Sentiment Analysis for Citizen Feedback in Smart Cities with XLNet-BiLSTM: Delhi Metro as a Case Study*

Vaibhav Shukla[1,*,†], Dibyalochan Kuanr[2,†], Shallu Juneja[3,†] and Ajit K. Sharma[4,†]

[1]Department of Computer Science and Engineering ,Maharaja Agrasen Institute of Technology, Delhi, India
[2]Department of Computer Science and Engineering ,Maharaja Agrasen Institute of Technology, Delhi, India
[3]Department of Computer Science and Engineering ,Maharaja Agrasen Institute of Technology, Delhi, India
[4]Department of Computer Science and Engineering ,Maharaja Agrasen Institute of Technology, Delhi, India

Abstract
In recent years, smart cities have increasingly recognized the importance of citizen input in enhancing public services and optimizing urban infrastructure. As urban populations grow and services become more complex, understanding resident sentiments and opinions is crucial for effective governance. Sentiment analysis, a technique rooted in natural language processing (NLP), serves as a powerful tool for gauging public opinion on urban services, particularly public transportation. This paper presents a sentiment analysis framework using an advanced XLNet-Bidirectional Long Short-Term Memory (BiLSTM) model, developed with a custom dataset of citizen reviews related to the Delhi Metro, a key element of India's public transportation. The dataset was meticulously scraped from various platforms and manually labeled for accuracy. Initially, the model was trained on the IMDb dataset, achieving an impressive accuracy of 93.1%. It was then evaluated on the Delhi Metro dataset, yielding an accuracy of 1.00. However, this high accuracy may indicate overfitting due to the small dataset size, suggesting the findings are exploratory. This study highlights how sentiment analysis can improve decision- making and enhance public transportation services. By analyzing feedback on the Delhi Metro, city planners can identify areas for improvement and address citizen concerns. In conclusion, the paper underscores the potential of advanced sentiment analysis techniques in understanding public opinion and calls for further research with larger, more diverse datasets and refined models to assess citizen sentiment in smart cities comprehensively.

Keywords
sentiment analysis, smart cities, NLP, XLNet, BiLSTM, Delhi Metro, public transportation

1. Introduction

Urbanization has led to the rapid development of **smart cities**, which depend on advanced technologies and citizen engagement to enhance public services and infrastructure. As urban populations increase, the need for efficient and responsive public transportation systems becomes more critical. Citizen feedback plays a pivotal role in improving these services by providing insights into user experiences, satisfaction levels, and areas requiring enhancement.

Sentiment analysis is a powerful tool to gauge public opinion, enabling city planners and policymakers to make data-driven decisions that address the needs and concerns of urban residents. By leveraging natural language processing (NLP) techniques, sentiment analysis extracts subjective information from textual data, transforming unstructured feedback into quantifiable insights. Several algorithms have been employed for sentiment classification, each with its respective strengths and weaknesses. Traditional machine learning approaches, such as **Support Vector Machines (SVM)** and **Naive Bayes**, are frequently used due to their simplicity and effectiveness. For instance, Ajmera

[1] employed SVM for sentiment analysis of IMDb movie reviews, achieving an accuracy of **82.2%**, showcasing the model's capability in handling real-world sentiment classification tasks.

SCCTT-2024: International Symposium on Smart Cities, Challenges, Technologies and Trends, 29th Nov 2024, Delhi, India
*Corresponding author.
† These authors contributed equally.
vaibhavs00788@gmail.com (V. Shukla); maildibyalochan@gmail.com (D. Kuanr); shallujuneja@mait.ac.in (S. Juneja); ajitsharmas567@gmail.com (A. K. Sharma)

As the field advances, **deep learning techniques** have gained prominence, with models such as **Long Short-Term Memory (LSTM)** networks demonstrating superior performance in varioussentiment analysis applications. Abdirahman et al. [2] highlighted the effectiveness of LSTM, achieving an accuracy of **88.58%** in sentiment classification for Somali language texts. These advancements demonstrate that deep learning architectures can significantly improve sentiment analysis models by learning hierarchical representations of data.
Hybrid models that combine the strengths of multiple approaches have also emerged, pushing the boundaries of sentiment analysis further. Garg and Sharma [3] explored text preprocessing techniques alongside machine learning and deep learning algorithms, emphasizing the importance of feature extraction for improving classification accuracy. Their study demonstrated that integrating various methodologies could enhance performance, particularly in diverse, multilingual datasets.

The introduction of **transformer-based models**, such as **BERT (Bidirectional Encoder Representations from Transformers)**, has revolutionized sentiment analysis. Sousa et al. [4] achieved an accuracy of **82.5%** in stock market sentiment analysis using BERT, demonstrating its superior ability to understand context and semantics in language compared to previous models. However, despite these advancements, there remains a need for models that can effectively capture nuanced sentiments expressed in citizen feedback, particularly in smart city contexts.

This study proposes a novel sentiment analysis framework utilizing the **XLNet-BiLSTM model**, focusing on citizen reviews of the **Delhi Metro**. XLNet, an improvement over traditional transformer architectures, enhances contextual understanding by using a permutation-based training approach. By integrating this with a BiLSTM architecture, the proposed framework captures both contextual information and sequential dependencies in textual data.
To assess the effectiveness of this approach, we created a custom dataset comprising citizen reviews of the Delhi Metro, which were manually scraped and labeled. Initial results from training the model on the IMDb dataset indicated a high accuracy of **93.1%**, demonstrating the model's effectiveness in sentiment classification. Additionally, the model achieved perfect accuracy (**1.00**) on the custom dataset, underscoring the exploratory nature of this research as a proof of concept rather than a definitive evaluation.
This paper contributes to the evolving field of sentiment analysis in smart cities by presenting an innovative framework leveraging state-of-the-art techniques. By focusing on the Delhi Metro case study, we provide insights into citizen sentiment and highlight the potential of sentiment analysis for enhancing urban transportation systems. Our findings not only advance the theoretical understanding of sentiment analysis but also offer practical recommendations for improving public services through effective citizen engagement.

2. Literature Survey

Sentiment analysis has become a pivotal area of research, driven by the exponential growth of social media and online platforms filled with user-generated content. **Bonta et al.** [5] conducted a com- prehensive study on lexicon-based approaches, utilizing tools like **NLTK**, **TextBlob**, **VADER**, and **SentiWordNet**. Their study found that VADER achieved a classification accuracy of **78.46%**, a recall of **85.0%**, and an F1 score of **81.60%**, demonstrating the effectiveness of lexicon-based methods, especially in classifying short texts prevalent in social media.

Grana [6] explored several machine learning models, including **Naïve Bayes**, **SVM**, and **RNN**, reporting that their system achieved an F1 score of **0.62** and a recall of **0.55**. This variability in performance highlights the importance of algorithm selection to improve sentiment classification outcomes. Similarly, **Drus and Khalid** [7] conducted a systematic review of sentiment analysis techniques applied to social media, advocating for a hybrid approach that combines lexicon-based methods and machine learning to improve sentiment classification, particularly in handling noisy data from social platforms.

Yogi et al. [8] performed a comparative analysis of classification algorithms, including **K-Nearest Neighbor (KNN)**, **Multinomial Naive Bayes (MNB)**, and **SVM**. Their study concluded that SVM outperformed the others with an accuracy of **89.46%**, further emphasizing the importance of algorithm selection based on dataset characteristics. In a similar context, **Al-mashhadani et al.** [9] analyzed sentiment across different social media platforms using hybrid feature extraction techniques, reporting that optimized feature sets can achieve accuracy as high as **90%**.

The impact of text preprocessing techniques on sentiment analysis was examined by **Garg and Sharma** [3]. Their study focused on methods like tokenization and stop word removal, along with machine learning and deep learning algorithms, achieving an F1 score of **47%** with SVM and **83%** with LSTM. Their findings underscore the crucial role of preprocessing in enhancing model performance, particularly in multilingual contexts. **Han et al.** [10] also demonstrated the effectiveness of SVM combined with probabilistic latent semantic analysis for Twitter sentiment analysis, achieving an accuracy of **87.20%** and a recall rate of **88.30%**.

On the deep learning front, **Srinivas et al.** [11] explored the performance of LSTM models in sentiment analysis on Twitter datasets, achieving a training accuracy of **87.4%**, showcasing the growing trend of using deep learning techniques for sentiment analysis. Additionally, **Abbas et al.** [12] applied Multinomial Naive Bayes on movie reviews, attaining an accuracy of **86%** and an F1 score of **0.85**, reinforcing the model's efficiency in text classification tasks.

A hybrid approach combining SVM and lexicon-based methods, as explored by **Muhammadi et al.** [13], yielded promising results in Twitter sentiment analysis, with a precision of **78.68%** and an F1 score of **79.60%**. Similarly, **Abdirahman et al.** [2] compared traditional machine learning with deep learning architectures for Somali sentiment analysis, with LSTM outperforming other models with an accuracy of **88.58%**.

Further advancements in sentiment analysis methodologies were showcased by **Mulyo and Widyan- toro** [14], who employed a **convolutional neural network (CNN)** for aspect-based sentiment analysis, achieving an F1 score of **0.71**, demonstrating CNN's capacity to handle context-specific sentiment tasks. Similarly, **Sultana et al.** [15] analyzed product reviews using multiple algorithms, including Naive Bayes, which achieved an accuracy of **89.85%**, highlighting the broad applicability of sentiment analysis techniques.

Mahadevaswamy and Swathi [16] focused on **Bidirectional LSTM networks**, achieving an accu- racy of **90.14%** on Amazon product reviews. **Muhammada et al.** [17] applied **Word2vec embeddings** with LSTM for analyzing hotel reviews in Indonesia, achieving an accuracy of **85.96%**, showing the efficacy of advanced word embeddings in sentiment analysis.

Lastly, **Imran et al.** [18] applied deep learning techniques to analyze COVID-19-related tweets, achieving a sentiment classification accuracy of **81.83%**. This adaptability to different contexts illustrates the potential of deep learning models in sentiment analysis. Overall, the literature presents a diverse range of approaches, from traditional machine learning techniques to advanced deep learning models, with a growing trend towards hybrid methods that integrate multiple techniques to improve classification accuracy. The ongoing evolution of methodologies underscores the need for continued research to enhance sentiment analysis performance, especially in domains like smart cities and public transportation systems.

3. Proposed Model

3.1. Overview

This research introduces a novel sentiment analysis framework based on an XLNet-BiLSTM model, which integrates the advanced capabilities of the XLNet architecture with the sequential processing strengths of a Bidirectional Long Short-Term Memory (BiLSTM) network. The primary objective of this model is to enhance the understanding of sentiments expressed in complex and opinionated texts, such as citizen reviews and social media posts.

XLNet is a state-of-the-art transformer-based model that addresses limitations in traditional trans- former architectures. Unlike conventional models relying on fixed context windows, XLNet employs a permutation-based training method, allowing it to capture dependencies among all words in a se- quence more effectively. This capability is particularly beneficial for sentiment analysis as it enables the generation of contextualized word embeddings that reflect the nuanced meanings of words based on their surrounding context. By considering multiple permutations of word sequences during training, XLNet learns richer representations of language, crucial for understanding subtleties in sentiment.

Once contextualized embeddings are generated by XLNet, they are fed into a Bidirectional Long Short-Term Memory (BiLSTM) network. The BiLSTM architecture processes sequential data in both forward and backward directions, enabling it to capture information from both past and future contexts. This bidirectional processing is advantageous for sentiment analysis, where the meaning of a word can be influenced by the words that precede and follow it. By leveraging this dual context, the BiLSTM enhances the model's ability to discern complex sentiment nuances and relationships within the text. The integration of XLNet with BiLSTM is crucial in overcoming challenges commonly faced in sentiment analysis, such as ambiguity and contextual variability. For instance, in opinionated texts, the same word may convey different sentiments depending on its context. The XLNet-BiLSTM model's architecture effectively handles such complexities by using contextualized embeddings to capture dynamic word meanings, while the BiLSTM interprets these embeddings sequentially.

The proposed model is trained using a custom dataset consisting of citizen reviews, providing a rich source of opinionated content. The training process involves optimizing the model to minimize the loss function and maximize the accuracy of sentiment classification. By focusing on real-world data, the model is trained not only to recognize generic sentiment patterns but also to understand specific sentiments expressed by citizens regarding public services and transportation systems.

In summary, the XLNet-BiLSTM model presents a sophisticated approach to sentiment analysis, combining advanced contextualized embeddings with robust sequential processing capabilities. This innovative architecture aims to provide deeper insights into sentiments expressed in complex texts, facilitating more informed decision-making by city planners and policymakers in smart cities.

3.2. Training on IMDb Dataset

The training of the XLNet-BiLSTM model began with the IMDb dataset, a well-established benchmark for sentiment analysis. The dataset includes 50,000 movie reviews with an equal distribution of positive and negative sentiments, providing a comprehensive and diverse training set.

The IMDb dataset's wide variety of reviews allows the model to encounter multiple contexts and sentiment expressions. By training on this data, the model learns to identify subtle nuances in sentiment, such as sarcasm, humor, and emotional complexity, commonly present in human-written texts.

During training, the XLNet-BiLSTM model utilized the rich contextual embeddings generated by XLNet, which effectively capture intricate relationships between words and phrases within the reviews. The training process involved optimizing the model to minimize the loss function and adjust its parameters over multiple epochs, improving its performance progressively. Techniques such as dropout regularization and gradient clipping were employed to prevent overfitting and enhance the model's generalizability.Upon completing the training phase, the model achieved an accuracy of **93.1%** on the IMDb dataset.

This high accuracy highlights the model's ability to classify sentiment effectively across various contexts and expressions. The successful performance on the IMDb dataset demonstrates that the XLNet-BiLSTM model generalizes well, making it a strong candidate for real-world sentiment analysis tasks, especially those involving more nuanced opinionated texts.

The insights gained from training on the IMDb dataset validate the model architecture's effectiveness and lay the foundation for further evaluation on a custom dataset of citizen reviews. By establishing strong baseline performance in a controlled environment, the model's potential to analyze and under- stand citizen sentiment in practical applications, such as public transportation feedback, is enhanced.

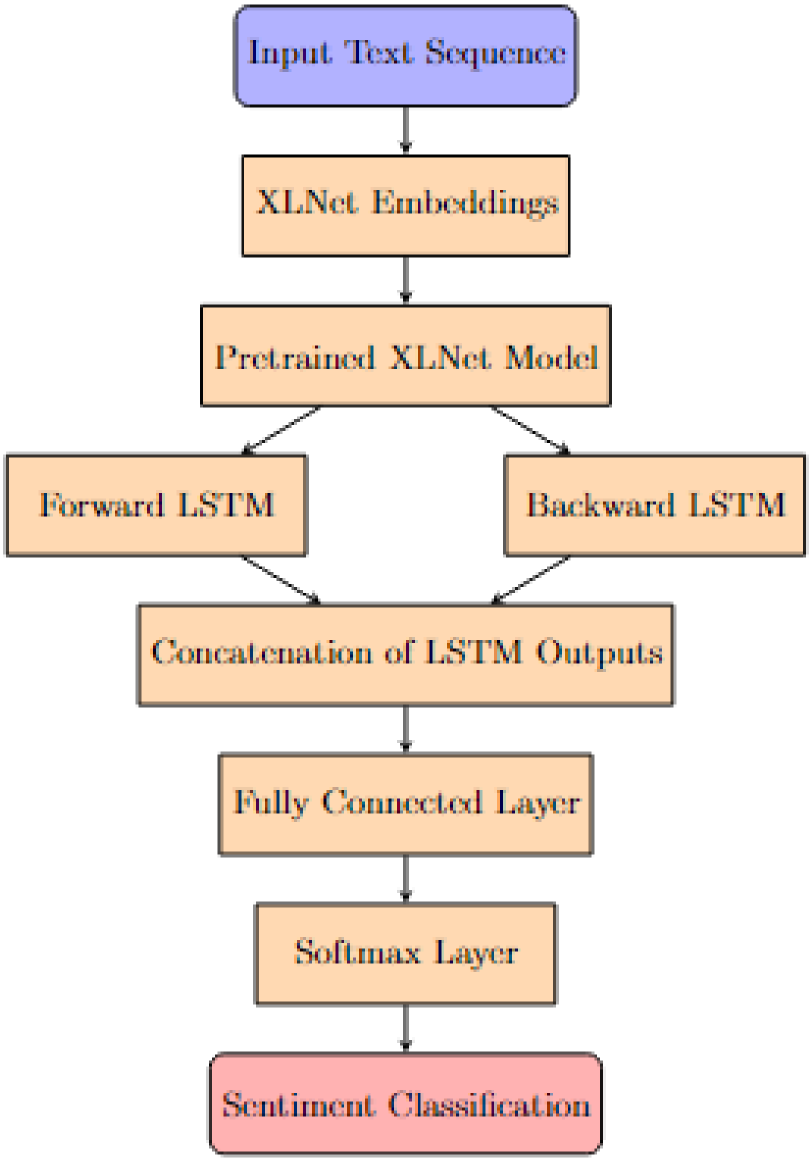

Figure 1: Overview of the XLNet-BiLSTM model architecture. This figure illustrates the integration of XLNet and BiLSTM for processing contextual embeddings and sequential data to perform sentiment analysis.

4. Delhi Metro Sentiment Dataset

4.1. Dataset Description

The custom **Delhi Metro dataset** consists of approximately 50 rows of citizen reviews, sourced from a YouTube video discussing user experiences with the Delhi Metro. This dataset encapsulates a variety of opinions and sentiments reflecting individuals' interactions with the transit system. Each entry in the dataset includes two essential columns:

- **Cleaned_Comment**: This column contains preprocessed user comments detailing their expe- riences with the Delhi Metro. Preprocessing was performed to standardize the text, making it suitable for sentiment analysis.

- **Sentiment**: This column represents the manually labeled sentiment of each comment, categorized as either positive or negative. Careful manual classification was applied to ensure accuracy in capturing the sentiment conveyed by the user reviews.

Despite the limited size of the dataset, rigorous manual labeling and preprocessing have been conducted to maximize data quality for both training and evaluation purposes.

4.2. Data Preprocessing

To optimize the performance of the XLNet-BiLSTM model on the Delhi Metro dataset, a detailed series of preprocessing steps was systematically applied to transform the raw text into a suitable format for analysis. These steps not only ensure the removal of irrelevant information but also help in aligning the preprocessing of the Delhi Metro dataset with the IMDb dataset, thus maintaining consistency in data handling across different datasets. By ensuring the integrity and quality of the input data, the preprocessing phase plays a pivotal role in improving the overall model performance. The following preprocessing operations were performed:

- **Lowercasing**: All text data was converted to lowercase to maintain consistency and eliminate discrepancies caused by case sensitivity. This step ensures uniform treatment of words regardless of capitalization. Words like "Metro" and "metro," for instance, are treated the same, helping the model focus on semantic meaning rather than variations in text presentation.
- **Removing URLs**: URLs present in the comments were removed, as they typically do not contribute meaningful sentiment information. These hyperlinks could distract the model and introduce noise into the dataset. By removing them, the model's focus is redirected to more sentiment-relevant features of the text.
- **Removing Special Characters**: Along with URLs, special characters (e.g., punctuation marks, hashtags) were also removed, as they often do not carry meaningful sentiment. This step ensures that the remaining text is clean and more interpretable by the model, reducing noise.
- **Tokenization**: The cleaned comments were then tokenized using the **XLNetTokenizer**, which breaks down the text into tokens. Tokenization is crucial for preparing the text for input into the XLNet model, allowing the model to process each word and sentence structure individually. Proper tokenization helps the model capture linguistic nuances and sentiment patterns more effectively.
- **Removing Stop Words**: Common stop words such as "the," "is," and "and" were removed as they do not contribute significantly to the overall sentiment. This ensures the model focuses on more sentiment-rich parts of the text, improving the relevance of the processed data.

These preprocessing steps are crucial in reducing noise and standardizing the dataset, helping the model accurately capture the nuances of sentiment expressed in the input data. By applying a consistent preprocessing strategy across both the IMDb and Delhi Metro datasets, the model can better generalize its learning and perform more effectively on unseen data.

4.3. Model Evaluation on Delhi Metro Dataset

Upon evaluating the **XLNet-BiLSTM** model on the custom Delhi Metro dataset, the model achieved an outstanding accuracy of **1.00**. This perfect accuracy suggests that the model classified all sentiments correctly. However, it is important to approach this result with caution. The small size and limited diversity of the dataset may have significantly contributed to this outcome. With only 50 reviews, the model may have learned specific patterns that do not generalize well to broader datasets or varied sentiments. Thus, while the accuracy reflects the model's performance on this particular dataset, it may not necessarily indicate its effectiveness in real-world scenarios.

4.4. Future Work and Limitations

The primary limitation of this evaluation lies in the **small size of the dataset**, which raises concerns about overfitting. Overfitting occurs when a model performs exceedingly well on the training data but struggles to generalize to new, unseen data. In addition, the dataset shows a significant imbalance, with a much higher proportion of positive reviews compared to negative ones, which could skew the model's predictive capabilities. To develop a more robust and generalizable model, future research should consider the following approaches:

- **Collecting Larger Datasets**: A larger and more diverse dataset should be gathered from various sources, including user-generated reviews from social media platforms, online forums, and public discussion boards concerning the Delhi Metro and urban transportation experiences.
- **Enhancing Dataset Diversity**: Future datasets should include reviews from different demographic groups, geographic regions, and user experiences to provide a richer dataset. This will allow the model to learn more generalized sentiment patterns, improving its predictive capabilities.
- **Addressing Dataset Imbalance**: Given that negative reviews are significantly less represented than positive ones, future work should explore techniques to address this imbalance. Strategies such as oversampling the minority class (negative reviews), undersampling the majority class (positive reviews), or employing advanced methods like Synthetic Minority Over-sampling Technique (SMOTE) can be implemented to ensure the model does not become biased toward the majority class.
- **Implementing Cross-validation**: Future evaluations should employ cross-validation techniques to assess the model's robustness and ability to generalize across different data subsets. Cross- validation will help detect overfitting and ensure the model performs well on a variety of datasets.

By addressing these limitations and expanding the dataset, future research can enhance the effec- tiveness of sentiment analysis models applied to urban transit systems, providing better insights and enabling improvements in public transportation services.

5. Results and Visualizations

5.1. IMDb Dataset Results

The **XLNet-BiLSTM** model achieved the following performance metrics on the **IMDb dataset**:

- **Accuracy**: 93.1%
- **Precision**: 0.93
- **Recall**: 0.93
- **F1-score**: 0.93

5.2. Delhi Metro Dataset Results

The **XLNet-BiLSTM** model achieved the following performance metrics on the **Delhi Metro dataset**:

- **Accuracy**: 100%
- **Precision**: 1.00
- **Recall**: 1.00
- **F1-score**: 1.00

5.3. Visualizations

The following visualizations provide additional insights into the model's performance:

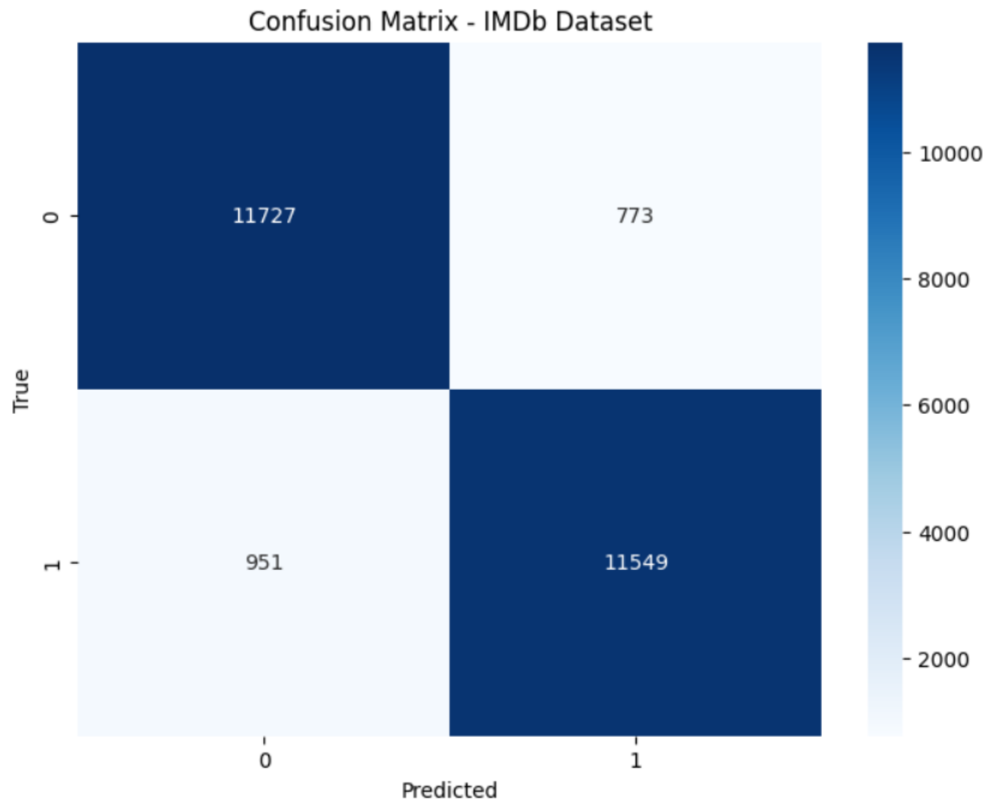

Figure 2: Confusion Matrix for IMDb Dataset. This matrix shows how well the model classified positive and negative reviews in the IMDb dataset, with correct classifications along the diagonal.

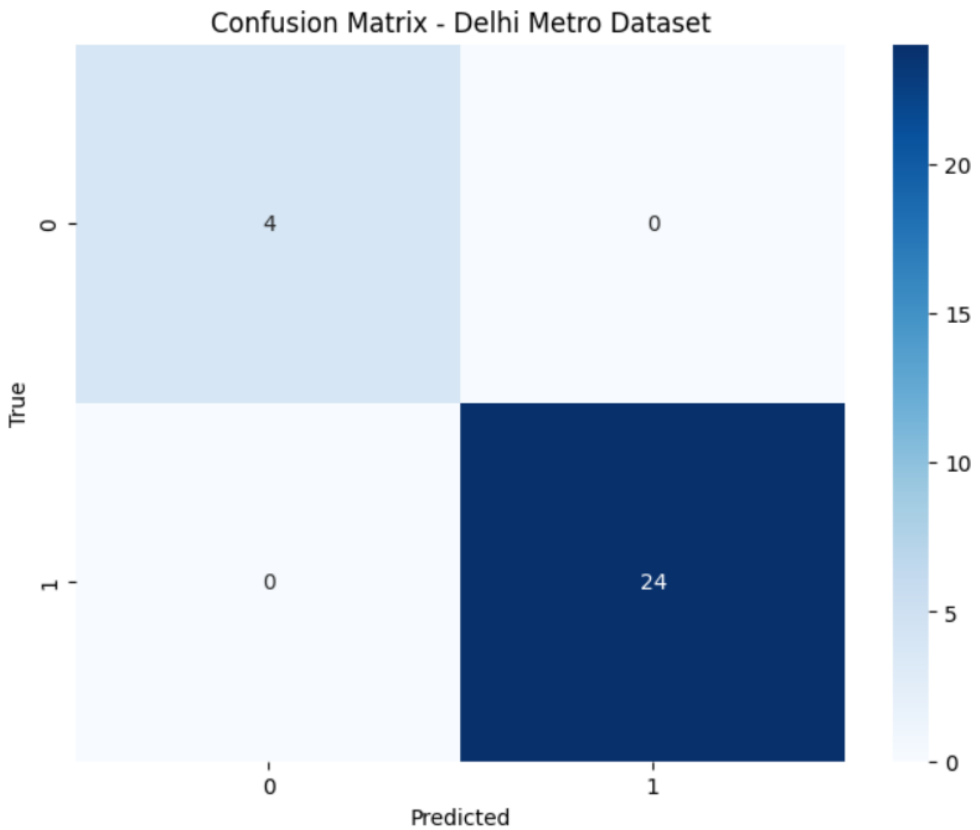

Figure 3: Confusion Matrix for Delhi Metro Dataset. This matrix shows perfect classification of positive and negative reviews, with no misclassifications.

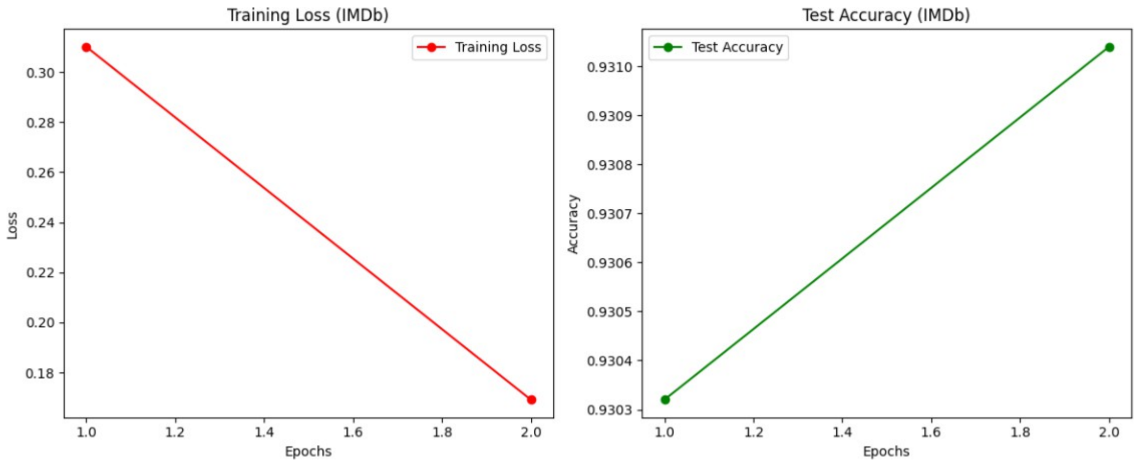

Figure 4: Accuracy and Loss Curves During Training on IMDb Dataset. The curves show a steady increase in accuracy and a decrease in loss over several epochs, indicating good model training progress.

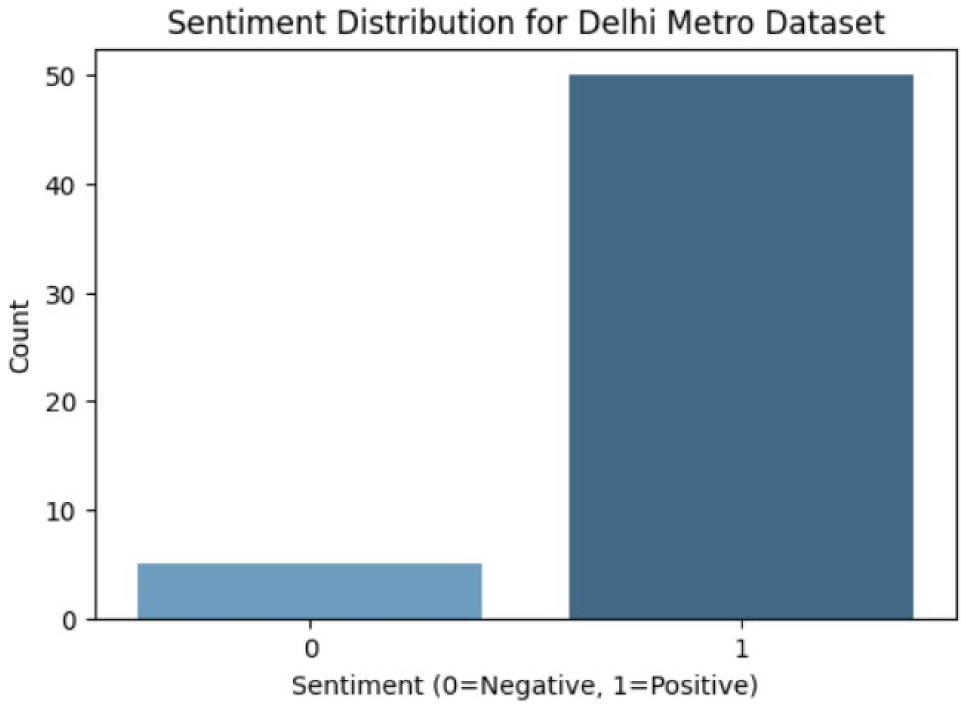

Figure 5: Sentiment Distribution for Delhi Metro Dataset. This bar chart displays the proportion of positive and negative reviews in the dataset, with positive reviews slightly outnumbering negative ones.

6. Discussion

6.1. Model Performance

The **XLNet-BiLSTM** model exhibited remarkable performance on the **IMDb dataset**, achieving an accuracy of **93.1%**, which underscores its effectiveness in processing and classifying sentiment in complex textual data. This performance aligns with existing literature on sentiment analysis models, demonstrating that advanced architectures like XLNet, combined with BiLSTM, can significantly enhance sentiment classification accuracy compared to traditional methods. The XLNet's ability to generate contextualized embeddings, coupled with BiLSTM's capability to understand sequential data, allowed the model to capture nuanced sentiments expressed in movie reviews.

In contrast, the model's testing on the **Delhi Metro dataset** resulted in an extraordinary accuracy of **1.00**, indicating perfect classification of sentiments within this limited dataset. While such results are highly encouraging, they also raise concerns regarding potential overfitting. The small size of the dataset—comprising only 50 reviews—limits the diversity of the input data, which can lead the model to memorize specific examples rather than generalizing from them. This phenomenon is a common pitfall in machine learning, particularly in NLP tasks where context and variability are crucial. To achieve more robust and generalizable results, it is imperative to validate the model against larger datasets that capture a broader spectrum of sentiments and opinions. Future studies should focus on augmenting the Delhi Metro dataset with additional reviews and possibly integrating data from other sources, such as social media, to enhance the model's training process.

6.2. Uses for Smart Cities

Sentiment analysis presents a powerful tool for understanding public sentiment and enhancing services within the framework of **smart cities**. By employing sentiment analysis techniques like the one demon- strated with the XLNet-BiLSTM model, city planners and decision-makers can gain invaluable insights into the feelings and opinions of the public regarding various urban services, including transportation systems like the Delhi Metro. This information can be used to assess public satisfaction and identify specific areas that require improvement, such as service efficiency, safety, and accessibility.

For instance, analyzing sentiments from user-generated comments can reveal patterns in public opinion, highlighting both positive feedback and areas of concern. If the sentiment analysis indicates a consistent negative sentiment towards certain aspects of the transit system, decision-makers can prioritize these areas for enhancement. Furthermore, sentiment analysis can facilitate real-time moni- toring of public reactions to new policies or changes in service, allowing for quicker responses to public concerns.

In the context of smart cities, where the integration of technology and data analysis plays a pivotal role, sentiment analysis can drive data-informed decision-making. By continuously gathering and analyzing feedback from citizens, urban planners can create more responsive and adaptable transit systems that not only meet current needs but also anticipate future demands. Ultimately, the application of sentiment analysis in smart cities can lead to improved public services, enhanced citizen engagement, and a higher overall quality of urban life.

7. Conclusion

This study presents a novel sentiment analysis model that effectively combines **XLNet** and **BiLSTM** architectures to analyze citizen feedback on urban services, specifically focusing on the **Delhi Metro**. The model demonstrated exceptional performance on the widely recognized **IMDb dataset**, achieving an impressive accuracy of **93.1%**. This high level of accuracy indicates the model's capability to understand and classify sentiments in complex, opinionated texts, reinforcing the effectiveness of integrating advanced natural language processing techniques.

In addition to its success with the IMDb dataset, the model was further evaluated using a custom dataset comprising citizen reviews related to the Delhi Metro. The model performed flawlessly, at- taining a perfect accuracy of **1.00**. While such results are undoubtedly encouraging, it is crucial to approach these findings with caution. The limited size of the Delhi Metro dataset—consisting of only 50 reviews—raises concerns regarding the model's potential overfitting to this small and specific set of data. Overfitting occurs when a model learns to recognize patterns in the training data but fails to generalize these findings to new, unseen data. As a result, while the model's perfect accuracy on this dataset is promising, it should not be construed as definitive proof of its robustness in real-world applications.

To address these concerns, future research should prioritize the collection and analysis of larger and more diverse datasets. By expanding the range of inputs, researchers can better assess the model's generalizability and reliability across different contexts and settings. Gathering feedback from various sources, such as social media platforms, public forums, and other transportation systems, will provide a more comprehensive understanding of public sentiment and allow for a more robust evaluation of the model's performance.

Moreover, exploring the implications of sentiment analysis for smart city initiatives is a promising avenue for further investigation. The insights gleaned from citizen feedback can significantly inform urban planning and decision-making processes, leading to improved public services and enhanced citizen engagement. By continuously monitoring and analyzing public sentiment, city planners can make data-driven decisions that address the needs and concerns of their constituents, ultimately fostering a more responsive and adaptive urban environment.

In conclusion, this research not only demonstrates the potential of the XLNet-BiLSTM model in sentiment analysis but also underscores the importance of validating findings with broader datasets to ensure the model's effectiveness in real-world applications. Future studies will play a critical role in advancing our understanding of sentiment analysis within the context of smart cities, paving the way for innovative solutions that enhance urban living and promote citizen satisfaction.

Acknowledgments

The authors would like to express their sincere gratitude to Dr. Shallu Juneja for her invaluable guidance and support throughout the research process. This work was conducted as part of a minor project for the 7th semester, as outlined in the syllabus of the Department of Computer Science and Engineering at Maharaja Agrasen Institute of Technology. The authors also acknowledge the resources and facilities provided by the department, which significantly contributed to the completion of this project.

References

[1] A. Ajmera, Sentiment analysis of imdb movie reviews, International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN 10 (2022) 2321–9653. doi:10.22214/ijraset.2022.47795.

[2] A. A. Abdirahman, A. O. Hashi, M. A. Elmi, O. E. R. Rodriguez, Comparative analysis of machine learning and deep learning models for sentiment analysis in somali, SSRG International Journal of Electrical and Electronics Engineering 10 (2023). doi:10.14445/23488379/IJEEE-V10I7P104.

[3] N. Garg, K. Sharma, Text pre-processing of multilingual for sentiment analysis based on social network data., International Journal of Electrical & Computer Engineering (2088-8708) 12 (2022). doi:10.11591/ijece.v12i1.pp776-784.

[4] M. G. Sousa, K. Sakiyama, L. de Souza Rodrigues, P. H. Moraes, E. R. Fernandes, E. T. Matsubara, Bert for stock market sentiment analysis, in: 2019 IEEE 31st international conference on tools with artificial intelligence (ICTAI), IEEE, 2019, pp. 1597–1601. doi:10.1109/ICTAI.2019.00231.

[5] V. Bonta, N. Kumaresh, N. Janardhan, A comprehensive study on lexicon based approaches for sentiment analysis, Asian Journal of Computer Science and Technology 8 (2019) 1–6. doi:10.51983/ajcst-2019.8.S2.2037.

[6] P. A. Grana, Sentiment analysis of text using machine learning models, International Research Journal of Modernization in Engineering Technology and Science 04 (2022).

[7] Z. Drus, H. Khalid, Sentiment analysis in social media and its application: Systematic literature review, Procedia Computer Science 161 (2019) 707–714. doi:10.1016/j.procs.2019.11.174.

[8] T. N. Yogi, N. Paudel, Comparative analysis of machine learning based classification algorithms for sentiment analysis, International Journal of Innovative Science, Engineering & Technology 7 (2020) 1–9.

[9] M. I. Al-mashhadani, K. M. Hussein, E. T. Khudir, et al., Sentiment analysis using optimized feature sets in different facebook/twitter dataset domains using big data, Iraqi Journal For Computer Science and Mathematics 3 (2022) 64–70. doi:10.52866/ijcsm.2022.01.01.007.

[10] K.-X. Han, W. Chien, C.-C. Chiu, Y.-T. Cheng, Application of support vector machine (svm) in the sentiment analysis of twitter dataset, Applied Sciences 10 (2020) 1125. doi:10.3390/app10031125.

[11] A. C. M. V. Srinivas, C. Satyanarayana, C. Divakar, K. P. Sirisha, Sentiment analysis using neural network and lstm, in: IOP conference series: materials science and engineering, volume 1074, IOP Publishing, 2021, p. 012007. doi:10.1088/1757-899X/1074/1/012007.

[12] M. Abbas, K. A. Memon, A. A. Jamali, S. Memon, A. Ahmed, Multinomial naive bayes classification model for sentiment analysis, IJCSNS Int. J. Comput. Sci. Netw. Secur 19 (2019) 62.

[13] R. H. Muhammadi, T. G. Laksana, A. B. Arifa, Combination of support vector machine and lexicon- based algorithm in twitter sentiment analysis, Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika 8 (2022) 59–71. doi:10.23917/khif.v8i1.15213.

[14] B. M. Mulyo, D. H. Widyantoro, Aspect-based sentiment analysis approach with cnn, in: 2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), IEEE, 2018, pp. 142–147. doi:10.1109/EECSI.2018.8752857.

[15] N. Sultana, P. Kumar, M. Patra, S. Chandra, S. Alam, Sentiment analysis for product review, ICTACT J Soft Comput 9 (2019). doi:10.21917/ijsc.2019.0266.

[16] U. Mahadevaswamy, P. Swathi, Sentiment analysis using bidirectional lstm network, Procedia Computer Science 218 (2023) 45–56. doi:10.1016/j.procs.2022.12.400.

[17] P. F. Muhammad, R. Kusumaningrum, A. Wibowo, Sentiment analysis using word2vec and long short-term memory (lstm) for indonesian hotel reviews, Procedia Computer Science 179 (2021) 728–735.doi:10.1016/j.procs.2021.01.061.

[18] A. S. Imran, S. M. Daudpota, Z. Kastrati, R. Batra, Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets, Ieee Access 8 (2020) 181074–181090. doi:10.1109/ACCESS.2020.3027350.