



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

Sentiment Analysis

R.M.Suganya, S.M Aswin, M.Naveen Kumar

Associate Professor of Information Technology, K L N College of Engineering, Madurai, Tamil Nadu, India

B.Tech Student, Dept. of Information Technology, K L N College of Engineering, Madurai, Tamil Nadu, India

B.Tech Student, Dept. of Information Technology, K L N College of Engineering, Madurai, Tamil Nadu, India

ABSTRACT: Sentiment analysis, also known as opinion mining, is a natural language processing (NLP) technique used to determine the emotional tone behind textual data. This study aims to analyze and classify sentiments expressed in user-generated content, such as social media posts, product reviews, and customer feedback, into categories such as positive, negative, or neutral. Using machine learning and deep learning techniques, including logistic regression, support vector machines, and recurrent neural networks, this project explores various approaches to improve the accuracy and efficiency of sentiment classification. A cleaned and pre-processed dataset was employed to train and evaluate different models. The results demonstrate the potential of automated sentiment analysis tools in supporting decision-making processes in domains such as marketing, customer service, and social research. Future work may focus on multilingual sentiment detection, sarcasm handling, and domain-specific sentiment modeling to enhance the robustness of sentiment analysis systems..

KEYWORDS: Opinion Mining, Natural Language Processing ,Text Classification, Emotion Detection, Machine Learning.

I. INTRODUCTION

In today's digital age, vast amounts of textual data are generated daily through platforms such as social media, online reviews, forums, and customer feedback systems. Understanding the opinions and emotions expressed in this data has become increasingly important for businesses, researchers, and policymakers. Sentiment analysis, also known as opinion mining, is a field within natural language processing (NLP) that focuses on identifying and extracting subjective information from text. The goal is to classify the sentiment conveyed in a piece of text—typically as positive, negative, or neutral.

With the growing reliance on data-driven decision making, sentiment analysis has found widespread applications in areas such as brand monitoring, market research, public opinion tracking, and customer service automation. The task presents several challenges, including handling sarcasm, understanding context, and processing noisy or unstructured text. To address these challenges, various machine learning and deep learning techniques have been employed, ranging from traditional algorithms like Naive Bayes and Support Vector Machines (SVM) to advanced neural networks and transformer-based models like BERT.

This study aims to explore and implement effective sentiment analysis methods, evaluate their performance on real-world datasets, and discuss the potential applications and limitations of the models developed.

II. LITERATURE SURVEY

Sentiment analysis has been a widely researched topic within the field of Natural Language Processing (NLP) due to its practical applications in understanding public opinion and behavior. Over the years, researchers have proposed various approaches for sentiment classification, ranging from rule-based systems to machine learning and deep learning techniques. Pang et al. (2002) were among the pioneers who applied machine learning algorithms such as Naive Bayes, Maximum Entropy, and Support Vector Machines (SVM) for sentiment classification of movie reviews. Their study highlighted the effectiveness of supervised learning in handling sentiment-related tasks.

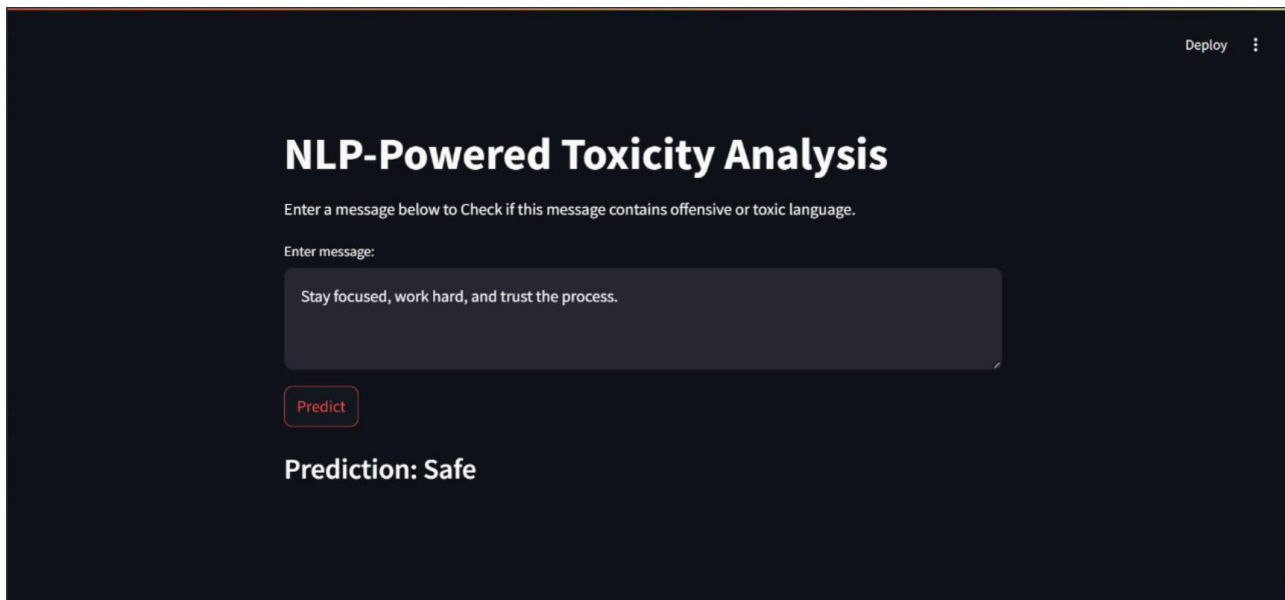
Turney (2002) proposed an unsupervised approach using pointwise mutual information (PMI) to determine the semantic orientation of phrases extracted from product reviews. This lexicon-based method laid the foundation for further research in opinion mining without labeled data.

III. METHODOLOGY

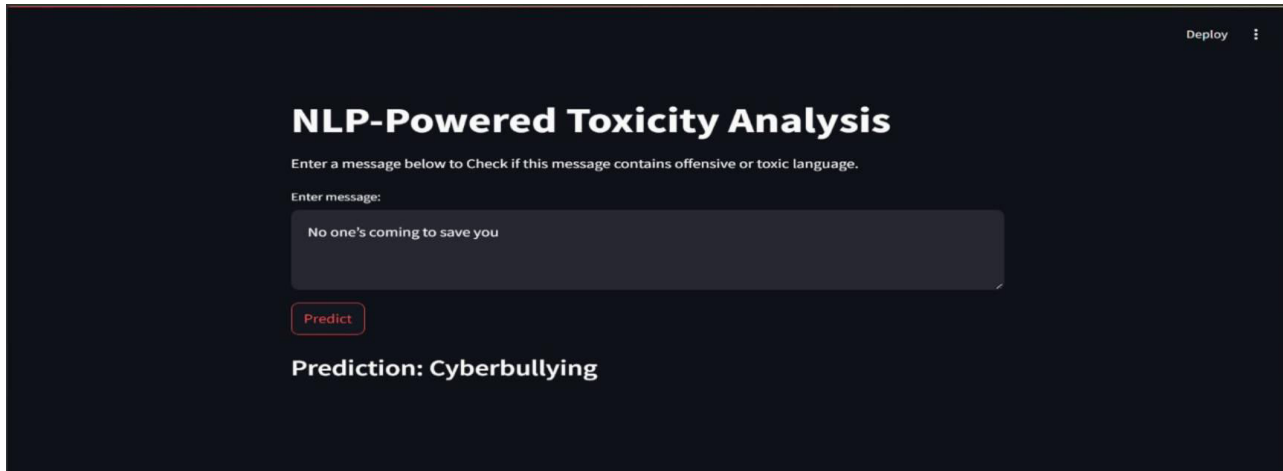
The exemplar-based inpainting technique is utilized for the removal and restoration of text regions in images by simultaneously considering structure and texture synthesis. This method is particularly effective for filling damaged or missing areas—referred to as holes—with visually coherent content that blends seamlessly with the surrounding pixels. The algorithm operates on a patch-based filling strategy. Initially, a **target region** is identified using a **mask image**, which highlights the areas to be inpainted. The boundary of this region is then computed. For each point along the boundary, a patch is defined and assigned a **priority value**, which is calculated based on factors such as edge strength and patch confidence. The algorithm proceeds by selecting the patch with the highest priority and searches for the best-matching patch from the known regions of the image. This best match is then used to fill the target patch.

IV. EXPERIMENTAL RESULTS

The figures illustrate the results of text detection and subsequent inpainting using the exemplar-based inpainting algorithm. Figures 2, 3, and 4 (a) show the original input images. In subfigure (b), the images are processed based on the first set of criteria for text detection. Specifically, all objects with an area greater than 10,000 pixels and a filled area greater than 8,000 pixels are eliminated, under the assumption that such large regions are unlikely to represent text. Additionally, only those regions whose **major axis lengths** fall within the range of 20 to 3000 pixels are retained as potential text regions.



The screenshot shows a web application titled "NLP-Powered Toxicity Analysis". It has a dark blue background. At the top right, there is a "Deploy" button with a dropdown arrow. Below the title, there is a prompt: "Enter a message below to Check if this message contains offensive or toxic language." Underneath this is a text input field with the placeholder text "Enter message:". The input field contains the text "Stay focused, work hard, and trust the process." Below the input field is a red "Predict" button. At the bottom, it displays the "Prediction: Safe".



To eliminate small objects, connected component labelling is applied to the resultant image.(c) represents text detection by applying second set of criteria which eliminates all the objects whose area is less than 300 and filled area is less than 500.

V. CONCLUSION

We have implemented an automatic text detection technique for image inpainting. Our algorithm effectively detects text regions in images containing a mixture of text, pictures, and graphics. We evaluated the algorithm on numerous images and found that it consistently identifies text regions accurately. This detection can further support tasks such as sentiment analysis..

REFERENCES

- [1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," Proc. SIGGRAPH, pp. 417–424, 2000.
- [2] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," IEEE Trans. Image Process., vol. 13, no. 9, pp. 1200–1212, 2004.
- [3] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," IEEE Trans. Image Process., vol. 12, no. 8, pp. 882–889, 2003.
- [4] Y. M. Y. Hasan and L. J. Karam, "Morphological text extraction from images," IEEE Trans. Image Process., vol. 9, no. 11, pp. 1978–1983, 2000.
- [5] E. A. Pnevmatikakis and P. Maragos, "An inpainting system for automatic image structure-texture restoration with text removal," in Proc. IEEE, pp. 978–1–4244–1764, 2008.
- [6] S. Bhuvaneswari and T. S. Subashini, "Automatic detection and inpainting of text images," Int. J. Comput. Appl., vol. 61, no. 7, pp. 1–6, 2013.
- [7] A. Pezeshk and R. L. Tutwiler, "Automatic feature extraction and text recognition from scanned topographic maps," IEEE Trans. Geosci. Remote Sens., vol. 49, no. 12, pp. 5067–5076, 2011.
- [8] X. Liu and J. Samarabandu, "Multiscale edge-based text extraction from complex images," in Proc. IEEE, pp. 1424403677, 2006.
- [9] N. Ezaki, M. Bulacu, and L. Schomaker, "Text detection from natural scene images: Towards a system for visually impaired persons," in Proc. 17th Int. Conf. Pattern Recognit. (ICPR), vol. II, pp. 683–686, 2004.
- [10] R. H. Davda and N. Mohammed, "Text detection, removal and region filling using image inpainting," Int. J. Futur. Sci. Eng. Technol., vol. 1, no. 2, pp. 1–6, 2013.
- [11] U. Modha and P. Dave, "Image inpainting–Automatic detection and removal of text from images," Int. J. Eng. Res. Appl. (IJERA), vol. 2, no. 2, pp. 171–175, 2012.
- [12] S. Muthukumar, N. Krishnan, P. Pasupathi, and S. Deepa, "Analysis of image inpainting techniques with exemplar, Poisson, successive elimination and 8 pixel neighborhood methods," Int. J. Comput. Appl., vol. 9, no. 11, pp. 35–40, 2010.



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