mathematics

MDPI

Article

Web-Informed-Augmented Fake News Detection Model Using Stacked Layers of Convolutional Neural Network and Deep Autoencoder

Abdullah Marish Ali [1], Fuad A. Ghaleb [2,3,*], Mohammed Sultan Mohammed [4], Fawaz Jaber Alsolami [1] and Asif Irshad Khan [1]

[1] Department of Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia; ammali@kau.edu.sa (A.M.A.); falsolami1@kau.edu.sa (F.J.A.); aikhan@kau.edu.sa (A.I.K.)
[2] Department of Computer Science, Faculty of Computing, Universiti Teknologi, Malaysia, Johor Bahru 81310, Malaysia
[3] Department of Computer Engineering and Electronics, Sanaá Community College, Sanaá 5695, Yemen
[4] Faculty of Electrical Engineering, Universiti Teknologi, Malaysia, Johor Bahru 81310, Malaysia; samohammed@utm.my
* Correspondence: abdulgaleel@utm.my

Abstract: Today, fake news is a growing concern due to its devastating impacts on communities. The rise of social media, which many users consider the main source of news, has exacerbated this issue because individuals can easily disseminate fake news more quickly and inexpensive with fewer checks and filters than traditional news media. Numerous approaches have been explored to automate the detection and prevent the spread of fake news. However, achieving accurate detection requires addressing two crucial aspects: obtaining the representative features of effective news and designing an appropriate model. Most of the existing solutions rely solely on content-based features that are insufficient and overlapping. Moreover, most of the models used for classification are constructed with the concept of a dense features vector unsuitable for short news sentences. To address this problem, this study proposed a Web-Informed-Augmented Fake News Detection Model using Stacked Layers of Convolutional Neural Network and Deep Autoencoder called ICNN-AEN-DM. The augmented information is gathered from web searches from trusted sources to either support or reject the claims in the news content. Then staked layers of CNN with a deep autoencoder were constructed to train a probabilistic deep learning-base classifier. The probabilistic outputs of the stacked layers were used to train decision-making by staking multilayer perceptron (MLP) layers to the probabilistic deep learning layers. The results based on extensive experiments challenging datasets show that the proposed model performs better than the related work models. It achieves 26.6% and 8% improvement in detection accuracy and overall detection performance, respectively. Such achievements are promising for reducing the negative impacts of fake news on communities.

Keywords: fake news detection; web-informed; augmented information; misinformation; two-stage classification; deep learning; stacked learning; CNN; deep autoencoder

MSC: 68-00; 68T50; 68T07

Citation: Ali, A.M.; Ghaleb, F.A.; Mohammed, M.S.; Alsolami, F.J.; Khan, A.I. Web-Informed-Augmented Fake News Detection Model Using Stacked Layers of Convolutional Neural Network and Deep Autoencoder. *Mathematics* **2023**, *11*, 1992. https://doi.org/10.3390/math11091992

Academic Editors: Nebojsa Bacanin and Catalin Stoean

Received: 12 March 2023
Revised: 15 April 2023
Accepted: 19 April 2023
Published: 23 April 2023

1. Introduction

Fake news has emerged as a major challenge for journalism, democracy, and the freedom of expression. Its impact on public trust in government is a cause for concern, and its potential influence on critical events like the 2016 U.S. presidential election and the disputed "Brexit" vote remains unclear [1]. The issue has been exacerbated by social media platforms like Facebook and Twitter, which provide users with fast, inexpensive, and

less regulated information than traditional news channels like newspapers and television. According to the study in [2], fake news on Twitter often receives much more user retweets than true news and spreads much more quickly, especially when it comes to political news. During the 2016 U.S. presidential election, the top 20 fabricated stories shared on Facebook generated over eight million shares, reactions, and comments [3]. False information not only exacerbates division and polarization but also deepens the divide as individuals and groups become more entrenched in their own beliefs, unwill-ing to engage in constructive discourse with those who have differing perspectives [4]. Furthermore, fake news can have serious consequences for public health and safety, such as spreading misinformation about COVID-19 or promoting unproven treatments [5]. Therefore, it is crucial to effectively detect and prevent the spreading of fake news.

Several approaches have been suggested to identify fake news. Numerous issues have been studied related to feature extraction [6–10], representation [8,11–17], classification [6,12,18–25], and model design [10,16,21,22,26–32]. Based on the representative features that are employed, fake news detection techniques can be can be categorized into four groups: content-, knowledge-, users-, and propagation-based features. The content- and knowledge-based features are widely used for detection due to their simplicity and effectiveness compared to the users- and propagation-based features. Although content-based features have been reported as the most effective if proper representation, classification, and model design are used, this view is built on the assumption that fake news is highly distinguishable when compared to true news. The type of the selected features plays an essential role in determining the representation technique, the training algorithm, and the model design. Thus, representation techniques such as TF/IDF, Word2Vec, and GloVe have been investigated [33]. Word-embedding techniques have been extensively discussed in the literature as these techniques take into account the semantics of news content. Various approaches, including both conventional machine learning and deep learning techniques, have been thoroughly investigated [1,34–37]. Deep learning techniques have been predominantly utilized for classification tasks due to their capability to automatically extract latent features without manual intervention. However, most of the existing solutions are poorly performed for non-traditional fake news, such as the items tweeted and posted on social media. Such news is short and contains insufficient and sparse features resulting in low performance of the constructed deep-learning models.. As a result, a suitable model design that can accurately identify fake news has not yet been proposed. Detecting fake news remains challenging despite various approaches that have been developed. The content-based approach is commonly used due to its simplicity, but it is less effective for short sentences, as seen in the LIAR dataset. This approach also faces the challenge of fake news agents mixing facts with false information to appear more credible. Relying solely on content-based features results in poor detection, highlighting the need for relevant external features. Context-based detection algorithms are believed to improve performance, but few recent studies have attempted to use them effectively (as in [23,31]). However, these studies blindly extract information from web searches, resulting in irrelevant and noisy features that impede effective learning and increase the complexity of detection.

Two key factors need to be considered for the receiver to accurately detect fake news: effective news representative features and appropriate model design. Despite this, many of the current solutions rely solely on content-based features, which are inadequate and can overlap. Additionally, most of the models used for classification are based on the concept of dense feature vectors, which are unsuitable for short news sentences. To overcome this issue, this study introduces a Web-Informed-Augmented Fake News Detection Model using Stacked Layers of Convolutional Neural Networks and Deep Autoencoder (ICNN-AEN-DM). The contribution of this study lies in the development of a novel Web-Informed Augmented Fake News Detection Model, which addresses the limitations of existing models by enriching content-based features with additional features gathered from trusted sources through a web search. The augmented information thereby gathered will either support or reject the claims in the news content. Then, a stacked layer of CNN with a

deep autoencoder is trained using probabilistic deep learning. The probabilistic outputs of the stacked layer are used to train decision-making stacked layers based on multilayer perceptron (MLP). After conducting numerous experiments on difficult datasets, it was demonstrated that the model proposed in this study outperformed the state-of-the-art models. The research provided the following contributions.

1. An augmented information-gathering algorithm is proposed to support the news sentences by enriching the representative features. The augmented information is retrieved from reliable sources using Google search techniques, and only the relevant news features are utilized as new knowledge to reduce the sparsity and improve the detection accuracy.
2. A model based on probabilistic deep-learning architecture has been developed for detecting fake news. The model involves stacking convolutional neural network layers with deep autoencoder layers. The last layer's probabilistic output serves as a new representative feature for the second stage of learning, which involves stacking multilayer perceptron layers into the CNN and the autoencoder layers.
3. The proposed model was put through extensive experiments to test its validity and effectiveness. Two datasets, specifically LIAR [38] and ISOT [39], were used as benchmarks. The outcomes were then compared to those of other leading models currently available.

The rest of this paper is organized as follows: Section 2 covers related research; Section 3 describes the proposed model in detail; Section 4 outlines the experimental design and methodology employed to verify and assess the model; Section 5 elucidates the results and discussion; and Section 6 concludes the study.

2. Related Work

The detection of fake news has been a prominent focus of numerous studies in the academic literature. Several survey and review papers that contain detection challenges and potential open issues have been published recently, including but not limited to [34–37,40–43]. Many issues have been researched related to feature extraction [6–10], representation [8,11–17], classification [6,12,18–25], and model design [10,16,21,22,26–32]. Various solutions have been investigated using statical, traditional machine and deep learning and natural language-processing techniques. However, accurate detection of fake news is a complicated task, and there are still many challenges to overcome, such as poor detection performance, detection of non-traditional fake news such the partially incorrect statement, lack of a gold-standard dataset, and automatic rational analysis based on facts and domain expertise, early detection in social media, and understanding how it spreads [43].

In terms of features extraction, fake news detection approaches can be categorized into four groups: knowledge-based [11,23,44,45], content-based [6,13,23,46], user-based [33,47,48], and propagation-based [1,43,48]. The knowledge-based includes fact-checking and crowdsourcing. This approach assumes that true news is consistent with facts, while false news contradicts them. However, the knowledge-based approach can be limited by the availability and accuracy of external knowledge sources. Many expert-based fact-checking websites have been developed to improve the detectability of fake news. Datasets such as LIAR [38] and FakeNewsNet are annotated with the help of the PolitiFact (http://www.politifact.com/ accessed on 12 February 2023) and GossipCop (https://www.gossipcop.com, accessed on 4 April 2022) fact-checking websites, respectively. The content-based approach involves extracting linguistic and syntactical features, analyzing sentiment, and examining latent features. In the content-based approach, the content of news articles is analyzed to detect fake news. It involves using natural language processing (NLP) and machine-learning techniques to extract features such as sentiment, emotion, and linguistic patterns from the text. Such an approach can be effective, especially when combined with external knowledge sources, but it can also be limited by the ability of the models to detect sophisticated forms of disinformation. The user-based approach looks after the credibility of the users and their age, reputation, class, and followers, among many other criteria. It analyses social media

user profiles and the subjects' behavior to identify patterns that may indicate a propensity for sharing fake news. This approach can be useful, but it can also be limited by privacy concerns and the fact that not all users who share fake news exhibit the same patterns. The propagation-based approach focuses on the spreading patterns of the news, such as speed, number of shares, number of likes, comments, etc. It involves using graph theory and network analysis to track the spread of fake news and identify the most influential nodes in the network. This approach can be useful, especially when combined with other approaches, but it can also be limited by the difficulty of tracking and analyzing the massive volume of data produced by social networks. This study focuses on content- and knowledge-based solutions due to the complexity of user- and propagation-based approaches.

Features representation is another aspect that affects the detection performance. While knowledge-based features are represented in Knowledge-Graph, most of the content-based features are represented either using statistics or using features-embedding techniques. The N-gram technique is also used to create more textual features. It has been demonstrated in some previous studies [20,32,49] that TF-IDF is effective in detecting fake news. However, to gather additional semantic details, language representation models were created that were pre-trained, also known as context-independent models, such as Word2Vec, global vectors (GloVe), and Bi-Directional Encoder Representations from Transformers (BERT) [33]. These models captured semantic patterns effectively, but they did not capture much contextual information [50]. Word2Vec is a technique to train pre-trained word embedding models using a neural network based on a given text corpus. GloVe is an extension of Word2Vec, a representation combined with global statistics from matrix-factorization techniques such as latent semantic analysis (LSA). In their study, Samadi and Mousavian [8] examined several deep contextualized models for text representation and suggested multiple deep-learning classifiers. Different pre-trained models, such as Funnel, GPT2, BERT, and RoBERTa, were analyzed, and the embedding layer was linked to various classifiers such as CNN, SLP, and MLP for classification purposes. The results revealed that the combination of Funnel and CNN was superior to other existing models. However, the above approach yielded low accuracy in predicting outcomes for the LIAR dataset, achieving only 48%. Ali and Ghaleb [20] demonstrate that TF-IDF with a proper detection design model outperforms the other complex designs and representations such as Funnel, GPT2, BERT, and RoBERT that were used in [8]. However, the representation of short news sentences leads to sparse feature vectors. The sparsity handles effective learning by machine learning algorithms.

In the quest to identify fake news, a variety of solutions have been developed, including those that leverage machine learning and deep learning techniques. [1,34]. Various ensemble methods have also been explored, combining classic and deep machine learning techniques. Nasir and Khan [27] utilized the ISOT and FA-KES datasets to train and evaluated a fake news classification model that integrated CNN and RNN models. However, feature extraction was solely based on the news content, which may not be practical for effective fake news detection. They used the GloVe pre-trained word embedding model for feature extraction. An ensemble fake news detection model was proposed by Hakak and Alazab [10]. A total of 26 features were derived from news content, encompassing various statistics related to words, characters, and sentences. Hakak and Alazab [10] also used named entity recognition to extract statistical features related to entities like people, organizations, and dates. Although the results exhibited an improvement in accuracy compared to existing models, the extracted features caused overfitting and were not generalizable to short news sentences in the LIAR dataset. In Ref. [50], The researchers performed a comparative analysis of conventional machine learning techniques and deep learning methods to identify false information.

Context-based detection algorithms are believed to be effective in improving the performance of automatic false news detection. However, a few recent studies research endeavors have been exploring context-based detection techniques, as seen in studies, such as [23,31]. Vishwakarma and Varshney [31] suggested a model that measures the reliability of links identified via Google search results to detect fake news. Their approach

is remarkable because it makes effective use of link information to identify fake news. However, such a model relies on a whitelist of "reliable links," which needs to be created manually. As web links can pertain to both genuine and fake news, relying solely on a whitelist may yield inadequate results. Moreover, the method in [31] involves building the whitelist manually, which requires expertise and labor, making it prone to bias and errors. Shimand Lee [23] presented a model for fake news detection based on the link2vec embedding technique which is extended from word2vect. A web search was performed for each news sentence and the resulting web links were used for constructing the link2vec embedding model. Although such an approach has outperformed the model that was proposed in [31], whether the results of the search has dynamic relevancy to the original news sentence depends on the availability of similar keywords in the stored links. As the same web links can relate to both genuine and fake news, such a technique may not be effective for short news sentences or news with partial correctness.

To sum up, although several approaches have been devised to detect false news, fake news detection is still challenging. The content-based approach is the most reported in the literature due to its simplicity of extracting features. It is also effective for obvious false information such as in the ISOT dataset. However, for short sentences like in the LIAR dataset, most of the existing solutions are poorly performed. The problem of fake news detection can be attributed to two main challenges: first, fake news has overlapped features with true news and proper model design; second, Fake news agents usually mix facts with false information to be more convincing and seemingly rational. Accordingly, solely depending on content-based features leads to poor detection performance. Therefore, extracting external yet relevant features related to the news content helps in reducing the sparsity and increasing the distinguishability. Context-based detection algorithms are believed to be effective in improving the performance of automatic false news detection. However, few recent studies attempt to use context-based detection such as in [23,31]. These studies blindly extract the information from the web search, thereby extracting irrelevant, noisy features that hinder effective learning and increase the detection complexity. To address this gap, in this study, information related to the news content is gathered from a set of reliable sources based on the Google search engine. Then, only results with high similarity to the news are retrieved. Then, convolutional neural network layers stacked with deep autoencoders and multilayer perceptron layers were designed with two stages of learning for the multi-class classification task. The contribution of this study is the proposal of a Web-Informed-Augmented Fake News Detection Model called ICNN-AEN-DM, which addresses the limitations of existing solutions by gathering augmented information from trusted sources to support or reject the claims in news content. The model uses stacked layers of Convolutional Neural Network and Deep Autoencoder to train a probabilistic deep-learning base classifier and multilayer perceptron (MLP) layers for decision-making. A comprehensive overview of the proposed model is presented in the subsequent section.

3. The Proposed Fake News Detection Model

The components of the proposed fake news model (ICNN-AEN-DM) are shown in Figure 1. The proposed model comprises five phases, specifically, Web-based augmented information collection to support the claims in the original news sentences, data cleaning, pre-processing, features representation using GloVe, probabilistic deep learning, which contains CNN layers stacked with deep autoencoders for probabilistic learning, and the decision-making phase, which encompasses multilayer perceptron (MLP). The comprehensive explanation of each phase is presented in the following sections.

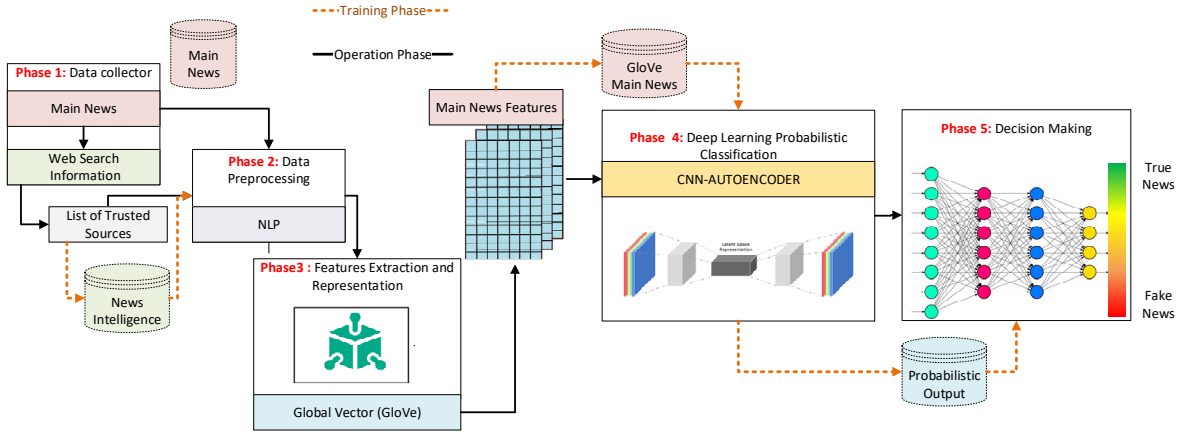


Figure 1. The proposed ICNN-AEN-DM model.

3.1. Augmented Information-Gathering Phase

In this phase, information about the news is gathered from trusted web sources. The claim is extracted from the main news, and then augmented information about the claim is collected from different sources. The augmented information is collected based on the similarity of the claim with what is presented in the external sources. The augmented information is used to find patterns from the sources that either support or reject the claim. The main challenge in this phase is finding whether a news source supports or rejects the claim. Text summarizations, language models, sentiment analysis, and expert systems are useful for this task. In this study, Algorithm 1 lists the steps on how the augmented information was gathered related to each news text. Assuming l denotes a trusted news source and L is the set of trusted news sources, in line 1 for each news weblink l in the set of weblinks of the trusted news sources L. For each sentence s in the news body S, research the related news sentence on the source l. The search results are cleaned and then for each sentence r in the results set R, use a pre-trained word embedding technique, namely, the Global Vectors for Word Representation *GloVe* to extract the representative numerical form of the sentences (see lines 3 to 6). The *GloVe* is used here to capture the semantic relationships of the researched sentence with the original news sentence. The Euclidian distance between the embedded vector of the claim s' and the extracted news r' are used to compare their similarities. If the similarity is less than the specific threshold, then the sentence is correlated and will be taken to either support or reject the claim. The threshold T is used here to reduce the noise by removing the irrelevant text. The trusted news sources can be determined dynamically based on the origin of the topic sentence in the news post. Examples of credible sources that can be used are Google News, the BBC, *The Guardian*, *The New York Times*, *The Wall Street Journal*, *The Washington Post*, and *The New Yorker*, among many others, and these should be determined by the user. In this study, the list of trusted news searches was determined dynamically using the Google search engine, i.e., the top 10 websites with high similarity with the original news were included. The idea behind collecting such information is to find the patterns of how different news classes can be expressed in different news sources.

Algorithm 1: Augmented Information Collecting Algorithm

$\forall\, l \in L\ do:$
  $\forall\, s \in S\ do:$
    $R \overset{search}{\leftarrow} g(l,\, c(s))$
    $\forall\, r \in R\ do:$
      $r' \overset{get}{\leftarrow} GloVe(r)$
      $s' \overset{get}{\leftarrow} GloVe(s)$
      $d(r',\, s') \overset{calculate}{\leftarrow} \sqrt{\sum_{i=1}^{n} (r' - s')^2}$
      $if(d(r',\, s') > T):$
        $I \overset{Append}{\leftarrow} r$
$return\ \ I$

3.2. Data Pre-Processing Phase

This phase is crucial to ensure that the learning algorithm is effective and efficient in reducing noise and dimensionality while retaining important information [51]. This process typically begins by removing unnecessary characters and symbols, followed by removing stop words—common words like "the", "and", and "in" that do not contribute much to the text's meaning. Tokenization is then used to separate words into individual tokens for easier processing, and each token is converted to lowercase using case folding. Characters that are not alphabetic and do not have a substantial impact on the analysis are eliminated from the token. Finally, the text is subjected to stemming and lemmatization, which involves reducing words to their root form. Stemming removes suffixes to create a base form, while lemmatization uses a dictionary to convert words to their base form.

3.3. Word Embedding Phase

In this phase, the text sample is converted to numerical forms in a process called vectorization. There are several approaches to vectorization, including bag-of-words, TF-IDF, and word embeddings. The main disadvantage of bag-of-words and TF-IDF is the sparse representations of the text and the failure to include semantic meaning. Word embedding techniques can capture the semantic meaning and reduce the computational complexity of text classification tasks.

Word embedding techniques are widely used in natural language processing to embody words in a numerical form. There are various techniques for creating word embeddings, with some of the most popular being Word2Vec [52], GloVe [53], and FastText [54]. Word2Vec and GloVe use co-occurrence statistics to capture the meaning of words, whereas FastText also considers subword information. Because fake news has high overlapping features with true news, GloVe word embedding can be more effective in this context than Word2Vec and FastText. GloVe is based on matrix factorization and is designed to capture the co-occurrence statistics of words in a corpus. It has been shown to perform well in tasks such as word analogy and word similarity and is especially good at capturing global word relationships.

GloVe constructs feature vectors for words by leveraging co-occurrence statistics. To create these vectors, GloVe starts by constructing a co-occurrence matrix that records the frequency of each word's occurrence in the context of all other words in a vast collection of text. This matrix is then used to derive a set of weighted co-occurrence statistics that were used to construct the neural network model. The neural network learns to predict the probability of observing a certain co-occurrence count for a pair of words, given their respective feature vectors. The training process iteratively updates the feature vectors until the predicted co-occurrence counts are close to the observed co-occurrence counts. The resulting feature vectors encode semantic relationships between words, such that words that appear in similar contexts and have similar feature vectors. The GloVe feature vector can be thought of as a mathematical representation of a word that encodes its semantic and

syntactic relationships with other words in a corpus of news text. These vectors are then used as input to the deep-learning model. The dimensionality of the GloVe feature vector is a hyperparameter that can be tuned, based on the specific task and corpus of text. However, for simplicity, pre-trained word vectors were utilized, which can be downloaded from the following URL (https://nlp.stanford.edu/projects/glove/, accessed on 21 January 2023).

In this study, each news sample in the dataset is concatenated with the corresponding augmented information collected from the Google search. Figure 2 presents an example of the process of preparing the news sample for generating the represented impeded sequence. It also shows how the text news samples are transformed into a dense vector using the pre-trained model. In Figure 2, the news sample was extracted from the LIAR dataset (it belongs to the half-true class) for clarification. In Figure 2, the dotted lines represent the operations that were done before training or testing, while the solid lines represent operations performed during either training or testing. For example, constructing the dictionary of unique words was performed before creating the sequences, and constructing the embedding matrix was performed before the sequence mapping. The dictionary (lexicon) is constructed to contain all the unique words (tokens) in the dataset samples. An index has been assigned to each word in the dictionary ordered based on the frequency of the words in the whole dataset, i.e., small indices go to the most frequent words. Next, each news sample in the dataset is fed to the tokenizer to be converted to a sequence of integers (the integers are the equivalent word indices in the constructed dictionary). And then, the embedding matrix is constructed using the pre-trained embedding weights of GloVe (see Figure 2). That is, each word in the constructed dictionary is mapped to its equivalent word vector that is extracted from the pre-trained GloVe. The words that do not exist in the pre-trained GloVe are represented by zero vectors (see Figure 2). Then to unify the size of the sequences, the size of the longest vector was used as the unified size of the input features. Thus, each sequence is padded based on the maximum length of the input vectors. Consequently, all samples are represented as integer sequences with the same size. These sequences are inputted to the embedded layer along with the embedding matrix (GloVe weights) generated using the GloVe embedding technique. Therefore, each sample is represented by a set of vectors with 100 dimensions size each extracted from the embedding matrix. That is each word in the original sample is represented by its equivalent word vector extracted from the embedding matrix using the pre-trained GloVe model.

3.4. Probabilistic Classification Phase

In this phase, a convolutional neural network with an autoencoder model is designed for pre-classification. Figure 3 illustrates that the proposed model which comprises 11 layers, with the initial layer serving as the input layer. It receives the feature vector that was generated from the data collection phase and is represented as explained in the word embedding phase (see Section 3.3). The text samples were converted to sequences of integer numbers that were extracted from the constructed dictionary. To make all the samples have the same length, the size of the input features is assigned based on the maximum length of one sequence in the dataset. That is, the length of the input features vector is the size of the longest news sample in the dataset after the preprocessing (it is 476 with the used LIAR dataset). The length could be changed based on the dataset and the augmented information (e.g., ISOT dataset max length news contains 2682 unique words). Thus, all the sequences are padded to the size of the maximum length news sample. The second layer of the model is known as the embedding layer, which utilizes GloVe pre-trained model to transform each word or character into a fixed-length vector (each word is represented by a vector with 100 dimensions in this study). This layer maps the input features vector (See Figure 2 the padded sequence) to a dense vector space where the proximity between vectors represents the semantic similarity between corresponding inputs. This allows the network to acquire input representations that contain the pertinent information for the task at hand. Following the embedding layer is the convolutional layer, which utilizes learnable filters to extract hidden features from the embedded layer. Each filter generates a feature map that captures

specific features in the input set. The size of the filter in the convolutional layer is set to the total number of vectors subtracted from it the size of the kernel minus one. In the case of the LAIR dataset, the number of embedded vectors for each sequence is 476. Thus, with a kernel size is 3×3, the size of the filter is set to $467 - 2 = 474$. The number of filters was set to 128 selected heretically by trial and error. The subsequent layer is a max pooling layer that is used to decrease the dimensionality of the feature maps while maintaining the significant features. The size of the pooling layer is set to two to reduce the dimensions by half. The next two layers are a stacked convolutional and max pooling layer which are implemented to learn increasingly complex and abstract features from the input features and to further decrease the dimensionality of the feature maps produced by the convolutional layers. The feature maps created by the convolutional and max pooling layers are multi-dimensional. Consequently, a flattened layer is employed to transform the multi-dimensional output into a one-dimensional vector which can be processed by a fully connected layer (see Figure 3).

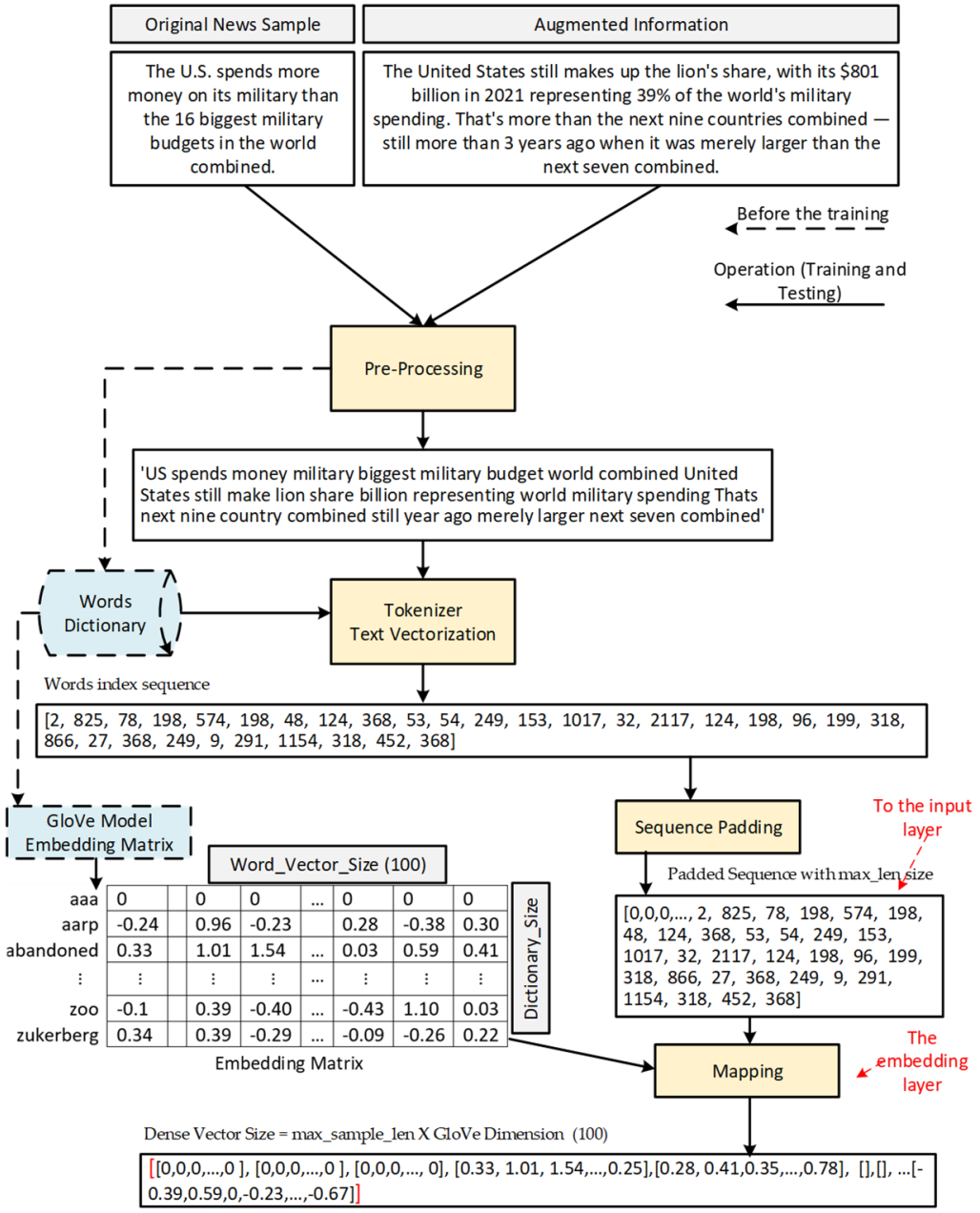


Figure 2. Example of the Process of Preparing the News Sample for generating the represented Impeded Sequence.

```
_________________________________________________________________
Model: "model_66"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_33 (InputLayer)       [(None, 476)]             0
 embedding_31 (Embedding)    (None, 476, 100)          18107900
 conv1d_28 (Conv1D)          (None, 474, 128)          38528
 max_pooling1d_24 (MaxPooling1D)  (None, 237, 128)     0
 conv1d_29 (Conv1D)          (None, 235, 64)           24640
 max_pooling1d_25 (MaxPooling1D)  (None, 117, 64)      0
 flatten_31 (Flatten)        (None, 7488)              0
 dropout_20 (Dropout)        (None, 7488)              0
 encoder_122 (Dense)           (None, 64)              479296
 decoder_123 (Dense)           (None, 476)             30940
 dropout_21 (Dropout)        (None, 476)               0
 dense_124 (Dense)           (None, 6)                 2862
=================================================================
Total params: 18,684,166
Trainable params: 576,266
Non-trainable params: 18,107,900
_________________________________________________________________
```

Figure 3. The architecture of the proposed CNN-Based Autoencoder.

To prevent overfitting, a dropout layer is added to drop out some neurons during the training. Then, an encoding layer is used to learn a compressed representation of features extracted from previous layers. The proposed architecture (see Figure 3) has an encoder network consisting of two layers, namely, encoding and decoding layers. These layers map the input data to a lower-dimensional latent space and a decoder network that reconstructs the input data from the compressed representation. The autoencoder network is trained using an unsupervised learning approach to learn a compressed representation of the flattened features. The output of the decoder layer is also fed to another dropout layer to improve the generalization by reducing the network's ability to memorize the training data. The last layer is the output layer which contains the SoftMax function to produce the probabilistic values of the membership of the sample in each target class. The Adaptive Moment Estimation optimization algorithm (Adam optimizer) was used to train the proposed model. Adam optimizer is an extension of the stochastic gradient descent (SGD). It maintains a running estimate of the mean and variance of the gradients of the model parameters to update the network weights iterative based on training data.

3.5. Decision Making

To aid in decision-making, the multilayer perceptron (MLP) has been employed, which is a widely used technique for classification and regression tasks. The MLP classifier utilizes the probabilistic outputs $p(c)$ of the preceding layer as new features for training. The suggested MLP has four layers: an input layer, two hidden layers, and an output layer with different numbers of neurons, primarily for multi-class classification tasks. The activation of the neurons is accomplished using ReLu functions, while SoftMax functions are employed for classification. The stacked layer of the MLP is trained utilizing the features extracted from the input features, specifically those for the MLP classifier. The prediction of a specific class, such as fake news, was denoted by $p(c)$ and can be calculated as follows. Let w_i denotes to the neuron i weight, θ represents the weights of the deep learners that were trained in the prior phase, and x_i denotes the corresponding output of the previous layer.

Then $p(c)$ which represents the score of correctly forecasting a particular class (such as fake news) can be calculated as follows.

$$p(class\_label = c) = \sum_{i=1}^{n} w_{ci} * x_i \; + \; \theta \tag{1}$$

Then, the logistic function is calculated for each predicted class as follows.

$$logit(p(class_{label} = c)) = x_c = \log\left(\frac{p(class_{label} = c)}{1 - \; p(class_{label} = c)}\right) \tag{2}$$

The x_c logit is the logistic function of the predicted class which maps the values from the range $(-\infty, +\infty)$ into $[0, 1]$. Let $\vec{x} = \{x_1, x_2, x_3, \; \ldots\ldots x_c)\}$ vectors contain the scores (x_c) predicted by the MLP for class c, then the final predicted class label is calculated based on the SoftMax function as follows.

$$\forall \; x_c \; \in \; \vec{x} \; calculate \; \frac{e^{x_c}}{\sum_{c=0}^{k} e^{x_c}} \; \overset{append}{\rightarrow} \; \vec{V} \tag{3}$$

where $\vec{V}$ is a vector that contains the probability of the input features belonging to the class by finding the maximum probability, the predicted class label is determined.

$$predicted \; class = index\_of\_\max(\vec{V}) \tag{4}$$

4. Performance Evaluation

The datasets and the performance measures that were used for validation are described in the following subsections.

4.1. Datasets

This research utilized two widely used datasets, LIAR [38] and ISOT [39] to assess and validate the proposed model. These datasets are popular among researchers who evaluate fake news detection solutions [8,10,20,38,55]. Below is a brief overview of these datasets.

4.1.1. LIAR Dataset

The LIAR dataset is an openly accessible dataset that has been created to assess the efficacy of fact-checking and fake news detection algorithms. It is referred to as "LIAR: A Benchmark Dataset for Fake News Detection" and contains a collection of brief statements that have been manually annotated for detecting fake news. The dataset was gathered by Kaliyarand Goswami [30] from POLITIFACT.COM over a decade and includes more than 12,800 statements covering diverse contexts. Each statement in the dataset has been assessed for its veracity by an editor from POLITIFACT.COM, who provides a comprehensive analysis report and links to source documents for each case. LIAR dataset is publicly available online on the URL (https://www.cs.ucsb.edu/~william/data/liar_dataset.zip, accessed on 1 February 2023). The dataset consists of a collection of statements that are labeled with one of six possible classes based on their truthfulness:

1. Pants-Fire: also referred to Pants on Fire: The statement is not true, and it's ridiculous. Examples of the pants on fire class include, "Evidence shows Zika virus turns fetus brains to liquid," and "In the 2012 election, there were more votes cast than registered voters in St. Lucie County, and Palm Beach County had 141 percent turnout."
2. False: The statement is not true, and evidence exists that proves it false. For example, the sentence, "Wisconsin is on pace to double the number of layoffs this year," is known to be false.
3. Barely-true: The statement contains an element of truth but is still mostly false. For example, the sentence, "The majority of people traveling to these resort destinations are not going for the primary purpose of gambling," is barely-true.

4. Half True: The statement is partially true, but also partially false. An example of a half-true sentence is, "Studies have shown that in the absence of federal reproductive health funds, we are going to see the level of abortion in Georgia increase by about 44 percent."
5. Mostly True: The statement contains an element of falsehood but is still mostly true. An example of the mostly true sentence is, "The sex-offender registry has been around for a long time, and the research that's out there says that it has no positive impact on public safety."
6. True: The statement is true, and evidence exists that proves it true. An example of true sentence is, "Before World War II, very few people had health insurance."

Each statement in the dataset is also labeled with metadata, such as the speaker, the context in which the statement was made, and the subject matter. The LIAR dataset was created to facilitate research on automatic fake news detection, and it has been used in numerous studies to evaluate the effectiveness of various machine-learning algorithms and natural language processing techniques. The sample distribution of the LIAR dataset, including those used for training, validation, and testing, is presented in Table 1 and Figure 4.

Table 1. LIAR Dataset samples distribution.

	Training	Validation	Testing	Total
Pants-fire	839	116	92	1047
FALSE	1994	263	249	2506
Barely-true	1654	236	212	2102
Half-true	2114	248	265	2627
Mostly-true	1962	251	241	2454
TRUE	1676	169	207	2052
Total	10,239	1283	1266	12,788

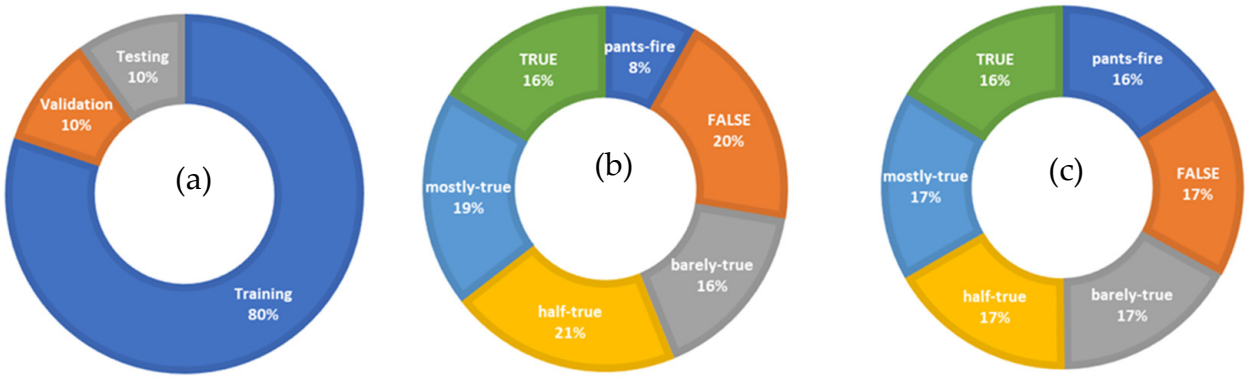


Figure 4. Samples Distribution Based on (**a**) Training, Validation, and Testing Sets, (**b**) News Categories, (**c**) Balanced News Categories Using SMOTE.

4.1.2. ISOT Dataset

The ISOT dataset was produced by gathering news articles from trustworthy sources such as Reuters.com, as well as from unreliable sources containing fabricated information. The dataset encompasses a total of 44,898 news statements, out of which 21,417 are authentic and 23,481 are counterfeit. A variety of researchers [8,10,28,56] have used the ISOT dataset in their studies. The distribution of the samples in the dataset, including those used for training, validation, and testing, is provided in Table 2.

Table 2. ISOT Dataset Sample Distributions.

News Category	Training	Validation	Testing	Total
FAKE	12,850	2142	6425	21,417
TRUE	14,089	2348	7044	23,481
Total	26,939	4490	13,469	44,898

4.2. Performance Measure

To evaluate and validate the proposed fake news detection model, four commonly used evaluation measures, as utilized in similar studies [5,11,18,20,44], were employed. These measures include accuracy, precision, recall, and F-measure, which are calculated using the concepts of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP represents the number of samples with false class labels that are correctly classified, while TN represents the number of negative samples with true class labels that are correctly classified. FP refers to the number of samples with true class labels that are misclassified, and FN refers to the number of samples with true class labels that are misclassified. The confusion matrix for binary and multiclass classification can be seen in Figure 5a,b. Accuracy is the proportion of correctly classified samples to the total number of samples, and it can be expressed as follows:

$$detection\ accuracy(acc) = = \frac{\mathrm{TP} + \mathrm{TN}}{TP + TN + FP + FN} \tag{5}$$

Detection accuracy indicates how well the model performs in terms of achieving true positive and true negative classification altogether. It doesn't matter which–either the false negative or false positive–costs much. The precision is calculated based on the number of false news items detected, divided by the number of samples classified as false, as follows:

$$Precession = \frac{\mathrm{TP}}{TP + FP} \tag{6}$$

The recall is calculated based on the number of correct classifications of false news divided by the number of fake news in the dataset false as follows:

$$Recall\ (detection\ rate) = \frac{\mathrm{TP}}{TP + FN} \tag{7}$$

F-measure (F1) is the harmonic means of the precession and recall, and can be calculated as follows:

$$F - measure\ (F1) = \frac{2 \times Precession\ \times Recall}{Precession + \ Recall} \tag{8}$$

F-measure, also called the F1 score, is the overall performance when the dataset is unbalanced and false positives and false negatives are crucial. F-measure does not account for the true negative rate. It is important when the goal is to balance between the false positives and the false negatives.

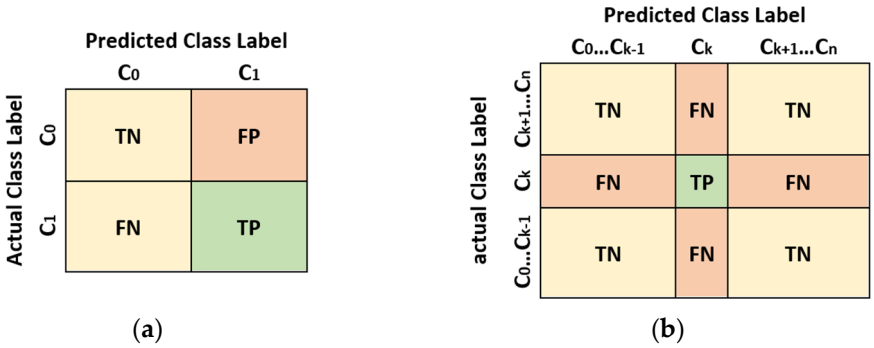


Figure 5. Confusion matrix for (**a**) binary classification and (**b**) multi-class classification.

4.3. Evaluation Procedure

The fake news model proposed in this study has been compared to other state-of-the-art models in the field. The proposed architecture ICNN-AUTOENC was used for multiple binary classifications and ICNN-AUTOENC-MD, which consists of the ICNN-AUTOENC with the decision making, was used for the multi-class classification task. These two models have been compared mainly with the work presented recently in [20] because it has been reported to achieve the best performance compared with the other related works and use a close design with the proposed classifier in this study. However, the main difference from the proposed model is that it does not collect information as a new, trusted feature and does not employ embedding techniques for feature representation or CNN architecture.

5. Results and Discussion

The research was conducted on a computer system consisting of a 2.5 GHz i7 processor, 8 GB of RAM, and 4 CPUs, and the proposed model was developed using Python 3.7 programming language. The performance of the classifier that was constructed using the LIAR dataset was evaluated and presented in Figures 6 and 7. These figures show the accuracy, precision, recall, and F1 scores for both binary and multi-class classification. Figure 6 displays the performance for binary classification using the proposed ICNN-AEN-DM model, while Figure 7 shows the multi-class classification performance using the ICNN-AEN-DM model.

Figure 6. Performance of the proposed detection model—Binary Classification using ICNN-AEN-DM using LIAR dataset.

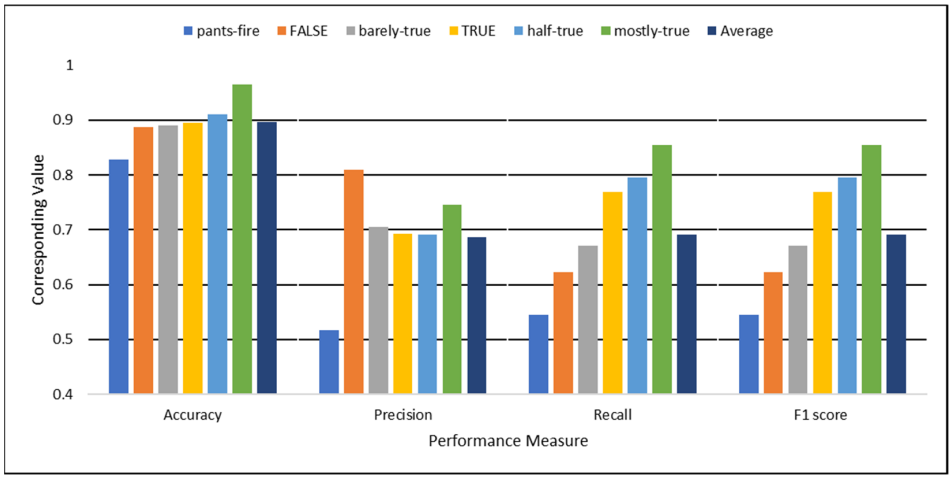


Figure 7. The performance of the proposed ICNN-AUTOENC-DM model in multi-class classification using the LIAR dataset.

As shown in Figure 6, the proposed ICNN-AEN-DM multiple binary classifiers achieved accuracy higher than 75% with greater than 85% overall performance among all developed binary classifiers. In Figure 6, the news classified as pants fire achieved a high detection rate of 97.8% in comparison with other types of news. This is because pants-fire contains obvious false facts, and with the proposed augmented information-gathering algorithm, such a category can be easily classified. Meanwhile, the news categorized false achieved the lowest performance (70% detection rate) and an accuracy of 61%. This is because this category has highly overlapped features with other categories.

The performance of the proposed ICNN-AEN-DM model in multi-class classification is illustrated in Figure 7, while the F-measure average performance of the proposed ICNN-AENC-DM model is presented in Figure 8 and compared with relevant studies. As depicted in Figure 7, while the accuracy performance of all classifiers surpasses 80%, the overall performance, as measured by the F-measure stands at 69%. This occurs because the f-measure does not count as the true negative, for the majority of the tested samples are considered negative to each news category. Accordingly, for multi-class, the F-measure is the most important measure for evaluation.

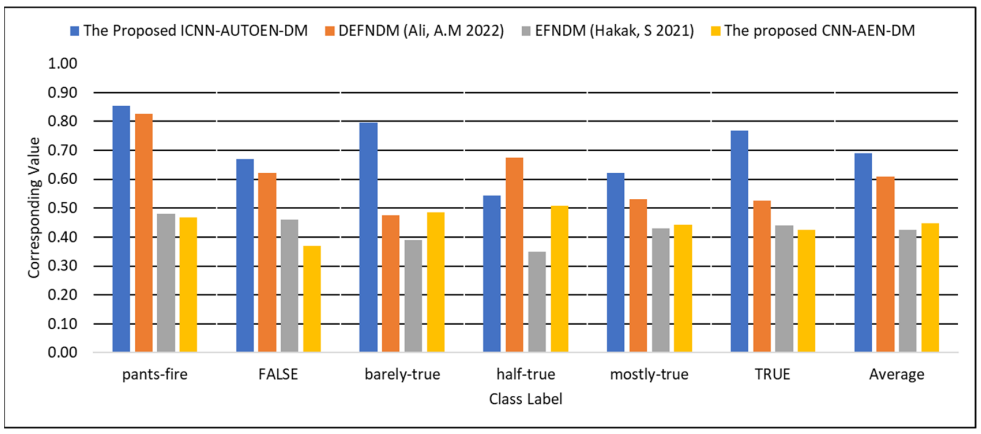


Figure 8. Performance Comparison of the proposed model Multi-Class Classification with EFNDM (Hakak, S. 2021) [10], DEFNDM (Ali, A.M. 2022) [20] using LIAR dataset.

As shown in Figure 8, the proposed model demonstrated performance superior to that of the EFNDM [10] and DENDM [20] studies. In particular, the proposed model achieved an overall F-Measure performance improvement of 8% and 24.4% compared with EFNDM [10] and DENDM [20], respectively. The main reason for such improvement can

be attributed to the proposed augmented features extracted from trusted web sources. This is reasonable as the true news sentences will have support from different news sources, unlike the features extracted solely from the news itself, as what has been done in the CNN-AEN-DM without the augmented features and the related studies EFNDM [10] and DENDM [20]. It can be also noted that pants-fire is more distinguished than other classes by the proposed ICNN-AEN-DM and also the DEFNDM proposed in [20]. Both models achieved more than 80% performance. However, the proposed model demonstrated superior performance in all other categories when compared to the studied models. The significant improvement in the performance of the proposed model can be attributed to the augmented information-gathering algorithm used for feature extraction. This is supported by comparing the results of the proposed model with and without the augmented features. The proposed architecture ICNN-AEN-DM achieved 69.1% performance, compared to 44.7% performance for CNN-AEN-DM. Moreover, the proposed model performance in classifying the half-true categories is lower than the performance of the DENDM [20] because the augmented information collected has highly overlapped features that belong to either false or true. Another reason can be attributed to the representation technique. The proposed ICNN-AEN-DM uses features embedding GloVe while DENDM [20] uses TF/IDF for representation. Unlike TF/IDF, GloVe does not rely solely on local statistics, but rather it generates more global statical vectors, and TF/IDF does not include semantic meaning. However, GloVe uses a limited number of words in the representation while TF/IDF generates more words, as argued in [20].

The performance of the proposed fake news detection model is compared to other relevant models in Figure 9 and Table 3. The evaluation measures used for comparison include accuracy, precision, recall, and F1 score, and their averages are calculated for each model. On average, the proposed model outperforms other models, achieving significant improvements of 26.59%, 7.09%, and 7.09% in terms of accuracy, recall, and F1 score, respectively. However, DEFNDM and EFNDM achieved higher precision than the proposed model but traded off with the lowest recall among the studied models. RoBERTa achieved the best F1 score on the average compared with Conv-HAN, DEFNDM, and EFNDM. Nevertheless, the proposed model outperformed RoBERTa with significant improvements in accuracy, precision, recall, and F1 score. The proposed model uses more representative features extracted from trusted websites for constructing the classifier, as illustrated in Figure 9 and Table 3.

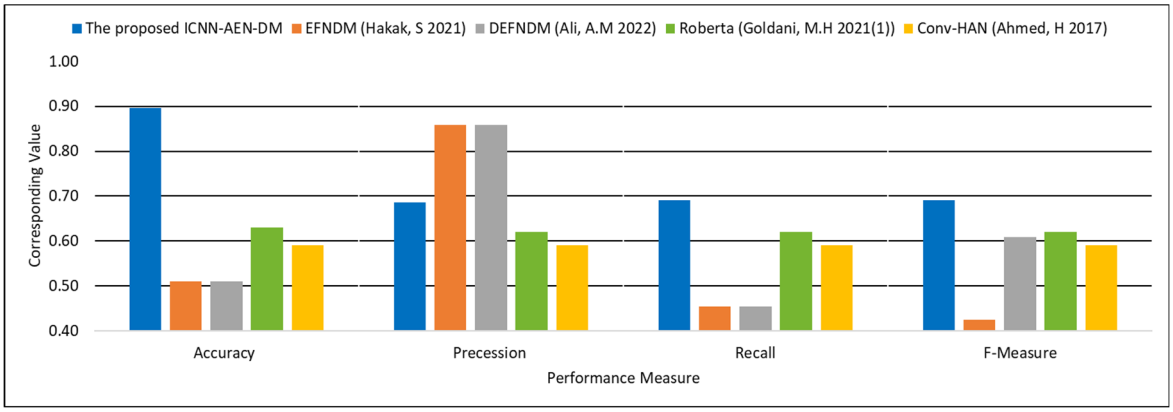


Figure 9. Performance Comparison with EFNDM (Hakak, S. 2021) [10], DEFNDM (Ali, A.M. 2022) [20], RoBERTa (Goldani, M.H. 2021(1)) [28], and Conv-HAN (Ahmed, H. 2017) [39] using LIAR dataset.

Table 3. Performance Comparison using LIAR dataset.

Model	Notes	Accuracy	Precession	Recall	F-Measure
RoBERTa [28]	GloVe.6B.300d + CNN	63.00%	62.00%	62.00%	62.00%
Conv-HAN [39]	GloVe + LSVM	59.00%	59.00%	59.00%	59.00%
DEFNDM [20]	Capsule neural networks	51.05%	85.86%	45.47%	60.90%
EFNDM [10]	Statistical + RF	51.05%	85.86%	45.47%	42.50%
The proposed	ICNN-AEN-DM	89.59%	68.59%	69.09%	69.09%

The ISOT dataset's performance evaluation of the proposed model is presented in Table 4 and Figure 10. As demonstrated by Table 4 and Figure 10, the proposed model, along with most of the related work, achieved an F-measure of 100%, indicating a perfect overall detection performance. The lowest performance reported in [39] achieves 92%, which is far from the performance of the other solutions. Generally, the ISOT data set is less challenging than the LIAR dataset. The news samples comprise more and longer sentences than the news samples in the LIAR data set, which consists of short and single sentences. This makes the features vectors extracted using the LIAR dataset more sparse, containing less information that affects the learning performance. In addition, the news in the ISOT dataset is collected from social media that may be posted by layman users, while the ones in LIAR are written by politicians. Accordingly, the overlapped features between different news categories are higher than that in the LIAR dataset.

Table 4. Performance Comparison using ISOT dataset.

Author and Year	F-Measure	Notes
Goldani, et al. [28], 2021,	99.90%	Glove.6B.300d + CNN
Ahmed, et al. [39], 2017,	92.00%	GloVe + LSVM
Goldani, et al. [56], 2021,	99.80%	Word2Vec + Capsule neural networks
Hakak, et al. [10], 2021,	100%	Statistical + RF
Samadi, et al. [8], et al., 2021,	99.96%	Funnel + CNN
Marish et al. [20], 2022	100%	TF/IDF,+ SDL + MLP
CNN-Autoencoder	100%	CNN-Autoencoder
The proposed	100%	Glove.6B.100d, ICNN-AEN-DM

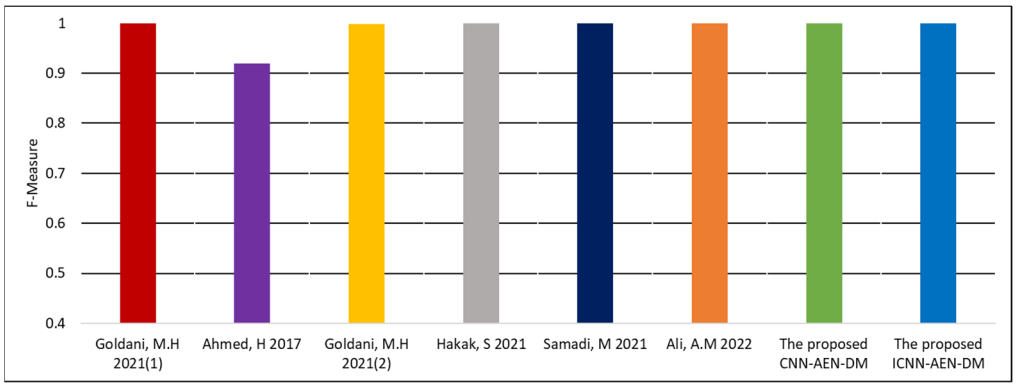
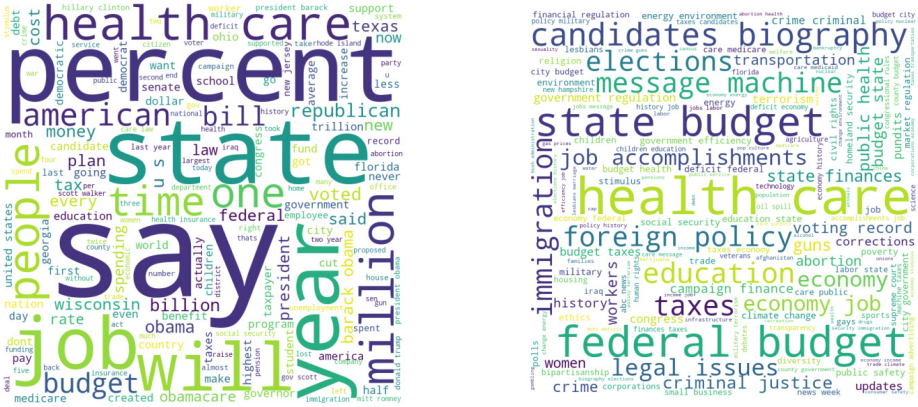


Figure 10. Performance Comparison with Goldani, M.H. 2021(1) [28], Ahmed, H. 2017 [39], Goldani, M.H. 2021(2) [56], Hakak, S. 2021 [10], Samadi, M. 2021 [8], and Ali, A.M. 2022 [20] using ISOT dataset.

To gain some insights about the features that were extracted from the original news dataset and the features extracted using the augmented dataset gleaned from a web search, the term cloud-based analysis has been considered. The difference between density and size variance can give insights into the performance of the feature extraction and learning. Figure 11 shows how the distribution of the features changed, based on the source of the

information. Figure 11a presents the word cloud of the original dataset while Figure 11b presents the word cloud of the augmented dataset (web search features have been included). The word cloud of the original dataset is sparse compared with the word cloud of the augmented dataset, which has more features. By combining these two text features, more informative features were extracted, resulting in improved classification performance of the proposed model. This can be observed from the comparison between the ICNN-AEN-DM and CNN-AEN-DM models in Figures 7 and 8, where the integration of original features with augmented features extracted from web search results in performance gains.



(**a**) (**b**)

Figure 11. Words important within (**a**) the original dataset (**b**) the augmented dataset.

To sum up, it is a difficult undertaking to accurately detect classified fake news when relying only on the text features extracted from news articles due to their overlapping nature with true news and lack of sufficient distinguishing features. Extracting news features is also costly and subject to various types of inaccuracies and falsehoods, such as those stemming from political or financial motivations. Researchers have tried to improve text classification accuracy by developing more advanced features based on syntax, semantics, and contextual information. However, to our knowledge, extracting similar information from reliable sources has not yet been investigated in the literature. This study shows that searching for similar information (e.g., facts) in reliable news sources boosts the performance of the classification. Features representation and model design also play important roles in the performance of the detection model. For example, TF/IDF with the Ensemble classifier achieved higher accuracy than the combined GloVe with CNN as reported by previous studies [20].

6. Conclusions

This study proposed a Web-Informed-Augmented Fake News Detection Model using Stacked Layers of Convolutional Neural Networks with impeded deep Autoencoder layers called (ICNN-AEN-DM). The study highlights the importance of effective news representative features and proper model design for the accurate detection and prevention of fake news. The contribution of this study lies in the development of a novel Web-Informed Augmented Fake News Detection Model, which addresses the limitations of existing models by enriching content-based features with additional features gathered from trusted sources through a web search. Two key factors were considered to accurately detect fake news: effective news representative features and appropriate model design. The insufficient features extracted from news content have been augmented by features gathered through web searches from trusted sources. The collected information is examined, and only the news with high similarity with the original one is retrieved from the list of reliable sources. The model leverages the GloVe word-embedding technique and stacked convolutional layers embedded with autoencoder layers to extract hidden features and reduce noises. Then, stacked convolutional layers were embedded with autoencoder layers to extract the

hidden features and reduce unwanted noises through a probabilistic deep learning-based classifier. Finally, the probabilistic outputs from the preceding layers were fed into stacked multilayer perceptron (MLP) layers for decision-making. The proposed model surpasses state-of-the-art models, demonstrating a notable improvement of 26.6% in detection accuracy and 8% in overall performance. This presents a promising solution to reduce the negative consequences of fake news, such as influencing public opinion, shaping political narratives, and inciting violence. The proposed model can also contribute to improving digital society by protecting democracy, reducing the spread of misinformation, preserving trust in media, and preventing harm.

This study focuses on short news sentences where the news content lacks sufficient features for learning. For longer news content, more investigations are required to collect the augmented features through a web search. The long news content may be split into smaller sentences by applying the proposed web information-collecting algorithm on each sentence to obtain probabilistic pre-decision. The final decision about the news class can be inferred by aggregating the results of the pre-decision. An in-depth investigation will be conducted in future investigations. The classification task of identifying fake news can be approached as a regression problem where the correctness of news can be represented as a value ranging from 0 to 1. As fake news often contains some fact, this representation may be more realistic. Hence, the existing models, including the proposed one in this study, have shown insufficient overall performance, and further improvement is required. This issue is currently being investigated, and the findings will be presented in future publications. Additionally, while this study mainly focused on extracting features from content and web sources, other sources of information such as knowledge-based, context-based, user-based, and propagation-based ones should also be explored.

Author Contributions: Conceptualization, F.A.G. and A.M.A.; methodology, F.A.G.; software, A.I.K.; validation, A.M.A., F.A.G. and M.S.M.; formal analysis, A.M.A., F.J.A. and A.I.K.; investigation, M.S.M. and A.I.K.; resources, F.J.A.; data curation, M.S.M.; writing—original draft preparation, F.A.G. and A.M.A.; writing—review and editing, F.A.G. and A.M.A.; visualization, F.A.G. and M.S.M.; supervision, F.J.A.; project administration, A.M.A. and F.J.A.; funding acquisition, A.M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by Institutional Fund Projects under grant no. (IFPRC-024-611-2020) from the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets that were used in this study are available online on the following links: 1. LIAR Dataset: https://www.cs.ucsb.edu/~william/data/liar_dataset.zip, last accessed on 1 February 2023. 2. ISOT Dataset: https://paperswithcode.com/dataset/isot-fake-news-dataset, last accessed on 10 December 2022.

Acknowledgments: The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number IFPRC-024-611-2020 and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhou, X.; Zafarani, R. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Comput. Surv.* **2020**, *53*, 1–40. [CrossRef]
2. Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false news online. *Science* **2018**, *359*, 1146–1151. [CrossRef] [PubMed]
3. Silverman, C. This Analysis Shows How Viral Fake Election News Stories Outperformed Real News on Facebook. Available online: https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook (accessed on 18 February 2023).

4. Nistor, A.; Zadobrischi, E. The Influence of Fake News on Social Media: Analysis and Verification of Web Content during the COVID-19 Pandemic by Advanced Machine Learning Methods and Natural Language Processing. *Sustainability* **2022**, *14*, 10466. [CrossRef]
5. Alhakami, H.; Alhakami, W.; Baz, A.; Faizan, M.; Khan, M.W.; Agrawal, A. Evaluating Intelligent Methods for Detecting COVID-19 Fake News on Social Media Platforms. *Electronics* **2022**, *11*, 2417. [CrossRef]
6. Choudhary, A.; Arora, A. Linguistic feature based learning model for fake news detection and classification. *Expert Syst. Appl.* **2021**, *169*, 114171. [CrossRef]
7. Sahoo, S.R.; Gupta, B.B. Multiple features based approach for automatic fake news detection on social networks using deep learning. *Appl. Soft Comput.* **2021**, *100*, 106983. [CrossRef]
8. Samadi, M.; Mousavian, M.; Momtazi, S. Deep contextualized text representation and learning for fake news detection. *Inf. Process. Manag.* **2021**, *58*, 102723. [CrossRef]
9. Sheikhi, S. An effective fake news detection method using WOA-xgbTree algorithm and content-based features. *Appl. Soft Comput.* **2021**, *109*, 107559. [CrossRef]
10. Hakak, S.; Alazab, M.; Khan, S.; Gadekallu, T.R.; Maddikunta, P.K.R.; Khan, W.Z. An ensemble machine learning approach through effective feature extraction to classify fake news. *Future Gener. Comput. Syst.* **2021**, *117*, 47–58. [CrossRef]
11. Koloski, B.; Perdih, T.S.; Robnik-Šikonja, M.; Pollak, S.; Škrlj, B. Knowledge Graph informed Fake News Classification via Heterogeneous Representation Ensembles. *Neurocomputing* **2022**, *496*, 208–226. [CrossRef]
12. Chauhan, T.; Palivela, H. Optimization and improvement of fake news detection using deep learning approaches for societal benefit. *Int. J. Inf. Manag. Data Insights* **2021**, *1*, 100051. [CrossRef]
13. Kumari, R.; Ekbal, A. AMFB: Attention based multimodal Factorized Bilinear Pooling for multimodal Fake News Detection. *Expert Syst. Appl.* **2021**, *184*, 115412. [CrossRef]
14. Song, C.; Ning, N.; Zhang, Y.; Wu, B. A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks. *Inf. Process. Manag.* **2021**, *58*, 102437. [CrossRef]
15. Song, C.; Ning, N.; Zhang, Y.; Wu, B. Knowledge augmented transformer for adversarial multidomain multiclassification multimodal fake news detection. *Neurocomputing* **2021**, *462*, 88–100. [CrossRef]
16. Zeng, J.; Zhang, Y.; Ma, X. Fake news detection for epidemic emergencies via deep correlations between text and images. *Sustain. Cities Soc.* **2021**, *66*, 102652. [CrossRef] [PubMed]
17. Wang, Y.; Wang, L.; Yang, Y.; Lian, T. SemSeq4FD: Integrating global semantic relationship and local sequential order to enhance text representation for fake news detection. *Expert Syst. Appl.* **2021**, *166*, 114090. [CrossRef]
18. Kirn, H.; Anwar, M.; Sadiq, A.; Zeeshan, H.M.; Mehmood, I.; Butt, R.A. Deepfake Tweets Detection Using Deep Learning Algorithms. *Eng. Proc.* **2022**, *20*, 2.
19. Sastrawan, I.K.; Bayupati, I.P.A.; Arsa, D.M.S. Detection of fake news using deep learning CNN–RNN based methods. *ICT Express* **2022**, *8*, 396–408. [CrossRef]
20. Ali, A.M.; Ghaleb, F.A.; Al-Rimy, B.A.S.; Alsolami, F.J.; Khan, A.I. Deep Ensemble Fake News Detection Model Using Sequential Deep Learning Technique. *Sensors* **2022**, *22*, 6970. [CrossRef]
21. Souza Freire, P.M.; Matias da Silva, F.R.; Goldschmidt, R.R. Fake news detection based on explicit and implicit signals of a hybrid crowd: An approach inspired in meta-learning. *Expert Syst. Appl.* **2021**, *183*, 115414. [CrossRef]
22. Meel, P.; Vishwakarma, D.K. HAN, image captioning, and forensics ensemble multimodal fake news detection. *Inf. Sci.* **2021**, *567*, 23–41. [CrossRef]
23. Shim, J.-S.; Lee, Y.; Ahn, H. A link2vec-based fake news detection model using web search results. *Expert Syst. Appl.* **2021**, *184*, 115491. [CrossRef]
24. Yuan, H.; Zheng, J.; Ye, Q.; Qian, Y.; Zhang, Y. Improving fake news detection with domain-adversarial and graph-attention neural network. *Decis. Support Syst.* **2021**, *151*, 113633. [CrossRef]
25. Abu Salem, F.K.; Al Feel, R.; Elbassuoni, S.; Ghannam, H.; Jaber, M.; Farah, M. Meta-learning for fake news detection surrounding the Syrian war. *Patterns* **2021**, *2*, 100369. [CrossRef]
26. Gupta, A.; Li, H.; Farnoush, A.; Jiang, W. Understanding patterns of COVID infodemic: A systematic and pragmatic approach to curb fake news. *J. Bus. Res.* **2022**, *140*, 670–683. [CrossRef]
27. Nasir, J.A.; Khan, O.S.; Varlamis, I. Fake news detection: A hybrid CNN-RNN based deep learning approach. *Int. J. Inf. Manag. Data Insights* **2021**, *1*, 100007. [CrossRef]
28. Goldani, M.H.; Safabakhsh, R.; Momtazi, S. Convolutional neural network with margin loss for fake news detection. *Inf. Process. Manag.* **2021**, *58*, 102418. [CrossRef]
29. Probierz, B.; Stefański, P.; Kozak, J. Rapid detection of fake news based on machine learning methods. *Procedia Comput. Sci.* **2021**, *192*, 2893–2902. [CrossRef]
30. Kaliyar, R.K.; Goswami, A.; Narang, P.; Sinha, S. FNDNet—A deep convolutional neural network for fake news detection. *Cogn. Syst. Res.* **2020**, *61*, 32–44. [CrossRef]
31. Vishwakarma, D.K.; Varshney, D.; Yadav, A. Detection and veracity analysis of fake news via scrapping and authenticating the web search. *Cogn. Syst. Res.* **2019**, *58*, 217–229. [CrossRef]
32. Agarwal, V.; Sultana, H.P.; Malhotra, S.; Sarkar, A. Analysis of Classifiers for Fake News Detection. *Procedia Comput. Sci.* **2019**, *165*, 377–383. [CrossRef]

33. Chiang, T.H.C.; Liao, C.-S.; Wang, W.-C. Investigating the Difference of Fake News Source Credibility Recognition between ANN and BERT Algorithms in Artificial Intelligence. *Appl. Sci.* **2022**, *12*, 7725. [CrossRef]
34. Ansar, W.; Goswami, S. Combating the menace: A survey on characterization and detection of fake news from a data science perspective. *Int. J. Inf. Manag. Data Insights* **2021**, *1*, 100052. [CrossRef]
35. Meel, P.; Vishwakarma, D.K. Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Syst. Appl.* **2020**, *153*, 112986. [CrossRef]
36. Saquete, E.; Tomás, D.; Moreda, P.; Martínez-Barco, P.; Palomar, M. Fighting post-truth using natural language processing: A review and open challenges. *Expert Syst. Appl.* **2020**, *141*, 112943. [CrossRef]
37. Zhang, X.; Ghorbani, A.A. An overview of online fake news: Characterization, detection, and discussion. *Inf. Process. Manag.* **2020**, *57*, 102025. [CrossRef]
38. Wang, W.Y. "liar, liar pants on fire": A new benchmark dataset for fake news detection. *arXiv* **2017**, arXiv:1705.00648.
39. Ahmed, H.; Traore, I.; Saad, S. Detection of online fake news using n-gram analysis and machine learning techniques. In Proceedings of the Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments: First International Conference, ISDDC 2017, Vancouver, BC, Canada, 26–28 October 2017; pp. 127–138.
40. Khan, T.; Michalas, A.; Akhunzada, A. Fake news outbreak 2021: Can we stop the viral spread? *J. Netw. Comput. Appl.* **2021**, *190*, 103112. [CrossRef]
41. Bondielli, A.; Marcelloni, F. A survey on fake news and rumour detection techniques. *Inf. Sci.* **2019**, *497*, 38–55. [CrossRef]
42. Domenico, G.D.; Sit, J.; Ishizaka, A.; Nunan, D. Fake news, social media and marketing: A systematic review. *J. Bus. Res.* **2021**, *124*, 329–341. [CrossRef]
43. Rastogi, S.; Bansal, D. A review on fake news detection 3T's: Typology, time of detection, taxonomies. *Int. J. Inf. Secur.* **2022**, *22*, 177–212. [CrossRef]
44. Xu, J.; Zadorozhny, V.; Zhang, D.; Grant, J. FaNDS: Fake News Detection System using energy flow. *Data Knowl. Eng.* **2022**, *135*, 101985. [CrossRef]
45. Ko, H.; Hong, J.Y.; Kim, S.; Mesicek, L.; Na, I.S. Human-machine interaction: A case study on fake news detection using a backtracking based on a cognitive system. *Cogn. Syst. Res.* **2019**, *55*, 77–81. [CrossRef]
46. Trueman, T.E.; Kumar, A.; Narayanasamy, P.; Vidya, J. Attention-based C-BiLSTM for fake news detection. *Appl. Soft Comput.* **2021**, *110*, 107600. [CrossRef]
47. Shu, K.; Wang, S.; Liu, H. Understanding user profiles on social media for fake news detection. In Proceedings of the 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, USA, 10–12 April 2018; pp. 430–435.
48. Silva, A.; Han, Y.; Luo, L.; Karunasekera, S.; Leckie, C. Propagation2Vec: Embedding partial propagation networks for explainable fake news early detection. *Inf. Process. Manag.* **2021**, *58*, 102618. [CrossRef]
49. Hakim, A.A.; Erwin, A.; Eng, K.I.; Galinium, M.; Muliady, W. Automated document classification for news article in Bahasa Indonesia based on term frequency inverse document frequency (TF-IDF) approach. In Proceedings of the 2014 6th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 7–8 October 2014; pp. 1–4.
50. Alghamdi, J.; Lin, Y.; Luo, S. A Comparative Study of Machine Learning and Deep Learning Techniques for Fake News Detection. *Information* **2022**, *13*, 576. [CrossRef]
51. Ozbay, F.A.; Alatas, B. Fake news detection within online social media using supervised artificial intelligence algorithms. *Phys. A Stat. Mech. Its Appl.* **2020**, *540*, 123174. [CrossRef]
52. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.
53. Pennington, J.; Socher, R.; Manning, C.D. Glove: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
54. Joulin, A.; Grave, E.; Bojanowski, P.; Douze, M.; Jégou, H.; Mikolov, T. Fasttext. zip: Compressing text classification models. *arXiv* **2016**, arXiv:1612.03651.
55. Jadhav, S.S.; Thepade, S.D. Fake news identification and classification using DSSM and improved recurrent neural network classifier. *Appl. Artif. Intell.* **2019**, *33*, 1058–1068. [CrossRef]
56. Goldani, M.H.; Momtazi, S.; Safabakhsh, R. Detecting fake news with capsule neural networks. *Appl. Soft Comput.* **2021**, *101*, 106991. [CrossRef]