transportation quality in both cities. Transjakarta and BRT Semarang was chosen because of their transportation modes, which is bus, that exist in both cities.

This paper has several objectives, which are to analyze the factors affecting the low awareness among the public about using public transportation and public perceptions regarding public transportation; and to collect data on public opinions about public transportation. After the analysis & the evaluation is completed, hopefully this research can assist the government in identifying the factors influencing low public awareness of public transportation usage, understand public perceptions of the quality of public transportation, and aid the government in gathering necessary data to develop strategies that can increase public transportation usage. This paper will be divided into previous research, methods, results and discussion; and conclusion parts.

PREVIOUS RESEARCH

Several previous researches related to the public's interest in public transportation can be found globally. First, the research conducted by Pertiwi [10] presents a model for analyzing sentiment regarding transportation for homecoming using Naive Bayes, Neural Network, KNN, and SVM algorithms from Twitter data. The model with the KNN algorithm achieved the highest accuracy of 90.76%, followed by the SVM algorithm with an accuracy of 89.03%, Naive Bayes with an accuracy of 78.16%, and Neural Network with an accuracy of 52.73%. This research found that positive sentiment was more dominant during the 2019 homecoming period. However, negative sentiment was more frequently found after it. Rachman et al. [7] present a model for analyzing public opinion sentiment regarding various public transportation modes in Jakarta, using Twitter data collected with Twint. This research used LDA for topic the extraction method and achieved the smallest mean error of 8.79%. This research's result is KRL and Transjakarta tended to have negative sentiments, while MRT and LRT had positive sentiments.

Krisnawati et al. [11] offer a model for analyzing Twitter user sentiment toward online transportation services by clustering tweets using K-Means and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) methods. The study shows that K-Means and DBSCAN are less suitable for clustering tweets about online transportation services due to silhouette coefficients below 0.5. However, this research also implied that more careful attention is needed in the text preprocessing process as it can affect the clustering results. Ilhamsyah et al. [12] tried to do case study research for PT. PLN (PERSERO) APD East Java's Dashboard,focusing on Human-Centered Design and Key Performance Indicators evaluation and improvement. This research evaluated the design through several stages. The first stage involved determining the target respondents, which in this study were staff members from PLN APD East Java. Next, task scenarios were defined based on the activities or tasks performed by the respondents. Following this, a survey was conducted to gather respondents' perspectives on the dashboard. Finally, system usability was assessed using the SUS score from the respondents' survey results.

There are also several previous research related to public transportation in other countries. In Singapore, [13] tried to analyze the relation of public transportation with the house pricing and noise complaints from residents using natural language processing. The research’s result if the house is closer to the public transportation routes for every 100 meter, the noise complaints will increase 10%, hence impacted the house pricing. [14] also researched about the Uber service in India-Pakistan through the customer’s review in Roman Urdu/Hindi on Facebook. However, [14] methods are a little different because other than using sentiment analysis, the researcher used customized APIs and lexicon-based approach to translate the words into English. The research’s result is the majority of people are complaining, especially for the “Driver” and “Ride” aspects. Lastly, [15] researched about the global public’s opinion related to public transportation by collecting 15.776 tweets in English language and analyzed it with WordClouds and LDA. The result of this research is people tends to choose private vehicles such as car and bicycle, and even walking for their mobility. The critical aspect that has been found for increasing the public’s interest in public transportation is reshaping and safe opening of the transit system.

METHODS

The method’s flow diagram is shown in **FIGURE 2**. This research’s method encompasses sequential stages starting from data collection, proceeding through information extraction, information extraction evaluation, dashboard creation, and dashboard evaluation.

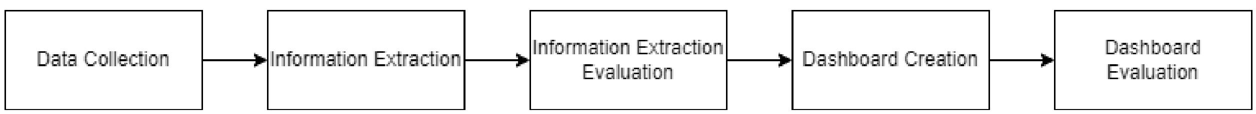

FIGURE 2. Methods flow diagram

A. DATA COLLECTION

The dataset used in this solution is a collection of tweets that was scraped from the X platform. To gather tweets related to Transjakarta, the researchers used the keyword by tagging the official TransJakarta account, @PT_TransJakarta, excluding tweets from the account itself. The data collection period ranged from 2022 to 2024, excluding replies. However, due to scrolling limitations on the X platform, the data obtained was only from 2024. The researchers followed the same process to collect tweets related to Trans Semarang using the official Trans Semarang account, @Transsemarang. The result was 574 tweets for Transjakarta and 505 tweets for Trans Semarang.

B. INFORMATION EXTRACTION

The methods are done by four following steps, which are data cleaning, preprocessing, topic extraction, and sentiment classification. The figure describing these methods can be seen on **FIGURE 3**.

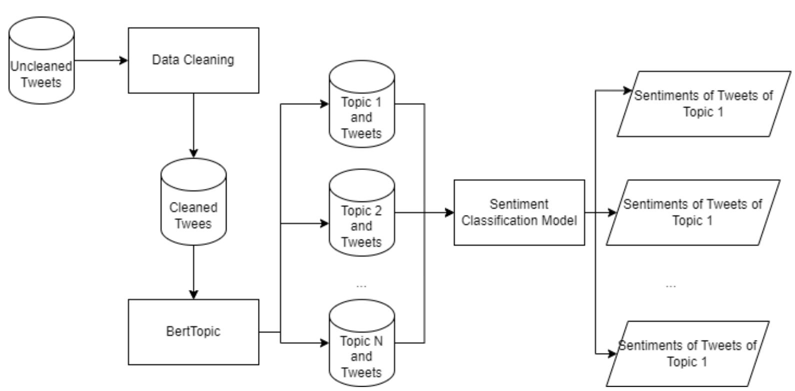

FIGURE 3. Diagram illustrating Information extraction flow

First is the data cleaning step by conducting to ensure the quality and integrity of the dataset used. This step included duplicate tweets removal to ensure that no identical tweets were present, deleting irrelevant columns to simplify the dataset, such as the image URL column, and ensuring that there were no entries with empty tweet content. By following these processes, the resulting analysis will be more accurate and relevant.dataset used in this solution is a collection of tweets that was scraped from the X platform.

Second step is the data preprocessing. During this step, several important steps were taken to prepare the data for processing. For starter, all text was converted to lowercase to maintain consistency. Then, the special character such as mentions, symbols, hyperlink, stopwords, and numbers were removed to avoid misinterpretation and to ensure that all remaining words were relevant to the analysis process. Some abbreviations were also expanded into full words

Third step is the topic extraction. This research will use BERTopic model for the topic extraction process, which was called from the bertopic library in Python. BERTopic is a model that uses clustering process to identify topics and provides accurate results by leveraging BERT embeddings [16]. Before performing topic extraction, the first process was to set suitable hyperparameters for the data. This was done by experimenting with various hyperparameter values and finding the best ones based on the coherence score. After that, the fitting process was carried out to extract topics according to the data's characteristics.

Last step is the sentiment classification. For sentiment classification, the researchers utilized the pre-trained model "bert-base-indonesian-1.5G-sentiment-analysis-smsa" from the HuggingFace website. This model was imported using the Transformers library and applied to each tweet to determine its sentiment, whether positive, negative, or neutral. This sentiment classification was then combined with various features to provide further insights related to the quality of Transjakarta and Trans Semarang.

C. INFORMATION EXTRACTION EVALUATION

The coherence score can be used to ease of interpretation, ensure the relevance, and evaluate the quality of topics that were obtained from topic modeling. After completing the topic extraction process, the coherence score is used as a metric to assess the relationship between words within each extracted topic [17]. A score of 0 indicates equivalent topic relevance, a score of -1 indicates poor topic relevance, and a score of 1 indicates perfect topic relevance. Generally, a score above 0.5 indicates good performance for a topic extraction model.

Then, the Cohen's Kappa Score will be used to measure the level of agreement between two or more labeller. The formula of Cohen Kappa can be seen at equation (1).

$$k = \frac{p_0 - p_e}{1 - p_e} \tag{1}$$

In this study, Cohen's Kappa Score was applied to the manual sentiment classification conducted by 3 people, which was then used to evaluate the pre-trained model. This metric assesses the alignment of the manual sentiment classification and its reliability as a benchmark. Cohen established that a score below 0 indicates very poor alignment, 0.01–0.20 poor, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good, and 0.81–1.00 very good.

One metric that can be used to measure the quality of a model is the confusion matrix. This metric presents a comparison between the model's predictions and actual values in the form of a table. The confusion matrix consists of four components: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). This metric facilitates the interpretation of several other metrics such as accuracy, precision, recall, and F1-Score. In this research, these metrics will be used to provide a comprehensive overview of the sentiment classification model's performance.

D. DASHBOARD'S CREATION

This research will design an interactive dashboard that provides a comprehensive analysis of tweets related to Transjakarta and Trans Semarang with several key features. First, the dashboard will show the three tweets with the most likes for each mode of transportation. This feature allows users to see the content of tweets that have garnered significant approval from the public. Second, the dashboard will show the distribution of tweet sentiments divided by day, making it easier for users to observe sentiments on different days. The data that will be shown on this features is the tweets about TransJakarta or Trans Semarang that are most frequently shared on weekdays and weekends. Lastly, the dashboard will show the sentiment distribution of the ten most relevant topics that have been previously extracted and show the calculated sentiment's score from the number of tweets with neutral (0), positive (+1), or negative (-1) sentiments. This feature is separated between TransJakarta and Trans Semarang to clearly compare the two. The dashboard also provides sentiment, day, and topic filters so users can interact with the dashboard and obtain the desired information. These features provide an overview of public perception toward the two transportation services. It is hoped that this will make it easier for the government to improve public services in line with public perceptions.

This paper proposed solution differs in several ways from previous research, such as the research by Pertiwi [10]. One of the main differences is that this research focuses solely on analyzing tweets about Trans Semarang and Transjakarta, whereas the previous research had a broader scope, analyzing transportation used for homecoming. Additionally, the previous research used classification algorithms such as KNN, SVM, Naive Bayes, and Neural Networks, while this research used a pre-trained model from HuggingFace. Lastly, this research also performed topic extraction using BERTopic to make the analysis more specific.

D. DASHBOARD'S EVALUATION

Lastly, the user experience survey will be used as a measure of the designed dashboard's quality. Some evaluation's aspects include usability and the alignment of the UI/UX design of the dashboard. This survey will provide user's perspectives on the dashboard's quality such as the user feedback on navigation, visual design, functionality, and information delivery.

RESULTS AND DISCUSSION

A. DASHBOARD RESULT

The dashboard's result can be accessed through Looker Studio on https://lookerstudio.google.com/s/lOfwOGtmJaE. The dashboard's pages can be seen in **FIGURE 4**.

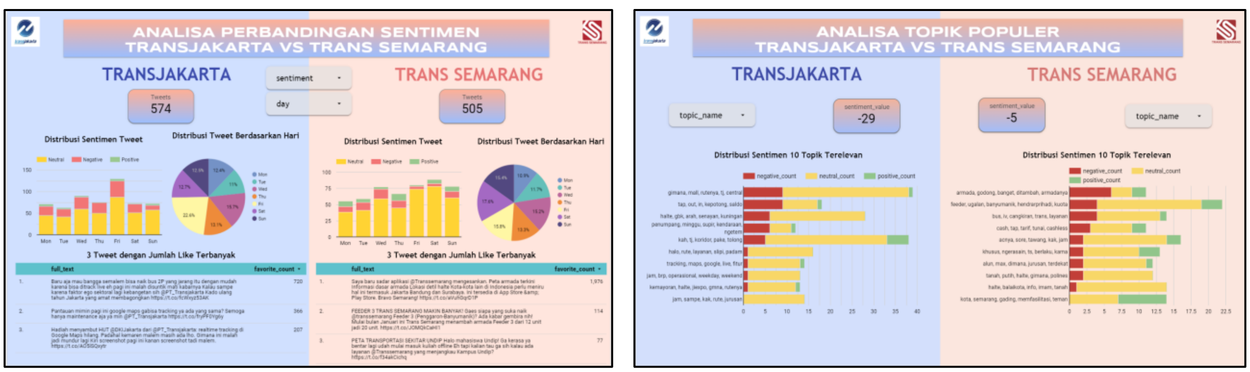

FIGURE 4. Dashboard created based on the topics extracted and sentiment classification

The dashboard includes several key features. First, it displays the three tweets with the most likes for each mode of transportation. This feature allows users to see the content of tweets that have garnered significant approval from the public. Second, the dashboard can show the distribution of tweet sentiments divided by day, making it easier for users to observe sentiments on different days. It also displays the distribution of tweets by day, allowing users to see when tweets about TransJakarta or Trans Semarang are most frequently shared, such as on weekdays or weekends. Lastly, the dashboard can display the sentiment distribution of the ten most relevant topics that have been previously extracted and show the sentiment score calculated from the number of tweets with neutral (0), positive (+1), or negative (-1) sentiments. This feature is separated between TransJakarta and Trans Semarang to clearly compare the two. The dashboard also provides sentiment, day, and topic filters so users can interact with the dashboard and obtain the desired information.

Several insights can be extracted from the sentiments regarding TransJakarta and Trans Semarang. Negative sentiment tweets about TransJakarta are most frequently posted on Fridays. This could be due to the fact that Friday is the last working day of the week, when workers are already tired after five days and expect better service. With this in mind, the government might consider implementing promotional strategies or hosting enjoyable events on Fridays to improve public sentiment toward TransJakarta. In the case of Trans Semarang, there are significantly more positive sentiments compared to TransJakarta, particularly on Thursdays.

Looking at the dashboard's display of the three most-liked tweets for TransJakarta, all three tweets are negative, while none of the top three tweets for Trans Semarang are negative. This indicates that Trans Semarang has won public sentiment in quantity and quality, as many people agree on the advantages of Trans Semarang's facilities and services. In terms of the topics displayed, it is evident that the sentiment toward Trans Semarang is more positive, as shown by the sentiment score for Trans Semarang (-5) compared to TransJakarta (-29), even though both have negative sentiments. For TransJakarta, topics that could be used for evaluation include balance deductions, long waiting times, route clarity, and operational hours. Regarding the live tracking feature, there is an equal number of positive and negative sentiments. For Trans Semarang, improvements could be made regarding fleet shortages, service, and route clarity. Additionally, there are positive sentiments regarding the facilities.

B. INFORMATION EXTRACTION EVALUATION

As an evaluation of the topic modeling, a coherence score of 0.71 was obtained for the topic extraction model of Transjakarta, while the model for Trans Semarang scored 0.70. Both of these scores exceed 0.5, indicating that the results of the topic extraction for Transjakarta and Trans Semarang can be considered reliable. The higher the coherence score, the better the correlation of words within the topics. The results above suggest that the topic model designed for Transjakarta provides slightly better topics compared to Trans Semarang, as seen from the higher score and the greater number of tweets.

To evaluate the sentiment classification model, the researchers obtained Cohen's Kappa Score from the manual classification results by the three annotators. The agreement between the annotators in classifying the sentiment of tweets about Trans Semarang and Transjakarta is considered relatively high, as seen from the Cohen's Kappa Score results: 0.92 between annotator 1 and 2, 0.84 between annotator 1 and 3, and 0.84 between annotator 2 and 3. These scores indicated that the agreements between the three annotators are consistent and can be used as a benchmark for the evaluation.

To measure the performance of the pre-trained sentiment classification model, the calculation the accuracy, precision, recall, F1-score, and the confusion matrix were calculated. For the numerical metrics, the model has the accuracy of 0.78, precision of 0.79, recall of 0.78, and F1-Score of 0.77. These statistics show that this sentiment classification model performs quite well in predicting sentiments in tweets about Transjakarta and Trans Semarang. Based on the confusion matrix's result, there is still significant confusion between sentiment classes 0 (negative) and 1 (neutral), particularly in classifying negative sentiments as neutral. There are also relatively many false positives in class 2 (positive), which indicates the model often incorrectly predicts other classes as positive. However, the model's accuracy is already satisfactory for the overall score. The details of the confusion matrix's results can be seen in **FIGURE 5.**

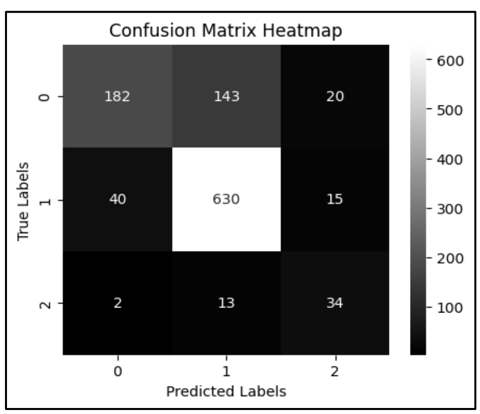

FIGURE 5. Confusion Matrix of Sentiment Classification Prediction Model's Result

D. DASHBOARD'S USER EXPERIENCE EVALUATION

The survey was conducted using Google Forms to evaluate the dashboard's user experience. The researchers chose Google Form to facilitate the data collection process, making it quick and efficient to gather perspectives from respondents. The survey was divided into two parts which are the respondent's information as the first part and the

User Experience Questionnaire (UEQ) to see the usability score, informative score, dashboard's element's contribution for user's understanding score, and the user's challenges as the second part. The survey was successfully obtained a total of 32 respondents, with an age range between 18 and 24 years. They were from various cities in Indonesia, with 96.8% coming from the two cities that the researchers focused on, namely Jakarta and Semarang. With this demographic, it is expected that the respondents will provide critical feedback and have sufficient information about the sentiment comparison between Transjakarta and Trans Semarang.

From **TABLE 1**, the level of contribution of each feature to the interpretation of information by dashboard users can be seen. The scale ranges from one to five, where one means the feature is not contributive, and five means it is highly contributive. The survey's results indicated that the dashboard's performance is quite satisfactory. The average score for the ease of interpreting the user interface was 4.34 out of 5, with 15 respondents giving a score of 5. The average score for the informativeness of the dashboard was 4.62 out of 5, with 20 respondents giving a score of 5. Overall, all features on the dashboard are considered contributive by the respondents, Additionally, there is an evaluation related to the placement of the dropdown features for topic names, sentiment types, and days. Thirty-one out of 32 respondents indicated that the placement is appropriate within the dashboard. The average user satisfaction level while using the dashboard is 8.56 out of 10. As for the challenges, most respondents did not experience any difficulties while using the dashboard. However, some respondents felt that our dashboard needed more detail and that the bar charts were difficult to understand, especially the tweet's sentiment distribution and the sentiment distribution of the 10 most relevant topics that could be displayed using a clearer visualization method to present information more intuitively.

TABLE 1. Contribution Scale of Dashboard's Features

Feature/Contribution Scale	1	2	3	4	5	Average
Tweets Sentiment Distribution	1	0	5	14	12	4,12
Tweets Distribution (per Day)	0	0	1	14	17	4,5
3 Tweets with the most number of Like	0	1	1	10	20	4,53
Total Tweets	0	0	3	6	23	4,62
10 Relevant Topics Tweets Sentiment	0	0	3	18	11	4,35

CONCLUSIONS

Based on the analysis conducted, the result is that the public sentiment towards Trans Semarang is more favorable compared to Transjakarta. In this study, a dashboard was presented as a resource that can help improve the quality of Transjakarta and Trans Semarang through the integration of topic modeling with BERTopic, which has coherence scores of 0.71 and 0.70, and a sentiment classification model with an accuracy of 0.78. The proposed dashboard has proven to provide valuable information on the sentiment comparison between Transjakarta and Trans Semarang, with an average user satisfaction score of 8.56 out of 10. Future research can focus on improving the accuracy of the sentiment classification model and refining the dashboard's interface design to make information delivery more intuitive and high-quality.

REFERENCES

[1] Badan Pusat Statistik Indonesia, "Jumlah Penduduk Pertengahan Tahun - Tabel Statistik - Badan Pusat Statistik Indonesia." Accessed: Aug. 21, 2024. [Online]. Available: https://www.bps.go.id/id/statistics-table/2/MTk3NSMy/jumlah-penduduk-pertengahan-tahun--ribu-jiwa-.html

[2] D. D. Putra, R. D. Aufaa, H. Luthfiyah, and S. Sahara, "Peningkatan Mutu Transportasi Umum Demi Kenyamanan dan Keamanan Pengguna," *MIMBAR ADMINISTRASI FISIP UNTAG Semarang*, vol. 20, no. 1, pp. 112–119, 2023.

[3] A. (Avi) Ceder, "Urban Mobility and Public Transport: Future Perspectives and Review," *International Journal of Urban Sciences*, vol. 25, no. 4, pp. 455–479, 2021, doi: 10.1080/12265934.2020.1799846.

[4] N. Rohman, "Pengaruh Transportasi Umum Terhadap Pelestarian Lingkungan Di Kota Jakarta," 2019. [Online]. Available: https://www.researchgate.net/publication/333718194

[5] S. Das, A. Boruah, A. Banerjee, R. Raoniar, S. Nama, and A. K. Maurya, “Impact of COVID-19: A radical modal shift from public to private transport mode,” *Transp Policy (Oxf)*, vol. 109, pp. 1–11, Aug. 2021, doi: 10.1016/j.tranpol.2021.05.005.

[6] H. Fadhil, “Peringkat Kemacetan Jakarta Naik ke Posisi 29 Dunia pada 2022.” Accessed: Jul. 22, 2024. [Online]. Available: https://news.detik.com/berita/d-6570551/peringkat-kemacetan-jakarta-naik-ke-posisi-29-dunia-pada-2022

[7] F. F. Rachman, R. Nooraeni, and L. Yuliana, “Public Opinion of Transportation integrated (Jak Lingko), in DKI Jakarta, Indonesia,” in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 696–703. doi: 10.1016/j.procs.2021.01.057.

[8] A. R. Rahmatullah, D. I. K. Dewi, and C. D. T. Nurmasari, “Integrasi Antar Transportasi Umum di Kota Semarang,” *Jurnal Pengembangan Kota*, vol. 10, no. 1, pp. 36–46, Jul. 2022, doi: 10.14710/jpk.10.1.36-46.

[9] S. Sukarmi and D. Suwondo, “Kebijakan Penggunaan Moda Transportasi Umum untuk Kesejahteraan Ekonomi Masyarakat Kota Semarang,” *Masalah-Masalah Hukum*, vol. 48, no. 2, pp. 194–203, 2019, [Online]. Available: http://repository.umy.ac.id/handle/123456

[10] M. W. Pertiwi, “Inti Nusa Mandiri Analisis Sentimen Opini Publik Mengenai Sarana Dan Nusa Mandiri Analisis Sentimen Opini Publik Mengenai Sarana dan Transportasi Mudik Tahun 2019 Pada Twitter Menggunakan Algoritma Naïve Bayes, Neural Network, KNN dan SVM,” *Inti Nusa Mandiri*, vol. 14, no. 1, pp. 27–32, 2019, [Online]. Available: http://www.nusamandiri.ac.id

[11] D. Krisnawati, A. Astutik, A. Indrasetianingsih, and F. Fitriani, “Penerapan Text Mining pada Analisis Sentimen Pengguna Twitter Layanan Transportasi Online Menggunakan Metode Density Based Spatial Clustering of Applications with Noise (DBSCAN) dan K-Means,” *J Statistika*, vol. 15, no. 1, pp. 184–194, 2022, [Online]. Available: www.unipasby.ac.id

[12] Ilhamsyah and S. Rahmayuda, “Pemanfaatan Dashboard pada Olah Data Mining Berdasarkan Analisa Pola Hubungan Data Kecelakaan Lalu Lintas Kota Pontianak,” *CYBERNETICS*, vol. 4, no. 1, pp. 1–8, 2020.

[13] Y. Fan, H. P. Teo, and W. X. Wan, “Public transport, noise complaints, and housing: Evidence from sentiment analysis in Singapore,” *J Reg Sci*, vol. 61, no. 3, pp. 570–596, Jun. 2021, doi: 10.1111/jors.12524.

[14] S. Ali, G. Wang, and S. Riaz, “Aspect based sentiment analysis of ridesharing platform reviews for kansei engineering,” *IEEE Access*, vol. 8, pp. 173186–173196, 2020, doi: 10.1109/ACCESS.2020.3025823.

[15] M. A. Habib and M. A. H. Anik, “Impacts of COVID-19 on Transport Modes and Mobility Behavior: Analysis of Public Discourse in Twitter,” in *Transportation Research Record*, vol. 2677, no. 4, SAGE Publications Ltd, 2023, pp. 65–78. doi: 10.1177/03611981211029926.

[16] M. Grootendorst, “BERTopic: Neural topic modeling with a class-based TF-IDF procedure,” Mar. 2022, Accessed: Aug. 26, 2024. [Online]. Available: https://arxiv.org/abs/2203.05794v1

[17] M. Röder, A. Both, and A. Hinneburg, “Exploring the space of topic coherence measures,” *WSDM 2015 - Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, pp. 399–408, Feb. 2015, doi: 10.1145/2684822.2685324.