INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.625

Machine Learning Algorithms for Sentiment Analysis on Social Media

Dr. R Priya Vaijayanthi[1], Yamaka Manikanta[2], Yanadi Sai Vardhan[3], Yernagula Sai Praneeth[4],

Shinagam Hemanth[5]

Senior Professor, Department of CSE (AI&ML), NSRIT, Visakhapatnam, India[1]

U.G. Student, Department of Computer Engineering (AI&ML), NSRIT, Visakhapatnam, India[2,3,4,5]

ABSTRACT: This paper provides a comprehensive survey of sentiment analysis within the context of artificial intelligence (AI) and large language models (LLMs). Sentiment analysis, a critical aspect of natural language processing (NLP), has evolved significantly from traditional rule-based methods to advanced deep learning techniques. In this paper we present a system which collects Tweets from social networking sites, we'll be able to do analysis on those Tweets and thus provide some prediction of business intelligence. Then, we use text mining techniques to generate and process the variables. A user who is being followed by another user need not necessarily reciprocate by following them back, which renders the links of the network as directed. Results of trend analysis will be display as tweets with different sections presenting positive, Negative and neutral.

KEYWORDS: Sentiment Analysis, Deep Learning, Explainable AI, Large Language Model, Deep Learning, Artificial Intelligence.

I. INTRODUCTION

Image inpainting is a method for repairing damaged pictures or removing unnecessary elements from pictures. It recovers the missing or corrupted parts of an image so that the reconstructed image looks natural. In real world, many people need a system to recover damaged photographs, designs, drawings, artworks etc. damage may be due to various reasons like scratches, overlaid text or graphics etc.

This system could enhance and return a good looking photograph using a technique called image inpainting. Image inpainting modify and fill the missing area in an image in an undetectable way, by an observer not familiar with the original image. The technique can be used to reconstruct image damage due to dirt, scratches, overlaid text etc.

Some images contain mixed text-picture-graphic regions in which text characters are printed in an image. Detecting and recognizing these characters can be very important, and removing these is important in the context of removing indirect advertisements, and for aesthetic reasons. There are many applications of image inpainting ranging from restoration of photographs, films, removal of occlusions such as text, subtitle, logos, stamps, scratches, red eye removal etc.

Paper is organized as follows. Section II describes automatic text detection using morphological operations, connected component analysis and set of selection or rejection criteria. The flow diagram represents the step of the algorithm. After detection of text, how text region is filled using an Inpainting technique that is given in Section III. Section IV presents experimental results showing results of images tested. Finally, Section V presents conclusion.

II. RELATED WORK

The concept of image inpainting was first introduced by Bertamio et al. [1]. The method was inspired by the real inpainting process of artists. The image smoothness information interpolated by the image Laplacian is propagated along the isophotes directions, which are estimated by the gradient of image rotated by 90 degrees. Exemplar Based method proposed by Criminisi et al. [2] used a best exemplar patch to propagate target patch including missing pixels. This technique uses an approach which combine structure propagation with texture synthesis and hence produced very good results. In [3], the authors decompose the image into sum of two functions and then reconstruct

each function separately with structure and texture filling-in algorithms. Morphological technique is used to extract text from the images presented in [4]. In [5], the inpainting technique is combined with the techniques of finding text in images and a simple algorithm that links them. The technique is insensitive to noise, skew and text orientation. The authors in [6] have applied the CCL (connected component labelling) to detect the text and fast marching algorithm is used for Inpainting.

The work in this paper is divided in two stages. 1) Text- Detection 2) Inpainting. Text detection is done by applying morphological open-close and close-open filters and combines the images. Thereafter, gradient is applied to detect the edges followed by thresholding and morphological dilation, erosion operation. Then, connected component labelling is performed to label each object separately. Finally, the set of selection criteria is applied to filter out non text regions. After text detection, text inpainting is accomplished by using exemplar based Inpainting algorithm.

III. METHODOLOGY

Exemplar based Inpainting technique is used for inpainting of text regions, which takes structure synthesis and texture synthesis together. The inpainting is done in such a manner, that it fills the damaged region or holes in an image, with surrounding colour and texture. The algorithm is based on patch based filling procedure. First find target region using mask image and then find boundary of target region. For all the boundary points it defined patch and find the priority of these patches. It starts filling the target region from the highest priority patch by finding the best match patch. This procedure is repeated until entire target region is inpainted.

The algorithm automatically generates mask image without user interaction that contains only text regions to be inpainted.

IV. EXPERIMENTAL RESULTS

Machine Learning Algorithms
Machine learning, a branch of artificial intelligence, is a scientific discipline concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or databases. A learner can take advantage of examples (data) to capture characteristics of interest of their unknown underlying probability distribution

Naive Bayes algorithm
The Bayesian Classification represents a supervised learning method as well as a statistical method for classification. Assumes an underlying probabilistic model and it allows us to capture uncertainty about the model in a principled way by determining probabilities of the outcomes. It can solve diagnostic and predictive problems. This Classification is named after Thomas Bayes (1702-1761), who proposed the Bayes Theorem. Bayesian classification provides practical learning algorithms and prior knowledge and observed data can be combined. Bayesian Classification provides a useful perspective for understanding and evaluating many learning algorithms. It calculates explicit probabilities for hypothesis and it is robust to noise in input data.

V. CONCLUSION

We show that using emoticons as noisy labels for training data is an effective way to perform distant supervised learning. Machine learning algorithms (Naive Bayes, maximum entropy classification) can achieve high accuracy for classifying sentiment when using this method. Although Twitter messages have unique characteristics compared to other corpora, machine learning algorithms are shown to classify tweet sentiment with similar performance. punctuation and special characters then we change all tweets to lowercase. Then we prepare the data where we have converted the cleaned tokens to Python dictionary with words as keywords and True as values and split into training and test data in a ratio of 70:30. We then applied supervised machine learning algorithms to classify our tweets into positive and negative and compared them in terms of accuracies.

REFERENCES

1. J. Li, S. Fong, Y. Zhuang, and R. Khoury, " Hierarchical Classification in Text Mining for Sentiment Analysis," in 2014 International Conference on Soft Computing and Machine Intelligence, september. 2014,
p. 46-51, doi: 10.1109/ISCMI.2014.37.

2.H. Parveen and S. Pandey, "Sentiment analysis on Twitter Dataset using   NaiveBayes algorithm ," in 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology.

3. S. S. and P. K.v., " Sentiment analysis of techniques ," ICT Express, april. 2020, doi: 10.1016/j.icte.2020.04.003.

4. R. A. Tuhin, B. K. Paul, F. Nawrine,
M. Akter, and A. K. Das, " An Automated System of Sentiment Analysis from Bangla Text using Eryiğit, G. (2016). Semeval-2016 Task Aspect Based Sentiment Analysis. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016).

5.Wang Y, Zhang Y, Liu B. Sentiment lexicon expansion based on neural PU Learning, double dictionary lookup, and polarity association. In Proceedings of the Conference on Empirical Methods on Natural Language Processing (EMNLP 2017), 2017.

6. Wang Z, Zhang Y. Opinion recommendation using a neural model. In Proceedings of the Conference on Empirical Methods on Natural Language Processing (EMNLP 2017), 2017.

7. Du J, Xu R, He Y, Gui L. Stance classification with target-specific neural attention networks. In Proceedings of the Internal Joint Conference on Artificial Intelligence (IJCAI 2017), 2017.

8. Du J, Xu R, He Y, Gui L. Stance classification with target-specific neural attention networks. In Proceedings of the Internal Joint Conference on Artificial Intelligence (IJCAI 2017), 2017.

ISSN
INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE
SJIF Scientific Journal Impact Factor

doi crossref

NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

📱 9940 572 462   🟢 6381 907 438   ✉ ijircce@gmail.com

Scan to save the contact details