

A Survey on Deep Learning Techniques for Sentiment Analysis

Harinam ¹ and Mohammad Suaib ²

¹ PG Scholar, Department of Computer Science & Engineering, Integral University Lucknow, India

² Assistant Professor, Department of Computer Science & Engineering, Integral University Lucknow, India

harinam.yadav.2008@gmail.com, suaibcs09@gmail.com

Abstract: Social media is a rich source of information nowadays. If we look into social media, sentiment analysis is one of the challenging problems. Sentiment analysis is a substantial area of research in the field of Natural Language Processing. This survey paper reviews and provides the comparative study of deep learning approaches CNN, RNN, LSTM and ensemble-based methods.

Keywords: Natural Language Processing; Sentiment Analysis; Deep Learning Classifiers; LSTM; CNN

1. Introduction

SA is a branch of psychology that analyses people's thoughts, feelings, and emotions derived through customer script automatically. Sentiment analysis is a hot topic in natural language processing, and it's still getting much attention in data mining because emotions are powerful drivers of social behaviour. In a way, sentiment analysis started as a research topic in Natural Language Processing across the world. The most important application of NLP, computational linguistics, and text processing is sentiment analysis. Sentiment analysis is an opinion mining task that can be used to ascertain the writer's or speaker's emotions, attitude toward a specific task, such as product reviews, film reviews, or the overall tone of the document. The era of digitalization has resulted in the exponential growth of data. Data is stored in a variety of formats, including structured, semi-structured, and unstructured. The difficult task is to discover useful information through data analysis.

2. Deep Learning Models

It is a subset of Machine Learning in which the number of layers (multilayers) used to achieve the desired results is increased. It is a term used to refer to an extensive neural network. Nodes in a neural network communicate with one another via interconnected layers of nodes. Deep learning models incorporate dozens, if not hundreds, of hidden layers to create complex models that continuously learn and improve. This article discusses the various deep learning-based models that have been used to analyze sentiments.

2.1 Recurrent Neural Network

The Recurrent Neural Network is a deep learning technique used for sentiment analysis and is based on sequential data. It generates the output using sequential information based on previous computations. It can accept multiple input vectors and output multiple vectors. Traditionally, neural networks have relied on independent inputs, making them unsuitable for specific tasks in Natural Language Processing. Consider the following example: word

prediction within a given sentence. The RNN model is a highly efficient model for sentiment analysis. RNNs make use of memory cells that are capable of storing data about lengthy sequences. To begin, we must comprehend sequential information. Sequential data is data that is ordered in such a way that similar items follow one another.

2.2 Long Short-Term Memory

Sepp Hochreiter and Juergen Schmidhuber presented another model for sequential information called long short-term memory. LSTM networks are a form of RNN that may learn order dependence in situations of sequence prediction. The LSTM architecture is an RNN that remembers the information at regular periods. It is employed in the solution of the vanishing gradient problem. It can learn long-term dependence. At observation time, RNN has only two gates: an input gate and an output gate from the last hidden state, and there is no knowledge about the past to remember. RNNs can remember their inputs for a long time due to LSTMs. That is why long short-term memory uses its memory to accumulate information over a long period. This memory cell is known as a gated cell because it represents whether or not to store or delete information dependent on the relevance of the information. LSTM is made up of three gates. The input gate is used for new data input, the forget gate is used to determine whether or not the data should be deleted, and the output gate is used to determine the output at the current time step.

2.3 Convolutional Neural Network

The neural network image processing group was the first to create the convolutional neural network. ConvNets are developed to accommodate data in the form of several arrays, such as a colour image made up of three 2D arrays representing pixel elevations in each of the colour channels. As attribute extractors, a CNN uses two operations as convolution and pooling. As in a multi-layer perceptron, the output of this series of operations is bound to a completely connected layer. Convolutional neural

networks are often used on text in Natural Language Processing. There are two types of pooling used: max-pooling and average-pooling. When we use CNN for text instead of images, we display the text with a 1-Dimensional string. CNN is mostly used in sentence classification in NLP tasks.

3. Related Work

Numerous analysts have tried to consolidate machine learning and deep learning ideas concept in ongoing years for the accurate classification of sentiments. This section quickly portrays the various investigations related to sentiment analysis of web contents about user’s sentiments, emotions, opinion toward various matters like motion pictures and products using machine learning techniques. The authors have been presented the hybrid model for text classification that gives better accuracy than traditional models. They combined two well-known neural network models, CNN and LSTM, in this paper. The experimental result demonstrated an increase in text classification accuracy. They evaluate with an accuracy of 87.31 per cent [1]. The authors propose an efficient model for sentiment classification, which calculates the accuracy of 82.53% on Bengali text. They used two deep neural network models, such as deep RNN with BiLSTM [2], to evaluate. It has achieved significant results in the field of text classification. The authors have been presented an innovative approach for target-based emotion analysis that reduces training time of the proposed model through regional Long Short-Term Memory [3].

Deep learning models are frequently used in Natural Language Processing applications. An efficient approach has been proposed for a multi-domain system that is based on word embedding. The tool named NeuroSent gives the accuracy 85.15% by using the amazon website dataset for multi-domain [4]. Some of the machine learning models are based on the content classification in NLP. An ensemble method proposed for Vietnamese text for sentiment analysis by some researchers. In this approach, they combine the traditional method with a deep learning algorithm. This proposed approach gives an accuracy of 89.19% by using the voting rule [5]. The author has been presented a paper on the deep learning approach for text classification [6]. A novel approach has been proposed for IMDB movie review sentiment analysis using the Deep CNN-LSTM model that gives the accuracy 89% [7]. Some authors give a survey on various models for analysis. A comparative study is given by using deep learning models and classifiers [8]. Researchers have been proposed an approach for analysis [9]. The ensemble approach has been performed better than traditional models in the field of text classification. The authors have been presented a machine learning method for performing analysis. They used Long Short-Term Memory, Naïve Bayes and SVM for analysis using reviews on Google Play in Chinese [10].

The authors have been proposed a model named Sent WordNet that is dependent on Word2Vec to perform

sentiment analysis [11]. A novel approach proposed by researchers named ECNN is to identify opinion, polarity and emotions in microblogs [12]. To perform sentiment analysis, the authors have been presented the model related to text classification. They used word embedding at word level and sentence level using the skim gram model [13]. The authors have been proposed a model for text analysis base on CNN and SVM [14]. The authors have been proposed an efficient method to perform sentiment analysis on the IMDB review dataset. They found that the deep learning model RNN is effective in terms of words semantic and evaluate the accuracy of 89.8% [15]. The author has been building a model that is based on heterogeneous feature [16]. The authors have been used word embedding for sentiment analysis. An efficient approach has been proposed for sentiment analysis by using word embedding. The author has suggested providing human behavioural and trust security by using blockchain and verifying sentiment analysis [17]. Some of the authors have been presented an overview of sentiment analysis. This paper presents the different feature selection methods and machine learning algorithms [23]. A combined framework for sentence classification has been proposed that is based on CNN and RNN. On a movie review dataset, the proposed framework achieves an accuracy of 93.3 per cent [24]. The authors have proposed an analysis strategy. This article discusses the accuracy of the ConvLstm at 88.3 per cent for fine-grained data [25]. This approach is certain to succeed because both operate under distinct conditions: one between users and the data centre and another between memory in the data centre [26].

Table 3.1: Summary of the Machine Learning Approaches for Sentiment Analysis

Ref. No.	Methodology	Deep Learning Models	Results
[1]	Text Classification, Nature Language Processing, LSTM, Long Text Sequences, CNN	CNN-LSTM	87.31%
[2]	Bengali text, Deep learning, Sentiment Classification, RNN, LSTM, BiLSTM, Facebook, NLP	BiLSTM	85.67%
[3]	Deep Learning, Sentiment analysis, Target-based sentiment analysis, Convolutional neural network, Long short term memory network, Deep neural network model	CNN-RLSTM	94.35%
[4]	Sentiment Analysis, Natural Language Processing, Neural Networks, Multi-domain	NeuroSent	84.60%

	Sentiment Analysis, Deep Learning		
[5]	Sentiment analysis, deep learning, ensemble Learning	Vote-rule	92.80%
[6]	sentiment analysis, natural language processing, deep Learning, long- short term memory	AraFT	93.5%
[7]	IMDb, sentiment analysis, text classification, neural network, CNN, LSTM	Base-Model	84.98%
[8]	Deep Learning, Sentiment Analysis, Consumer Review, Recurrent Neural Network (RNN), Long Short Term Memory (LSTM)	Deep Learning with Bi-LSTM	94.00%
[9]	Convolutional neural network, long short- term memory, recurrent neural network	ConvLstm	88.3%
[10]	Deep Learning, Tibetan Microblog, Word vector, Sentiment Analysis	CNN-LSTM	86.21%
[11]	Sentiment analysis, Word2Vec, Word embeddings, SentiWordNet	Senti2vec	85.4%
[12]	Natural language processing, sentiment analysis, deep learning, convolution neural network, emoticons.	ECNN	72.55%
[13]	Word Embedding	MS-skip, CS-skip	87.89%, 89.62%
[14]	Deep learning model, Text classification, Natural Language Processing, CNN, RNN	TextRNN+Attention	84.56% (precision)
[15]	NLP, sentiment analysis, deep learning, machine learning, text classification	Deep Learning	89.8%
[16]	NLP, Text Mining, Data Mining, Big Data, Sentiment Analysis,	Naïve Bayes, Linear SVM	84%, 79%
	Opinion Mining, Machine Learning, Deep Learning, SentiWordNet (SWN)		
[17]	Word embedding, Word2Vec, Machine Learning, Bag-of-words.	Word2Vec	81%
[23]	Sentiment Analysis, Machine Learning	-	-

[24]	Convolutional neural network, recurrent neural network, natural language processing, deep Learning, sentiment analysis, long-term dependencies.	Deep Learning	93.2%
[25]	convolutional neural network; long short- term memory; recurrent neural network	ConvLstm	88.3%

4. CONCLUSION

Sentiment classification is the method of extracting a user’s view as positive or negative for a specific task. Social media is a rich source of information nowadays. If we look into social media, sentiment analyses are one of the challenging problems. Sentiment analysis is a substantial area of research in the field of Natural Language Processing. This survey paper reviewed and provided the comparative study of well-known deep Learning approaches CNN, RNN, LSTM and ensemble.

REFERENCES

- [1]. Zhang J, Li Y, Tian J, Li T. LSTM-CNN Hybrid Model for Text Classification. In 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) 2018 Oct 12 (pp. 1675-1680). IEEE.
- [2]. Gope M, Hashem MM. Knowledge Extraction from Bangla Documents: A Case Study. In 2018 International Conference on Bangla Speech and Language Processing (ICBSLP) 2018 Sep 21 (pp. 1-6). IEEE.
- [3]. Chen S, Peng C, Cai L, Guo L. A Deep Neural Network Model for Target-based Sentiment Analysis. In 2018 International Joint Conference on Neural Networks (IJCNN) 2018 Jul 8 (pp. 1-7). IEEE.
- [4]. Dragoni M, Petrucci G. A neural word embeddings approach for multi-domain sentiment analysis. IEEE Transactions on Affective Computing. 2017 Oct 1;8(4):457-70.
- [5]. Nguyen HQ, Nguyen QU. An Ensemble of Shallow and Deep Learning Algorithms for Vietnamese Sentiment Analysis. In 2018 5th NAFOSTED Conference on Information and Computer Science (NICS) 2018 Nov 23 (pp. 165-170). IEEE.
- [6]. Alwehaibi A, Roy K. Comparison of Pre-Trained Word Vectors for Arabic Text Classification Using Deep Learning Approach. In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) 2018 Dec 17 (pp. 1471-1474). IEEE.
- [7]. Yenter A, Verma A. Deep CNN-LSTM with combined kernels from multiple branches for IMDB

- Review Sentiment Analysis. In 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON) 2017 Oct 19 (pp. 540-546). IEEE.
- [8]. Day MY, Lin YD. Deep Learning for sentiment analysis on google plays consumer review. In 2017 IEEE International Conference on Information Reuse and Integration (IRI) 2017 Aug 4 (pp. 382-388). IEEE.
- [9]. Yenter A, Verma A. Deep CNN-LSTM with combined kernels from multiple branches for IMDB Review Sentiment Analysis, 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON) 2017 Oct 19 (pp. 540-546). IEEE.
- [10]. Day MY, Lin YD. Deep Learning for sentiment analysis on google plays consumer review. In 2017 IEEE International Conference on Information Reuse and Integration (IRI) 2017 Aug 4 (pp. 382-388). IEEE.
- [11]. Alshari EM, Azman A, Doraisamy S, Mustapha N, Alkeshr M. Effective Method for Sentiment Lexical Dictionary Enrichment Based on Word2Vec for Sentiment Analysis. In 2018 Fourth International Conference on Information Retrieval and Knowledge Management (CAMP) 2018 Mar 26 (pp. 1-5). IEEE.
- [12]. Yang G, He H, Chen Q. Emotion-Semantic-Enhanced Neural Network. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2019 Mar;27(3):531-43.
- [13]. Zhang Z, Lan M. Learning sentiment-inherent word embedding for word-level and sentence-level sentiment analysis. In 2015 International Conference on Asian Language Processing (IALP) 2015 Oct 24 (pp. 94-97). IEEE.
- [14]. Cai J, Li J, Li W, Wang J. Deep learning Model Used in Text Classification. In 2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP) 2018 Dec 14 (pp. 123-126). IEEE.
- [15]. Zharmagambetov AS, Pak AA. Sentiment analysis of a document using a deep learning approach and decision trees. In 2015 Twelve International Conference on Electronics Computer and Computation (ICECCO) 2015 Sep 27 (pp. 1-4). IEEE.
- [16]. Bandana R. Sentiment Analysis of Movie Reviews Using Heterogeneous Features. In 2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech) 2018 May 4 (pp. 1-4). IEEE.
- [17]. Yadav, A.S. and Kushwaha, D.S., 2021. Digitization of Land Record Through Blockchain-based Consensus Algorithm. IETE Technical Review, pp.1-18.
- [18]. <https://www.datascience.com/blog/understanding-ai-machine-learning-deep-learning>
- [19]. <https://medium.com/@chethankumargn/artificial-intelligence-definition-types-examplestechnologies-962ea75c7b9b>
- [20]. [https://en.wikipedia.org/wiki/Hadamard_product_\(matrices\)](https://en.wikipedia.org/wiki/Hadamard_product_(matrices))
- [21]. <https://towardsdatascience.com/understanding-gru-networks-2ef37df6c9be>
- [22]. <https://medium.com/@rgrgrajat1/sentence-classification-using-cnn-with-deep-learning-studio-fe54eb53e24>
- [23]. Mejora Y. Sentiment analysis: An overview. Comprehensive exam paper. Computer Science Department. 2009:
- [24]. Hassan A, Mahmood A. Convolutional recurrent deep learning model for sentence classification. IEEE Access. 2018; 6:13949-57.
- [25]. Hassan A, Mahmood A. Deep learning approach for sentiment analysis of short texts. In 2017 3rd international conference on control, automation and robotics (ICCAR) 2017 Apr 24 (pp. 705-710). IEEE.
- [26]. Yadav, R. K. S. A. S., & Khare, M. B. M. D. An Cost-Effective Euclidean Steiner Tree-based Mechanism for Reducing Latency in Cloud.