Bulletin of Pure and Applied Sciences
Zoology (Animal Science), Vol.43B, No.1s,
January-June 2024: P.1059-1075

Print version ISSN 0970-0765
Online version ISSN 2320-3188

Original Research Article

Content Available online at: www.bpasjournals.com

EFFICIENT FEATURE EXTRACTION WITH OPTIMIZATION BASED HYBRID CLASSIFICATION MODEL FOR FAKE NEWS DETECTION

[1*]R. Uma Maheswari, [2]Dr. N. Sudha
[1*]Research Scholar, Department of Computer Science, Bishop Appasamy College of Arts and Science, Coimbatore, Tamil Nadu 641018.
Email id- umasuresh.agni@gmail.com
[2]Associate Professor Department of Computer Science, Bishop Appasamy College of Arts and Science, Coimbatore, Tamil Nadu 641018.
Email id- sudhanatarajan105@gmail.com

Abstract: A massive volume of information is circulated daily over online and print media; however, it is difficult to say whether the data contains truth or is fake. Due to their lower costs, simple accesses, and speedy distributions, social media have recently become one of the primary worldwide news sources. Dubious News also have a cost involved even though reliability and sizable possibility of reading "fake news" that was purposefully released to misinform readers [1,2]. Social media has grown in popularity as a means of consuming news over the last ten years because it is simple to use, spreads information quickly, and is inexpensive. Hence, fake news detections from social media have gained traction in research studies. Social media introduces new characteristics and issues for detecting false news, rendering traditional news industry's existing detection algorithms ineffective or irrelevant [3]. Recent research compares fake and authentic news strips using critical keywords in hybrid classification algorithms. The three steps described in this study can assist in identifying false news from social media information. The unstructured data from the media are transformed into structured form using a number of pre-processes. The Lexicon Model extracts the features and unknowable properties of misleading news in the second stage. This research project is currently in its third stage. a feature selection method by WOA (whale optimization Algorithm) for weight value to tune the classification part. Finally, a hybrid classification model called **WOA-HYC, a** hybrid fuzzy based CNNs (Convolution Neural Networks) and kernel based SVM (Support Vector Machine) to identifies false information and its effectiveness is measured in terms of values obtained for accuracy, precision, recall and F1 score. Experiment results show the suggested model achieves higher results amongst methods considered in this study.

Keywords: Fake news; pre-processing, feature extraction Method, Lexicon Model, Fuzzy weight based Convolutional Neural Network , Support Vector Machine, fake news detection.

1. Introduction

Increasing social media has recently expedited the spread of incorrect information. Because of how fast, broadly, and diversely information is distributed through social networks, online erroneous information may have a large detrimental impact on the entire society. Fake news can take a variety of forms, such as inadvertent mistakes made by news aggregators, overtly fraudulent stories, or tales created with the purpose to deceive and sway readers' opinions [4]. Although false news may take many different forms, because it departs from the facts, it typically has a negative impact on individuals, governments, and organisations.

The continuing spread of fake information and news have made periodical evaluations necessary for improving quality of digital contents and also a hot topic for research in society at large and new industry. Societies frequently struggle with truths in both textual and digital media. Throughout history of information transmissions, sensationalism of not quite accurate eye-catching and thrilling headlines attempt to capture audiences for sales. The increased reach and expedited truthfulness of information transmissions on social networking sites also have damaging influences on millions of people within minutes.

Many of data mining algorithms have been introduced in recent years to identify bogus news. Various scholars are striving to identify bogus news. The content lengths of online false reviews and rumours are frequently shorter than those of online fake news since they are always condensed and information-intensive. Typical language processing and embedding techniques, such as bag-of-words or n-grams, useful for reviewing information fail to discover fake information. Hence, the assess essential viewpoints and semantic sequential orders in news contents, sophisticated embedding algorithms which can identify false information in news are needed.

RNNs, for example, are helpful for encoding underlying semantic links, remembering important semantic sequential orders, and embedding real language. Recent developments in DL (Deep Learning) methodologies are to blame for this. DL algorithms are now being used more often to identify internet's false information [6]. Additionally, the majority of research exclusively uses content-based analysis in the field of "rumour or satire detection". The majority of internet research on false news, however, incorporates data from the item's author, its substance, and its social environment. Different algorithms are being used by researchers to identify bogus news. SVM, NB (Naive Bayes), LSTM (Long Short-Term Memory), and DL models are just a few of the models that are used for fake news detection.

A hybrid classification strategy for spotting fake news is described in this study. Data pre-processes, feature extractions and selections followed by classifications are used in this work. These detection methods aim to anticipate purposely misleading news based on an assessment of genuine and false news. Thus, it is crucial to have access to high-quality and huge amounts of training data. So, this study effort focuses on enhancing the detection accuracy by introducing the feature selection and hybid model in order to address this issue. Section 2 of the study paper discusses some of the most current techniques for spotting fake news. Section 3 describes the fake news detecting algorithm. Section 4 discusses the results. Section 5 covers conclusion and future work.

2. Literature Review

Discussing several methods for spotting false news in various forms of data is done in this section. The literature on false news identification has been reviewed in terms of previous and related efforts.

Nasir et al [2021] [7] developed a novel hybrid recurrent DL model of CNNs for classifying false news and verified on datasets ISO and FA-KES for fake news where its detections were superior to earlier non-hybrid baseline methods by a wide margin. Additional trials on the suggested model's generalisation across various datasets yielded promising results.

Yang et al [2018] [8] examined "fake news detection" problem. Automatic fake news identification is difficult since there are few models that can manage the problem of pure model-based news fact-checking. Following a thorough examination of the fake news data, several useful explicit pieces from the language and photographs utilised in the bogus news are discovered. In addition to the explicit characteristics, our model's multiple convolution layers may be utilised to recover certain latent features that can be used to find hidden patterns in the language and imagery used in fake news. The TI-CNN (Text and Image information based CNN) model was suggested for application. By combining explicit and latent characteristics into a single feature space, TI-CNN is trained on both text and picture data at once. The effectiveness of TI-CNN in addressing the false news identification issue has been demonstrated by extensive experiments on real-world fake news datasets.

Kaliyar et al [2020] [9] proposed deep CNNs called FNDNets for detecting bogus news. Rather than relying on manual feature constructions, FNDNets automatically learnt discriminations for recognising false news and by usage of hidden layers in deep NNs. To extract a number of characteristics at each layer, we build a deep CNN. We contrast the effectiveness of the suggested technique with a number of reference models. The suggested model produced cutting-edge findings with accuracy on the test data after being trained and tested on benchmark datasets. Several performance evaluation metrics, including Wilcoxon, false positives, true negatives, precisions, recalls, F1 scores, and accuracies of methods were used in evaluations. The study's results demonstrate considerable improvements in recognising false news compared to other methods, illustrating practicality of the strategy for identifying false news on social media. This work can be of assistance to researchers in better comprehensions of potential usages of CNNs based deep models for false news identification.

Balwant [2019] [10] analysed 12836 brief news pieces from various sources, including social media from Liar-Liar dataset. The proposed architecture tagged news articles with POS (part of speech) and speaker attributes from CNNs using Bidirectional LSTM. The results revealed that their hybrid architecture significantly improved its ability to recognise fake news in the dataset.

Saleh et al [2021] [11] identified false information using both ML (machine learning) and DL (deep learning) techniques with improved accuracy. Their schema called OPCNN-fraudulent identified false information using improved model of CNNs and evaluated against RNNs (Recurrent Neural Networks), LSTM, and six traditional ML algorithms namely DTs (Decision Trees), LRs (Logistic Regressions), KNNs (K-Nearest Neighbours), RFs (Random Forests), SVM, and NB on four fake news datasets. Their use of grid searches and hyperopt optimisations tuned parameters for ML and DL. Glove word embeds encoded features for DL models as feature matrces, while N-grams and TF-IDFs (Term Frequency—Inverse Document Frequencies) were employed to extract features from datasets by classical ML. They validated their schema's performances in terms of values obtained for accuracy, precision, recall, and the F1-measures where their suggested OPCNN-FAKE model outperformed other models on the considered datasets.

Krishna andAdimoolam [2022] [12] employed ML techniques to research the detection of fake news, including the DT and SVM algorithms (N = 311). Results: We tested and assessed the DT and SVM algorithms' accuracy in spotting false news on social media. The DT machine algorithm appears to be more accurate than the SVM approach, with a reported accuracy of 91.74%. The difference between the research groups is statistically significant, with significance values for accuracy and precision of 0.092 and 0.825, respectively, for a confidence interval (CI) of 91%.

Gravanis et al [2019] [13] presented the "UNBiased" (UNB) dataset, a new text corpus that incorporates a variety of news sources and complies with a number of norms and guidelines to prevent biassed classification results. Our test results demonstrate the excellent classification accuracy of false news using an expanded linguistic feature set including word embeddings, ensemble methods, and SVM.

Albahar [2021] [14] identified crucial details utilised for spotting fake news using hybrid model which combined RNNs and SVM. News articles and comments were transformed into numerical feature vectors using RNNs with bidirectional gated recurrent units. These encoded attributes were inputs for radial basis function SVM which classified true information from fake news. The suggested schema outperformed other related techniques in its performances.

Ahmed et al [2022] [15] Passive Aggressive, NB, and SVM classifiers were used. Simple classification is not totally accurate since classification algorithms are not especially designed for detecting bogus news. Fake news can be identified by train classifiers on news data where ML and text-based processing can be combined. The main goal of text classifications are to separate different textural characteristics and then include them into classifications.

Meesad [2021] [16] proposed three basic components proposed namely IR (information retrieval), NLP (natural language processing), and ML . The two stages of this project are data collecting and ML model construction. In order to extract useful elements from web data, we used natural language processing algorithms to analyse material retrieved from Thai online news sources. Models like NB, LR, KNN, MLP, SVM, DT, RF, Rule-Based Classifier, and LSTM were used for comparison. We developed an automated online web application for false news identification after comparison analysis on the test set that showed LSTM to be the best model.

3. Proposed Methodology

The motivation for this research is to determine how a news item links to its headline. This section contains information on the proposed method for spotting fake news which encompasses four steps as detailed below:

- The pre-process in the first stage transforms unstructured information to structured information.
- In the second phase, unidentified aspects of fake news features are extracted using a lexical model.
- In the third step, WOA based feature selections are introduced for weight calculations and tuning.
- Finally, this work builds an Hybrid Model to learn the detect of the fake news in an efficient manner. . In the proposed work fuzzy weight value based CNN with kernel based SVM used to improve the classification part.

Figure 1 depicts the proposed methodology's processes.

3.1. Data Collection

The datasets that were utilised in this research are publicly accessible and may be downloaded at no cost online. The information comes from a variety of sources and contains both fake and genuine news stories. While fake news websites make untrue claims, legitimate news reports accurately describe events. Politifact.com and Snopes.com are useful for fact-checking, many of those publications' political statements may be carefully examined for consistency.

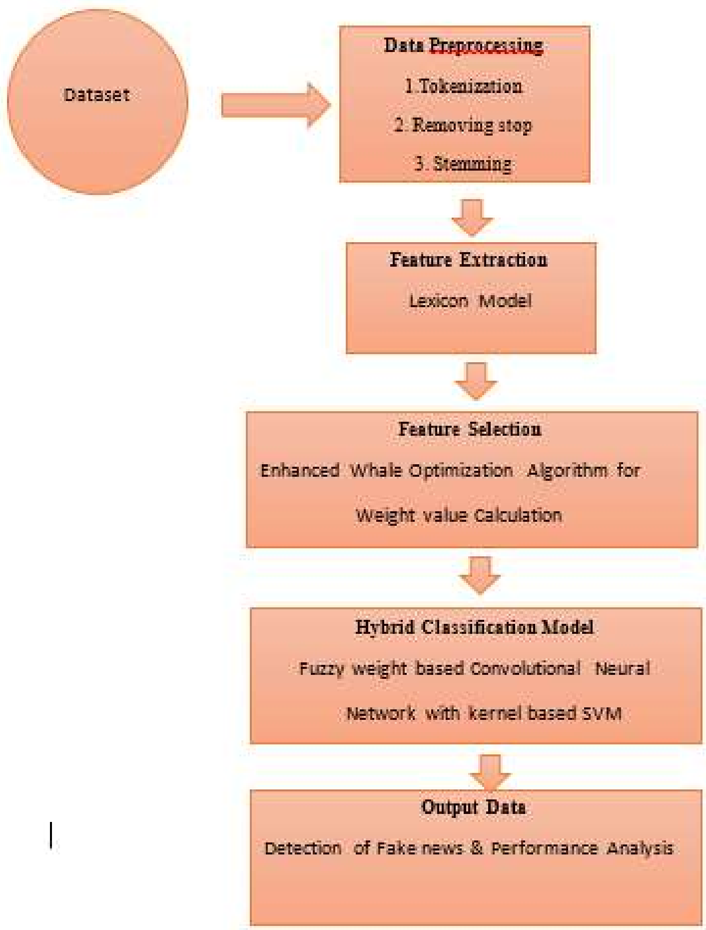

Figure .1. The Overall process of the proposed Hybrid Feature Extraction model with Capsule Neural Network (HFEM-CNN) The "ISOT Fake News Dataset" , which includes both real and fraudulent items pulled from the Internet, is the dataset utilised in this research. The original stories originated from the well-known news site reuters.com, while the fake ones were obtained from a number of sources, primarily from sites that politifact.com has flagged as fake. There are 44,898 articles in all, 21,417 of which are true, and 23,481 of which are fake. Articles from a variety of topics are included in the corpus overall, but political news is the most prevalent focus.

3.2. Data Pre-processing

Before training and analysing the data using a model, pre-processing the data is often the initial step. The quality of ML algorithms depends on the data they are fed. For there to be enough consistency to provide the best outcomes, it is essential that data be structured correctly and that important characteristics be included. As shown in, normalising picture inputs and dimensions reductions are only two of the several stages involved in pre-processing the data for computer vision ML algorithms [17]. They are intended to reduce some of the insignificant details that serve as visual cues for various images. The job of identifying the picture does not benefit from features like brightness or darkness. Similar to this, certain text passages are useless for determining if a passage is authentic or not.

- **Tokenization**

By eliminating all punctuations and tokenizing inputs, textual data is fragmented, or divided into smaller pieces. Multiple filters were used exclude words with numeric characters. Text data was converted to lowercases or uppercases and words were changed to lowercases in documents. Finally, words which had characters fewer than N were excluded using N-chars filter.

- **Stop-words removal**

Despite the fact that they are regularly employed to conclude sentences and link phrases, stop words like prepositions, pronouns, and conjunctions are not required terminologies. These sentences (linguistic words), do not provide any information [18]. There are four to five hundred stop words in English where a, an, about, by, but, that, does, on, above, once, after, till, too, again, when, where, what, all, am, and, any, against are certain examples.

- **Stemming**

This method transforms a words' grammatical forms into their basic forms. Stemming is used to determine underlying word forms with equal meanings like "Connect" can be formed from terms such as "connection," "connections," "connective," "connected," and "connecting," among others. Algorithm 1 contains data pre-processing techniques [19].

Algorithm 1. Preprocessing process
Input: Given textual data
Output: Preprocessed data

1. Textual data without numbers
2. Remove all punctuation from text data
3. Characters with the character < N should be filtered out
4. Convert textual data's case
5. Delete all stop words
6. Stem textual data

3.3. Feature Extraction

The sheer quantity of variables makes analysis tough and one of the most difficult jobs. Classification algorithms that involve numerous variables and require a lot of memory and processing resources may overfit training data and perform badly on new samples. By combining variables, feature extractions assist in overcoming these difficulties and accurately characterising data. The focus lexicon model for feature extraction serves as the foundation for the suggested investigation.

3.3.1. Lexicon Model

About 6300 terms make up the lexicon that was created. Vader Sentiment was used as a baseline while it was hand created [20]. Each word in the lexicon has an emotion score that ranges from 100 (the most positive) to 100 (the least positive). According to empirical research, some positive and negative words can occasionally appear in sentences with a neutral meaning. We calculated a conditional probability (P), as shown in Eq. 1, for each word in the lexicon.

$$P(positive \mid w) \; for \; positive \; w$$
$$P(negative \mid w) \; for \; negative \; w \tag{1}$$

For each positive phrase, the likelihood that a random message containing this word is positive was assessed using a collection of labelled data.[21] The odds for each negative term were assessed in the same way. Additionally, it sought to determine whether using such data in the sentiment categorization process may aid in handling messages with conflicting (both positive and negative) sentiment. Random selections were used to select 100,000 positive and negative fake news reports as samples. The frequencies of selected words in selected positive and negative messages (w) were computed. Depending on whether the word was positive or negative, equations 2 and 3 were used to calculate the conditional probability.

$$P\,(positive|w) = \frac{P(positive \cap w)}{P(w)} = \frac{\#w_P}{\#w} \tag{2}$$

$$P\,(negative|w) = \frac{P(negative \cap w)}{P(w)} = \frac{\#w_N}{\#w} \tag{3}$$

where #wP and #wN represent, respectively, the positive and negative message counts from the sample that contain the word "w." The probability for positive and negative terms were estimated using the two formulae, respectively [22]. The procedure was performed several times to provide a more accurate result, and the lexicon now contains the average probability that was achieved for each term. The next hybrid classification model uses data probabilities.

3.4. Feature Selection Method

Only a portion of the redundant feature extracted input data used for this work presents diverse profiles for various classes of samples. In order to save time, using high discriminative features from data has been increasingly intriguing in the classification sector. In this study, the weight value is computed using WOA for feature selection. A unique population-based stochastic optimisation approach called the WOA was just recently created [23]. The WOA employs a collection of search agents to identify the ideal response to an optimisation issue. The WOA employs a technique known as "bubble-net hunting" to mimic the actions taken by humpback whales as they pursue prey. Encircling the prey, using a bubble net to assault, and hunting for the best prey are the three basic phases of the WOA.

Whales enclose preys and update their positions for obtaining optimum solutions. Its mathematical depiction is shown in Eqations (4) and (5).

$$X(t+1) = X^*(t) - A.|C.X^*(t) - X(t)| if p < 0.5 \tag{4}$$
$$X(t+1)|C.X^*(t) - X(t)|.e^{bl}\cos(2\pi t) + X^*(t) if p \geq 0.5 \tag{5}$$

If t stands for index of iteration counts, X* stands for best current solutions, and X stands for vectors containing whales's locations. A=2a. (r-a); C=2.r; l represents random values between -1 and 1 and are used to toggle between (7) and (8) while updating whales' locations. a denotes a coefficient vector that drops linearly from 2 to 0 in iterations. The relative probabilities of 50% in Eqs. (7) and (8) show that whales may equally choose either options at random in optimisations. The form of the logarithmic spiral is affected by the constant value b, which varies depending on the route. Its value is set at 1 in this study. Vector A's random value is [-1, 1] during the bubble-net phase, but during the searching phase, it can also be more or less than 1. The search algorithm is displayed in equation (6)

$$X(t+1) = X_{rand} - A\,.\,|C.X_{rand} - X(t)| \quad (6)$$

This random search mechanism requires the WOA algorithm to conduct a global search and emphasises the searching operation with a value of |A| larger than one. At the start of the WOA search processes, randomized solutions are generated [24]. The procedure is then used to update these solutions iteratively. The search will continue up to a predetermined maximum number of iterations. The features with the highest fitness value will be chosen as the ideal features weight value once this new solution's fitness has been compared.

Algorithm for Whale Optimization

START

1. importing data
2. establish the locations of the whale population X
3. determine whales' fitnesses
4. Compute A and C after initialising a and r.
5. establish X* as ideal positions for hunting whales.
6. Initialize X** as the top choice.
7. initialise t to 7
8. **while** $t \leq$ max iterations **do**
9. **for** each hunting whale **do**
10. **if** $p < 0.5$
11. **if** |**A**| < 1
12. update current hunting whales' location' using (4)
13. **else if** |**A**| ≥ 1
14. randomly choose different search agents
15. update current hunting whales' locations using (5)
16. **end if**
17. **else if** $p \geq 0.5$
18. update current hunting whales' location' using (6)
19. **end if**
20. **end for**
21. Update X* if a better solution becomes available.
22. Obtain a fresh set of answers
23. determine each whale's fitness
24. Update X** if a better solution becomes available. $t = t + 1$
25. **end while**
26. output **X****

END

3.5. Hybrid neural network model

The main concept of the hybrid technique is to combine a number of models, each of which serves the same starting aim, in order to generate a superior composite global model with more precise and dependable estimates or conclusions than can be reached by using a single model. The hybrid classification model consists of CNNs with fuzzy weights and KSVM (kernel-based SVM), as described in the section below.Therefore, integrating their forecasts is probably going to increase the ability to detect bogus news. Hybrid classifiers therefore outperformed solo classifiers in terms of performance.

3.5.1. Fuzzy Weight based CNN

In the suggested study project, fake news classification is done by using fuzzy weight based CNN (FWCNN). Convolution layers, sub sample or pooling layers, and fully linked layers make up this category of networks. These three layers make up the majority of the structure of this sort of network [25]. The network consists of an input layer that uses the chosen characteristics as an input, an output layer from which it obtains the trained output, and intermediary layers that are referred to as the hidden layers and are seen in figure 2.

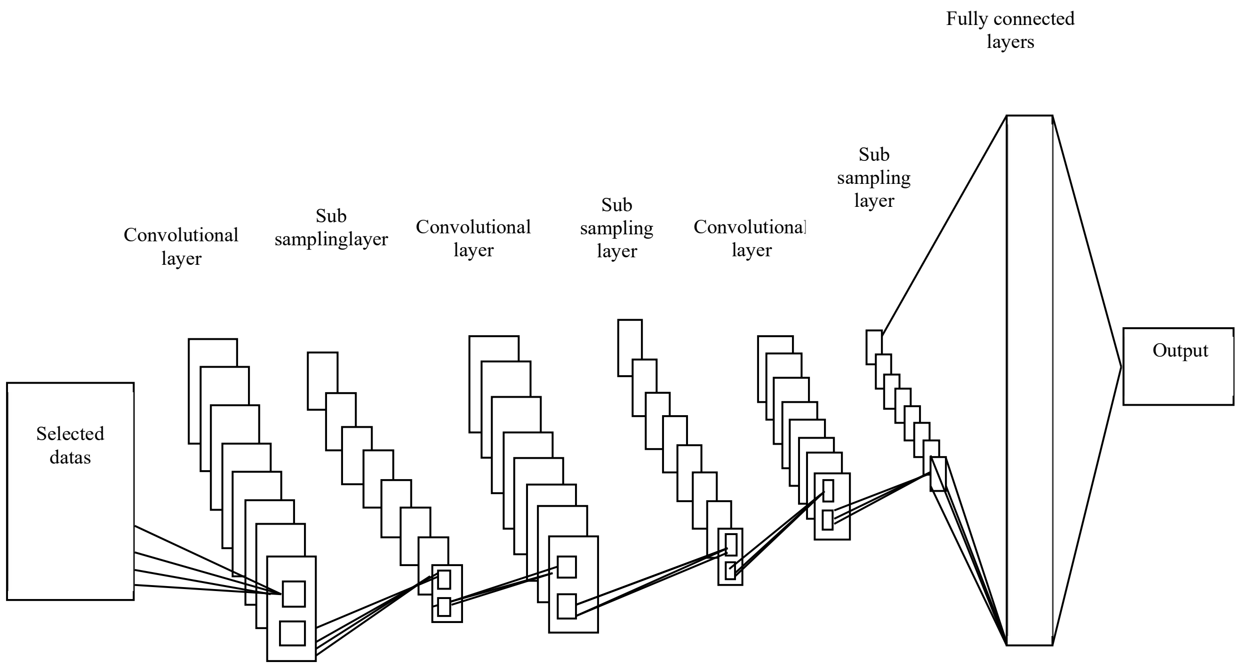

Figure 2: CNN (CNN)

Convolution layer

In the work that is being suggested, some data will be used as input. The CNN network starts off with a convolution layer as its very first layer. Convolution occurs in this layer by combining the input characteristics with a kernel (filter). The result is formed by separately convolving each feature of the input matrix with the kernel. The output, denoted by n, is produced by making use of the result of the convolution that was performed using the input and the kernel. In common parlance, a kernel of a convolution matrix is known as a filter, and the output that is created by convolving the kernel with the input is known as feature maps of size i*i.

The inputs and outputs of the subsequent convolutional layers are feature vectors, and the CNN may have many convolutional layers. Each convolution layer has a number of n filters. These filters are convolved with the input, and the number of filters used in the convolution process is equal to the depth of the resulting maps (n). Keep in mind that every filter map is regarded as a distinct feature at a particular point of the input.

The l-th convolution layer's output is indicated by the symbol $C_j^{(l)}$, is made up of maps. The calculation is

$$C_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{a_i^{(l-1)}} K_{i,j}^{(l-1)} * C_j^{(l)} \quad (7)$$

Where, $B_i^{(l)}$ is also known as the bias matrix$K_{i,j}^{(l-1)}$links the j-th feature map in layer using a convolution filter $(l-1)$ including the i[th] feature map into the same layer as the others the results of the work$C_i^{(l)}$ map features make up the layer in question. In (7), the first layer of convolutions in the network $C_i^{(l-1)}$ is a place for input, which is, $C_i^{(0)} = X_i$.

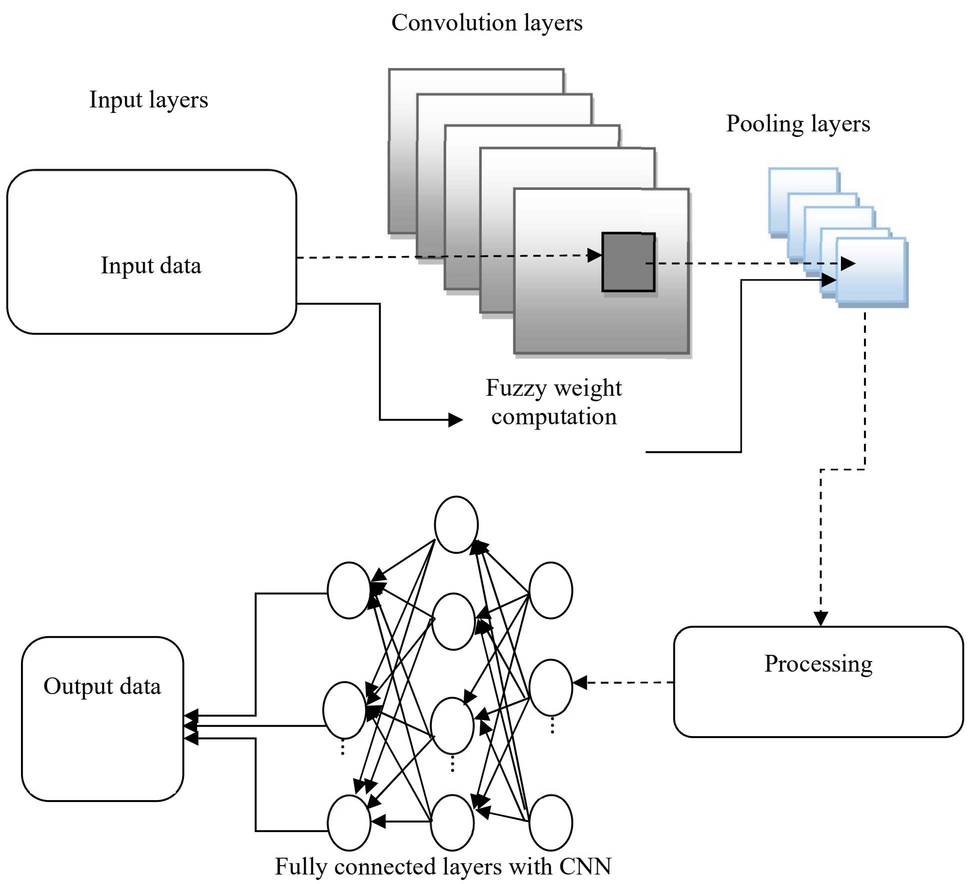

Figure 3: Fuzzyweight basedCNN (FWCNN)classifier It is the kernel that produces the feature map. The activation function may then be used to perform nonlinear transformations on the outputs of the convolutional layer after it has been implemented following the convolution layer:

$$Y_i^{(l)} = Y(C_i^{(l)}) \quad (8)$$

Where,

$Y_i^{(l)}$ - result of the activating factor being applied

$C_i^{(l)}$- the information that is fed into it.

Sigmoid, tanh, and rectified linear units are the three activation functions that are most often utilized (Sigmoid). Within the scope of this study, weighted Sigmoid, also known as $Y_i^{(l)} = w.\max(0, Y_i^{(l)})$ are used. Because of its effectiveness in mitigating the negative impacts of interaction and nonlinearity , this function is often used in DL models. To further enhance performance, the feature weight value has been included here. If the Sigmoiddetects that the input is negative, it will set the output to 0, but if the input is positive, it will return the same value as the positive input. This activating function has a benefit over other functions in that it allows for speedier learning as a direct result of the errors derivatives being extremely tiny in the saturation area. As a consequence of this, essentially no changes are made to the weights during training. This particular issue is referred to as the vanishing gradient problem. avoiding this issue should be your top priority add a weight for the fuzzy that has been adjusted in the convolution section. The weight value is then determined to those fuzzy weight values. Instead of considering the random weight values, convolution kernel value is computed based on the importance of the feature. The major advantage of using convolution weight is to reduce the error of the classifier and it considers the importance of features are extracted from the data. Since, the weight is generated based on the extracted feature from the kernel (figure 3).

Sub sampling or pooling Layer

The convolutional layer is the one that is followed by the sub sampling layer. The primary objective of this layer is to geographically lower the dimensionality of the feature maps that were recovered from the layer of convolution that came before it . It does an average after first dividing the features into blocks of size 2x2 and then doing the division. The relative information between features is what is preserved by the sub sample layer, not the precise connection between the features [26]. Take note that the presence of a sub sampling layer enables the convolution layer to better withstand rotation and translation among the input data.

Fully Connected layer

The soft max activation function is used on the output layer:

$$Y_i^{(l)} = f(z_i^{(l)}), \qquad (9)$$

where, $z_i^{(l)} = \sum_{i=1}^{m_i^{(l-1)}} w_{i,j}^{(l)} y_i^{(l-1)}$

Where, $w_{i,j}^{(l)}$ are the weight values that the whole fully linked layer must adjust in order to create the representation of each class, and f is the transfer function that symbolizes the nonlinearity.

3.5.2. Support Vector Machine

SVM accurately forecasts decision surfaces. The creation of probability distributions across the training set does not need the application of a model. Excellent performance indicators are present. The margin is defined as the separation between the hyper-planes [27]. The support vectors are determined by the closest outside and inside of class hyperplanes. The optimisation procedure is given structure by the structural risk minimization (SRM) concept. The hyperplane that maximises the margin while minimising the empirical risk is the optimum hyperplane. As a result, there is greater generality. The SVM algorithm's schematic is shown in Fig. 4.

The training examples are used to define an SVM classifier. Data that can only be separated using a nonlinear decision surface is used in real-world classification [28]. The usage of a kernel-based transformation is included in the optimization of the input data in this situation.

$$K(X_i, X_j) = \Phi(X_i).\Phi(X_j) \quad (10)$$

Kernels compute dot products in a higher dimensional space without explicitly mapping the data into these spaces. The following is a decision function:

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i K(X, X_i) + b \quad (11)$$

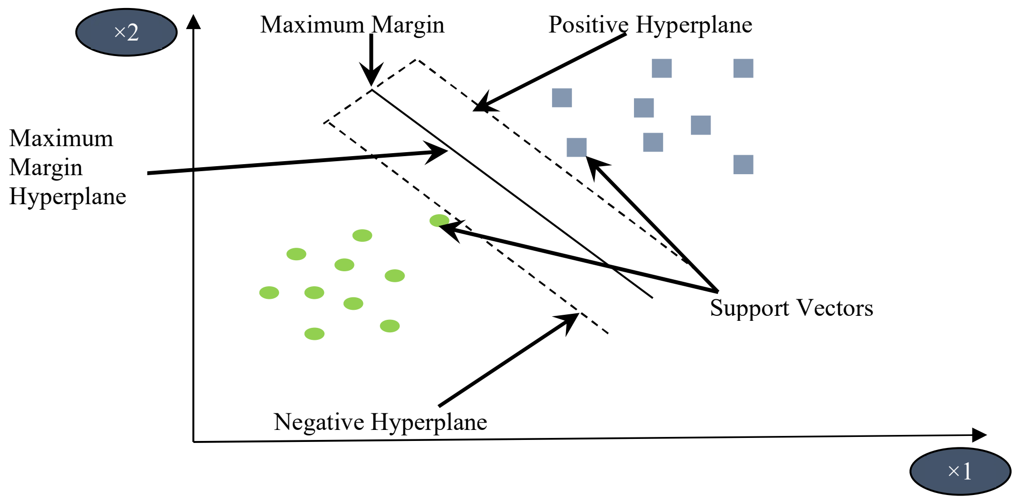

Fig 4. Structure of SVM Algorithm This work explores commonly used kernel function is,

$$K(x, y) = exp\{-\gamma|x - y|^2\} \quad - radial\ basis\ function \quad (12)$$

A data-dependent kernel known as a radial basis function (RBF) kernel has evolved as a potent alternative tool. RBF kernels have a slower convergence time than polynomial kernels, but they provide superior results. Dot products are used in the classification process. The number of support vectors scales linearly with the classification job. For non-separable data, soft margin classifiers are utilised. Slack variables are used to ease the separation requirements.

$$x_i.w + b \geq +1 - \xi_i, \quad for\ y_i = +1 \quad (13)$$

$$x_i.w + b \leq -1 + \xi_i, \quad for\ y_i = -1 \quad (14)$$

$$\xi_i \geq 0, \quad \forall\ i \quad (15)$$

According to the results of the preceding investigation, the proposed hybrid classification methodology is a very efficient method for assessing fakedata. So the hybrid model based classification approach has highly efficient technique for the detection of fake news in an efficient manner.

4. Results and Discussion

This model was implemented in Matlab. Several datasets have been proposed to detect fake news. This study uses the ISOT false news dataset to train hybrid models. This work considers the existing RFclassifier and proposed ECM algorithm for assessing performance measures, including f1 score, accuracy, precision, and recall. Real and fraudulent news pieces are both included in the collection.

4.1. The ISOT fake news dataset

This dataset includes news articles from Reuters.com and Kaggle.com, respectively, for the real news and false news categories. Each example in the collection has more than 200 characters. Each article, whether phoney or real, has metadata such as Article kind, Article

Text, Article Title, Article Date, and Article Label. Table 1 displays the size and type of articles for the real and fake categories, respectively.

4.2. Evaluation metrics

The effectiveness of this study was assessed using the provided dataset in conjunction with current methodologies like Capture, Score, and Integrate (CSI), Capsule Neural Network, and Hybrid Feature Extraction Model with Capsule Neural Network (HFEM-CNN), and a hybrid classification model ((WOA-HYC) was proposed. The section below defines the evaluation measure used in the trials. Then, by assessing the true positive (TP), false positive (FP), true negative (TN), and false negative (FN) rates, a number of performance metrics are derived. The initial performance metric was precision, which is the proportion of relevant occurrences among those that were retrieved. The second performance metric was recall, or the percentage of pertinent occurrences that were recovered.

Despite potentially clashing with one another, accuracy and recall measurements are crucial for assessing a prediction strategy. These two metrics can be combined with equal weights to provide a single metric, the F-measure. The last performance metric to be established was accuracy, which was defined as proportions of properly anticipated occurrences to all predicted instances.

Precisions are ratios of correctly detected positive observations to total anticipated positive observations.

$$\text{Precision} = TP/TP+FP \qquad (16)$$

Sensitivities or recall ratios are ratios of correctly identified positive observations to all observations.

$$\text{Recall} = TP/TP+FN \qquad (17)$$

F-measures are weighed averages of Precisions and Recalls and hence consider false positives and false negatives.

$$\text{F1 Score} = 2*(\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision}) \qquad (18)$$

These are the positives and negatives used to calculate accuracy:

$$\text{Accuracy} = (TP+FP)/(TP+TN+FP+FN) \qquad (19)$$

Table .1. Article size and kind by category for the ISOT Dataset

News type	Size	Subject	
		Type	Size
Real-News	21 417	World-News	10 145
		Politics-News	11 272
Fake-News	23 481	Government-News	1570
		Middle-east	778
		US News	783
		Left-News	4459
		Politics	6841
		News	9050

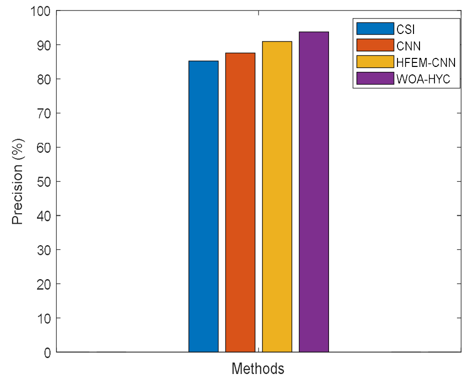

Fig.5. shows comparative precision findings of the suggested and current fake news detection models

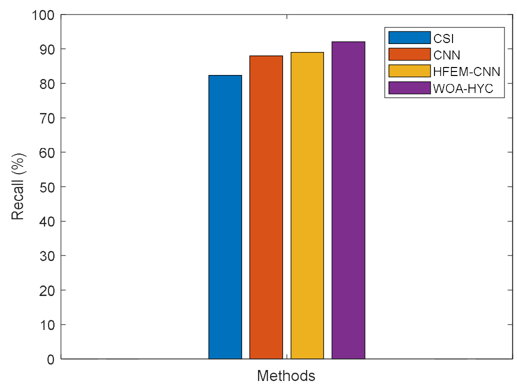

Fig.6. shows comparative recall findings of the suggested and current fake news detection models

The memory comparison between the suggested and existing false news detection techniques is shown in Figure 6. The findings demonstrate that the suggested WOA-HYC technique outperforms the already in use categorization algorithms in terms of recall.

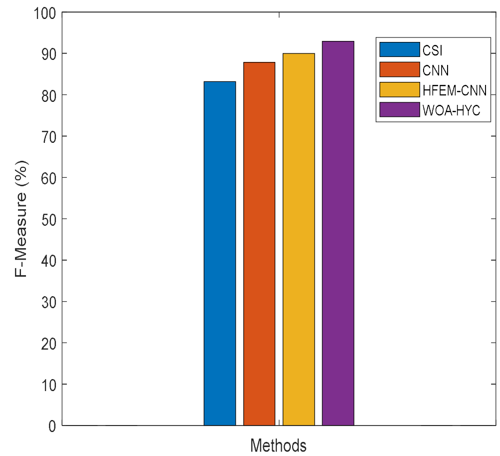

Fig.7. shows comparative F-measure values of the suggested and existing fake news detection models

Using F-measure values, Fig. 7 compares proposed and present false news detection approaches. The results show that the proposed WOA-HYC approach beats the currently used classification algorithms in terms of F-measure.

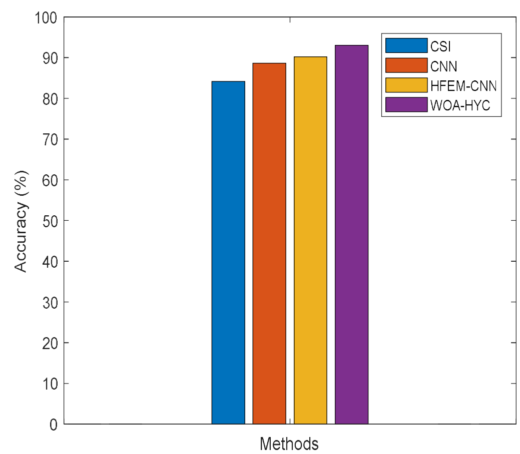

Fig.8. The proposed and current fake news detection models' accuracy comparison findings

Figure 8 displays the accuracy comparisons of the suggested and existing false news detection methods. The results demonstrate that the proposed WOA-HYC methodology provides good accuracy results when compared to the existing classification methodologies.

5. Conclusion

This paper describes a unique hybrid strategy to detecting fake news in social datasets. The data is pre-processed by removing any repeated words or symbols, such as stop words and numbers. The fake news data set was subjected to feature extraction in order to minimise the size of the feature space. A classifier based on a hybrid classification model is then constructed to detect fake news in the given datasets. The ISOT data set was used to calculate the accuracy, recall, precision, and F-measure scores for this combination model. The exploration of novel techniques and the use of clever hybrid classification algorithms for better outcomes may enhance the current work in following efforts. For the dataset, the suggested method has an average accuracy of 92.99%,.In future work, enhance the efficiency and accuracy of hybrid classification techniques and compare them with other well-established classification schemes over big data.

REFERENCES

1. Zhang, X. and Ghorbani, A.A., 2020. An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, *57*(2), p.102025.
2. Reis, J.C., Correia, A., Murai, F., Veloso, A. and Benevenuto, F., 2019. Supervised learning for fake news detection. *IEEE Intelligent Systems*, *34*(2), pp.76-81.
3. Thota, A., Tilak, P., Ahluwalia, S. and Lohia, N., 2018. Fake news detection: a deep learning approach. *SMU Data Science Review*, *1*(3), p.10.
4. Monti, F., Frasca, F., Eynard, D., Mannion, D. and Bronstein, M.M., 2019. Fake news detection on social media using geometric deep learning. *arXiv preprint arXiv:1902.06673*.
5. Sahoo, S.R. and Gupta, B.B., 2021. Multiple features based approach for automatic fake news detection on social networks using deep learning. *Applied Soft Computing*, *100*, p.106983.
6. Lu, Y.J. and Li, C.T., 2020. GCAN: Graph-aware co-attention networks for explainable fake news detection on social media. *arXiv preprint arXiv:2004.11648*.
7. Nasir, J.A., Khan, O.S. and Varlamis, I., 2021. Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights*, *1*(1), p.100007.
8. Yang, Y., Zheng, L., Zhang, J., Cui, Q., Li, Z. and Yu, P.S., 2018. TI-CNN: Convolutional neural networks for fake news detection. *arXiv preprint arXiv:1806.00749*.
9. Kaliyar, R.K., Goswami, A., Narang, P. and Sinha, S., 2020. FNDNet–a deep convolutional neural network for fake news detection. *Cognitive Systems Research*, *61*, pp.32-44.
10. Balwant, M.K., 2019, July. Bidirectional LSTM based on POS tags and CNN architecture for fake news detection. In *2019 10th International conference on computing, communication and networking technologies (ICCCNT)* (pp. 1-6). IEEE.
11. Saleh, H., Alharbi, A. and Alsamhi, S.H., 2021. OPCNN-FAKE: Optimized convolutional neural network for fake news detection. *IEEE Access*, *9*, pp.129471-129489.
12. Krishna, N.L.S.R. and Adimoolam, M., 2022, February. Fake News Detection system using Decision Tree algorithm and compare textual property with Support Vector Machine algorithm. In *2022 International Conference on Business Analytics for Technology and Security (ICBATS)* (pp. 1-6). IEEE.

13. Gravanis, G., Vakali, A., Diamantaras, K. and Karadais, P., 2019. Behind the cues: A benchmarking study for fake news detection. *Expert Systems with Applications*, *128*, pp.201-213.
14. Albahar, M., 2021. A hybrid model for fake news detection: Leveraging news content and user comments in fake news. *IET Information Security*, *15*(2), pp.169-177.
15. Ahmed, S., Hinkelmann, K. and Corradini, F., 2022. Development of fake news model using machine learning through natural language processing. *arXiv preprint arXiv:2201.07489*.
16. Meesad, P., 2021. Thai fake news detection based on information retrieval, natural language processing and machine learning. *SN Computer Science*, *2*(6), p.425.
17. Narra, M., Umer, M., Sadiq, S., Karamti, H., Mohamed, A. and Ashraf, I., 2022. Selective Feature Sets Based Fake News Detection for COVID-19 to Manage Infodemic. *IEEE Access*, *10*, pp.98724-98736.
18. Mridha, M.F., Keya, A.J., Hamid, M.A., Monowar, M.M. and Rahman, M.S., 2021. A comprehensive review on fake news detection with deep learning. *IEEE Access*, *9*, pp.156151-156170.
19. Felber, T., 2021. Constraint 2021: Machine learning models for covid-19 fake news detection shared task. *arXiv preprint arXiv:2101.03717*.
20. Zhou, X., Jain, A., Phoha, V.V. and Zafarani, R., 2020. Fake news early detection: A theory-driven model. *Digital Threats: Research and Practice*, *1*(2), pp.1-25.
21. Khan, J.Y., Khondaker, M.T.I., Afroz, S., Uddin, G. and Iqbal, A., 2021. A benchmark study of machine learning models for online fake news detection. *Machine Learning with Applications*, *4*, p.100032.
22. Alonso, M.A., Vilares, D., Gómez-Rodríguez, C. and Vilares, J., 2021. Sentiment analysis for fake news detection. *Electronics*, *10*(11), p.1348.
23. Nasiri, J. and Khiyabani, F.M., 2018. A whale optimization algorithm (WOA) approach for clustering. *Cogent Mathematics & Statistics*, *5*(1), p.1483565.
24. Aljarah, I., Faris, H. and Mirjalili, S., 2018. Optimizing connection weights in neural networks using the whale optimization algorithm. *Soft Computing*, *22*, pp.1-15.
25. Kundu, R., Basak, H., Singh, P.K., Ahmadian, A., Ferrara, M. and Sarkar, R., 2021. Fuzzy rank-based fusion of CNN models using Gompertz function for screening COVID-19 CT-scans. *Scientific reports*, *11*(1), p.14133.
26. Mohamed, B., Haytam, H. and Abdelhadi, F., 2022. Applying fuzzy logic and neural network in sentiment analysis for fake news detection: case of Covid-19. *Combating fake news with computational intelligence techniques*, pp.387-400.
27. Poddar, K. and Umadevi, K.S., 2019, March. Comparison of various machine learning models for accurate detection of fake news. In *2019 Innovations in Power and Advanced Computing Technologies (i-PACT)* (Vol. 1, pp. 1-5). IEEE.
28. Albahar, M., 2021. A hybrid model for fake news detection: Leveraging news content and user comments in fake news. *IET Information Security*, *15*(2), pp.169-177.