Sentiment Analysis Using Deep Learning Techniques: A Review

Qurat Tul Ain*, Mubashir Ali*, Amna Riaz†, Amna Noureen‡, Muhammad Kamran‡, Babar Hayat* and A. Rehman*
*Department of Computer Science and Information Technology,
The University of Lahore, Gujrat, Pakistan
†Department of Information and Technology,
University of Gujrat, Gujrat, Pakistan
‡Department of Computer Engineering,
EME, NUST, Islamabad, Pakistan

Abstract—**The World Wide Web such as social networks, forums, review sites and blogs generate enormous heaps of data in the form of users 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. The unstructured form of data from the social media is needed to be analyzed and well-structured and for this purpose, sentiment analysis has recognized significant attention. Sentiment analysis is referred as text organization that is used to classify the expressed mind-set or feelings in different manners such as negative, positive, favorable, unfavorable, thumbs up, thumbs down, etc. The challenge for sentiment analysis is lack of sufficient labeled data in the field of Natural Language Processing (NLP). And to solve this issue, the sentiment analysis and deep learning techniques have been merged because deep learning models are effective due to their automatic learning capability. This Review Paper highlights latest studies regarding the implementation of deep learning models such as deep neural networks, convolutional neural networks and many more for solving different problems of sentiment analysis such as sentiment classification, cross lingual problems, textual and visual analysis and product review analysis, etc.**

Keywords—*Sentiment analysis; recurrent neural network; deep neural network; convolutional neural network; recursive neural network; deep belief network*

I. INTRODUCTION

A. Sentiment Analysis

Sentiment analysis refers to the management of sentiments, opinions, and subjective text [1]. Sentiment analysis provide the comprehension information related to public views, as it analyze different tweets and reviews. It is a verified tool for the prediction of many significant events such as box office performance of movies and general elections [2]. Public reviews are used to evaluate a certain entity, i.e., person, product or location and might be found on different websites like Amazon and Yelp. The opinions can be categorized into negative, positive or neutral. The purpose of sentiment analysis is to automatically determine the expressive direction of user reviews [3]. The demand of sentiment analysis is raised due to increase requirement of analyzing and structuring hidden information which comes from the social media in the form of unstructured data [4].

1) Features of Sentiment Analysis: Sentiments contains a variety of featured values like tri-grams and bi-grams by means of polarities and combinations. So sentiments are being assessed both as negative and positive aspects through the numerous support vector machines, by using training algorithms. The neural networks are implemented in sentiment analysis to compute belongingness of labels. To help out data extraction at context level the conditional dependencies among several edges and nodes of acyclic graph operated by Bayesian networks are used. By optimizing words and sentences, learning and data accuracy can be attained on social media platform. At word root level, data tokenization is used to produce negative and positive aspects of data. Techniques are being used to decrease the errors in sentiment analysis to attain higher level of precision in data for social media. [5]

2) Sentiment Analysis as multidisciplinary Field: The sentiment analysis is multidisciplinary field, because it includes numerous fields such as computational linguistics, information retrieval, semantics, natural language processing, artificial intelligence and machine learning [6]. The classification for the approaches of sentiment analysis can be done in three extraction levels a) feature or aspect level; b) document level; and c) sentence level [5].

3) Techniques for Sentiment Analysis: Sentiment analysis relies on two types of techniques, i.e., lexicon based and machine learning based techniques [5].

a) Machine learning based techniques: This type of techniques are implemented by extracting the sentences and aspect levels. The features consist of Parts of Speech (POS) tags, n-grams, bi-grams, uni-grams and bag-of-words. Machine learning contains three flavors at sentence and aspect, i.e., Nave Bayes, Support Vector Machine (SVM) and Maximum Entropy.

b) Lexicon based or corpus based techniques: These techniques are based on decision trees such as k-Nearest Neighbors (k-NN), Conditional Random Field (CRF), Hidden Markov Model (HMM), Single Dimensional Classification (SDC) and Sequential Minimal Optimization (SMO), related to methodologies of sentiment classification.

Machine learning approach has three categories: i) supervised; ii) semi supervised; and iii) unsupervised. This approach

is capable of automation and can handle huge amount of data therefore these are very suitable for sentiment analysis.[6].

B. Deep learning

Deep Learning was firstly proposed by G.E. Hinton in 2006 and is the part of machine learning process which refers to Deep Neural Network [7]. Neural network is influenced by human brain and it contains several neurons that make an impressive network. Deep learning networks are capable for providing training to both supervised and unsupervised categories [8].

Deep learning includes many networks such as CNN (Convolutional Neural Networks), RNN (Recurrent Neural Networks), Recursive Neural Networks, DBN (Deep Belief Networks) and many more. Neural networks are very beneficial in text generation, vector representation, word representation estimation, sentence classification, sentence modeling and feature presentation. [9].

1) 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 be act good in semi supervised learning such as Deep belief network (DBN) and will attain prominent success in Natural language processing community[10]. The Deep Learning consists of improved software engineering, enhanced learning procedures and accessibility of computing power and training data [11]. It is inspired by neuro science and has splendid impact on a range of applications like speech recognition, NLP (Natural Language Processing) and computer vision. One of the basic challenge of deep learning research is the way of learning the structure of model and the quantity of layers and quantity of hidden variables for each layer [12].

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 [13]. 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[13].

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 [14].

C. 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 [16].

By following the effort of [17], many researchers has performed sentiment classification by using neural networks. For example, Kalchbrenner [18] anticipated a DyCNN (Dynamic Convolutional Neural Network) which use a pooling operation, i.e, dynamic k-max pooling on linear sequences. Kim [19] use CNN to learn sentiment-bearing sentence vectors.

Furthermore, Paragraph vector is proposed by Mikolov, *et al*. [20] which performs better sentiment analysis than bag-of-words model and ConvNets for learning the SSWE (sentiment specific word embedding). Recently [21] 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 [22].

In this paper, Section II covers the detailed literature review and Section III presents the conclusion whereas, Section IV shows the analysis phase been analyzed in literature review.

In the Literature Review section, various studies of sentiment analysis using Deep Learning techniques are discussed. This review is conducted on the basis of numerous latest studies in the field of sentiment analysis.

II. LITERATURE REVIEW

For the accurate classification of sentiments, many researchers have made efforts to combine deep learning and machine learning concepts in the recent years. This section briefly describes the numerous studies, related to sentiment analysis of web contents about users opinions, emotions, reviews toward different matters and products using deep learning techniques.

Sentiment analysis tasks can be performed efficiently by implementing different models such as deep learning models, which have been extended recently. These models include CNN (convolutional neural networks), RNN (recursive neural network), DNN (deep neural networks), RNN (recurrent neural networks) and DBN (deep belief networks). This section describes the efforts of different researchers toward implementing deep learning models for performing the sentiment analysis [23]. Several researchers have used more than one model in their study, and these are mentioned under the hybrid neural network section.

A. Convolutional Neural Networks (CNN)

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 [26] 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 [27] 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 by [28] was 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 network (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.

In study by [15], the researcher has represented a seven-layer framework to analyze the sentiments of sentences. This frame work depends on CNN (Convolutional neural network) and Word2vec for SA and to calculate vector representation, respectively. Word2vec have been proposed by Google. The Dropout technology, Normalization and Parametric Rectified Linear Unit (PReLU), have been used to progress the correctness and generalizability of proposed model. The framework was verified on the data set from rottentomatoes.com which contains movie review excerpts' corpus, the dataset consists of five labels positive, somewhat positive, neural, negative and somewhat negative. By comparing the proposed model with previous models such as Matrix-Vector recursive neural network (MV-RNN) and recursive neural network (RNN), the proposed model outperformed the previous models with the 45.5 % accuracy. In Table I, we have presented the analysis of best studies following the convolutional neural network.

B. Recursive Neural Network (RNN)

The recursive neural network (RNN) [24] lies in supervised learning. It contains a tree structure which is settled before training and the nodes can have different matrices. There is no need of reconstruction of input in RNN.

The proposed work [29] builds a Treebank for chines sentiments of social data to overcome the deficiency of labeled and large corpus in existing models. To predict the labels at sentence level i.e positive or negative, the Recursive Neural Deep Model (RNDM) was proposed and achieved high performance than SVM, Nave Bayes and Maximum Entropy. 2270 movie reviews were collected from the website and Chinese word segmentation tool ICTCLAS was used to segment these reviews. Five classes were settled for each sentence and standford parser applied for sentence parsing. The proposed model improved the prediction of sentiment labels of sentences by concluding 13550 chines sentences and 14964 words. ME and NB performs higher with contrastive conjunction structure than baselines with great margin.

In this study [30], a model comprising RNTN (Recursive Neural Tensor Network) and Sentiment Treebank has been proposed to correctly clarify the compositional effects at different levels of phrases, i.e., positive and negative phrases.The proposed model was compared with all the existing models.

TABLE I.   ANALYSIS OF CONVOLUTIONAL NEURAL NETWORKS

Researcher Name and Year	Model Used	Purpose	Data Set	Results
J.Islam and Y. Zhang 2016 [25]	Convolutional Neural Networks (CNN)	Visual SA	1269 images from twitter	GoogleNet gave almost 9 % performance progress than AlexNet.
A. Severyn and A. Moschitti, 2015 [26]	Convolutional Neural Networks (CNN)	Phrase level and message level task SA	Semeval-2015	Compared with official system ranked 1st in terms of pharse level subtask and ranked 2nd in terms of message level.
L. Yanmei and C. Yuda, 2015 [27]	Convolutional Neural Networks (CNN)	Micro-Blog SA	1000 micro-blog comments (HuaQiangu)	Proposed model can effectively improve the accuracy of emotional orientation, validation.
Q. You, J. Luo, H. Jin, and J. Yang, 2015 [28]	Convolutional Neural Networks (CNN)	Textual-visual SA	Getty Images, 101 keywords	Joint visual and textual model outperforms the early single fusions.
X. Ouyang, P. Zhou, C. H. Li, and L. Liu, 2015 [15]	Convolutional Neural Networks (CNN)	Sentiments of sentences	rottentomatoes.com (contains movie review excerpts)	The proposed model outperformed the previous models with the 45.5% accuracy.

TABLE II.   ANALYSIS OF RECURSIVE NEURAL NETWORKS

Researcher Name and Year	Model Used	Purpose	Data Set	Results
C. Li, B. Xu, G. Wu, S. He, G. Tian, and H. Hao, 2014 [29]	Recursive Neural Deep Model (RNDM)	Chines sentiments analysis of social data	2270 movie reviews from websites	Performs higher (90.8%) than baselines with a great margin.
R. Socher, A. Pere-lygin, and J. Wu, 2013 [30]	RNTN (Recursive Neural Tensor Network)	Semantic Com-positionality	11,855 single sentences from moview review (Pang and Lee 2005)	The RNTN achieved 80.7% accuracy in sentiment prediction , an improvement of 9.7 % over baselines (bag of features).
W. Li and H. Chen, 2014 [31]	Recursive Neural Network (RNN)	Identifying Top Sellers In Underground Economy	Russian carding Forum)	Results have been indicated that Deep learning techniques accomplish superior outcomes than shallow classifiers. Carding sellers have fewer ratings than malware sellers.

In existing models, the meaning of long phrases cannot be expressed effectively by semantic word spaces, so for sentiment detection, more rich and supervised evaluation and training resources are needed as it requires more influential composition models. The RNTN achieved 80.7% accuracy in sentiment prediction by performing fine-grained labeling over all the phrases and outperformed previous models.

This study [31] has contributed a generalized and scaled framework to recognize top carding/malware sellers. The model is based on deep learning for sentiment analysis and used in thread classification and snowball sampling to assess the quality of sellers service/product by analyzing the customer feedback. The evaluation of proposed model has been conducted on Russian carding forum and a web crawler was used to gather the conversation stuff from the forum. A sentiment tree bank has been used and it was trained by using recursive neural tensor network on online review corpus. For evaluating the validity and effectiveness, two experiments were conducted in which proposed model was compared with Nave Bayes, KNN and SVM based models. This study has searched out the sellers who are highly rated for malicious services/products and the effectiveness of deep learning for recognizing these sellers. Results have been indicated that Deep learning techniques accomplish superior outcomes than the shallow classifiers and it was established that the carding

sellers have fewer ratings than malware sellers. Table II, presenting the analysis of best approaches based on recursive neural network.

C. Deep Neural Networks (DNN)

In this study [32], author has proposed a model for sentiment analysis considering both visual and textual contents of social networks. This new scheme used deep neural network model such as Denoising auto encoders and skip gram. The base of the scheme was CBOW (Continuous Bag-Of-Words) model. The proposed model consisted of two parts CBOW-LR (logistic regression) for textual contents and was expanded as the CBOW-DA-LR. The classification was done according to the polarity of visual and textual information. Four datasets were evaluated, i.e., Sanders Corpus, Sentiment140, SemEval-2013 and SentiBank Twitter dataset. The proposed model outperformed the CBOWS+SVM and FSLM (fully supervised probabilistic language model). Perhaps the ESLAM (extended fully supervised probabilistic language model) in term of small training data had outperformed the current model. The feature learning and skip grams both required large datasets for best performance.

In this study [33], deep neural network architecture has been proposed to evaluate the similarity of documents. The architecture was trained by using several market news to

TABLE III.   ANALYSIS OF DEEP NEURAL NETWORKS

Researcher Name and Year	Model Used	Purpose	Data Set	Results
C. Baecchi, T. Uricchio, M. Bertini, and A. Del Bimbo, 2016 [32]	deep neural networks (CBOW-DA-LR)	Visual and Textual SA	4 datasets: Sanders Corpus, Sentiment140, SemEval-2013 and SentiBank Twitter Dataset	CBOW-DA-LR model obtained superior classification accuracy than previous models.
H. Yanagimoto, M. Shimada, and A. Yoshimura, 2013 [33]	Deep Neural Network (DNN)	Document Similarity Estimation	T&C News	The proposed method accomplished superior performance in terms of similarity estimation of articles according to polarity.

produce vectors foe articles. The T&C news have been used as dataset. The cosine similarity was calculated among labeled articles and the polarity of documents was considered but contents were not considered. The proposed method accomplished superior performance in terms of similarity estimation of articles according to polarity. In Table III, we have presented the summarized analysis of two approaches using the deep neural network.

D. Recurrent Neural Networks (Recurrent NN)

The Recurrent neural network (RNN) [24] is an influential model in language modeling because it doesnt represent the context of fixed-length that contaminate all history words .

In this study [34], 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 information, 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 [35] 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.

This author [36] proposed a sequence model to focus on the embedding of reviews having temporal nature toward products as these reviews had less focus in existing studies. The combination of gated recurrent units with recurrent neural network is used to learn dispersed representations of products and users. For sentiment classification these representations fed into machine learning classifier. The approach was evaluated on three datasets collected from Yelp and IMDB. Each review labeled according to rating score. To train the network the back-propagation algorithm with Adam stochastic optimization method has been used. Results show sequence modeling of dispersed product and user representation learning improves the performance sentiment classification of document-level and the proposed approach achieves high-tech results on the benchmark datasets. The result of proposed model compared with many baselines including recursive neural networks, user product neural network, word2vec, paragraph vector and algorithm JMARS. We have also made an analysis of some approaches based on Recurrent Neural Network and this analysis is presented in tabular form in Table IV.

E. Deep Belief Networks (DBN)

Deep belief networks (DBNs) [37] includes several hidden layers, composed by RBM (restricted Boltzmann machines). DBN has been proved efficient for feature representation. It utilizes the unlabeled data and fulfills the deficiencies of labeled analysis issues.

In this paper [38], a new deep neural network structure has been presented termed as WSDNNs (Weakly Shared Deep Neural Networks). The purpose of WSDNNs is to facilitate two languages to share sentiment labels. The features of language specific and inter language have been presented through building multiple weakly shared layers of features. The datasets from Prettenhofer and Stein have been used containing four languages French, German, English and Japanese. In comparison with existing studies the proposed work address the challenge of shortening overlap among feature spaces of both source and target language data through cross lingual information transfer process using backpropogation. DNNs used for transformation of information from source to target language. The experiments have been conducted for sentiment classification tasks of cross multilingual product reviews of Amazon. results concluded that the proposed approach is more effective and powerful in terms of cross lingual sentiment classification than the previous studies.

TABLE IV.   ANALYSIS OF RECURRENT NEURAL NETWORKS

Researcher Name and Year	Model Used	Purpose	Data Set	Results
R. Silhavy, R. Senkerik, Z. K. Oplatkova, P. Silhavy, and Z. Prokopova, 2016 [34]	HBRNN (hierarchical bidirectional Recurrent Neural Network)	Sentiment Analysis of Customer Reviews	150,175 labelled reviews from 1500 hotels (DBS text mining Challenge 2015)	The experimental results explored that that HBRNN outperformed all other methods.
A. Hassan, M. R. Amin, A. Kalam, A. Azad, and N. Mohammed, [35]	Deep Recurrent model especially LSTM (Long Short Term Memory)	Sentiment Analysis on Bangla and Romanized Bangla Text (BRBT)	9337 post Samples from different social sources	Ambiguous Removed with 78% accuracy. Ambiguous converted to 2 scored highest with 55% accuracy.
T. Chen, R. Xu, Y. He, Y. Xia, and X. Wang , 2016 [36]	Recurrent Neural Network (RNN-GRU)	Learning User and Product Distributed Representations	Three datasets collected from Yelp and IMDB.	Results have been indicated that proposed model outperformed many baselines including recursive neural networks, user product neural network , word2vec, paragraph vector and algorithm JMARS

Another study by [17] has used deep belief network with word vector for the political detection in Korean articles. The proposed model has used SVM for bias calculation, five stage pipeline for detection of political bias, python web crawler to gather news articles, KKMA for morpheme analysis, word2Vec and scikit-learn package. The dataset contained 50,000 political articles from 01 Jan, 2014 to 28 Feb, 2015. Results showed 81.8% accuracy by correctly predicting labels and the results contained mean square error of 0.120.

This research [37] has proposed a deep belief network with feature selection (DBNFS) to overcome the vocabulary problems, the network has used input corpus along with numerous hidden layers. Chi-Squared technique of feature selection was used to improve the Deep Belief Network (DBN) for the purpose of decreasing complexity of vocabulary input and for eliminating irrelevant features. By applying Chi-Squared technique, the learning phase of DBN was enhanced to DBNFS. In this work, two new tasks features, selection and reduction were used along with many other tasks of existing classification approaches, such as data portioning, feature extraction, model training and model testing. Performance of DBNFS was demonstrated and training time and accuracy of proposed DBNFS was also compared with other algorithms. Five Dataset of sentiment classification were used for estimation, datasets are books (BOO), electronics (ELE), DVDs (DVD), kitchen appliances (KIT) and movies reviews (MOV). For fair comparison, the parameters of learning were same as of existing works. Accuracy was evaluated by comparing the amount of features before and after the feature selection and reduction. The accuracy results were compared with the previous works and were proved better DBNFS than DBN. The training time was also lower in DBNFS than DBN. Training time was improved due to simple deep structure and proposed feature selection method. The only drawback of DBN was that it is costly plus time consuming. Summarized analysis of existing approaches for sentiment analysis by using DBN is presented in Table V.

F. Hybrid Neural Networks

This study [8] has proposed two deep learning techniques for the sentiment classification of Thai Twitter data, i.e., Convolutional Neural Network (DCNN) and Short Term Memory (LSTM). Data processing was conducted properly. Data was collected from the users and their followers of Thai Twitter. After filtering the data, only the users with Thai tweets and tweets with Thai characters were selected. Five experiments were conducted to achieve finest parameters for deep learning, to compare the deep learning with classical techniques and to achieve the words sequence importance. Three-fold cross validation was used to verify the process. The results concluded that the accuracy is high in DNN than LSTM and both techniques of deep learning are higher in accuracy than SVM and Nave Bayes but lesser than Maximum Entropy. Higher accuracies were found in original sentences than shuffled sentences so the words sequence is important.

In this research study [39] a hybrid model has proposed which consists of Probabilistic Neural Network (PNN) and a two layered Restricted Boltzmann (RBM). The purpose of proposing this hybrid deep learning model is to attain better accuracy of sentiment classification. The polarity, i.e., negative and positive reviews vary according to different context in order to solve this type of problem this model performs well, neutral reviews are not considered. Experiments have done with datasets of Pang and Lee and Blitzer, *et al.*, binary classification implemented on every dataset. The accuracy has been enhanced for five datasets by comparing with the existing state-of-the-art Dang, *et al.* [13]. There are no outer resources in proposed approach such as POS tagger and sentiment dictionary etc therefore it is faster too than competitor. To attain a reduced number of features the dimensionality reduction has been implemented as previous study used a complex strategy for feature selection. Approaches based on hybrid neural networks are presented in summarized form in Table VI.

TABLE V.   ANALYSIS OF DEEP BELIEF NETWORKS

Researcher Name and Year	Model Used	Purpose	Data Set	Results
G. Zhou, Z. Zeng, J. X. Huang, and T. He, 2016 [38]	WSDNNs (Weakly Shared Deep Neural Networks)	Cross-Lingual Sentiment Classification	Four languages reviews from amazon, each language consists of 1000 negative and 1000 positive reviews	Proposed approach is more effective and powerful than the previous studies by applying experiments on 18 tasks of cross lingual sentiment classification.
T. Mikolov, K. Chen, G. Corrado, and J. Dean, 2013 [17]	Deep Belief Networks Along with word vector	Political Detection in Korean articles	50,000 political articles	Results showed 81.8% accuracy by correctly predicting labels.
P. Ruangkanokmas, T. Achalakul, and K. Akkarajitsakul, 2016 [37]	Deep Belief Network with Feature Selection (DBNFS)	Feature Selection	Five sentiment classification datasets (1 is movie reviews and other four are multi-domain). Total 2,000 labeled reviews (1,000 negatives and 1,000 positives).	The accuracy results are compared with previous works and proved better DBNFS than DBN.

G. Other Neural Networks

In this study [40] to overcome the complexity in word-level models the character-level model have been proposed. The motivation of proposed model CDBLSTM is an existing model that is DBLSTM neural networks [41]. The focus of this work is only on textual content and on the polarity analysis of tweets in which a tweet is classified into two classes, i.e., positive and negative. There can be more options than positive and negative such as natural and finer but here the model is restricted only to positive and negative classes to compare with existing published results. The tweets are encoded from character level and trained by the use of CCE (categorical cross-entropy). Experiments were conducted on two datasets, first one is latest benchmark dataset for SemEval 2016 and the second one is provided by GO dataset. Adam algorithm was implemented to train all the models and the learning rate settled to 0.1. Final predictions were obtained by the model of logistic regression. Results demonstrate that proposed approach is competent for the polarity problems. By applying different experiments results shows that CDBLSTM performs better than DBLSTM and by comparing the results with Deep Convolutional Neural Network (DCNN) [42] which performs well on twitter SPC (sentiment polarity classification). The 85.86% accuracy was achieved on STS (Stanford Twitter Sentiment) corpus and 84.82% on SemEval-2016. In this study [41], author has proposed TF-IDF, GR, and RBFNN for sentiment classification on Hinglish text. Many studies have worked on sentiment analysis of various languages such as English, Turkish, Flemmish, Spanish, Arabic and Chinese but no work has been done on Indic language. To fill this gap, the sentiments were classified in Hinglish language as it contains Hindi words along with the English. Dataset has creep from Facebook comments and viz. news, five methods of feature

selection information gain, chi-square, t-statistics, association and gain ratio have been implemented on DTM (Document-Term Matrix) and TF-IDF (Term Frequency-Inverse Document Frequency). Many classifiers have used such as SVMs (Support Vector Machine), RBFNet (Random Forest, Radial Basis Function Neural Network), Naive Bayes, J48 (Decision Tree), CART, JRip, Logistic Regression (LR) and Multi-layer Perception (MLP) methods for the classification of data. Total 840 experiments were performed on datasets and best results were achieved. The proposed triumvirate approach was proved efficient for sentiment classification of Hinglish text.

This contribution [42] overcomes the problem that occurs in effectively analyzing the emotions of customers toward companies in blog sphere. A neural network (NN) based technique is proposed which subordinate the advantages of Semantic orientation index and machine learning methods for the classification of sentiments effectively and quickly. The input of neural network is semantic orientation indexes. While considering the fault tolerance ability the BPN (back-propagation neural network) is selected as basic learner of proposed method. Data was collected from the blogs of real world such as from "LiveJournal" and "Review Center". Segmentation of words, SO indexes calculations, neural network trainings and performance evaluations have been conducted. Training and test sets were settled of datasets. The overall accuracy (OA), performance evaluation matrices and F1 were used. Results concluded that proposed method has enhanced the performance of classification and saved training time as compared to traditional ML and IR.

This contribution [43] proposed a data driven supervised approach for the purpose of feature reduction and development of lexicon specific to twitter sentiment analysis about brand. Statistical analysis and n-gram were used for twitter specific

TABLE VI.   ANALYSIS OF HYBRID AND SOME OTHER NEURAL NETWORKS

Researcher Name and Year	Model Used	Purpose	Data Set	Results
P. Vateekul and T. Koomsubha , 2016 [8]	Convolutional Neural Network (DCNN) and Short Term Memory (LSTM).	Sentiment Analysis on Thai Twitter Data	3,813,173 tweets (33,349 negative tweets and 140,414 positive tweets)	Higher in accuracy than SVM and Nave Bayes lesser than Maximum Entropy Higher accuracies in original sentences than shuffled sentences
R. Ghosh, K. Ravi, and V. Ravi, 2016 [39]	Probabilistic Neural Network (PNN) and a two layered Restricted Boltzmann (RBM)	Better accuracy of sentiment classification	Pang and Lee and Blitzer, *et al.* (1000 negative and 1000 positive reviews on each of DVDs, Books (BOO), Kitchen appliances (KIT) and Electronics (ELE).	The proposed model attains accuracies in following manner: MOV = 93.3%, BOO = 92.7%, DVD = 93.1%, ELE = 93.2%, KIT = 94.9%
R. Goebel and W. Wahlster, 2011 [40]	Deep Bi-directional Long Short-Term Memory Neural Networks (DBLSTM)	Sentiment Analysis of Social Data	SemEval 2016 and the second one is provided by Go dataset (1.6 million tweets)	85.86% accuracy was achieved on STS (Stanford Twitter Sentiment) corpus 84.82% on SemEval-2016.
K. Ravi and V. Ravi, 2016 [41]	Radial Basis Function Neural Network (RBFNN)	Sentiment classification on Hinglish text	300 news articles from viz. news and facebook.	The proposed approach performed better sensitivity than specificity in terms of news dataset. The proposed approach performed better specificity than sensitivity in terms of fb dataset.
L.-S. Chen, C.-H. Liu, and H.-J. Chiu, 2011 [42]	Back Propagation Neural Network (BPN)	Sentiment classification in the blogosphere	LiveJournal and Review Center have been used to collect reviews	Results concluded that proposed method enhance performance of classification and save training time as compared to traditional ML and IR.
M. Ghiassi, J. Skinner, and D. Zimbra, 2013 [43]	Dynamic Artificial Neural Network (DANN)	Twitter brand sentiment analysis (for justinbieber brand)	Total 10,345,184 tweets related to justinbieber brand	Results concluded that: More than 80% tweets dont have sentiment. Reduced feature set with characterize 97.3% of all messages in the 10,345,184 Justin Bieber Twitter corpus. Only six expressions were found related to Justin Bieber brand out of 181 and other were found twitter specific. Facilitates the Justin Bieber brand to identify the issues and views about the brand.

lexicon and feature reduction accordingly. The existing models SVM and nave Bayes were used for twitter specific lexicon to compete with the existing studies. The classification was done in proposed model by using artificial neural networks for twitter specific lexicon and this difference outperformed the existing models. The input matrix was parse matrix. Datasets were divided into training and test datasets to achieve better accuracy. The feature engineering phase was done through preprocessing activities for transforming the documents in simple form and to produce the vector presentation. The data was collected from Twitter API v1.0. The proposed model was applied on Justin Bieber Twitter corpus and it was established that the emotions contain high explanatory power as compared to existing studies. The proposed model has reduced the features of Justin Bieber corpus and enhanced the classification accuracy and high amount of coverage. Only six expressions were found related to Justin Bieber brand out of 181 and others were found twitter-specific. The proposed model has facilitated the Justin Bieber brand to identify the issues and views about the brand.

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 unstructured 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.

TABLE VII.   ANALYSIS OF ALL THE BEST STUDIES DISCUSSED IN THIS REVIEW

Researcher Name and Year	Model Used	Purpose	Data Set	Results
J.Islam and Y. Zhang 2016 [25]	Convolutional Neural Networks (CNN)	Visual SA	1269 images from twitter	GoogleNet gave almost 9% performance progress than AlexNet.
R. Silhavy, R. Senkerik, Z. K. Oplatkova, P. Silhavy, and Z. Prokopova, 2016 [34]	HBRNN (hierarchical bidirectional Recurrent Neural Network)	Sentiment Analysis of Customer Reviews	150,175 labelled reviews from 1500 hotels (DBS text mining Challenge 2015)	The experimental results explored that that HBRNN outperformed all other methods.
G. Zhou, Z. Zeng, J. X. Huang, and T. He, 2016 [38]	WSDNNs (Weakly Shared Deep Neural Networks)	Cross-Lingual Sentiment Classification	4 languages reviews from Amazon, each language consists of 1000 negative and 1000 positive reviews	Proposed approach is more effective and powerful than the previous studies by applying experiments on 18 tasks of cross lingual sentiment classification
P. Vateekul and T. Koomsubha , 2016 [8]	Convolutional Neural Network (DCNN) and Short Term Memory (LSTM)	Sentiment Analysis on Thai Twitter Data	3,813,173 tweets (33,349 negative tweets and 140,414 positive tweets)	Higher in accuracy than SVM and Nave Bayes lesser than Maximum Entropy Higher accuracies in original sentences than shuffled sentences
R. Ghosh, K. Ravi, and V. Ravi, 2016 [39]	Probabilistic Neural Network (PNN) and a two layered Restricted Boltzmann (RBM)	Better accuracy of sentiment classification	Pang and Lee and Blitzeret al (1000 negative and 1000 positive reviews)	The proposed model attains accuracies in following manner: MOV = 93.3%, BOO = 92.7%, DVD = 93.1%, ELE = 93.2%, KIT = 94.9%
Q. You, J. Luo, H. Jin, and J. Yang, 2015 [28]	Convolutional Neural Networks (CNN)	Textual-visual SA	Getty Images, 101	keywords Joint visual and textual model outperforms the early single fusions
W. Li and H. Chen, 2014 [31]	Recursive Neural Network (RNN)	Identifying Top Sellers In Underground Economy	Russian carding Forum	Deep learning techniques accomplish superior outcomes than shallow classifiers carding sellers have fewer ratings than malware sellers.
H. Yanagimoto, M. Shimada, and A. Yoshimura, 2013 [33]	Deep Neural Network (DNN)	Document Similarity Estimation	T& C News	The proposed method accomplished superior performance in terms of similarity estimation of articles according to polarity.
T. Mikolov, K. Chen, G. Corrado, and J. Dean, 2013 [17]	Deep Belief Networks Along with word vector	Political Detection in Korean articles	50,000 political articles	Results showed 81.8% accuracy by correctly predicting labels

IV. ANALYSIS

This review has described ample of studies related to sentiment analysis by using deep learning models as summarized in Table VII. After analyzing all these studies, it is established that by using deep learning methods, sentiment analysis can be accomplished in more efficient and accurate way. As the sentiment analysis is used to predict the views of users and deep learning models are all about the prediction or mimic of human mind, so the deep learning models provide more accuracy than shallow models. Deep learning networks are better than SVMs and normal neural networks because they have more hidden layers as compared to normal neural networks that have one or two hidden layers. Deep learning networks are capable to provide training in both supervised/unsupervised ways. Deep-learning networks carry out automatic feature extraction and

doesnt involve human intervention therefore it can save time because feature engineering is not needed. Sentiment Analysis comprises different kinds of problem statements. The capability of settling in the task variations by having little alterations in system itself includes a feather in strength of Deep Learning standard. This method also has some limitations as well, as compared to previous models such as SVM. It requires large data sets and is tremendously costly to train. These complex models can obtain weeks to train by using machines equipped with expensive GPUs.

REFERENCES

[1] A. V. Yeole, P. V. Chavan, and M. C. Nikose, Opinion mining for emotions determination, ICIIECS 2015 - 2015 IEEE Int. Conf. Innov. Information, Embed. Commun. Syst., 2015.

[2]   B. Heredia, T. M. Khoshgoftaar, J. Prusa, and M. Crawford, Cross-Domain Sentiment Analysis: An Empirical Investigation, 2016 IEEE 17th Int. Conf. Inf. Reuse Integr., pp. 160165, 2016.

[3]   F. Luo, C. Li, and Z. Cao, Affective-feature-based sentiment analysis using SVM classifier, 2016 IEEE 20th Int. Conf. Comput. Support. Coop. Work Des., pp. 276281, 2016.

[4]   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.

[5]   J. Singh, G. Singh, and R. Singh, A review of sentiment analysis techniques for opinionated web text, CSI Trans. ICT, 2016.

[6]   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.

[7]   M. Day and C. Lee, Deep Learning for Financial Sentiment Analysis on Finance News Providers, no. 1, pp. 11271134, 2016.

[8]   P. Vateekul and T. Koomsubha, A Study of Sentiment Analysis Using Deep Learning Techniques on Thai Twitter Data, 2016.

[9]   Y. Zhang, M. J. Er, N. Wang, M. Pratama, and R. Venkatesan, Sentiment Classification Using Comprehensive Attention Recurrent Models, pp. 15621569, 2016.

[10] S. Zhou, Q. Chen, and X. Wang, Active deep learning method for semi-supervised sentiment classification, Neurocomputing, vol. 120, pp. 536546, 2013.

[11] 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.

[12] 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.

[13] L. Arnold, S. Rebecchi, S. Chevallier, and H. Paugam-Moisy, An Introduction to Deep Learning, Esann, no. April, p. 12, 2011.

[14] 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.

[15] X. Ouyang, P. Zhou, C. H. Li, and L. Liu, Sentiment Analysis Using Convolutional Neural Network, Comput. Inf. Technol. Ubiquitous Comput. Commun. Dependable, Auton. Secur. Comput. Pervasive Intell. Comput. (CIT/IUCC/DASC/PICOM), 2015 IEEE Int. Conf., pp. 23592364, 2015.

[16] R. Socher and C. Lin, Parsing natural scenes and natural language with recursive neural networks, Proc. , pp. 129136, 2011.

[17] T. Mikolov, K. Chen, G. Corrado, and J. Dean, Efficient Estimation of Word Representations in Vector Space, Arxiv, no. 9, pp. 112, 2013.

[18] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, A Convolutional Neural Network for Modelling Sentences, Acl, pp. 655665, 2014.

[19] Y. Kim, Convolutional Neural Networks for Sentence Classification, Proc. 2014 Conf. Empir. Methods Nat. Lang. Process. (EMNLP 2014), pp. 17461751, 2014.

[20] T. Mikolov, K. Chen, G. Corrado, and J. Dean, Distributed Representations of Words and Phrases and their Compositionality, Nips, pp. 19, 2013.

[21] Z. Wu, T. Virtanen, T. Kinnunen, E. S. Chng, and H. Li, Exemplar-based unit selection for voice conversion utilizing temporal Information, Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH, pp. 30573061, 2013.

[22] K. S. Tai, R. Socher, and C. D. Manning, Improved semantic representations from tree-structured long short-term memory networks, Proc. ACL, pp. 15561566, 2015.

[23] A. A. Al Sallab, R. Baly, and H. Hajj, Deep Learning Models for Sentiment Analysis in Arabic, no. November. 2015.

[24] C. Z. Jiajun Zhang, Neural Networks in Machine Translation: An Overview, IEEE Intell. Syst., pp. 17241734, 2015.

[25] J. Islam and Y. Zhang, Visual Sentiment Analysis for Social Images Using Transfer Learning Approach, 2016 IEEE Int. Conf. Big Data Cloud Comput. (BDCloud), Soc. Comput. Netw. (SocialCom), Sustain. Comput. Commun., pp. 124130, 2016.

[26] A. Severyn and A. Moschitti, Twitter Sentiment Analysis with Deep Convolutional Neural Networks, Proc. 38th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. - SIGIR 15, pp. 959962, 2015.

[27] L. Yanmei and C. Yuda, Research on Chinese Micro-Blog Sentiment Analysis Based on Deep Learning, 2015 8th Int. Symp. Comput. Intell. Des., pp. 358361, 2015.

[28] Q. You, J. Luo, H. Jin, and J. Yang, Joint Visual-Textual Sentiment Analysis with Deep Neural Networks, Acm Mm, pp. 10711074, 2015.

[29] C. Li, B. Xu, G. Wu, S. He, G. Tian, and H. Hao, Recursive deep learning for sentiment analysis over social data, Proc. - 2014 IEEE/WIC/ACM Int. Jt. Conf. Web Intell. Intell. Agent Technol. - Work. WI IAT 2014, vol. 2, pp. 13881429, 2014.

[30] R. Socher, A. Perelygin, and J. Wu, Recursive deep models for semantic compositionality over a sentiment treebank, Proc. , pp. 16311642, 2013.

[31] W. Li and H. Chen, Identifying top sellers in underground economy using deep learning-based sentiment analysis, Proc. - 2014 IEEE Jt. Intell. Secur. Informatics Conf. JISIC 2014, pp. 6467, 2014.

[32] C. Baecchi, T. Uricchio, M. Bertini, and A. Del Bimbo, A multimodal feature learning approach for sentiment analysis of social network multimedia, Multimed. Tools Appl., vol. 75, no. 5, pp. 25072525, 2016.

[33] H. Yanagimoto, M. Shimada, and A. Yoshimura, Document similarity estimation for sentiment analysis using neural network, 2013 IEEE/ACIS 12th Int. Conf. Comput. Inf. Sci., pp. 105110, 2013.

[34] R. Silhavy, R. Senkerik, Z. K. Oplatkova, P. Silhavy, and Z. Prokopova, Artificial intelligence perspectives in intelligent systems: Proceedings of the 5th computer science on-line conference 2016 (CSOC2016), vol 1, Adv. Intell. Syst. Comput., vol. 464, pp. 249261, 2016.

[35] A. Hassan, M. R. Amin, A. Kalam, A. Azad, and N. Mohammed, Bangla Text (BRBT) using Deep Recurrent models.

[36] T. Chen, R. Xu, Y. He, Y. Xia, and X. Wang, Using a Sequence Model for Sentiment Analysis, no. August, pp. 3444, 2016.

[37] P. Ruangkanokmas, T. Achalakul, and K. Akkarajitsakul, Deep Belief Networks with Feature Selection for Sentiment Classification, Uksim.Info, pp. 16, 2016.

[38] G. Zhou, Z. Zeng, J. X. Huang, and T. He, Transfer Learning for Cross-Lingual Sentiment Classification with Weakly Shared Deep Neural Networks, Proc. 39th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. - SIGIR 16, pp. 245254, 2016.

[39] R. Ghosh, K. Ravi, and V. Ravi, A novel deep learning architecture for sentiment classification, 3rd IEEE Int. Conf. Recent Adv. Inf. Technol., pp. 511516, 2016.

[40] R. Goebel and W. Wahlster, Integrated Uncertainty in Knowledge Modelling and Decision Making, Proc. Int. Symp. Integr. Uncertain. Knowl. Model. Decis. Mak. (IUKM 2011), vol. 1, pp. 362373, 2011.

[41] A. Graves, N. Jaitly, and A. R. Mohamed, Hybrid speech recognition with Deep Bidirectional LSTM, 2013 IEEE Work. Autom. Speech Recognit. Understanding, ASRU 2013 - Proc., pp. 273278, 2013.

[42] C. N. dos Santos and M. Gatti, Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts, Coling-2014, pp. 6978, 2014.

[43] K. Ravi and V. Ravi, Sentiment classification of Hinglish text, 2016 3rd Int. Conf. Recent Adv. Inf. Technol. RAIT 2016, pp. 641645, 2016.