



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 4, April 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com



Sentimental Analysis using Neural Network

Vishal Satpute, Vaibhav Singh, Dhananjay Gayke, Rushikesh Kamod, Prof. (Dr) Amol Potgantwar,
Prof. Rais Shaik

Department of Computer Engineering, Sandip Institute of Technology and Research Centre, Nashik, India

ABSTRACT: Our research explores how computers can understand human feelings in text using neural networks. Specifically, we focus on recurrent neural networks (RNNs) and their effectiveness in capturing subtle sentiments. Through careful data selection, preprocessing, and ethical considerations, we demonstrate practical uses in customer feedback analysis, brand monitoring, political sentiment understanding, content moderation, financial trend prediction, and educational feedback evaluation. Our study highlights the importance of neural networks in analyzing human sentiment across different areas.

KEYWORDS: Sentiment analysis, Neural networks, Custom dataset, RNNs, LSTMs, Dataset creation, Preprocessing.

I. INTRODUCTION

Sentiment analysis stands as a cornerstone in the realm of natural language processing, offering invaluable insights into public sentiment across digital platforms. With the exponential growth of social media and online communication channels, there's been an unprecedented demand for robust sentiment analysis tools. In response, neural networks, particularly Recurrent Neural Networks (RNNs), have emerged as powerful instruments for extracting nuanced sentiments from text.

In this paper, we delve into the practical application of neural networks in sentiment analysis, placing a special focus on RNNs. We embarked on a journey to explore their effectiveness, tackle associated challenges, and address pertinent ethical considerations. Through rigorous assessment, we evaluated various architectures, including RNNs, LSTMs, and transformers, aiming to discern complex sentiment patterns with precision.

Our research journey traversed critical aspects of model training, from meticulous dataset selection to robust bias mitigation strategies. Ethical deployment remained a paramount consideration throughout our exploration, underscoring the need for responsible AI practices. Our overarching objective was to furnish valuable insights for both researchers and practitioners navigating the landscape of sentiment analysis using neural networks.

The rationale behind our preference for RNNs stemmed from their innate capacity to capture sequential dependencies within textual data, thereby augmenting accuracy and contextual awareness. As we journeyed through our investigation, we illuminated how RNNs adeptly decipher intricate sentiment patterns embedded within diverse textual sources, thus enriching our understanding of human expression and sentiment in the digital age.

II. RELATED WORK

[1] The paper proposes a hybrid approach to sentiment analysis using 1D-CNN, LSTM, and CNN with an improved Support Vector Machine (SVM) for non-domain-specific data sets. The CNN architecture is used to learn sentiment incorporated vectors and obtain sentiment classification, with a supported vector machine (SVM) replacing the soft-max function in the output layer. The hybrid model improves the CNN's classification capabilities. The proposed method yielded better accuracy values for Kindle, Fire Tablet, and Fire TV data classes, with F1 scores of 83, 82, and 85 respectively. The paper focuses on research into RNN deep learning methods and incorporating existing data from urban areas, such as social and political information. [2] Novel CNN uses deep learning to detect human emotions through facial expression recognition. This method uses large filters and large filters to analyses images, resulting in high accuracy. Techniques used in training CNN include data augmentation, kernel recurrence, and batch normalization. The model categorizes images into six groups: poor, fair, average, good, best, and excellent. The determinant of emotion analysis is recognized as a dependent and independent variable. The proposed model improves the number of covenants for finding emotions based on customer reviews. In the training and testing variables, normalization occurs based on review comments, and feature extraction is

done using vector variables. The model uses binary trees to split comments into dimensions, mapping to feature labels connected to the six groups of classes. Customer reviews are crucial for building commercial businesses, and machine learning models can improve results in selling and customer-related issues. The proposed Novel-CNN model outperforms existing CNN models in customer reviews with 98.3% accuracy, making it more reliable. [3] Their proposed system analyses reviews to identify sentiment expressions. The dataset is collected from ai.stanford.edu and noise is removed by eradicating noisy chunks of data, such as emojis and least important words. Words are then reduced to their root form to reduce the number of features. This task is part of the Natural Language Processing (NLP) of text data, which can be achieved using NLP Tool Kits. The raw data undergoes pre-processing, making it noise-free and tokenizing to form lists of tokens or words. Words like prepositions, articles, and helping verbs are treated as non-subjective, and root words are used to calculate IDF. TFIDF values are normalized to the range [0, 1] to create a sparse matrix. Four classification algorithms are compared to decide the finest model for sentiment analysis. The Logistic Regression Classifier shows the best accuracy at a learning rate of 0.5, while the SVM Classifier has the highest accuracy at 89.4%. The study reveals that the sentiment of a review is dependent on the words used, with logistic regression classifier showing better performance than other classification algorithms. The Gini index and entropy factors contributed to the maximum accuracy of the Decision Tree model. Negation handling significantly improved sentiment analysis results, and this research can be extended to identify sentiment polarity at higher precision. [4] This paper presents a CNN model that was tested using a single standard dataset from Stanford AI Lab. The underlying arrangement of characteristics, i.e., vector, weights, and bias, are selected indiscriminately after a progressive cycle, and the underlying parameters achieve optimality after an iterative preparation procedure. The CNN model is tested according to the following criteria: minimal LOSS, accuracy, gradient descent, precision, recall, and f1-score. The results states that, the CNN model performs better in terms of classification accuracy and also provides an optimal balance of weights and biases. [5] A Deep Learning Model for Live Tweet Analysis has been developed, which generates sentiment scores for recent tweets of a search query. The model uses a large dataset from the Sentiment140 dataset for training. The model achieved training accuracy of 86.33%, validation accuracy of 79.61%, and test accuracy of 79.73%. This project aims to develop software that can obtain user sentiment on specific topics from Twitter. Further research is needed to train more advanced models for higher accuracy and ease of integration.

These studies highlight the versatility of neural networks in sentiment analysis, employing hybrid models, novel architectures, and varied preprocessing methods. However, there exists a notable gap in synthesizing these approaches into a unified framework. The absence of a comprehensive comparative analysis across diverse datasets and domains impedes a clear understanding of the most effective strategies.

III. METHODOLOGY

In our methodology, we started by carefully selecting datasets suitable for sentiment analysis, covering a range of sources like social media posts, product reviews, and news articles to ensure a comprehensive representation of sentiments. We paid close attention to maintaining a balance between positive and negative sentiments within these datasets.

Next, we preprocess the data rigorously before training our sentiment analysis model. This involved tasks such as standardizing the text format, removing stop words, and tokenization to enhance model performance. For the sentiment analysis task, we opted for LSTM - Recurrent Neural Networks (LSTM-RNNs) due to their ability to capture sequential dependencies in textual data from both forward and backward directions.

The architecture comprised layers of recurrent cells, enabling the model to remember previous words in a sentence from both directions, which proved beneficial in capturing subtle sentiment nuances. Our LSTM-RNN model underwent a training process using a carefully partitioned dataset, with a standard training-validation-test split to evaluate its performance. We fine-tuned training parameters like learning rate and batch size through iterative experimentation to achieve optimal results.

To assess the model's performance, we employed widely recognized evaluation metrics such as accuracy, precision, recall, and F1 score. These metrics provided a comprehensive understanding of the model's ability to classify sentiments accurately across different categories.

Additionally, we conducted a comparative analysis with other neural network architectures commonly used in sentiment analysis, including transformer models. This comparison helped highlight the strengths and limitations of each architecture in capturing sentiment patterns effectively.

IV. IMPLEMENTATION PLAN

a) Data Collection and Preprocessing:

In our data collection process, we tapped into a diverse array of sources, including social media platforms, product review websites, and news articles, to gather textual data pertinent to our research domain. Specifically, we accessed datasets from Kaggle, comprising two primary sources: one nearly containing 27,000 reviews and the other containing 150,000 reviews. By combining these datasets, we formed a comprehensive corpus to drive our analysis.

Subsequently, we engaged in meticulous data cleaning procedures to ensure the quality and relevance of our dataset. Techniques were employed to eliminate noise and extraneous information, encompassing the removal of special characters and HTML tags. Furthermore, we undertook measures to normalize the text, including tokenization, conversion to lowercase, and the elimination of stopwords.

Text normalization was a pivotal step in our preprocessing pipeline, where we implemented techniques such as stemming or lemmatization. This process aimed to reduce words to their base forms, fostering uniformity and consistency across the dataset. Through these efforts, we primed our textual data for subsequent analysis, laying a robust foundation for sentiment analysis using neural networks.

b) Feature Extraction:

In the feature extraction phase, we harnessed the power of TensorFlow's Keras library to preprocess our textual data. Specifically, we utilized the Tokenizer function to tokenize our text and collect frequency counts of words. This process enabled us to transform raw text into a numerical representation that could be ingested by our neural network model.

Following tokenization, we embarked on feature selection, a critical step in distilling relevant information from the text data. We identified and extracted key features, including sentiment-related words, phrases, and syntactic structures. By focusing on the most frequent words, we aimed to capture the essence of sentiment expression within our dataset.

Moreover, we recognized the importance of dimensionality reduction in optimizing our feature space. Through techniques such as principal component analysis (PCA) or singular value decomposition (SVD), we aimed to reduce the complexity of our data while retaining essential information. This process ensured that our neural network model could operate efficiently and effectively, enhancing its ability to discern subtle sentiment patterns within the text.

c) Model Selection and Architecture Design:

In selecting the model architecture, we carefully assessed the characteristics of our dataset and the requirements of our sentiment analysis task. Considering the sequential nature of textual data and the need to capture long-range dependencies, we opted for recurrent neural network (RNN) architectures.

Among RNN variants, we specifically considered Long Short-Term Memory Networks (LSTMs) due to their ability to retain context over long sequences. These architectures were deemed suitable for capturing nuanced sentiment patterns within the text.

To tailor the model to our specific needs, we engaged in model customization. This involved fine-tuning parameters and adjusting the architecture's layers to optimize its performance for sentiment analysis. By adapting the model to the unique characteristics of our dataset, we aimed to enhance its ability to accurately discern sentiment nuances.

Furthermore, we conducted hyperparameter tuning to optimize the model's learning process. Through iterative experimentation, we adjusted parameters such as the learning rate and batch size to maximize the model's performance. This meticulous tuning process was crucial for achieving optimal results in sentiment analysis tasks.

d) Training and Validation:

Dataset Splitting: Divide the dataset into training, validation, and test sets, ensuring a balanced distribution of samples across sentiment classes.



Model Training: Train the selected neural network architecture on the training data using appropriate loss functions and optimization algorithms.

Validation: Validate the trained model using the validation set to monitor performance metrics and prevent overfitting, adjusting model parameters as needed.

e) Evaluation and Fine-Tuning:

Performance Metrics: Evaluate the trained model's performance using metrics like accuracy, precision, recall, and F1-score on the test set.

Error Analysis: Analyze model errors to identify patterns or specific instances where the model struggles, providing insights for further improvement.

Fine-Tuning: Incorporate additional features or data augmentation techniques based on error analysis findings to enhance model performance.

f) Deployment and Integration:

For deploying our sentiment analysis model into a production environment, we chose Flask as our web framework due to its simplicity and flexibility. Flask allowed us to create a RESTful API that can handle incoming requests and provide predictions based on the sentiment analysis model.

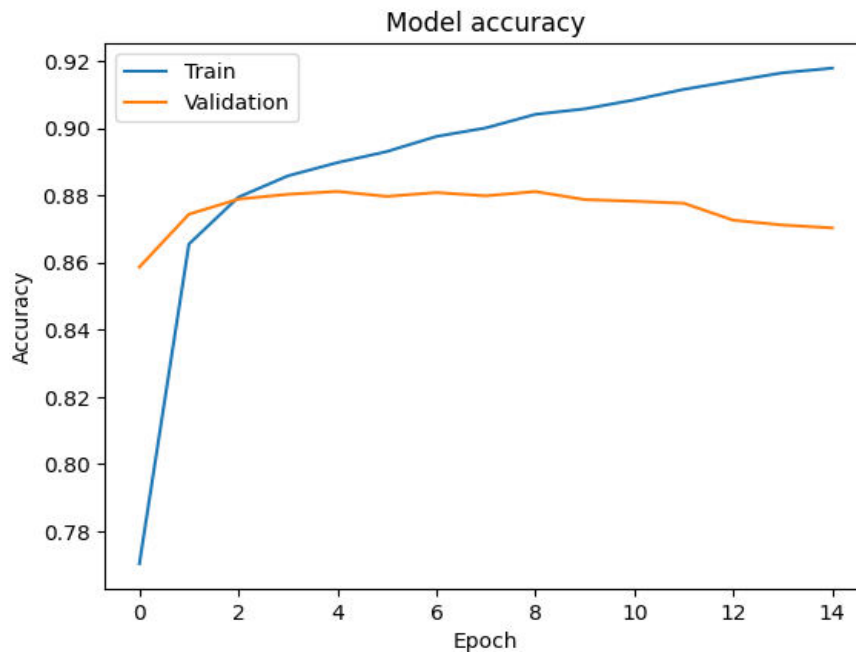
The deployment process involved setting up a server to host our Flask application.

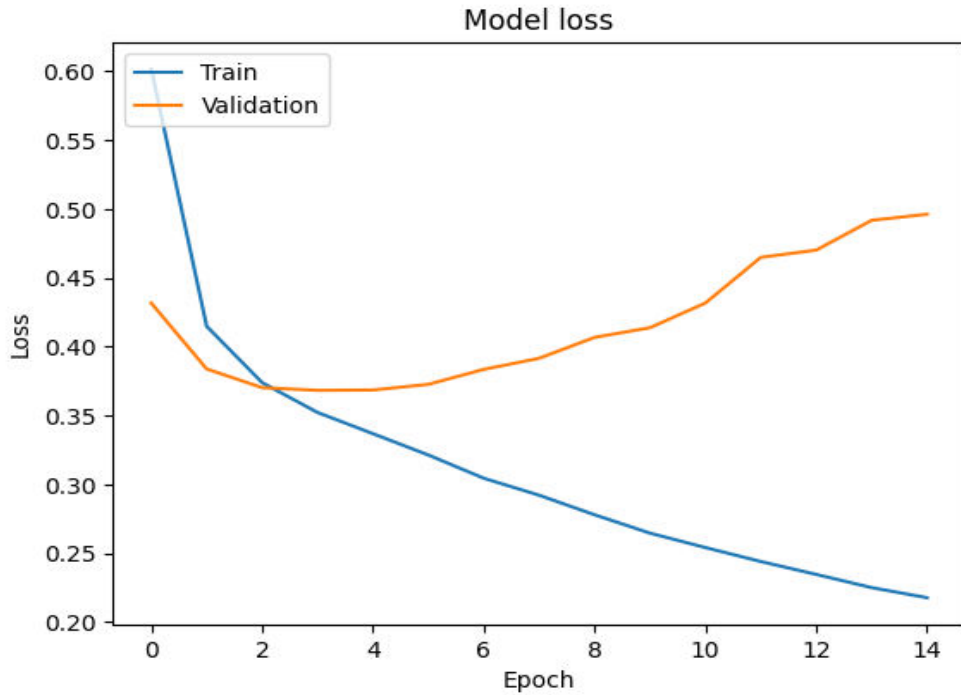
Once deployed, we established monitoring and maintenance procedures to continuously assess the model's performance. We implemented monitoring tools to track key metrics such as response times, error rates, and resource utilization. Regular maintenance involved periodic retraining of the model with new data to adapt to evolving sentiment patterns and maintain accuracy over time.

By employing Flask for model deployment and implementing robust monitoring and maintenance practices, we ensured that our sentiment analysis system remained reliable and effective.

V. RESULT

Traning Results:





Test Loss: 0.5022764205932617

Test Accuracy: 0.8711066842079163

Classification Report:

	precision	recall	f1-score	support
0	0.82	0.80	0.81	8726
1	0.91	0.87	0.89	16049
2	0.86	0.92	0.89	13303

Accuracy			0.87	38078
Macro Average	0.86	0.86	0.86	38078
Weighted Average	0.87	0.87	0.87	38078

Confusion Matrix :

$$\begin{bmatrix} 7008 & 843 & 875 \\ 908 & 13987 & 1154 \\ 588 & 540 & 12175 \end{bmatrix}$$

VI. CONCLUSION

In conclusion, our study has focused on sentiment analysis, particularly highlighting the effectiveness of Recurrent Neural Networks (RNNs). RNNs excel at capturing sequential patterns, making them invaluable for decoding complex sentiments. Our methodological approach, covering dataset selection, preprocessing, and ethical considerations, ensures practical application across domains like customer feedback and political analysis.

As sentiment analysis continues to influence decision-making, our research with RNNs contributes significant insights. We've not only advanced the field but also paved the way for future innovations. Our goal is to deepen our

understanding of human sentiment in the realm of artificial intelligence, driving progress in sentiment analysis for transformative applications ahead.

REFERENCES

- [1] Prachi Juyal, Amit Kundaliya, "A Comparative Study of Hybrid Deep Sentimental Analysis Learning Techniques with CNN and SVM." IEEE World Conference On Applied Intelligence and Computing, 2023.
- [2] N. Deepa, J. Sathya Priya, Devi T, "Sentimental Analysis Recognition in Customer Review using Novel- CNN." International Conference on Computer Communication and Informatics (ICCCI), 2023.
Nibedan Panda, Alok Jena, V Ramana Bendi, "Performance Metrix Assessment in Sentimental Analysis over Machine Learning Approaches." IEEE International Conference on Contemporary Computing and Communication, 2023.
- [3] K. Kavitha, Suneetha Chittineni, Naveen Kumar Andhavarpu, Vasundhra Chappa, "Textual Sentimental Classification using Convolutional Neural Network Algorithm." International Conference on Computing, Communication and Power Technology (IC3P), 2022.
- [4] Bhavani M., Shrijeeth S., Rohit M., Sanjeev Krishna R., Sharveshran R., "A detailed study on Sentimental Analysis using Twitter Data with an Improved Deep Learning Model." Fifth International Conference on I-SMAC, 2021.
- [5] Siyin Luo, Youjian Gu, Xingxing Yao, Wei Fan, "Research on Text Sentiment Analysis Based on Neural Network and Ensemble Learning", Revue d'Intelligence Artificielle Vol. 35, No. 1, February, 2021, pp. 63-70
Journal homepage: <http://iieta.org/journals/ria>.
- [6] Aytug Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks", Concurrency Computat Pract Exper. 2020;e5909.wileyonlinelibrary.com/journal/cpe © 2020 John Wiley & Sons, Ltd.
- [7] Lilis Kurniasari¹ and Arif Setyanto, "Sentiment Analysis using Recurrent Neural Network", 1st Bukittinggi International Conference on Education IOP Conf. Series: Journal of Physics: Conf. Series 1471 (2020) 012018.
- [8] Alpna Patel¹ and Arvind Kumar Tiwari, "Sentiment Analysis by using Recurrent Neural Network", 2nd INTERNATIONAL CONFERENCE ON ADVANCED COMPUTING AND SOFTWARE ENGINEERING (ICACSE-2019).
- [9] Merin Thomas¹, and Latha C.A, "Sentimental analysis using recurrent neural network", International Journal of Engineering & Technology, 7 (2.27) (2018) 88-92.
- [10] Sneh Paliwal, Sunil Kumar Khatri, Mayank Sharma, "Sentiment Analysis and Prediction using Neural Networks", Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA 2018) IEEE Xplore Compliant Part Number:CFP18N67-ART; ISBN:978-1-5386-2456-2.
- [11] Qionxia Huang, Xianghan Zheng, Riqing Chen, and Zhenxin Dong, "Deep Sentiment Representation Based on CNN and LSTM", International Conference on Green Informatics, 2017, pp. 30-33.
- [12] Brett Duncan, and Yanqing Zhang, "Neural Networks for Sentiment Analysis on Twitter", IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), 2015, pp. 275-278.
- [13] M. S. Neethu, and R. Rajasree, "Sentiment analysis in twitter using machine learning techniques", Computing Communications and Networking Technologies (ICCCNT), 2013, pp. 1-5.
- [14] Rong, B. Peng, Y. Ouyang, C. Li, and Z. Xiong, "Semi-supervised Dual Recurrent Neural Network for Sentiment Analysis," in Dependable, Autonomic and Secure Computing (DASC), 2013 IEEE 11th International Conference on, Dec 2013, pp. 438-445.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details