International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 4, July-Aug-2020, ISSN (Online): 2395-566X

Sentiment Analysis using Deep Learning: A Review

Digvijay Singh
Department of CSE
Anand International College of Engineering,
Jaipur (Rajasthan)-303012, India
ds93755@gmail.com

***Abstract* – Deep learning has evolved as a powerful technique that learns from the features present in multiple layers of the data and produces futuristic predictions. Deep Learning has been impacting many domains such as Computer Vision, Natural Language Processing and many more. As a part of Natural Language Processing, Deep Learning has been being used for the Sentiment Analysis, either to identify the purchase behavior of customers or in recommendation engine to provide them completely amazing user experience. Under this smart internet era World Wide Web such as Social media networks, Survey Forums, Review Sites generate huge amounts of data in the form of user-views, emotions, opinions and arguments about different social events, products, brands, and politics. Sentiments of users that are expressed on the web has great influence on the readers, product vendors and politicians. This unstructured data from the online resources is converted into well-structured meaningful data for Sentiment Analysis, Sentiment Analysis has recognized significant attention and has great impact on the industries. Sentiment analysis is referred to as identifying the feeling of people around us in the form of negative, positive, favorable, unfavorable, thumbs up, thumbs down, etc. One of the challenges in Sentiment Analysis is the lack of labelled data. And to rectify this issue, the sentiment analysis and deep learning techniques have been merged because deep learning models are very effective due to their automatic learning capability. This paper gives an overview of deep learning and then provides a comprehensive overview of its current applications in sentiment analysis.**

***Keywords* – Natural Language Processing, Deep Learning, Sentiment Analysis, Natural Language Understanding, Recurrent Neural Networks.**

I. INTRODUCTION

Sentiment analysis also referred to as the opinion-mining is the computational study of people's sentiments, opinions, appraisals, emotions and towards entities such as services, products, individuals, organizations, events, issues, topics and their attributes. The rapid growth of the field, deep learning, coincide with those of the social media on the Web, for example, blogs, microblogs, reviews, Twitter, social networks and forum discussions because in this technology era, we have a huge volume of data recorded in digital forms. Since early 2000, sentiment analysis has grown to be one of the most active research areas in natural language processing (NLP).

It is also widely studied in Web mining, text mining, data mining, and information retrieval. In fact, it has spread from computer science to management sciences and social sciences such as marketing, finance, political science, communications, health science, and even history, due to its importance to business and society as a whole. This accumulation is due to the fact that opinions are central to almost all human activities and are key influencers of our behaviors. This is not only true for individuals but also true for organizations. Organization perform the Sentiment Analysis task on the data that they have to better understand their customer needs. So that they can provide them completely customized experience. Nowadays, if one wants to buy a consumer product, is no longer limited to asking one's friends and family for opinions because there are many user reviews and discussions about the product in public forums on the Web. For an organization, it may no longer be necessary to conduct surveys, opinion polls, and focus groups in order to gather public opinions because there is an abundance of such information publicly available.

However, finding, monitoring and understanding opinion sites on the Web and capturing the information contained in them remains a formidable task because of the augmentation of diverse sites. Each site typically contains a huge volume of opinion text that is not always easily deciphered in long blogs and forum postings. The average human reader will have difficulty identifying relevant sites and extracting and summarizing the opinions in them. Automated sentiment analysis systems are thus needed.

1. Deep learning

Deep Learning was firstly proposed by G.E. Hinton in 2006 and is a subset of machine learning process which refers to Deep Neural Networks. Neural network is inspired by human brain and it contains several neurons that make an interconnected network. Deep learning networks are capable of providing training to both supervised and unsupervised categories of Deep Learning. Deep learning includes many networks such as ANN (Artificial Neural Networks), CNN (Convolutional Neural Networks), RNN (Recurrent Neural Networks), Recursive Neural Networks, DBN (Deep Belief Networks) and many more. Neural networks are very beneficial in word representation estimation, text generation, vector representation, sentence classification, sentence modeling and feature presentation.

Applications of Deep Learning: Deep architecture consists of numerous levels of non-linear operations. The capability of modeling the tasks of hard artificial intelligence makes the expectations that deep architecture will act as good in semi supervised learning such as Deep belief network (DBN) and will attain prominent success in Natural language processing community. Deep Learning consists of improved software engineering, enhanced learning procedures and accessibility of computing power and training data. It is inspired by neuro science and has splendid impact on a range of applications like NLP (Natural Language Processing), speech recognition, and computer vision. One of the basic challenges of deep learning research is the way of learning the structure of model and the quantity or number of layers and quantity of hidden variables for each layer.

While dealing with varying functions, the architecture of deep learning shows full potential and requires labeled samples in high amount for data capturing by deep architecture. Deep learning networks and techniques have been implemented widely in various fields such as in visual classification, pedestrian detection, off-road robot navigation, object categories, acoustic signals and Time series predictions tasks. A very motivating approach in natural language processing has explored that complex multi-tasking such as semantic labeling can be highly performed by using deep architectures.

In terms of data, deep learning efforts to learn high level abstractions by exploiting the hierarchical architectures. It is a promising approach and has been extensively applied in artificial intelligence field, like computer vision, transfer learning, semantic parsing, natural language processing and many more. Now a days, deep learning is prosperous because of three main and important reasons, i.e., improved abilities of chip processing (GPU units), extensively lower expenditure of hardware and significant enhancements in machine learning algorithms.

Combining Sentiment Analysis and Deep Learning
Deep learning is very influential in both unsupervised and supervised learning, many researchers are handling sentiment analysis by using deep learning. It consists of numerous effective and popular models and these models are used to solve the variety of problems effectively [15]. The most famous example Socher has used is the Recursive Neural Network (RNN) for the representation of movies reviews from the website rottentomatoes.com.
By following the effort of, many researchers have performed sentiment classification by using neural networks. For example, Kalchbrenner anticipated a Dy-CNN (Dynamic Convolutional Neural Network) which use a pooling operation, i.e, dynamic k-max pooling on linear sequences. Kim use CNN to learn sentiment-bearing sentence vectors.

Furthermore, Paragraph vector is proposed by Mikolov, et al. which performs better sentiment analysis than bag- of-words model and ConvNets for learning the SSWE (sen-timent specific word embedding). Recently has applied ConvNets on characters rather than directly applying on word embeddings. Another composition model for the classification of sentiments is LSTM which can simply take flow on one input direction.
In the Literature Review section, we will describe Neural Networks and various types of Neural Networks used in Deep Learning.

II. LITERATURE REVIEW

1. Neural Networks
Neural networks are a set of algorithms, model loosely after the human brain, that are designed to recognize patterns. They interpret sensory data through a kind of machine perception, labelling or clustering raw input. The patterns they recognize are numerical, contained in vectors, into which all real-world data, be it images, sound, text or time series, must be translated.

Neural networks help us cluster and classify. One can think of them as a clustering and classification layer on top of the data you store and manage. They help to group un-labelled data according to similarities among the example inputs, and they classify data when they have a labelled dataset to train on. (Neural networks can also extract features that are fed to other algorithms for clustering and classification; so one can think of deep neural networks as components of larger machine-learning applications involving algorithms for reinforcement learning, classification and regression.)
Artificial Neural Networks

Deep learning is the application of Artificial Neural Networks (ANNs) for learning tasks using networks of multiple layers. It can exploit much more learning (representation) power of neural networks, which once were deemed to be practical only with one or two layers and a small amount of data.
Inspired by the structure of the biological brain, neural networks consist of a large number of information processing units (called neurons) organized in layers, which work in unison. It can learn to perform tasks (e.g.,

classification) by adjusting the connection weights between neurons, resembling the learning process of a biological brain.

A simple example of a feedforward neural network is given in Figure 1, which consists of three layers L1, L2, and L3. L1 is the input layer, which corresponds to the input vector (x1, x2, x3) and intercept term +1. L3 is the output layer, which corresponds to the output vector (s1). L2 is the hidden layer, whose output is not visible as a network output. A circle in L1 represents an element in the input vector, while a circle in L2 or L3 represents a neuron, the basic computation element of a neural network. We also call it an activation function. A line between two neurons represents a connection for the flow of information. Each connection is associated with a weight, a value controlling the signal between two neurons. The learning of a neural network is achieved by adjusting the weights between neurons with the information flowing through them. Neurons read the output from neurons in the previous layer, process the information, and then generate output to neurons in the next layer. As in Figure 1, the neutral network alters weights based on training examples (x(i), y(i)). After the training process, it will obtain a complex form of hypotheses hw, b(x) that fits the data.

Diving into the hidden layer, we can see that each neuron in L2 takes input x1, x2, x3 and intercept +1 from L1, and outputs a value as Summation(Wixi + b) by the activation function f. Wi are weights of the connections; b is the intercept or i =1 bias; f is normally nonlinear. The common choices of f are sigmoid function, hyperbolic tangent function (tanh), or rectified linear function (ReLU).

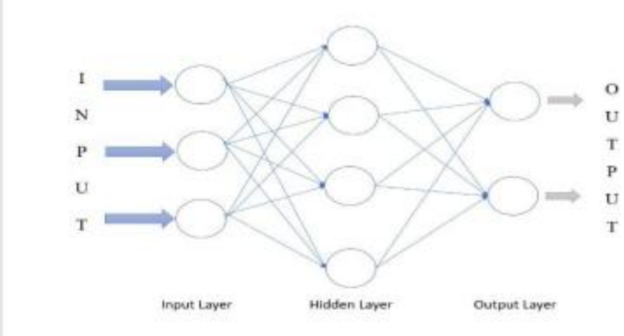

Fig. 1. Artificial Neural Network.

2. Convolutional Neural Networks

Convolutional neural network (CNN) is a special type of feedforward neural network originally employed in the field of computer vision. Its design is inspired by the human visual cortex, a visual mechanism in animal brain. The visual cortex contains a lot of cells that are responsible for detecting light in small and overlapping subregions of the visual fields, which are called receptive fields. These cells act as local filters over the input space. CNN consists of multiple convolutional layers, each of which performs the function that is processed by the cells in the visual cortex.

The CNN (convolutional neural network) [24] includes pooling layers and sophistication as it gives a standard architecture to map the sentences of variable length into sentences of fixed size scattered vectors.

This study [25] has proposed a novel convolutional neural network (CNN) framework for visual sentiment analysis to predict sentiments of visual content. CNN has been implemented using Caffe and Python on a Linux machine. Transfer learning approach and hyper-parameter has been used in biases and weights are utilized from pre-trained GoogLeNet. As CNN enhance its performance by increasing its size and depth, so a very deep CNN model, inspired by GoogLeNet is proposed with 22 layers for sentiment analysis. It is optimized by using SGD (Stochastic gradient descent) algorithm. The strategy with 60 epochs has been performed for training the network as GoogLeNet has performed 250 epochs. For experimental work, a dataset of twitter containing 1269 images is selected and back propagation is applied. Amazon Mechanical Turk (MTurk) and popular crowd intelligence is used to label the images. Five workers were involved to generate sentiment label in favor of every image. The proposed model was evaluated on this dataset and acquired better performance than existing systems. Results shows that proposed system achieve high performance without fine-tuning on Flickr dataset. However, AlexNet was used in previous works and GoogleNet provided almost 9% performance progress than AlexNet. By converting GoogLeNet in to visual sentiment analysis framework, the better feature extraction was achieved. Stable and reliable state were achieved by using hyper parameters.

The authors have proposed the system of deep learning for sentiment analysis of twitter. The main focus of this work was to initialize the weight of parameters of convolutional neural network and it is critical to train the model accurately while avoiding the requirement of adding new feature. A neural language is used to initialize the word embedding and is trained by big unsupervised group of tweets. For further refining the embedding on bulky supervised corpus, a conventional neural network is used. To initialize the network, previously embedded words and parameters were used, having same architecture and training on supervised corpus as of Semeval- 2015. The components used in proposed work are activations, sentence matrix pooling, softmax and convolutional layers. To train the network, stochastic gradient descent (SGD) and non-convex function optimization algorithms were used and to calculate the gradients back propagation algorithm was used. Dropout technique were used to enhance the neural networks regularization. The deep learning model is applied on two tasks: message level task and phrase level task from Semeval-2015 to predict polarity and achieve high outcomes. By applying six test-set, the proposed model lies at first rank in terms of accuracy.

A detailed research by has presented an overview of sentiment analysis related to Micro-blog. The purpose of this effort was to get the opinions and attitudes of users about hot events by using Convolutional Neural Network (CNN). The use of CNN overcomes the problem of explicit feature extraction and learns implicitly through training data. To collect the data from target, the input URL and focused crawler have been used, 1000 micro-blog comments were collected as a corpus and divided into three labels, i.e., 274 neutral emotions, 300 negative emotions and 426 positive emotions. The proposed model has been compared with the previous studies as those had studies used CRF, SVM and additional traditional algorithms to perform sentiment analysis with a high price. However, the performance proves that the proposed model is reasonable and sufficient to enhance the accuracy in terms of emotion analysis. Research, motivated through the need of controlling of comprehensive social multimedia content and employ both textual and visual SA techniques for combined textual- visual sentiment analysis. A convolutional neural net- work (CNN) and a paragraph vector model were used for both the image and textual SA accordingly. The proposed model was termed as rule-based sentiment classifier VADER. After conducting the wide range of experiments on manually labeled and weakly labeled visual tweets, it was concluded that mutual textual-visual features outperformed the sentiment analysis algorithms which were only depend on Visual contents. It was demonstrated that how models of one domain to a different domain can be transmitted easily. Getty Images had been selected to crawl data and Caffe was used to tune the CNN model. Tweets were gathered through Twitter API. To make sentiment labels for chosen visual tweets, the Mechanical Turk (AMT) and crowd intelligence had been employed. The results recommend that the joint textual-visual model has performed better than the both single visual and textual sentiment analysis models.

3. Recurrent Neural Networks

The Recurrent neural network (RNN) is an influential model in language modeling because it doesn't represent the context of fixed-length that contaminate all history words.

In this study, the HBRNN (hierarchical bidirectional recurrent neural network) has been developed to extract the reviews of customers about different hotels in a complete and concise manner. To model the sequential long term in- formation, HBRNN has used the terminology of RNN and the prediction process was done at review level by HBRNN. The experimental data was taken from DBS text mining Challenge 2015. HBRNN performance was improved through networks parameters along with the fine tuning and the model was compared with LSTM (long shot term memory) and BLSTM (Bidirectional LSTM). After performing the experiments, the evaluation recall, F1 scores and precision was made on highly biased data. The development, test set and train splits were used for comparing outcomes with benchmark systems, tenfold cross validation used to present the performance of HBRNN. The main challenges that was resolved is lack of online reviews with high quality and lack of high skewness in the reviewed data. Experimental Results on the dataset proved that HBRNN performed better than other methods. This model can be applied to other opinion mining activities which consists of huge data volume. This contribution has been done to overcome the issue of dataset of Bangla as it is standard and large for SA (Sentiment Analysis) tasks. The issue has been resolved by providing a significant dataset for sentiment analysis of 10,000 BRBT (Bangla and Romanized Bangla Text). The Deep Recurrent model especially LSTM (Long Short Term Memory) was used to test the dataset by using two loss functions, i.e., binary and categorical cross-entropy. Data were gathered from different sites like YouTube, Face book, Twitter and others. The experiments were conducted to prepare dataset of one mark for another (and the other way around) to investigate the fact whether it contributes towards the better outcomes.

III.CONCLUSION

Sentiment analysis refers to the management of sentiments, opinions, and subjective text. The demand of sentiment analysis is raised due to the requirement of analyzing and structuring hidden information, extracted from social media in form of un- structured data. The sentiment analysis is being implementing through deep learning techniques. Deep learning consists of numerous effective and popular models, these models are used to solve the variety of problems effectively. Different studies have been discussed in this review to provide a deep knowledge of the successful growing of deep learning applications in the field of sentiment analysis. Numerous problems have been resolved by having high accuracy of both fields of sentiment analysis and deep learning.

REFERENCES

[1]. M. Haenlein and A. M. Kaplan, An empirical analysis of attitudinal and behavioral reactions toward the abandonment of unprofitable customer relationships, J. Relatsh. Mark., vol. 9, no. 4, pp. 200228, 2010.
[2]. J. Singh, G. Singh, and R. Singh, A review of sentiment analysis techniques for opinionated web text, CSI Trans. ICT, 2016.
[3]. E. Aydogan and M. A. Akcayol, A comprehensive survey for sentiment analysis tasks using machine learning techniques, 2016 Int. Symp. Innov. Intell. Syst. Appl., pp. 17, 2016.

[4]. M. Day and C. Lee, Deep Learning for Financial Sentiment Analysis on Finance News Providers, no. 1, pp. 11271134, 2016.
[5]. P. Vateekul and T. Koomsubha, A Study of Sentiment Analysis Using Deep Learning Techniques on Thai Twitter Data, 2016.
[6]. Y. Zhang, M. J. Er, N. Wang, M. Pratama, and R. Venkatesan, Sentiment Classification Using Comprehensive Attention Recurrent Models, pp. 15621569, 2016.
[7]. S. Zhou, Q. Chen, and X. Wang, Active deep learning method for semi-supervised sentiment classification, Neurocomputing, vol. 120, pp. 536546, 2013.
[8]. L. Deng, G. Hinton, and B. Kingsbury, New types of deep neural network learning for speech recognition and related applications: an overview, 2013 IEEE Int. Conf. Acoust. Speech Signal Process., pp. 85998603, 2013.
[9]. S. Bengio, L. Deng, H. Larochelle, H. Lee, and R. Salakhutdinov, Guest Editors Introduction: Special Section on Learning Deep Architectures, IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 8, pp. 17951797, 2013.
[10]. L. Arnold, S. Rebecchi, S. Chevallier, and H. Paugam-Moisy, An Introduction to Deep Learning, Esann, no. April, p. 12, 2011.
[11]. Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, Deep learning for visual understanding: A review, Neurocomputing, vol. 187, pp. 2748, 2016.
[12]. X. Ouyang, P. Zhou, C. H. Li, and L. Liu, Sentiment Analysis Using Convolutional Neural Network, Comput. Inf. Technol. Ubiquitous Com- put. Commun. Dependable, Auton. Secur. Comput. Pervasive Intell. Com- put. (CIT/IUCC/DASC/PICOM), 2015 IEEE Int. Conf., pp. 23592364, 2015.
[13]. Cho, K., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine transla- tion. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2014).
[14]. Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modelling. arXiv preprint arXiv:, 1412, 3555.
[15]. Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. Journal of Machine Learning Research.
[16]. Dahou, A., Xiong, S., Zhou, J., Haddoud, M. H., & Duan, P. (2016). Word embeddings and convolutional neural network for Arabic sentiment classification. In Pro- ceedings of the International Conference on Computational Linguistics (COLING 2016).
[17]. Deng, L., & Wiebe, J. (2016). Recognizing opinion sources based on a new categorization of opinion types. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2016).