International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Social Media Trustable Reviews Based Sales Forecasting Using Polarity Parameters and Pattern Via Deep Learning Models

M. Priya Alagu Dharshini[1], Dr. S. P. Victor[2]

Submitted: 18/01/2024 **Revised**: 27/02/2024 **Accepted**: 04/03/2024

Abstract: Nowadays, a lot of people use product reviews when making decisions. Nevertheless, reviewers manipulate the method by publishing fictitious reviews to elevate or denigrate the products because they do it for financial gain. Fake review detection has received a greater attention in the past decade from academic and industrial communities alike. Sentiment analysis recognizes and separates the opinion from the provided review; the analysis procedure involves text analytics, natural language processing (NLP), computational linguistics, and classifying the opinion's polarity. The lack of labeling materials for supervised learning and evaluation, however, means that the problem will continue to be difficult to solve. The problem has been approached from both the reviewer's and the reviewer's perspectives in numerous current works. For fake review prediction, we employed GloVe embedding along with BERT. The sentiment polarity for real reviews are then determined using Co-Sensitive Weighted Fusion GCN. It has a positive, negative, or neutral outcome. As a result, we are using the previous quadrants to forecast the current quarter sales possibility. Sales forecasting uses a variety of parameters, including subjectivity, tweet rate, polarity, mean, standard deviation, variance, skew, and kurtosis score of each polarity. Along with this parameter this approach introduces a temporal based polarity pattern to increase the efficiency of sales forecasting with the help of BiLSTM based regression model. The proposed method for fake news prediction HE-CNN_TB yields an F1_score of 90.85% for GossipCop dataset and 93.58% for the PolitiFact dataset. The suggested technique CoSWFGCN obtains an F1_score of 90% for the Amazon dataset and 76.19% for the Twitter_15 dataset in terms of sentiment polarity. The suggested BiLSTM-Pattern prediction model yields an average 0f 0.6 R2 value for the forecasting of the laptop brands such as Asus, HP, Dell and Lenovo.

Keywords: *Natural Language Processing, Supervised Learning, Text Analytics, Computational Linguistics, Graph Neural Networks, Graph Convolutional Network.*

1. Introduction

The use of social media as a means for content exchange and social networking has increased in importance and popularity in the past couple of decades. The material created by the aforementioned websites is still available, though. We illustrate whether actual research results are able to be predicted using social media content. Specifically, we ascertain whether the news is authentic or not. We also used sentiment analysis with deep learning architecture, such as Graph Neural Networks, to ascertain whether the assessment is neutral, negative, or positive if the news is trustworthy. We show how the prediction power of social media for the specific product may be increased by utilizing sentiments collected from data. incorporating the applicable criteria that follow.

Fake news and internet misinformation are becoming a much bigger concern as more people use social media. Fake news is made up information intended to mislead readers. It has features similar to news media content but is different in organizational structure or goal [1]. It is challenging to rectify reliable news sources because fake news is continually expanding on social media, in online blogs, journals, forums, and newspapers. Social media has increased into a prime environment for the creation, manipulation, and dissemination of false information. Because it's so simple to create and spread information. Brexit, the Covid-19 outbreak, and the US presidential election are a few recent samples of the substantial influence fake news has had on important events [2]. Many techniques make advantage of the social context of the news to confirm its accuracy. These methodologies either examine the patterns of distribution of authentic and fraudulent news on social media and users' varied responses to the two classifications, or they leverage components like news writer, the publication's origin, or the dependability of its recurring users. By examining its content, other academics search for indicators of bogus news. Their creations can be divided into two separate groups: the first looks into fake news using knowledge-based techniques; in reality, these pieces analyze what is happening as a number of logical premises and try to determine whether or not these propositions fit with understanding that has come before. Using techniques from natural language processing, the

*[1] Research Scholar, Manonmaniam Sundaranar University,
Tirunelveli, Tamil Nadu, India.
preethi.murugiah10@yahoo.com
Scopus ID: 57541144300
[2] Associate Professor, Department of Computer Science & Engineering,
St. Xavier's College (Autonomous),
(Affiliated to Manonmaniam Sundaranar University)
Tirunelveli, Tamil Nadu, India.
drspvictor@gmail.com
Scopus ID : 55406344900*

second group of researchers tackles the problem as text categorization. The validity of the news was initially assessed using classification-based methods that attempted to manually extract a variety of features from the news text and combine them with traditional learning techniques. It has always been difficult, time-consuming, and energy-consuming for people to extract information from text as a complex input. There are numerous methods for resolving this issue, one of which is to use deep neural networks to determine fake news based on the style of its language. Transfer learning with transformers has recently become popular deep neural network architectures for NLP. BERT is a highly developed pre-trained word embedding model which employs transformer encoded architecture. One of the most intriguing transformers, BERT scores better than other models in numerous NLP benchmarks. Scholars have looked into a number of remedies for the problem of false news. In NLP, the BERT network, which we employ in this research work, has produced some of the greatest outcomes.

Customers' opinions of a business sector have a big impact on how well it is promoted. Many studies solved this problem by categorizing the sentiments in a more compositionally sound way. The research is still in progress due to excess content, grammatical errors, and sparse use of emotive phrases. The quick growth of e-commerce platforms has made it possible to acquire a variety of things in different designs right from your own home. Customers had numerous annoyances as a result of virtual environment purchases, including lower product quality, flaws, and minor product problems. Customers must be drawn to the business, if it is a physical store or an online marketplace. This industry is built on word-of-mouth and written reviews. Different stakeholders are going to compose reviews in different formats. The system is quite difficult to process due to the huge product reviews. Making a decision about a product may be immensely assisted by analyzing the sentiments and drawing conclusions about the product review based on these. This work makes use of sentiment analysis, a well-known text mining technique. A person searches for reviews before purchasing burger from a store. Burger was charred, the meal was of terrible quality, and the food delivery was 40 minutes late, to name a few sample reviews. The customer will be hesitant to get a Burger from that establishment in this situation. The proprietors might simultaneously work to resolve this problem in an effort to win back the confidence of the clientele. The problem may be resolved by management conducting an evaluation, Is the oven worth substituting or restoring, providing staff training, etc. Hence, the Co-Sensitive Weighted Fusion Graph Convolution Network (CoSWFGCN) is a novel framework proposed in this paper that is depend on the Graph Convolutional Network (GCN). This work uses the syntactic and co-sensitive specific semantic GCN (CS3GCN) features to provide the words' semantics and syntax additional weight for graph learning.

Given that social media can be perceived as a type of collective intelligence, we decided to look into its ability to forecast real-world occurrences. Interestingly, we discovered that online chatter can be used to produce quantitative forecasts that are more accurate than those produced by artificial markets. These data markets, if wide enough and properly constructed, usually facilitate the exchange of state-contingent commodities and are, on average, more accurate than alternative methods for gathering scattered data, such as opinion polls and surveys. In particular, it has been shown that prices in these markets are good indicators of the future because they significantly correspond with actual outcome frequency. When it pertains to social media, the volume and great diversity of data that circulates among big user bases offer a chance to transform the information into a format that enables accurate forecasts of particular results without the need to enact market mechanisms. It is also possible to build models that aggregate the views of the whole population, extract pertinent data regarding their actions, and predict future patterns. Furthermore, knowing how customers talk about specific products can be useful when creating marketing and advertising campaigns. Thus, we forecast a product's future using the data that has already been analyzed.

2. Related Work

Since fake news is purposefully designed to present false data, it can be difficult to identify. For identifying fake news, numerous methods based on machine learning have been developed. These techniques can be categories as: conventional approaches and deep-learning approaches. The conventional machine-learning techniques, including Decision Trees (DT), MNB, LR, LSVM, and extreme gradient boosting (XGBoost). The term frequency-inverted document frequency (TF-IDF) was employed with n-gram analysis to take features out for identifying false news [14]. Six distinct machine-learning algorithms were researched and compared. Similar to this, the same embedding technique is used to examine 5 various machine-learning models [15]. The most effective methods were LSVM and XGBoost. To extract features from texts, the term frequency (TF) weighting method and the document-term matrix were used [16]. The salp swarm optimization (SSO) and grey wolf optimizer (GWO) algorithms were adopted for the false news detection rather than machine learning techniques [17]. The writing style underneath using part-of-speech (POS) tags to identify fake news was also executed [18]. To generate the first model, the POS features are given into the XGBoost. The final prediction is obtained by ensembleing the multi-layer perception model with the first model and feeding it an average of TF-IDF weights and Word2Vec word embeddings. GloVe and BERT embeddings utilize 2 deep learning methods, CNN and

LSTM [21]. The fact that fake news is always changing makes it difficult for these techniques to keep up. When studying the identification of fake news using different classification techniques, the earlier theories [3] are helpful in providing direction. Approaches based on news content [1, 4, 5, 6] deal with the many writing styles used in published news stories. Our primary goal in using these techniques is to extract various informational and writing-style-related elements from phony news articles. In order to detect fraudulent articles by capturing the manipulators' writing style through linguistic elements, style-based approaches [7, 8, 9] are beneficial in these learning. As a result, it is challenging to identify false news with greater accuracy using solely news content-based features [4, 10, 11]. Social context-based techniques [4, 12, 13, 5, 6] deal with the latent information that exists among the reader and the news piece. Social interactions might be a useful feature for spotting fake news. In these methods, instance-based strategies [5] focus on how users interact with any social media post to encourage the integrity of original news articles. Furthermore, by disseminating credibility values among users, posts, and news, propagation-based approaches [5] consider the relationships between significant social media posts to guide the learning of validity scores. To combine news content with context-level information, the most popular and successful methods currently in use [4, 13, 5] employ unidirectional pre-trained word embedding models. Strong feature extraction capabilities and bidirectional pre-trained word embedding models have great potential. Every character is immediately accepted as the primary processing unit when dealing with Chinese text. A neural computing framework based on convolution is used to get feature representation for news texts. This is so because news texts are usually short and certain terms can stand out clearly in them. Processing effectiveness and detection ability in situations with short Chinese text can both be ensured by this kind of architecture [20]. Architecturally hybrid system that employs a meta-heuristic approach to select features and trains deep neural networks to detect false information on social media [21]. Using BERT, an encoder that accurately obtains the context of a sentence. BERT gets around the unidirectional limitation by employing a mask language model (MLM). Only the word's original vocabulary id is predicted; certain input tokens are randomly masked. MLM enhances the potential of BERT to perform when contrasted to prior embedding techniques. With simultaneous both the left and right hemispheres are conditioned; its highly bidirectional architecture allows it to handle unlabeled text [19]. Message Credibility (MCred) provides a mix of local text semantics based on N-gram characteristics and global text semantics depending on the connections among words in sentences by combining Bidirectional Encoder Representations from Transformers (BERT) and Convolutional Neural Networks (CNN) [23]. Previous works have used both BERT-based (bidirectional encoder representations from transformers) and the light gradient boosting machine (LightGBM) model [24]. In order to further enhance outcomes for class balancing and a two-step binary classification process that yields superior performance, each of the BERT-based models—SBERT, RoBERTa, and mBERT as part of contribution—as well as artificial data generated using ChatGPT [25]. the study of separating true information from rumors and false information on the internet using versions of the pre-trained, distilled BERT model. The results obtained using various techniques prior to and after combining two smaller, previously trained BERT models into a single architecture are utilized to evaluate how well distilled models learn the essential characteristics needed to discern among real and fake news [26].

Recent developments in deep learning frameworks have successfully used sentiment analysis. Socher et al. created a neural tensor network model in a backward-looking form. To classify binary emotions using a sentiment treebank [27].a convolutional neural network (CNN) with word2vec word embeddings paired as the input layer were employed [28]. It comes after softmax layers, max pooling, and convolution. Different filters are employed with the convolution layers. The parallel CNN in addition to CNN was introduced [29]. By varying the bag-of-words model, they developed feature vectors. Convolution layers are utilized in two or more parallel layers of a parallel CNN.CNN, which examined the sentiment in movie reviews using seven layers [30], was advancement. A neural network model [31] that captures both the user input and the review's semantics has been presented. A suggested unsupervised approach called paragraph vectors [32] learns variable length text, like sentences, paragraphs, etc. to create feature vectors with fixed length. A single-layered CNN that learns reviews of varied lengths and generates 300-dimensional vectors were suggested in the past [33]. The investigation used Yelp datasets and IMDB embeddings. Padding compensated for length variations. To extract the temporal data, a recurrent neural network (RNN) was employed. Cross-validation analysis of various machine learning techniques, including naive Bayes and the perceptron algorithm, was conducted to predict the rank of the review [34]. As a business model, the sentiment from customer reviews was analyzed using a multinomial Naive Bayesian [35] network. The reviews using a decision tree classifier and a naive Bayes classifier were classified [36]. In order to avoid more operationally focused engineering characterization, CNN to perform a semantic role categorization task is used [37]. For improved compositionality in sentiment recognition, RNN [38] was utilized. As a method for extracting the co-occurrence between the terms, Graph Neural Networks (GNN) [39, 40] was introduced. Globally, this network protects structured data. In order to use graph embeddings for classification, GNN's relational structure will be sufficiently complex to endure. The classification of writings also made use of this idea [41]. A novel model based on GNN known the Graph Convolutional Network (GCN) [42] was proposed and is used to build a single huge graph. This network's nodes are made up of words and documents from a large corpus that accurately reflects neighborhood information. The co-occurrence data of words creates the edges. Word frequency and word frequency in their documents are used to form the boundary among word and document nodes.

Models that aggregate the views of the whole population, extract pertinent data regarding their behavior, and predict future patterns can be developed. Designing marketing and advertising strategies can also benefit from learning about how people talk about specific products [43, 44]. The relationship between performance and mentions in blogs and reviews has been studied in the past. In order to forecast surges in book sales, the automated queries for mining blogs were constructed [45]. The linear regression to forecast movie revenue using text and metadata information were used prior [46]. Blog post attitudes and box office results are correlated [47]. The positive sentiment connections they discovered are rather weak and insufficient to be used as a prediction tool. To categorize films into groups ranging from "flop" to "blockbuster," In order to solve the issue of prediction as a classification problem, neural networks were employed [48]. Their model's greatest accuracy is rather poor, and they are forecasting ranges rather than precise figures. A news aggregation model and IMDB data were combined to forecast box office results for movies [49]. A deep neural network to forecast backorders, and it employs effective methods to deal with the disparity in data between backorders and filled orders [50]. The dataset is balanced using a range of techniques, such as randomly assigned oversampling, SMOTE oversampling, minority class weight boosting, and a mix of under- and oversampling. The predictive model calculates the likelihood of product cancellations based on the training data [50].

3. Social Media Review Based Sales Forecasting Framework

This framework consists of three modules including, initially finding whether a review is trustable or not with the help of the proposed Hybrid Embedding with CNN Transformer (HE-CNN-TransBlock). Once the review is trustable then the its polarity is detected using the extended version of Co-Sensitive Fusion Graph Convolution Network (CoSFGCN)[51] with additional introduced weightage and masking concept to improve the detection of polarity. The third module forecasts the expected sales of the products using the sentiment polarity score and its statistical measurements. A pattern based polarity correlation with temporal information is proposed in this module along with statistical measurements of the polarity scores using deep learning RNN regression model.

3.1 Fake Detection using Hybrid Embedding with CNN Transformer Block

Google announced the Bidirectional Encoder Representations from Transformers (BERT) in late 2018 [51]. It is a reliable pre-trained word embedding model that uses a transformer-encoded architecture [53,54]. BERT is therefore anticipated to be helpful for the objective of predicting review helpfulness scores. In earlier studies, the context-specific duties of the BERT Base and BERT Large model types were examined. In contrast to other language models [55, 56], BERT is fine-tuned for a variety of end goals. So, without altering its pre-trained language model, BERT can feed any of a variety of distinct downstream jobs. The most crucial step in the examination of review helpfulness is feature extraction. Previous research demonstrated that psychological and linguistic characteristics affect how beneficial a review is. Additionally, the small number of explanatory variables makes it difficult to measure these characteristics.

By using a Mask Language Model (MLM) [53], BERT removes the unidirectional restriction. Only the original semantic identification of the disguised word is guesswork; some of the input tokens are hidden at random. When compared to previous embedding techniques, BERT's ability to outperform has improved with MLM. By simultaneously conditioning on the right as well as the left context at all levels, this extremely bidirectional system is able to manage unlabeled text. We have aggregated the hidden-state string for the complete input sequence in this work, or we extracted the embeddings for a word or group of words. Both word embedding and the BERT embedding model are used in this study.

1) Word Embedding

The use of word embeddings is common in deep learning and machine learning models [57]. A contextual model for each word in an expression is created based on the words that come before and after it. Words can be converted into useful vectors using the two most popular word embedding methods currently in use: Word2Vec and GloVe [57,58,59]. In order to use pre-trained embed models for training, input embedding vectors are used in place of the processing layer's parameter values. Utilizing the word overlap frequencies found in the input vectors, the GloVe weighted least squares model is trained [60]. Utilizing training data derived from non-zero elements of a word-to-word co-occurrence matrix, it effectively leverages the benefits of statistical data. The GloVe is an unsupervised training model that may be used to determine how two words relate to one another given their position in a vector space [61]. There are many embedding vector sizes with processing ranging from 50 to 300 dimensions. In this study, we used the 300-dimensional variant. In this hybrid embedding model apart from Glove the BERT embedding is added to maintain the context based model which produce accurate feature representations. In this work a 768 feature representation is extracted from the given input sequence.

2) BiLSTM

The advantage of the Bi-directional Long-Short Term Memory Model (BiLSTM) is that it can take into account both past and future aspects. The ultimate categorized result is produced by connecting a fully contented layer and two separate directions of the Bi-LSTM simultaneously to provide a weighted-based outcome, which is often accomplished via Sigmoid Layer.

In this work, the ratio of useful votes to overall votes for every review is utilized to find the reality of product review data using a deep learning-based model that was created with BERT features rather than explanatory factors. This neural network-based program uses BERT characteristics to analyse the product review data. The suggested model examines the whether the customer reviews are genuine or not and assigns a ranking score. The overall architecture of the fake review prediction is shown in Figure 1.

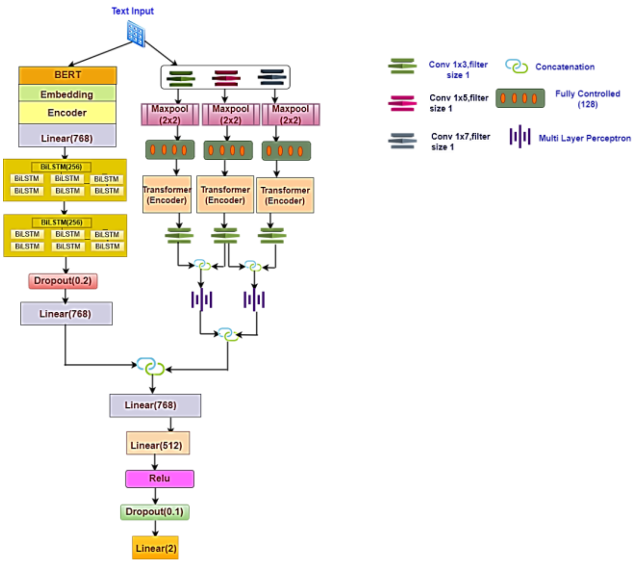

Fig. 1. Architecture of Fake Review Prediction using Hybrid Embedding CNN Transformer Block.

We simultaneously applied GloVe embedding and BERT to the input text. The embedding layer in BERT is supplied into the encoder and then the linear layer. The result is then twice sent into the BiLSTM layer and then fed into the Dropout and Linear layer. On the other hand, in GloVe embedding, three convolution layers of varying sizes such as Convolution 1*3, Convolution 1*5, and Convolution 1*7 are sent into Maximum Pooling Layer MP. The three Max Pooling layer results are subsequently routed through three distinct Fully connected layer and then into Transformer layers and ultimately into three distinct 1*3 convolution layers. Concatenation of the convolution layer's first and second results as well as its second and third results is performed. A pair of independent Multi-Layer Perceptron MLPs receive the two concatenated layers. MLP layer output is subsequently concatenated. This concatenation and linear layer result is then sent into a linear layer with size 768 and then into a linear layer with size 512 once more. Subsequently, the outcome is fed into Dropout and Relu. Lastly, a size 2 linear layer receives the output of Dropout. The equation of the proposed process is given by,

$$R1 = Linear(DropOut(BiLSTM(BiLSTM(BERT\ (TextInput) \quad (1)$$

$$R2=ConCat(Mlp_{1,2}(Concat(Con_{1,2}+Con_{2,3}(Trans_{1,2,3}(FullyCon_{1,2,3}(MaxPool_{1,2,3}(Conv_{1,2,3}(TextInput)))))))) \quad (2)$$

$$R3 = ConCat(R1+R2) \quad (3)$$

$$R4 = Linear(DropOut(Relu(Linear_{512}\ (Linear_{768}\ (R3) \quad (4)$$

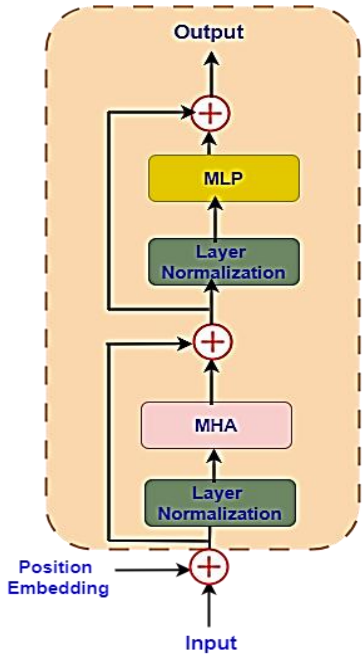

Fig. 2. Transformer Architecture's Encoder Structure

The Transformer Architecture's encoder used in our model structure is depicted in figure 2. I consist of six layers with two subdivisions. Sine and cosine functions with various frequencies are employed to create the vectors of positional encoding. Initially the input data is added with its positional embedding and further normalized in the layer wise. The normalized data passed to multi-head self-attention block with h heads. The liner projected values of the sequence in the format of query, key, value is received by each head and the attention mechanism is utilized to produce the outcome of the first sub layer and added in the form of residual addition. Once again the layer normalization is applied and followed by the fully linked feed-forward network along with residual addition to yield final outcome.

3.2 Sentiment Polarity using Co-Sensitive Weighted Fusion GCN

An enhanced version of previously designed deep GCN [51] model for sentiment score prediction is introduced in this paper. In this model Co-Sensitive Weighting and Dependency and Co-Sensitive masking is added in the Co-Sensitive Fused Graph Convolution Network.

3.2.1 Co-Sensitive Weighting

GCN is a network used widely in sentiment classification applications for calculating the dependency level and prominence of words. Co-Sensitive weight calculation is a part of Co-Sensitive Specific GCN to provide weight age the nodes before get into process of GCN learning. For this calculation, dependency level of the words is intensified further. Co-Sensitive weights (CoSW) of specific word is calculated in this step. Based on the Sensitive Dependency Graph Vector DCV the CoSW is created using the equation (5), where i=1,2,3,…,n(number of words).

$$CoSW_i = \begin{cases} 1, & if\ \boldsymbol{SDGV_i}\ is\ equal\ to\ 1 \\ 1, if\ \boldsymbol{either} & \boldsymbol{SDGV_{i+1}}\ or\ \boldsymbol{SDGV_{i-1}}\ is\ equal\ to\ 1 \\ 0.6, & otherwise \end{cases} \quad (5)$$

The following figure 3 shows the Co-Sensitive weighting of the above example using Sensitive Dependency Graph Vector shown in the figure 4(c) where w_5 and w_9 are having the value as one, hence weight for those hidden state (words) are assigned the weight as 1, as well as its predecessor and successor are also assigned as weight as one and others are assigned weight as 0.6. The hidden states are multiplied with its corresponding weights and then that node information is passed to Sensitive Specific Semantic GCN.

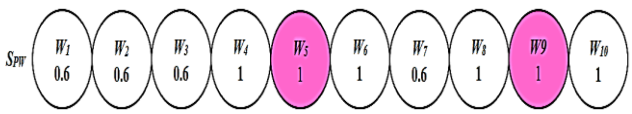

Fig. 3. Co-Sentive Weighting

3.2.2 Dependency and Co-Sensitive Masking(DCo-SM)

The output from the Sensitive Specific Semantic GCN model is further selected on the basis of masking process. Those selected hidden vectors only given as input for the next consequent graph convolutional layer or final output of the proposed CS^3 GCN model. The selection mask (DCo-SM) is formed using the equation 6.

$$DCo-SM_i = \begin{cases} 1, & if\ \boldsymbol{DGV_i}, \boldsymbol{SDGV_i}, \boldsymbol{SDGV_{i-1}}, or\ \boldsymbol{SDGV_{i+1}}\ is\ equal\ to\ 1 \\ & (most\ sensitive\ or\ dependent) \\ 0, & otherwise \end{cases} \quad (6)$$

The following figure 4 shows the generated dependency and co-sensitive masking vector. The hidden state of word 1, word 4, word 7 and word 10 is selected using the dependency graph vector and word 5 and word 9 representations are selected on the basis of sensitive dependency vector, whereas word 6 and 8 representations are selected on the basis of Co-Sensitive concept. The final output of both SynGCN and CS^3 GCN are multiplied with the Bi-LSTM hidden states to provide as final result of this two models.

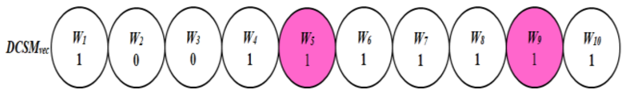

Fig. 4. Dependency and Co-Sensitive Masking Vector

3.3 Sales Forecasting Prediction

In this work the future sales forecasting is predicted for the laptop products in four different brands including dell, lenovo, asus and hp. For this analysis we collected the review of the each brands are collected from the twitter with the help of twitter API. Its corresponding sales details in quarter wise are collected from the Statista which is world wide data and business report platform. From this platform[62] we collected the overall shipment of the products in quarter wise from 2011 to 2023 (up to second quarter)

Subjectivity:

The subjectivity represents the opinion on review in two ends with respect to without any remarks on it.

$$\text{Subjectivity} = \frac{|\text{Positive and Negative Tweets}|}{|\text{Neutral Tweets}|} \quad (7)$$

Tweet Rate:

Tweet rate find the brand which is more discussed in social media compared to other competitive brands.

$$\text{Tweet} - \text{rate}(\text{product}) = \frac{|\text{tweets(particular brand product)}|}{|\text{Total Tweets of all brands(in hours)}|} \ldots \quad (8)$$

Polarity:

This parameter finds that whether a particular brands has more positive impacts than the negative impacts.

$$\text{PNratio} = \frac{|\text{Tweets with Positive Sentiment}|}{|\text{Tweets with Negative Sentiment}|} \ldots \quad (9)$$

Statistical Score:

This session calculate the overall of mean, standard deviation and variance for each individual score of the quarterly predicted polarity. Similarly the distribution is analysed with the help of skewness and kurtosis score for positive, negative and neutral score obtained for each product in three months wise. Totally it consists of 15 feature vector representation (5 statistical score for 3 polarity). Along with total positive count, negative count, neutral count, subjectivity, tweet rate and polarity the overall sentiment polarity based 21 feature vectors are generated.

Polarity Temporal Patten:

In this work a new polarity temporal based pattern is introduced. The construction of pattern is fully depends on mean positive score, mean negative score as well as mean neutral score for the reviews in the particular period time. In this work we are considering the period as three consecutive months. We assign three code word for each score representation such as for positive score it is 111, 110 for neutral and for negative it is 101. A nine bits pattern is generated based on the descending order of the score with its corresponding code word. For example if the score in the order of (mean positive score > mean negative score > mean neutral score) means then the nine bits pattern as '111101110'.

In this work next nine bits pattern is generated on the basis of temporal polarity. For example the initial pattern is generated for a particular quarter (i) means, the and operation of this pattern with its previous quarter (i-1) is considered as the second polarity based temporal pattern.

For example the 9 bits initial pattern of the second quarter of the year 2022 is as '111101110' and the 9 bits initial pattern of the first quarter of the year 2022 is '111110101' means, the second temporal 9 bits pattern is constructed by the AND operation of these two patterns which yields as '111100100'. Totally 18 bits of a pattern is constructed.

Combining all these feature, a 31 feature vector representation is used for the forecasting model. In this model, sales of the current quarter is used as target means the previous quarter feature vector is used as training transaction. Hence we can forecast the sales of a particular product in upcoming quarter with the help of the social media reviews in last three months. The following BiLSTM regression model shown in the figure 5, is designed to predict the future sales. From the figure 5, it is clearly noted initially two BiLSTM layer with 64 hidden size and followed by three kinds of linear layer with fully connected size as 1024, 256 and 32. The last sigmoid layer is act as the regression layer.

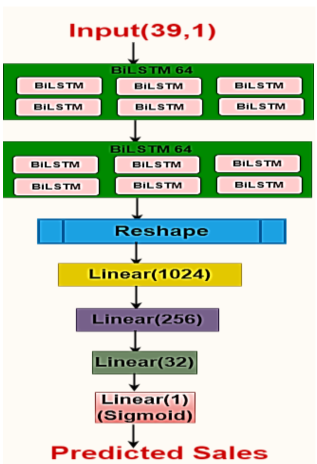

Fig. 5. Sales Forecasting Model

4. Results and Discussion

4.1 Fake Detection Dataset Description

4.1.1 Liar Dataset

Wang [63] produced the LIAR dataset. Researchers frequently utilize it to verify their suggested false news detection techniques. The dataset's statements are divided by into six categories: false, pants-fire false, true, half-true, mostly-true, and barely-true.

4.1.2 News Content Dataset

This experiment makes use of two datasets[64] that were gathered from the websites PolitiFact [65] and GossipCop [66]. GossipCop is a website for publishing the facts related to news on entertainment that are published in magazines and newspapers for the period of July 2000 to December 2018. PolitiFact is a website where political statements and reports about the United States are available for the period of May 2002 to July 2018. The two datasets with ground truth labels are statistically analysed. The statistics of features with text we have taken for our work are shown in Table 1.

Table 1. Statistics of the Fake News Net Repository

	Type	Features	PolitiFact		GossipCop	
			Fake	Real	Fake	Real
News Content	Linguistic	#News articles with text	420	528	4947	16694

4.2. Dataset for Sentiment Polarity

Utilizing datasets like TWITTER_15 [28] and Amazon [67], the proposed work is performed. Twitter posts are used by TWITTER_15 [28]. Table 2 is a list of datasets along with their details. The tabular data contains the count and sample information. The dataset from Amazon [67] comprises 11,754 labeled review sentences and 1.1 million weakly labelled sentences. The beginning pre-processing of all the data involves converting all uppercase and lowercase cases, eliminating special characters, non-English words, stop words, non-alphabetic first letters, and substituting emoij with the appropriate text. How to Lemmatize and Tokenize.

Table 2: Datasets with Class wise Samples

Dataset	Category	TWITTER_15	Amazon
Train	Positive Samples	1561	4737
	Negative Samples	1560	4665
	Neutral Samples	3127	-
Test	Positive Samples	173	876

	Negative Samples	173	1476
	Neutral Samples	346	-

4.3 Evaluation Metrics

A. Accuracy

This shows how well the classified model performed overall. The effectiveness of the model is assessed by comparing the proportion of accurate predictions to the total cases.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

B. Precision

A total of positive samples is correctly or wrongly identified as positive depending on the proportion of actually positive patterns.

$$Precision = \frac{TP}{TP+FP} \quad (11)$$

C. F1_Score

The term used to identify the harmonic mean values among recall and precision values is F1 score.

$$F1\_Score = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (12)$$

D. Recall

Calculating the percentage of positive patterns that are correctly categorized requires the measurement of recall.

$$Recall = \frac{TP}{TP+FN} \quad (13)$$

4.4 Analysis on Fake News Detection

The evaluation results of Valid and Test on the LIAR dataset are shown in Table 3.

Table 3. Evaluation Results on the LIAR Dataset

	Valid	Test	dataset count
SVMs [68]	25.8	25.5	vaild-1284, test-1283
Logistic Regression[68]	25.7	24.7	
Bi-LSTMs [69]	22.3	23.3	
CNNs [28]	26	27	
Text + All [62]	24.7	27.4	
Naive Bayes classifier [70]	-	70.2	Total true news - 6400, Total fake news - 6400
K-nearest neighbor [70]	-	46.1	
Decision Tree [70]	-	62.1	
Logistic regression [70]	-	70.3	
Random Forest [70]	-	68.6	
LSTM [72]	-	58.6	
HE-CNN_TB	76.1	73.3	valid-1283, test-1266

From the above table 3, the proposed HE-CNN_TB gives 50.1% better result than CNNs [28] and 50.3% better result than SVMs [68] in Validation. Figure 6 shows the accuracy comparison of various methods of LIAR Dataset in Testing. From that it is found that the proposed HE-CNN_TB gives 3.0% better result than Logistic Regression [68] approach and 3.1 better results than Naive Bayes classifier [70].

Table 4 shows the comparison metrics such as Accuracy, Precision, Recall, and F1-score of various methods for GossipCop. The proposed method HE-CNN_TB gives 3.91% better result compared to TieFake [86] in terms of accuracy, 0.06% better outcomes in terms of Precision, 3.56% better results in terms of Recall, and 1.65% better results in terms of F1_score. The proposed HE-CNN_TB gives 6.71% better accuracy than SpotFake+[83].

Table 4. Comparison of metrics for GossipCop Dataset

	Accuracy	Precision	Recall	F1-score
LIWC text [71]	83.60%	87.80%	31.70%	46.60%
SAFE\V text [63]	80.20%	85.30%	88.30%	86.80%
RST[72]	53.10%	-	-	-
LIWC[71]	73.60%	-	-	-
HAN[73]	74.20%	-	-	-
CNN-text[28]	73.90%	-	-	-
LSTM-ATT[74]	79.30%	-	-	-
RoBERTa-MWSS[75]	80.30%	-	-	-
BERT[51]	85.00%	-	-	-
XLNet[76]	85.50%	-	-	-
TM	84.20%	-	-	-
TMconv	85.82%	-	-	-
TMconv(max)[77]	86.28%	-	-	-
GRU[78]	79.30%	77.90%	80.10%	79.00%
VGG-19[79]	44.30%	47.80%	46.20%	45.00%
ResNet-50[80]	45.40%	46.90%	45.80%	46.30%
MVAE[81]	78.20%	80.20%	75.10%	77.60%

att-RNN[82]	77.40%	79.80%	82.10%	80.90%
SpotFake[83]	81.20%	80.70%	82.20%	81.40%
EANN[84]	83.30%	84.20%	83.50%	83.80%
SpotFake+[85]	86.40%	85.90%	88.20%	87.00%
SAFE[63]	83.10%	84.30%	89.40%	86.80%
TieFake[86]	89.20%	88.70%	90.20%	89.40%
HE-CNN_TB	**93.11%**	**88.76%**	**93.76 %**	**90.85%**

Table 5 shows the comparison metrics such as Accuracy, Precision, Recall, and F1-score of various methods for PolitiFact. The proposed method HE-CNN_TB gives 2.48% better result compared to TieFake [86] in terms of accuracy, 0.61% better results in terms of Precision, 2.57% better results in terms of Recall, and 1.58% better results in terms of F1_score.

Table 5. Comparison of metrics for PolitiFact Dataset

	Accuracy	Precision	Recall	F1-score
LIWC text [71]	82.20%	78.50%	84.60 %	81.50%
SAFE\V text [63]	72.10%	74.00%	83.10 %	78.20%
RST[72]	60.70%	-	-	-
LIWC[71]	76.90%	-	-	-
HAN[73]	83.70%	-	-	-
CNN-text[28]	65.30%	-	-	-
LSTM-ATT[74]	83.30%	-	-	-
RoBERTa-MWSS[75]	82.50%	-	-	-
BERT[51]	88.00%	-	-	-
XLNet[76]	89.50%	-	-	-
TM	87.10%	-	-	-
TMconv	90.27%	-	-	-
TMconv(max)[77]	91.21%	-	-	-
GRU[78]	68.10%	66.70%	63.20 %	64.40%
VGG-19[79]	45.80%	49.20%	47.30 %	48.20%
ResNet-50[80]	48.50%	47.80%	50.10 %	48.90%
MVAE[81]	72.60%	76.10%	67.80 %	71.70%
att-RNN[82]	74.10%	72.60%	81.30 %	76.70%
SpotFake[83]	77.00%	75.30%	79.50 %	77.00%
EANN[844]	79.50%	81.30%	76.10 %	78.60%
SpotFake+[85]	85.60%	87.80%	84.60 %	86.20%
SAFE[63]	87.20%	88.30%	89.70 %	89.00%
TieFake[86]	91.20%	93.10%	90.90 %	92.00%
HE-CNN_TB	**93.68%**	**93.71%**	**93.47 %**	**93.58 %**

4.5 Analysis of Sentiment Polarity Detection

The performance based on the metrics Accuracy and F1-Score of the proposed model along with the existing models for sentiment analysis on Twitter dataset is show in the following Table 6.

Table 6. Accuracy and F1-Score Analysis with Existing Models

Model	TWITTER_15	
	Acc.	F1S
SVM [87]	63.4	63.3
LSTM [88]	69.56	67.7
MemNet [89]	71.48	69.9
AOA [90]	72.3	70.2
IAN [91]	72.5	70.81
TNet-LF [92]	72.98	71.43
ASCNN [93]	71.05	69.45
DT ASGCN [93]	71.53	69.68
DG ASGCN [93]	72.15	70.4
DualGCN[94]	75.43	74.24
CoSWFGCN	77.17	76.19

From the table 6 it is finding that the proposed better result that TWITTER_15 and Amazon datasets. The proposed CoSWFGCN gives better accuracy result than +1.73 in DualGCN[94] for TWITTER_15. From the above table 6, it is find that the proposed better result that F1-score for Twitter-15 dataset. The proposed CoSWFGCN attains +1.95 improvements compared DualGCN [94].

Table 7. Accuracy and F1-Score Analysis for Amazon Dataset

	Amazon	
Method	Accuracy	F1-score
Lexicon [95]	77.2	72.1
SVM [96]	81.8	81.8
NBSVM [97]	82.6	82.5
SSWE [98]	83.5	83.4
SentiWV [99]	80.8	80.7
MemNet [89]	83.9	83.8
CNN-rand [100]	84.7	84.7
CNN-rand11m [100]	84.9	84.8

CNN-weak [100]	77.1	77.1
LSTM-rand [100]	84.5	84.5
LSTM-rand11m [100]	85	84.9
WDE-CNN [101]	87.7	87.6
WDE-LSTM [101]	87.9	87.9
DualGCN[94]	89.3	88.9
CoSFGCN	90.4	90

Table 7 provides the analysis on Accuracy and F1-score for Amazon dataset from that it is found that the CoSWFGCN is +1.1% better accuracy than DualGCN [94]. From the above table 7, it is find that the proposed better result that F1-score for Amazon dataset .The proposed CoSWFGCN attains +1.1 improvements compared DualGCN [94].

4.8 Analysis on Sales Forecasting

The review data collected for the laptop brands Asus, HP, Dell and Lenovo from twitter with the help of twitter API is used for analysis. Its corresponding sales collected from the Statista from 2011 to 2023 second quarter are used for analysis. Here the data from 2011 to 2021 is utilized for training and 2022 data is utilized for validation as prediction checking. The forecasting is done for the last two quarter as well as for the first quarter of the 2024. The following table 8 shows the regression metrics such as RMSE, MSE, MAE, MSLE, MPD and R2 for all brands sales prediction for the year 2022. The performance of the proposed BiLSTM-Pattern Prediction model compared with the BiLSTM and CNN models.

Table 8. Regression Metrics Analysis for Real Time Dataset for Asus, HP, Dell, Lenovo Brands

	BiLSTM-Pattern Predication				BiLSTM-Prediction				CNN-Prediction			
	Asus	HP	Dell	Lenovo	Asus	HP	Dell	Lenovo	Asus	HP	Dell	Lenovo
RMSE	0.53	0.73	0.78	0.46	0.64	1.55	1.53	0.48	0.71	2.00	1.65	0.52
MSE	0.28	0.54	0.60	0.21	0.40	2.42	2.33	0.23	0.50	3.98	2.74	0.27
MAE	0.52	0.71	0.74	0.44	0.57	1.46	1.36	0.43	0.67	1.85	1.45	0.45
MSLE	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00
MPD	0.04	0.03	0.05	0.01	0.06	0.15	0.18	0.01	0.07	0.23	0.19	0.01
R2	0.45	0.93	0.78	0.24	0.22	0.66	0.17	0.18	0.03	0.45	0.02	0.05

From the table 8, it is found that the proposed BiLSTM-Pattern model perform well in all terms of metrics compared to other deep regression models. The following figure 6 to 9 shows the predicted sales for the year 2022 as validation for each brands such as Asus, HP, Dell and Lenovo.

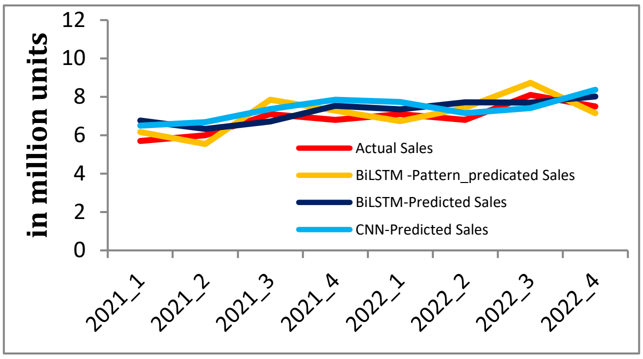

Fig 6. Predicted Sales vs Actual Sales for Asus

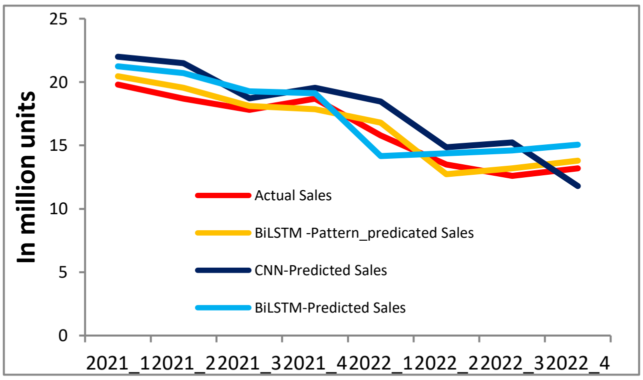

Fig 7. Predicted Sales vs. Actual Sales for HP

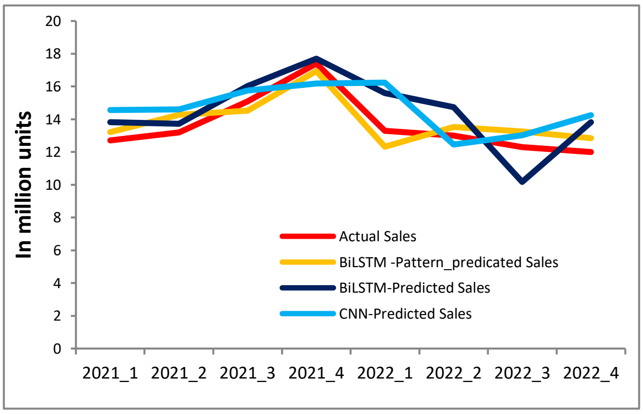

Fig 8. Predicted Sales vs Actual Sales for Dell

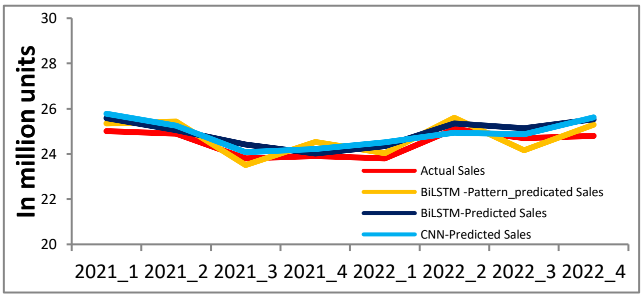

Fig 9. Predicted Sales vs Actual Sales for Lenovo

The following Figure 10 shows the sales trend from the year 2021 to 2023 second quarters. The forecasting of sales on last two quarters of 2023 as well as first quarter of 2024 is shown clearly in the figure 10. From the analysis it is found that the Lenovo occupies the higher market place, which is followed by HP. The sales on dell may fall in the coming 2024 year starting.

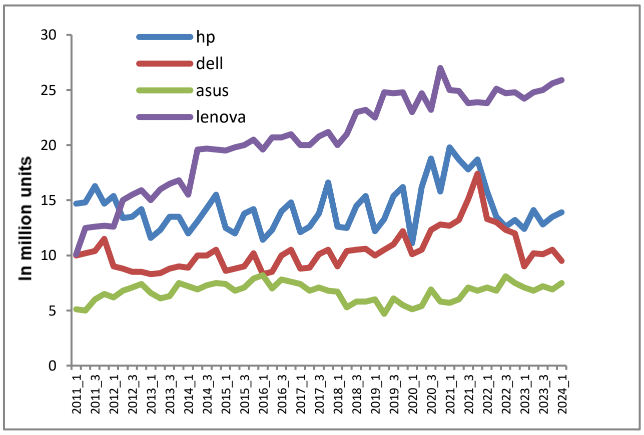

Fig 10. Sales Trend Analysis and sales forecasting in 2023 last two quarters and forthcoming 2024 first quarter.

5. Conclusion

In order to predict the sales of new individual products in upcoming seasons, this study investigates the application of deep learning to sales forecasting. We suggested methods for fake news prediction, sentiment polarity and then sales forecasting for better evaluation. The trustworthy of the review is effectively estimated using the designed hybrid embedding BiLSTM and CNN Transformer block model. With 73.3% of accuracy on test data, the HE-CNN_TB method is proposed for identifying fake news on the LIAR dataset. The proposed method obtains an accuracy of 93.11% for the GossipCop dataset and 93.68% for PolitiFact dataset. The extended version of CoSFGCN model strengthen the polarity estimation of the real review with the additional module of weighting and masking scheme. The proposed method CoSWFGCN achieves 77.17% accuracy in Sentiment Polarity findings for Twitter_15 dataset. Moreover, the accuracy for the Amazon dataset is 90.4%. The sales forecasting BiLSTM model trained with the statistical polarity features as well as the newly introduced temporal polarity pattern

References

[1] Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, Metzger MJ, Nyhan B, Pennycook G, Rothschild D, Schudson M, Sloman SA, Sunstein CR, Thorson EA, Watts DJ, Zittrain JL. The science of fake news. Science. 2018;359(6380):1094–1096. doi: 10.1126/science.aao2998.

[2] P. Meel, D.K. Vishwakarma, Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities, Expert Syst. Appl. 153 (2020). https://doi.org/10.1016/j.eswa.2019.112986.

[3] Ahmed H, Traore I, Saad S (2017) Detection of online fake news using N-gram analysis and machine learning techniques. In: International conference on intelligent, secure, and dependable systems in distributed and cloud environments. Springer, Cham, pp 127–138.

[4] Ghosh S, Shah C (2018) Towards automatic fake news classification. ProcAssocInfSciTechnol 55(1):805–807.

[5] Zhang X, Zhao J, LeCun Y (2015) Character-level convolutional networks for text classification. In: Advances in neural information processing systems, pp 649–657.

[6] Zhou X, Zafarani R (2018) Fake news: a survey of research, detection methods, and opportunities. arXiv:arXiv-1812.

[7] Fazil M, Abulaish M (2018) A hybrid approach for detecting automated spammers in twitter. IEEE Trans Inf Forensics Secur 13(11):2707–2719.

[8] Ruchansky N, Seo S, Liu Y (2017) Csi: A hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on conference on information and knowledge management. ACM, pp 797–806.

[9] Zhou X, Zafarani R (2018) Fake news: a survey of research, detection methods, and opportunities. arXiv:arXiv-1812.

[10] Reema A, Kar AK, VigneswaraIlavarasan P (2018) Detection of spammers in twitter marketing: a hybrid approach using social media analytics and bio inspired computing. Information Systems Frontiers 20(3):515–530.

[11] Vosoughi S, ’Neo Mohsenvand M, Roy D (2017) Rumor gauge: Predicting the veracity of rumors on Twitter. ACM Trans KnowlDiscov Data (TKDD) 11(4):1–36.

[12] Gupta M, Zhao P, Han J (2012) Evaluating event credibility on twitter. In: Proceedings of the 2012 SIAM international conference on data mining. Society for industrial and applied mathematics, pp 153–164.

[13] Shu K, Cui L, Wang S, Lee D, Liu H (2019) defend: Explainable fake news detection. In: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery and data mining, pp 395–405.

[14] Ahmed H, Traore I, Saad S (2017) Detection of online fake news using n-gram analysis and machine learning techniques. In: International conference on intelligent, secure, and dependable systems in distributed and cloud environments, pp 127–138

[15] Wijeratne Y (2021) How much bullshit do we need? benchmarking classical machine learning for fake news classification. LIRNEasia, [Online]. https://lirneasia.net/2021/07/how-much-bullshit-do-we-need-benchmarking-classical-machine-learning-for-fake-news-classification/. Accessed 24 Aug 2022

[16] Ozbay FA, Alatas B. Fake news detection within online social media using supervised artificial intelligence algorithms. Phys A. 2020;540:123174. doi: 10.1016/j.physa.2019.123174. [CrossRef] [Google Scholar]

[17] Ozbay FA, Alatas B. Adaptive salp swarm optimization algorithms with inertia weights for novel fake news detection model in online social media. Multimed Tools Appl. 2021;80(26):34333–34357. doi: 10.1007/s11042-021-11006-8. [CrossRef] [Google Scholar]

[18] Kansal A. Fake news detection usingpos tagging and machine learning. J Appl Secur Res. 2021 doi: 10.1080/19361610.2021.1963605. [CrossRef] [Google Scholar]

[19] Kaliyar, Rohit Kumar, AnuragGoswami, and Pratik Narang. "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach." Multimedia tools and applications 80.8 (2021): 11765-11788.

[20] Zhang, Qin, et al. "A deep learning-based fast fake news detection model for cyber-physical social services." Pattern Recognition Letters 168 (2023): 31-38.

[21] Kishwar, Azka, and AdeelZafar. "Fake news detection on Pakistani news using machine learning and deep learning." Expert Systems with Applications 211 (2023): 118558.

[22] Kumar, Sanjay, et al. "OptNet-Fake: Fake News Detection in Socio-Cyber Platforms Using Grasshopper Optimization and Deep Neural Network." IEEE Transactions on Computational Social Systems (2023).

[23] Verma, Pawan Kumar, et al. "MCred: multi-modal message credibility for fake news detection using BERT and CNN." Journal of Ambient Intelligence and Humanized Computing 14.8 (2023): 10617-10629.

[24] Essa, Ehab, Karima Omar, and Ali Alqahtani. "Fake news detection based on a hybrid BERT and LightGBM models." Complex & Intelligent Systems (2023): 1-12.

[25] Shushkevich, Elena, Mikhail Alexandrov, and John Cardiff. "Improving Multiclass Classification of Fake News Using BERT-Based Models and ChatGPT-Augmented Data." Inventions 8.5 (2023): 112.

[26] Rana, Vineet, et al. "Compact BERT-Based Multi-Models for Efficient Fake News Detection." 2023 3rd International Conference on Intelligent Technologies (CONIT). IEEE, 2023.

[27] Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., Potts, C. (2013) “Recursive deep models for semantic compositionality over a sentiment treebank”, Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2013.

[28] Kim, Y. (2014) “Convolutional neural networks for sentence classification”, Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1746–1751, October 25-29, 2014, Doha, Qatar.

[29] Johnson, R. & Zhang, T. (2015) “Effective use of word order for text categorization with convolutional neural networks”, Proceedings of Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL, pp 103–112, Denver, Colorado, May 31 – June 5, 2015.

[30] Ouyang, X. Zhou, P., Li, C. H., & Liu, L. (2015) “Sentiment analysis using convolutional neural network. Proceedings of IEEE International Conference on Computer and Information Technology;Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, pp. 2359-2364, October 2015.

[31] Tang, D., Qin, B., & Liu, T. (2015) Learning semantic representations of users and products for document level sentiment classification. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, pp. 1014–1023, Beijing, China, July 26-31, 2015.

[32] Le, Q. V. &Mikolov, T. (2014) “Distributed representations of sentences and documents”, Proceedings of the 31st International Conference on Machine Learning, Beijing, China, 2014.

[33] Chen, T., Xu, R., He, Y., Xia, Y., & Wang. X. (2016) “Learning user and product distributed representations using a sequence model for sentiment analysis”, IEEE Computational Intelligence Magazine, 11(3):34-44, 2016.

[34] Y. Xu, X. Wu, and Q. Wang. Sentiment analysis of yelps ratings based on text reviews, 2015.

[35] M. S. Elli and Y.-F. Wang. Amazon reviews, business analytics with sentiment analysis.

[36] C. Rain. Sentiment analysis in amazon reviews using probabilistic machine learning. Swarthmore College, 2013.

[37] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa. Natural language processing (almost) from scratch. Journal of Machine Learning Research, 12(Aug):2493–2537, 2011.

[38] Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., Potts, C. (2013) “Recursive deep models for semantic compositionality over a sentiment treebank”, Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2013.

[39] Battaglia, P. W.; Hamrick, J. B.; Bapst, V.; Sanchez-Gonzalez, A.; Zambaldi, V.; Malinowski, M.; Tacchetti, A.; Raposo, D.; Santoro, A.; Faulkner, R.; et al. 2018. Relational inductive biases, deep learning, and graph networks. arXiv preprint arXiv:1806.01261.

[40] Cai, H.; Zheng, V. W.; and Chang, K. 2018. A comprehensive survey of graph embedding: problems, techniques and applications. IEEE Transactions on Knowledge and Data Engineering 30(9):1616– 1637.

[41] Kipf, T. N., and Welling, M. 2017. Semi-supervised classification with graph convolutional networks. In ICLR.

[42] Yao, Liang, Chengsheng Mao, and Yuan Luo. "Graph convolutional networks for text classification." Proceedings of the AAAI conference on artificial intelligence. Vol. 33. No. 01. 2019.

[43] Jure Leskovec, Lada A. Adamic and Bernardo A. Huberman. The dynamics of viral marketing. In Proceedings of the 7th ACM Conference on Electronic Commerce, 2006.

[44] B. Jansen, M. Zhang, K. Sobel, and A. Chowdury. Twitter power: Tweets as electronic word of mouth. Journal of the American Society for Information Science and Technology, 2009.

[45] Daniel Gruhl, R. Guha, Ravi Kumar, Jasmine Novak and Andrew Tomkins. The predictive power of online chatter. SIGKDD Conference on Knowledge Discovery and Data Mining, 2005.

[46] Mahesh Joshi, Dipanjan Das, Kevin Gimpel and Noah A. Smith. Movie Reviews and Revenues: An Experiment in Text Regression NAACL-HLT, 2010.

[47] G. Mishne and N. Glance. Predicting movie sales from blogger sentiment. In AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs, 2006.

[48] Ramesh Sharda and DursunDelen. Predicting box-office success of motion pictures with neural networks. Expert Systems with Applications, vol 30, pp 243–254, 2006.

[49] W. Zhang and S. Skiena. Improving movie gross prediction through news analysis. In Web Intelligence, pages 301304, 2009.

[50] Shajalal, Md, PetrHajek, and Mohammad ZoynulAbedin. "Product backorder prediction using deep neural network on imbalanced data." International Journal of Production Research 61.1 (2023): 302-319.

[51] M.Priya Alagu Dharshini, S. Antelin Vijila, S.P. Victor. “CoSFGCN: Co-Sensitive Fusion Graph Convolution Network for Sentiment Analysis,” SSRG International Journal of Engineering Trends and Technology, vol. 71, no. 10, pp314-325, 2023..

[52] Devlin J, Chang M-W, Lee K, Kristina T (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. In: NAACL-HLT (1).

[53] Tenney I, Das D, Pavlick E (2019) BERT rediscovers the classical NLP pipeline. In: Proceedings of the 57th annual meeting of the association for computational linguistics.

[54] Xu, H., Liu, B., Shu, L., Yu, P.S.: BERT post-training for review reading comprehension and aspect-based sentiment analysis. arXiv preprint arXiv:190402232 (2019).

[55] Peters, M.E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., Zettlemoyer, L.: Deep contextualized word representations. arXiv preprint arXiv:180205365 (2018).

[56] Howard, J., Ruder, S.: Universal language model fine-tuning for text classification. arXiv preprint arXiv:180106146 (2018).

[57] Peters ME, Neumann M, Iyyer M, Gardner M, Clark C, Lee K, Zettlemoyer L (2018) Deep contextualized word representations. In: Proceedings of NAACL-HLT, pp 2227–2237.

[58] Young T, Hazarika D, Poria S, Cambria E (2018) Recent trends in deep learning based natural language processing. IEEE Comput Intell Mag 13(3):55–75.

[59] Qi Y, Sachan D, Felix M, Padmanabhan S, Neubig

G (2018) When and why are pre-trained word embeddings useful for neural machine translation? In: Proceedings of the 2018 conference of the north american chapter of the association for computational linguistics: human language technologies, vol 2 (short papers), pp 529–535.

[60] Asparouhov T, Muthen B (2010) Weighted least squares estimation with missing data. Mplus Technical ´ Appendix 2010: 1–10.

[61] Qi Y, Sachan D, Felix M, Padmanabhan S, Neubig G (2018) When and why are pre-trained word embeddings useful for neural machine translation? In: Proceedings of the 2018 conference of the north american chapter of the association for computational linguistics: human language technologies, vol 2 (short papers), pp 529–535.

[62] https://www.statista.com/statistics/298950/pc-shipments-worldwide-asus/

[63] Wang, W.Y. “Liar, liar pants on fire”: A new benchmark dataset for fake news detection. arXiv 2017, arXiv:1705.00648.

[64] Zhou X, Wu J, Zafarani R, SAFE: similarity-aware multi-modal fake news detection. pp.1-12,2020,doi:10.48550/arXiv.2003.04981.

[65] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, H. Liu, Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media,vol.8,issue.3,pp.171-188, 2018, doi: 10.1089/big.2020.0062.

[66] Danah m. Boyd, Nicole B. Ellison, Social Network Sites: Definition, History, and Scholarship, Journal of Computer-Mediated Communication, Vol.13, Issue.1, pp.210–230,1 October 2007, doi:10.1111/j.1083-6101.2007. 00393.x.

[67] W. Zhao et al., "Weakly-Supervised Deep Embedding for Product Review Sentiment Analysis," in IEEE Transactions on Knowledge and Data Engineering, vol. 30, no. 1, pp. 185-197, 1 Jan. 2018, doi: 10.1109/TKDE.2017.2756658.

[68] Koby Crammer and Yoram Singer. 2001. On the algorithmic implementation of multiclass kernel-based vector machines. Journal of machine learning re- search 2(Dec):265–292.

[69] Sepp Hochreiter and Jurgen Schmidhuber. 1997. ¨ Long short-term memory. Neural computation 9(8):1735–1780.

[70] Sharma, D.K., Garg, S. IFND: a benchmark dataset for fake news detection. Complex Intell. Syst. (2021). https://doi.org/10.1007/s40747-021-00552-1

[71] Pennebaker, J.W., Boyd, R.L., Jordan, K., Blackburn, K.: The development and psychometric properties of LIWC2015. Tech. rep. (2015)

[72] Rubin, V. L., Conroy, N., and Chen, Y. (2015). Towards news verification: Deception detection methods for news discourse. In Proceedings of the Rapid Screening Technologies, Deception Detection and Credibility Assessment Symposium, at the 48th An- 3770 nual Hawaii International Conference on System Sciences.

[73] Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., and Hovy, E. (2016). Hierarchical attention networks for document classification. In Proceedings of HLTNAACL.

[74] Lin, J., Tremblay-Taylor, G., Mou, G., You, D., and Lee, K. (2019a). Detecting fake news articles. In Proceedings of 2019 IEEE International Conference on Big Data.

[75] Shu, K., Zheng, G., Li, Y., Mukherjee, S., Awadallah, A. H., Ruston, S. W., and Liu, H. (2020b). Leveraging multi-source weak social supervision for early detection of fake news. arXiv preprint:2004.01732.

[76] Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., and Le, Q. V. (2019a). Xlnet: Generalized autoregressive pretraining for language understanding. In Proceedings of NIPS.

[77] Bimal Bhattarai, Ole-Christoffer Granmo, Lei Jiao.,ConvTextTM: An Explainable Convolutional Tsetlin Machine Framework for Text Classification.,2022.

[78] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, and M. Cha, “Detecting rumors from microblogs with recurrent neural networks,” in 25th International Joint Conferences on Artificial Intelligence(IJCAI), 2016, pp. 3818–3824.

[79] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in International Conference on Learning Representations(ICLR), 2015.

[80] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition(CVPR), 2016, pp. 770–778.

[81] D. Khattar, J. S. Goud, M. Gupta, and V. Varma, “Mvae: Multimodal variational autoencoder for fake news detection,” in The world wide web conference(WWW), 2019, pp. 2915–2921.

[82] Z. Jin, J. Cao, H. Guo, Y. Zhang, and J. Luo, “Multimodal fusion with recurrent neural networks

for rumor detection on microblogs," in Proceedings of the 25th ACM international conference on Multimedia(ACM MM), 2017, pp. 795–816.

[83] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, and S. Satoh, "Spotfake: A multi-modal framework for fake news detection," in 2019 IEEE fifth international conference on multimedia big data(BigMM). IEEE, 2019, pp. 39–47.

[84] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, and J. Gao, "Eann: Event adversarial neural networks for multi-modal fake news detection," in Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining(KDD), 2018, pp. 849–857.

[85] S. Singhal, A. Kabra, M. Sharma, R. R. Shah, T. Chakraborty, and P. Kumaraguru, "Spotfake+: A multimodal framework for fake news detection via transfer learning (student abstract)," in Proceedings of the AAAI conference on artificial intelligence(AAAI), vol. 34, no. 10, 2020, pp. 13 915–13 916.

[86] Quanjiang Guo1 , Zhao Kang1† , Ling Tian1 , Zhouguo Chen2.,TieFake: Title-Text Similarity and Emotion-Aware Fake News Detection.,2023.

[87] Li Dong, Furu Wei, Chuanqi Tan, Duyu Tang, Ming Zhou, and KeXu. 2014. Adaptive recursive neural network for target-dependent twitter sentiment classification. In Proceedings of the 52nd annual meet ing of the association for computational linguistics (volume 2: Short papers), volume 2, pages 49–54.

[88] Duyu Tang, Bing Qin, XiaochengFeng, and Ting Liu. 2016a. Effective lstms for target-dependent sentiment classification. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pages 3298– 3307.

[89] Duyu Tang, Bing Qin, and Ting Liu. 2016b. Aspect level sentiment classification with deep memory net work. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Process ing, pages 214–224.

[90] Binxuan Huang, YanglanOu, and Kathleen M Car ley. 2018. Aspect level sentiment classification with attention-over-attention neural networks. In International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior epresentation in Modeling and Simulation, pages 197–206. Springer.

[91] Dehong Ma, Sujian Li, Xiaodong Zhang, and Houfeng Wang. 2017. Interactive attention networks for aspect-level sentiment classification. In Proceed ings of the 26th International Joint Conference on Artificial Intelligence, pages 4068–4074. AAAI Press.

[92] Xin Li, Lidong Bing, Wai Lam, and Bei Shi. 2018. Transformation networks for target-oriented sentiment classification. In Proceedings of the 56th An nual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 946– 956.

[93] Zhang, Chen & Li, Qiuchi& Song, Dawei. (2019). Aspect-based Sentiment Classification with Aspect-specific Graph Convolutional Networks. https://aclanthology.org/D19-1464.pdf.

[94] Ruifan Li1∗ , Hao Chen1 , Fangxiang Feng1 , Zhanyu Ma1 , Xiaojie WANG1 , and Eduard Hovy.,Dual Graph Convolutional Networks for Aspect-based Sentiment Analysis.,pp.6319-6329, 2021.

[95] X. Ding, B. Liu, and P. S. Yu. A holistic lexicon-based approach to opinion mining. In WSDM, pages 231–240, 2008.

[96] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. Liblinear: A library for large linear classification. JMLR, 9:1871–1874, 2008.

[97] S. Wang and C. D. Manning. Baselines and bigrams: Simple, good sentiment and topic classification. In ACL, pages 90–94, 2012.

[98] D. Tang, F. Wei, N. Yang, M. Zhou, T. Liu, and B. Qin. Learning sentiment-specific word embedding for twitter sentiment classification. In ACL, volume 1, pages 1555–1565, 2014.

[99] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts. Learning word vectors for sentiment analysis. In ACL, pages 142–150, 2011.

[100] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In NIPS, pages 3111–3119, 2013.

[101] W. Zhao et al., "Weakly-Supervised Deep Embedding for Product Review Sentiment Analysis," in IEEE Transactions on Knowledge and Data Engineering, vol. 30, no. 1, pp. 185-197, 1 Jan. 2018, doi: 10.1109/TKDE.2017.2756658.