Vol. 06, No. 2 (2024) 853-862, doi: 10.24874/PES06.02A.012

Proceedings on Engineering Sciences

www.pesjournal.net

SENTIMENT ANALYSIS USING NOVEL DEEP LEARNING METHODS

Nitendra Kumar
Priyanka Agarwal
Sanjeev Bansal
Vinay Kumar Yadav
Dhrubajyoti Bhowmik[1]

Received 08.09.2023.
Received in revised form 01.03.2024.
Accepted 19.03.2024.
UDC – 004.85

Keywords:

Sentiment Analysis, Machine learning, Deep learning, Natural Language Processing, Deep Neural networks (DNN), Convolutional Neural Networks, Recurrent Neural Networks (RNN)

A B S T R A C T

In the current digital era, the humongous amount of data being generated has been impacting public lives in one or the other ways. Sentiment analysis, also known as opinion mining, is related to contextual mining of texts which helps in identification and extraction of subjective information from the source material. Sentiment analysis is being used for brand monitoring and reputation management across different market segments. It helps to understand how the public perceive a particular brand, product or service that is highly useful for different tech companies, marketing agencies, media organizations, fashion brands etc. In today's scenario we have been suffering with data overload which makes it impossible to analyze public sentiments without any sort of error or bias. Sentiment analysis provides better insights into the public reviews as it can be automated which ultimately helps in decision making. There are various deep learning and machine learning methods and models as well as natural language processing tools which help in examining and analyzing public opinions with low time complexity. However, deep learning methods have become highly popular in recent times as these models provide high efficiency and accuracy. In this review paper we have provided a complete overview of the common deep learning frameworks being employed for sentiment classification and analysis. This paper discusses various learning models, evaluation, text representations and other metrics in deep learning architectures. The key findings of different authors have been discussed in detail. This paper will help other researchers in understanding the deep learning techniques being used for sentiment analysis.

1. INTRODUCTION

The web of information has been growing at an unprecedented pace, implying that we have been living in the age of data. Interestingly, the data being generated has changed the lives of most of the people and the corporations around the world. Currently, tons of new data and related services sprout out every day leading to the creation of new research fields for example business intelligence, big data, data science, text mining, data mining etc.

There are different forms of communication media such as sign language, spoken words, written texts that

[1] Corresponding author: Dhrubajyoti Bhowmik
Email: babanbhowmik@gmail.com

communicate or perhaps give objective or subjective information. Impressions such as feelings, emotions, attitudes, evaluations, and opinions associated with any form of communication is termed as sentiment. Analysis of sentiments conveyed through different forms of communication is also referred to as opinion mining which deals with classification of sentiments related with any spoken, written or sign language. Through different methods and tools such as Natural Language Processing (NLP), deep learning, machine learning, artificial intelligence, statistics, linguistic features etc., sentiment analysis is carried out to evaluate whether the contents of the written texts, spoken words or sign language is favorable, unfavorable, or neutral. It can also be analyzed up to what degree the sentiments are favorable, unfavorable, or neutral. Sentiment Analysis has various applications such as information extraction through product reviews, marketing of products, in politics to know voter's opinion, business analytics, law and policy making, psychology, sociology (to find like-minded individuals or communities).

Political, corporate, and individual daily decision-making processes are significantly influenced by the opinions of the public. These decisions range from buying a product such as a smart TV to choosing the best school for kids to making investments or running an election from a particular constituency, all these decisions affect various aspects of our daily life. When the internet was not there, people would seek opinions on products and services through various sources such as relatives, friends, consumer reports, and through various magazines and newspapers. However, in the current internet era, it has become much easier to collect varied opinions from different social media platforms being posted from people around the world. People search review sites (Epinion.com, CNET etc.), online opinion sites (Yelp, Rotten Tomatoes, TripAdvisor etc.), e-commerce sites (Flipkart, Amazon etc.) and social media (Twitter, Facebook, Instagram etc.) to get feedback to know as to how a particular product or service is being perceived in the market. Likewise, different organizations use surveys, opinion polls, and social media as a mechanism to find feedback on their products and services.

The use of sentiment analysis is becoming more extensively leveraged because the information provided by the analysis could result in the monetization of products and services. For example, obtaining consumer feedback on a product can be helpful in building better products, which can help generate higher revenue. It will also help these organizations to compare competitor offerings.

1.1. Sentiment Classification Levels

The sentiment classification level can be categorized as:

a) Document level
b) Sentence level
c) Aspect/feature level

a. Document Level Classification

In document level classification, sentiment is extracted from the entire review, subsequently the entire opinion is classified based on the overall sentiment of the opinion holder. The main objective of this process is to classify a review as positive, negative, or neutral. This level of classification works best when the document is written by a single person to express an opinion or sentiment on a single entity.

b. Sentence Level Classification

The sentence level classification generally involves two steps:

- Subjectivity classification of a sentence into two classes such as objective and subjective.
- Subjective sentences are classified into two classes such as positive and negative.

Objective sentences give some information, on the other hand subjective sentences express personal views, feelings, beliefs, or emotions. Often, subjective sentences contain multiple opinions having subjective and factual clauses.

c. Aspect/Feature Level Classification

In this level of classification, the objective is to identify and extract object features based on the comments made by opinion holder, and subsequently it is determined if the opinion is positive, negative, or neutral. Synonyms of features are grouped together to produce feature-based summary of multiple reviews.

1.2. Converting Unstructured Text into Structured Opinions

Public sentiments are generally expressed in an unstructured format. Analysis of texts is carried out by collecting unstructured data and then converting the same into a more structured format that enables the analysis of the data and finding the sentiments associated with the data. Various techniques are being employed to convert unstructured data to a structured format.

For example, an opinion can be expressed as a quintuple (e_j, f_{jk}, SO_{ijkl}, h_i, t_l)

Where,

e_j is a target entity.
f_{jk} is a feature of the entity e_j.
SO_{ijkl} is the sentiment value of the opinion holder h_i on feature f_{jk} of entity e_j at time t_l. SO_{ijkl} is positive, negative, or neutral.

h_i is an opinion holder.
t_l is the time when the opinion is expressed.

An entity e is a product, person, event, organization, or topic that is represented as a hierarchy of components, or sub-components or so on. Each node is generally represented as a component which is associated with a set of attributes for the component.

The main objective of the quintuple is to transform unstructured data to more structured data enabling further analysis of the same. The first step is to determine all quintuples and then find attributes which are required by the quintuplet. Sentiment analysis becomes much easier once the data is transformed into a more structured form. Subsequently, the extracted quintuples are fed to visualization and analysis tools.

1.3. Aspect Based Sentimental Analysis (ABSA)

Aspect based sentiment analysis (ABSA) relies on identification of aspects target entities and subsequently estimating the sentimental polarity for every mentioned aspect. The process is carried out in two steps, the first being aspect extraction and finally aspect sentiment classification. Aspect extraction refers to recognizing aspects of the entity, it usually involves an information extraction task. The classification enables us to determine if the opinions on different aspects are positive, negative, or neutral.

a. Aspect Extraction

Aspect extraction includes one of the basic tasks of sentiment analysis, which is defined as the process of identification and extraction. The aspects having explicit or implicit nature are found from the opiniated texts posted by the public on different social media platforms. This step enables the extraction of instances of modifiers and product aspects, which then describes the opinion of a particular aspect. Based on syntactic dependency paths, the dependency parser tree present in Python’s spacy package is used to extract the pair of words. A group of nouns, verbs, adverbs, adjectives, and pairs as output are produced at this step to be followed by the next step. The noun is used to identify people, places, or things while adjectives are used to identify an attribute of a noun. Verb indicates an action, occurrence, or a state of being while on the other hand an adverb modifies an adjective, verb, preposition, sentence, or clause. Noun aspect extraction can differentiate the subject and the competitor enabling appropriate sentiment classification. In the case of review texts, the syntactic grammatical dependency relation present between words is returned by the dependency parsing using Stanford Parser. The identification of sentiment word, noun phrase aspect, and the aspect sentiment word pairs are exploited using proper dependency relation.

2. BACKGROUND

Proliferation of multitude of internet websites has made it difficult to find and monitor better pinion site on the web which can help in inferring the information contained in them. Various related websites generally contain a large volume of opinion texts which make it difficult to decipher the sentiments in these long blogs and forum postings. The readers have difficulty in identifying the related websites thereby making it difficult to extract and analyse the opinions in them. Sentiment analysis based on machine learning and artificial intelligence systems are thus needed. There are generally two types of methods are being employed such as supervised and unsupervised methods. Supervised machine learning methods include, Support Vector Machines (SVM), Maximum Entropy, Naïve Bayes etc. In unsupervised machine learning methods, sentiment lexicons, syntactic patterns, and grammatical analysis are being employed for sentiment analysis.

This review paper will discuss about the various deep learning and machine learning techniques and subsequently, provides a comprehensive survey of the sentiment analysis based on machine learning methods.

2.1 Deep Learning

Deep learning uses a cascade of multilayer approach to the hidden layers of the neural network for feature extraction and transformation. Traditional machine learning methods employ manual or feature selection methods for defining and extracting the features. On the other hand, deep learning models help in automatic feature learning and extraction, providing higher accuracy and performance. Figure 1 shows the differences between machine learning and deep learning classification approaches of sentiment polarity. Currently, artificial neural networks and deep learning techniques are being employed in the field of image and speech recognition as well as in natural language processing providing higher accuracies.

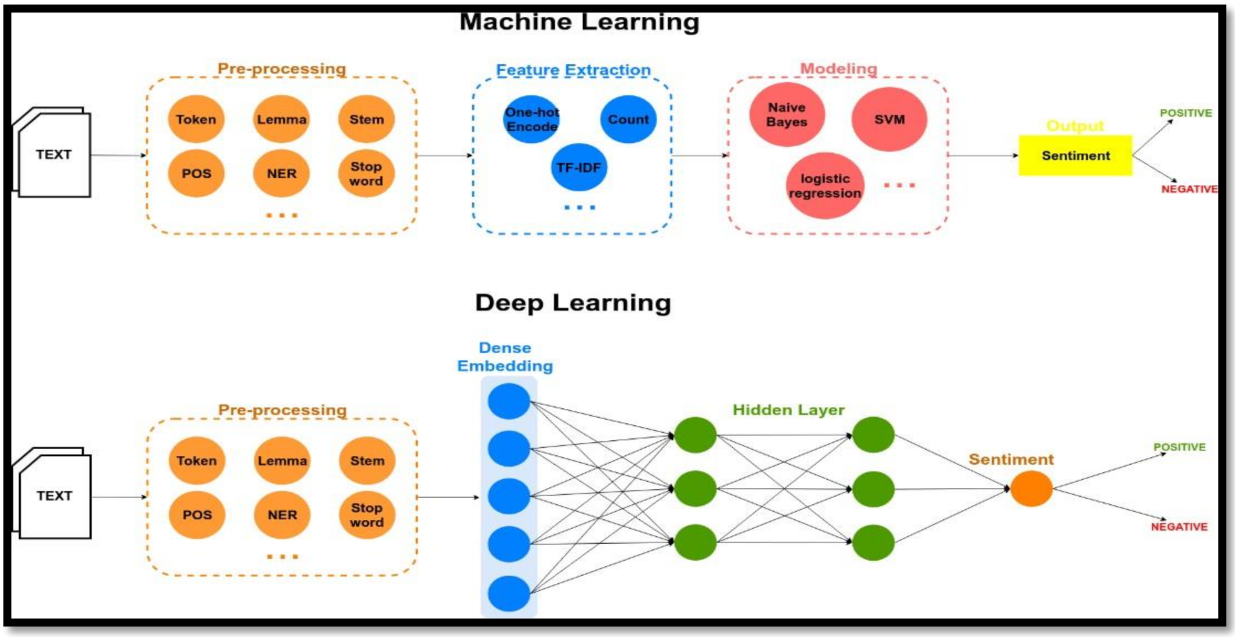

Figure 1. Differences between machine learning (top) and deep learning (bottom) classification approaches of sentiment polarity

2.1.1. Deep Neural Networks (DNN)

A deep neural network is a neural network having more than two layers, few of them are hidden layers as it is depicted in figure 2. Deep neural networks make use of highly sophisticated mathematical modelling to process data in various ways. Data is filtered through a cascade of several layers wherein each successive layer uses the output from the previous layer to infer its results. Large availability of datasets for data processing by deep neural network models give higher accuracy, as these models learn from previous results to refine their ability to make correlations and connections.

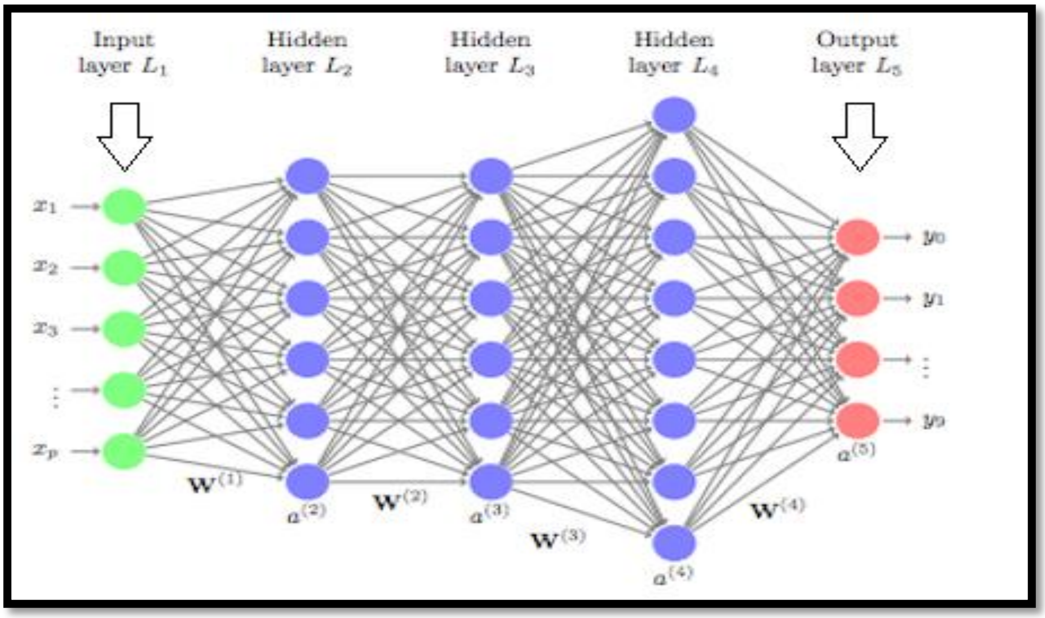

Figure 2. An architecture of deep neural network (DNN)

2.1.2. Convolutional Neural Networks

A convolutional neural network (CNN) is a type of feed-forward neural network, generally used to analyse visual images by employing data processing with grid-like topology. An CNN architecture consists of convolutional and pooling layers wherein they provide inputs to a fully connected classification layer. A CNN has multiple hidden layers where they perform extraction of information from any available image. The convolutional layer consists of various filters which perform convolution operation. Here, every image is considered as a matrix of pixel values. After the extraction of feature maps, they are sent to the rectified linear unit (ReLU) layer, where element wise operation is performed, and all the negative pixels are assigned 0 values. The output of ReLU layer is known as rectified feature map, which is then sent through a pooling layer that results in a pooled feature map. The pooling layer employs several filters to identify different parts of an image such as corners, edges, body, feathers, eyes etc as shown in figure 3.

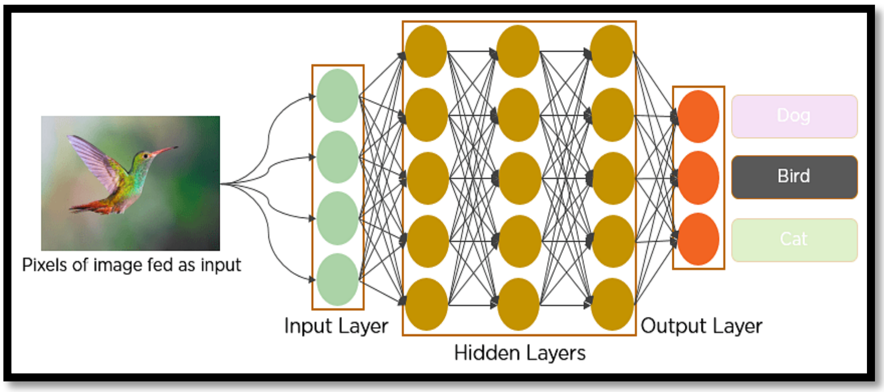

Figure 3. An architecture of CNN to identify the image of a bird.

2.1.3. Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNN) are as one of the basic and most powerful neural networks being used for various applications, thereby gaining immense popularity. RNN can process sequential data with more accuracy and higher efficiency. RNN is different from traditional neural networks as it can have internal memory which essentially enables it to remember important and essential information about an input. This is one of the most crucial aspects of RNN for predicting outcomes more precisely.

Generally, in a traditional neural network the input data is processed and sent to the next level without considering any sequence, while in RNN the sequential data is processed in a particular order that is essential to understand them clearly and precisely.

A Recurrent Neural Network architecture consist of an input layer, a hidden layer, and an output layer wherein all these layers work in a standard sequence. The data fetching responsibility is carried out by the input layer. The input layer performs the data pre-processing, the filtered data is passed onto the hidden layer. The hidden layer consists of neural networks, activation functions, and algorithms which enable retrieving useful information out of the given data. Subsequently, this information is passed to the output layer which provides the expected outcome.

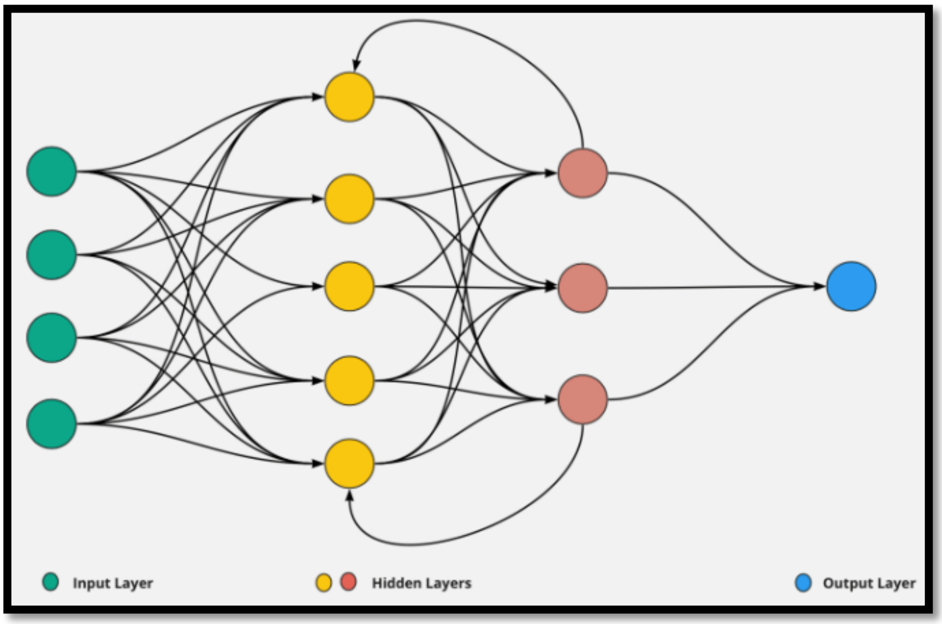

Figure 1. An architecture of RNN

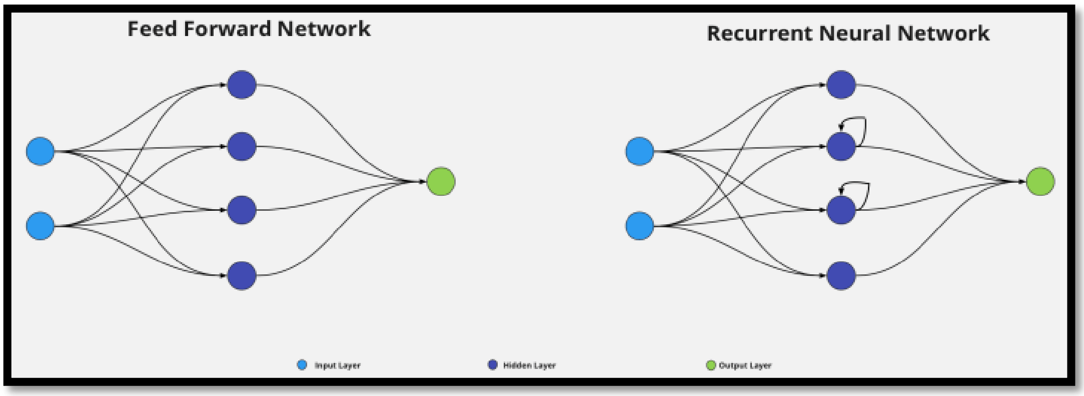

Figure 2. Neural network comparison

Figure 5 illustrates the difference between a recurrent neural network and a feed forward neural network. In a feed-forward neural network, the information after data processing moves only in one direction such as from the input layer to the hidden layer and finally to the output layer. One of the drawbacks of feed-forward neural network is that it is unable to remember the past except the training procedures it has gone through. On the other hand, in case of recurrent neural network, the passed information goes through a loop. Each input is dependent on the previous one for making decisions. In recurrent neural network, each layer is assigned the same and equal weight and bias. Hence, all the independent variables are converted to dependent variables. The loops present in RNN architecture ensures the information is stored and preserved in its memory.

3. RELATED WORK

Basiri et al. (2021) have proposed a advanced method by combining four deep learning and a classical supervised machine learning model (CNN, BiGRU, FastText, NBSVM, DistilBERT) for sentiment analysis of general public on coronavirus related tweets from eight different countries. They collected the tweets between 24th January 2020 to 21st April 2020 from these eight countries as well as from Google Trends users by employing coronavirus related keywords between the same period. The performance of the proposed model has been assessed by evaluating the five base models and proposed model by using Standard Sentiment 140 Twitter dataset with 1600,000 labelled tweets. The proposed model performed more accurately than the base deep classifiers which has shown 85.8% accuracy and 85.8% F1-score. The main observations of this study are firstly, there is correlation between rise in information about coronavirus and the reported first infected case. Secondly, different countries have shown unique sentiment patterns, and finally negative sentiment values are maximum whenever there is rise in the number of COVID-19 cases or increasing number of deaths.

Malla et al. (2021) have proposed a Majority Voting technique-based Ensemble Deep Learning (MVEDL) model for sentiment analysis related to COVID-19 tweets. They have used combination of pre-trained transformers such as RoBERTa, BERTweet, and CT-BERT. A dataset of 226,668 COVID-19 tweets have been collected by them between December 2019 to May 2020. The model proposed by them has shown 91.75% accuracy and 91.14% F1-score thereby outperforming the traditional machine learning and deep learning methods. They have used the latest TextBlob algorithm for the sentiment analysis of COVID-19 tweets. The researchers have shown that the proposed model is useful during live tweet sentiment analysis and classification, in understanding the depression status of patients as well as COVID-19 outbreak statistics.

Chakraborty et al. (2020) have proposed a fuzzy rule-based model guided by a set of seven disjunctive fuzzy rules for sentiment analysis of COVID-19 tweets. They have analysed two types of tweets firstly, 3000 re-tweets gathered between 1st Jan 2019 to 23rd March 2020 wherein the result shows that the maximum number of tweets reflect neutral or negative sentiments. On the other hand, the other dataset contains 226,668 tweets gathered between December 2019 to May 2020 wherein the analysis result has revealed that the maximum number of tweets have positive or neutral sentiments associated with them. The accuracy of the proposed model has been validated by using Doc2Vec model. Three algorithms related to Doc2Vec such as Distributed Bag-of-words, Distributed Memory Concatenated, and Distributed Memory Mean have been implemented further to check the accuracy. The proposed model has shown 81% accuracy. They have also proposed the implementation of a Gaussian membership function based fuzzy rule model for sentiments analysis of tweets which has shown 79% accuracy.

Ren et al. (2021) have proposed a multi-label personality detection model based on pre-trained BERT model along with Convolutional Neural Network and Recurrent Neural Network that combine the semantic and emotional features in the text. The researchers have used SenitcNet5 dictionary for feature extraction. The researchers have used two datasets, the first one being MBTI dataset which is one of the largest publicly published personality datasets. This dataset is divided into four different dimensions that is collected through Personality Café forum. This dataset contains 50 tweet texts from 8675 volunteers and their personality labels. This gives 422,845 labelled points. On the other hand, the second one is Big Five dataset. The big five dataset contains the extent of 2468 articles along with author tagged Big Five personality dimensions such as EXT, NEU, AGR, CON, and OPN. They have achieved better results by using BERT+CNN model. The proposed model has performed better compared to BERT model obtaining an average accuracy improvement of 6.91% and 6.04% on the given two datasets. There are seven machine learning algorithms such as Logistic Regression (LR), Decision, Tree), Random Forest (RF), Naïve Bayes, Gradient-Boosted Tree, Support Vector Machine, and the Neural Network Multilinear Perceptron. The research shows that the performance of the personality trait detection model that is based on semantic representation and lexical statistics is outperforms traditional language feature based and neural network methods.

Alwehaibi et al. (2021) have proposed a robust deep learning framework for Arabic short text sentiment analysis. They have used a dataset of total 15000 tweets that consist of MSA and dialectal Arabic tweets, subsequently the dataset is used for training and testing. They have used a combined RNN-LSTM framework for word level classification, wherein the feature has been extracted by employing FastText pre-trained word vectors. Convolutional Neural Network (CNN) has been used for character-level classification, wherein the features have been extracted from a 136-character alphabet set. They have used an ensemble neural model of CNN and LSTM to train by combining embedding features at the word and character levels. The result shows that the ensemble neural model of CNN and LSTM for word embedding anc character embedding has shown the highest accuracies of 96.7% by using 10-fold cross validation. CNN model is helpful in detecting the important features from a given text while LSTM is helpful in learning the sequential data which ultimately helps the ensemble model to perform better than the traditional deep learning models.

Salur et al. (2020) have proposed a novel hybrid (CNN + BiLSTM) deep learning model that deliberately combines different word embeddings such as Word2Vec, FastText, character level embedding by employing LSTM, GRU, BiLSTM, and CNN deep learning models. A dataset of 17,289 Turkish tweets have been collected from shared user tweets about a GSM operator in Turkey between 2011 and 2017. A total of 13,832 tweets have been used for training models while 3,457 tweets have been used for validation and testing of model. Figure 6 depicts the architecture of the proposed hybrid deep learning model proposed by the authors.

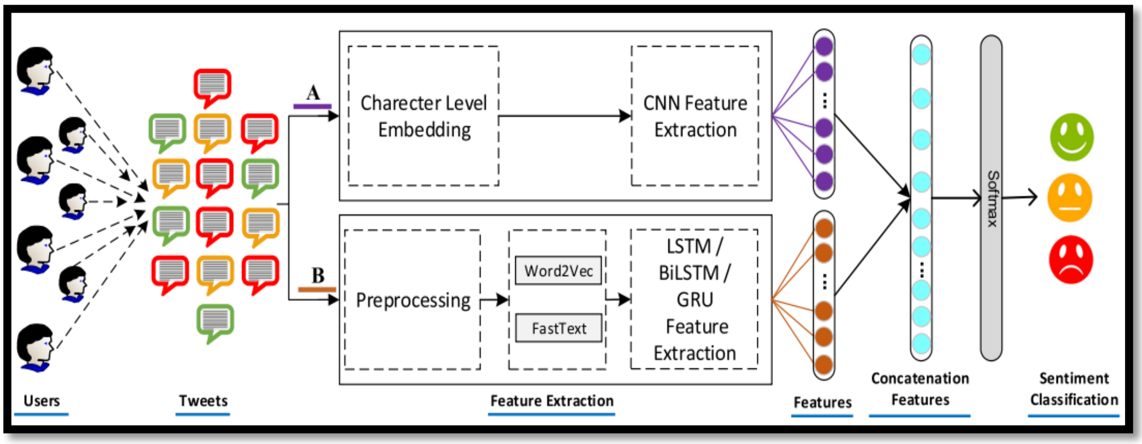

Figure 6. Architecture of the proposed hybrid deep learning model

They have achieved 82.14% classification accuracy by using the proposed hybrid model (CNN + BiLSTM). On the other hand, they have achieved 80.44% classification accuracy with BiLSTM through word embedding approach and 75.67% classification accuracy by using CNN through character embedding.

Saeed et al. (2019) have proposed a stacking ensemble classifier for spam detection in Arabian opinion texts. They have shown that the stacking ensemble classifier performs 12% better than other methods for DOCS dataset. They have used two different publicly available opinion datasets, the first one being Deceptive Opinion Spam Corpus containing 1600 opinion reviews in English of 20 highly popular hotels in Chicago, that is translated to Arabic language through a translator. The second dataset consists of Hotel Arabic Reviews Dataset (HARD) having 94,052 opinion reviews in Arabic language of 1858 hotels. The stacking ensemble classifier has been compared with related works which shows that it gives 28% more accuracy in detecting Arabic spam reviews.

Abdelgwad et al. (2021) have proposed a deep learning-based method using two models on Gated Recurrent Units (GRU) neural networks for Arabic aspect-based sentiment analysis. The first deep learning (BGRU-CNN-CRF) method uses word and character embeddings by combining bidirectional GRU, Convolutional Neural Network and Conditional Random Field models to extract the main opiniated aspects. On the other hand, the second models are interactive attention network based on bidirectional GRU (IAN-BGRU) to identify sentiment polarity of extracted aspects. The data consists of Arabic hotel reviews dataset containing 24,028 annotated aspect-based sentiment analysis tuples that has been divided into 19,226 tuples for training and 4,802 tuples for training. The results indicate that the models proposed by the authors have performed better than tradition deep learning methods. The models have achieved a 39.7% increase in F1-score for opinion target extraction and 7.58% in accuracy for aspect-based sentiment classification. For another task the model achieves 70.67% F1-score and an accuracy of 83.98%.

Brahimi et al. (2021) have proposed three deep learning-based models for sentiment classification of Arabic movie reviews. The first method uses n-grams model along with skip n-grams, the second approach employs Part of Speech (POS) tagging, while the third one is aimed at extracting review summaries and opinion conclusions of different Arabic movie reviews. Two types of datasets have been used, the first being OCA that is available freely and the second one being ARMD. The proposed models have shown more accuracy in terms of opinion classification. With OCA dataset, n-gram and POS tagging approaches combined with feature reduction (96% F1-score) can give best classification results. On the other hand, ARMD dataset provides 87.47% F1-score by integrating POS tagging and review summary approaches in the classification model.

Alassaf et al. (2021) have employed one-way ANOVA for improving sentiment analysis of Arabic tweets. The dataset consists of a real-world Twitter dataset which contains 8144 tweets related to Qassim University of Saudi Arabia. The ANOVA has been used as a feature selection method to significantly reduce the number of features while analysing opinions expressed through Arabic tweets. The result shows that Support Vector Machine (SVM) and Naïve Bayes (NB) have been able to achieve the best results with one-way ANOVA compared to baseline experimental results. The authors have proposed that one-way ANOVA combined with Support Vector Machine represents an excellent combination as it performs better than other reported methods. Figure 7 shows the ANOVA architecture for sentiment analysis of Arabic tweets.

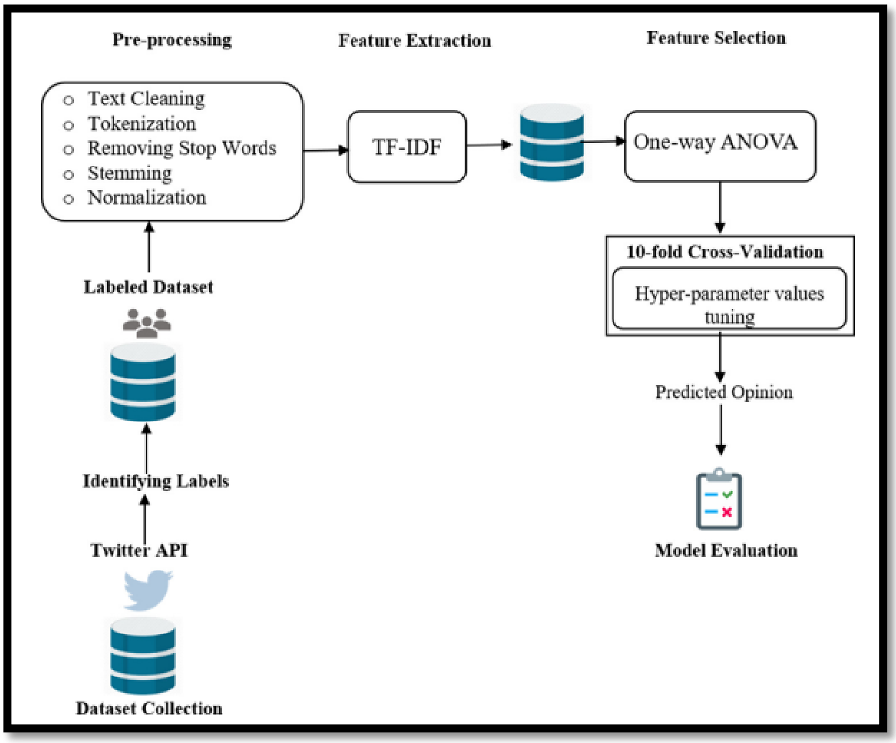

Figure 7. ANOVA framework for Arabic tweet sentiment analysis

Al-Dabet et al. (2021) have proposed two deep learning models for aspect-based sentiment analysis for hotel reviews in Arabic language. The proposed models have been given two different tasks such as aspect-category identification and aspect-sentiment classification. The first model consists of Convolutional Neural Network (CNN) combined with stacked independent long-short term memory (LSTM) which performs the task of aspect-category identification. Contrary to this, the second model consists of stacked bidirectional independent long-short term memory (BiLSTM) combined with a position-weighing mechanism and multiple attention mechanism layers. The result shows that the proposed models perform better than the traditional deep learning models wherein the C-IndyLSTM model has achieved 58.08% F1-score, and the second model has achieved an accuracy of 87.31%. Figures 8 and 9 show the proposed C-IndyLSTM and MBRA models respectively.

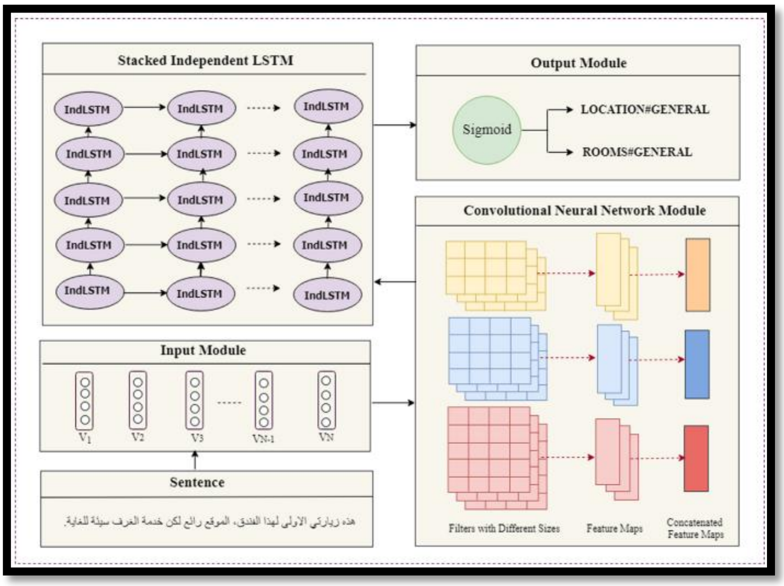

Figure 8. Architecture of C-IndyLSTM model

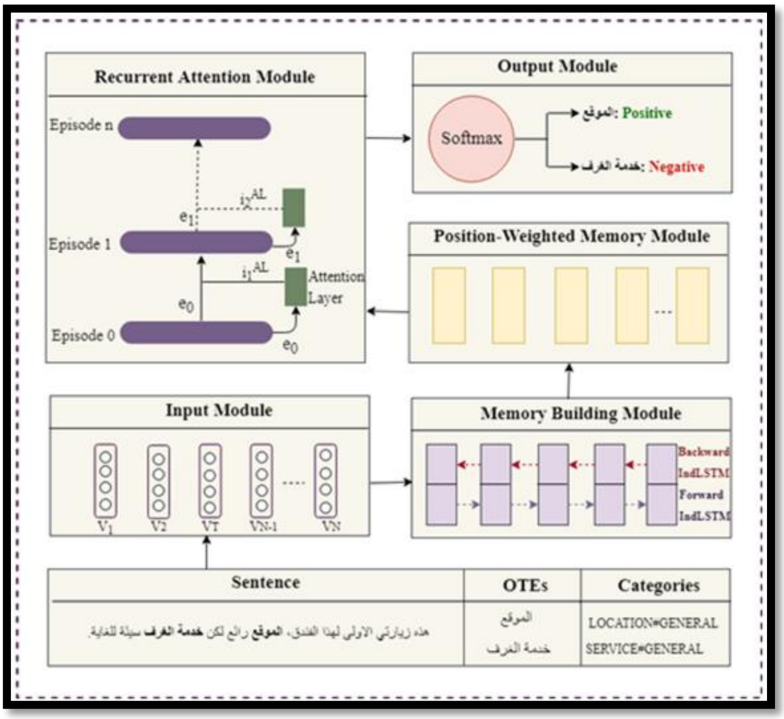

Figure 9. Architecture of MBRA model

4. CONCLUSION

This review article systematically discusses different deep learning tools and techniques being employed for sentiment analysis. Sentiment analysis of textual data by using deep learning techniques for various categories has been discussed in detail. Data generated on different platforms of social media, text-based sentiment analysis is very important which plays a vital role in any organization's branding, product, and service upgradation. We can conclude that deep learning techniques have been able to provide higher accuracies than other traditional sentiment analysis methods. Coherence and semantic in sentence can be handled by knowledge-based approach but produces lower accuracy as well as has time and space complexity. Though machine learning based supervised tools are faster and more accurate, they fail to handle negation, intensifier as well as modifier clause in each sentence. On the other hand, unsupervised knowledge-based approach and deep learning techniques provide better results in comparison to other methods.

References:

Abdelgwad, M. A., Soliman, T. H. A., Taloba, I., & Farghaly, M. F. (2022). Arabic aspect-based sentiment analysis using bidirectional GRU based models. *Journal of King Saud University - Computer and Information Sciences, 34*(9), 6652-6662. https://doi.org/10.1016/j.jksuci.2021.08.030

Alassaf, M., & Qamar, A. M. (2022). Improving sentiment analysis of Arabic tweets by one-way ANOVA. *Journal of King Saud University - Computer and Information Sciences, 34*(6, Part A), 2849-2859. https://doi.org/10.1016/j.jksuci.2020.10.023

Al-Dabet, S., Tedmori, S., & AL-Smadi, M. (2021). Enhancing Arabic aspect-based sentiment analysis using deep learning models. *Computer Speech & Language, 69*, 101224. https://doi.org/10.1016/j.csl.2021.101224

Alwehaibi, A., Bikdash, M., Albogmi, M., & Roy, K. (2022). A study of the performance of embedding methods for Arabic short-text sentiment analysis using deep learning approaches. *Journal of King Saud University - Computer and Information Sciences, 34*(8, Part B), 6140-6149. https://doi.org/10.1016/j.jksuci.2021.07.011

Basiri, M. E., Nemati, S., Abdar, M., Asadi, S., & Acharrya, U. R. (2021). A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets. *Knowledge-Based Systems, 228*, 107242. https://doi.org/10.1016/j.knosys.2021.107242

Brahimi, B., Touahria, M., & Tari, A. (2021). Improving sentiment analysis in Arabic: A combined approach. *Journal of King Saud University - Computer and Information Sciences, 33*(10), 1242-1250. https://doi.org/10.1016/j.jksuci.2019.07.011

Chakraborty, K., Bhatia, S., Bhattacharyya, S., Platos, J., Bag, R., & Hassanien, A. E. (2020). Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers—A study to show how popularity is affecting accuracy in social media. *Applied Soft Computing, 97*, 106754. https://doi.org/10.1016/j.asoc.2020.106754

Malla, S., & P J A, A. (2021). COVID-19 outbreak: An ensemble pre-trained deep learning model for detecting informative tweets. *Applied Soft Computing, 107*, 107495. https://doi.org/10.1016/j.asoc.2021.107495

Ren, Z., Shen, Q., Diao, X., & Xu, H. (2021). A sentiment-aware deep learning approach for personality detection from text. *Information Processing & Management, 58*(3), 102532. https://doi.org/10.1016/j.ipm.2021.102532

Saeed, R. M. K., Rady, S., & Gharib, T. F. (2022). An ensemble approach for spam detection in Arabic opinion texts. *Journal of King Saud University - Computer and Information Sciences, 34*(1), 1407-1416. https://doi.org/10.1016/j.jksuci.2019.10.002

Salur, M. U., & Aydin, I. (2020). A novel hybrid deep learning model for sentiment classification. *IEEE Access, 8*, 58080–58093. https://doi.org/10.1109/ACCESS.2020.2982538

Nitendra Kumar
Amity Business School, Amity University, Noida,
Uttar Pradesh – 201303,
India
nkshukla.kumar4@gmail.com
ORCID 0000-0001-7834-7926

Priyanka Agarwal
Amity Business School, Amity University, Noida,
Uttar Pradesh – 201303,
India
pagarwal5@amity.edu
ORCID 0000-0002-2943-2507

Sanjeev Bansal
Amity Business School, Amity University, Noida,
Uttar Pradesh – 201303,
India
nkshukla.kumar4@gmail.com
ORCID 0000-0001-6629-1451

Vinay Kumar Yadav
Glocal School of Science and Technology, Delhi-Yamunotri Marg (State Highway 57), Mirzapur Pole,
Saharanpur, U.P., 247121,
India
vint1983@gmail.com
ORCID 0009-0009-9596-0939

Dhrubajyoti Bhowmik
Department of Computer Science and Engineering, National Institute of Technology, Agartala,
India
babanbhowmik@gmail.com
ORCID 0000-0003-3051-8137