Fake news spreading rapidly worldwide is considered one of the most severe problems of modern technology that needs to be addressed immediately. The remarkable increase in the use of social media as a critical source of information combined with the shaking of trust in traditional media, the high speed of digital news dissemination, and the vast amount of information circulating on the Internet have exacerbated the problem of so-called fake news. The present work proves the importance of detecting fake news by taking advantage of the information derived from friendships between users. Specifically, using an innovative deep temporal convolutional network (DTCN) scheme assisted using the tensor factorization non-negative RESCAL method, we take advantage of class-aware rate tables during and not after the factorization process, producing more accurate representations to detect fake news with exceptionally high reliability. In this way, the need to develop automated methods for detecting false information is demonstrated with the primary aim of protecting readers from misinformation.

1. Introduction

Man learns and forms consciousness within the social context. He adopts opinions and constructs a large part of his perceptions based on the data given to him as undisputed knowledge by people he trusts, such as teachers, parents, friends, and information sources. This social transmission of knowledge is at the heart of human civilization. However, often, the information and “knowledge” are wrong, intentionally or not. They are usually written and published to mislead [1], harm an organization, a legal or physical person, or even for economic or political benefits, often using impressive titles, or entirely made to increase readability [2]. Fake news is a yellow press or propaganda done with deliberate misinformation or pranks spread through traditional or social media [3, 4].

Considerable research is being conducted on tactics for combating and suppressing fake news, including disinformation, which is the deliberate distribution of false narratives for political reasons or to undermine social cohesiveness in targeted populations [5]. On the other hand, numerous solutions must be adjusted to individual types of fake news, based, for example, on whether the fake news is actively manufactured or produced inadvertently or unintentionally [6].

The difficulty of detecting fake news on social media is mainly because fake news is deliberately written to deceive readers [3, 7]. The content of the information can vary in terms of subject matter and writing style. For example, an actual event may be presented in a misleading context to support something false [8]. As a result, a machine learning model that detects fake news solely through content may not deliver the best results. Instead, we can take advantage of additional information generated by the network, such as the profile of users who interact with the news on social media, to get more accurate results [9].

The problem of detecting fake news [10] on social media has been studied and categorized into three main approaches: content, network, and content - network (hybrid approach) [11]. Content-only strategies focus on analyzing
Computational Intelligence and Neuroscience

the content of the news (e.g., vocabulary, syntax) and detecting patterns through natural language processing methods. However, to produce satisfactory results, a predefined scope is required, which is difficult to achieve in the case of fake news because of its diversity [12]. Network-based approaches extract information from different networks created by users who interact with the news, such as diffusion or relationship networks. In general, network-based methods perform better in dynamic environments and are more suitable for detecting fake news. Finally, hybrid approaches aim to combine the advantages of content and network-based models, integrating both language features and network information into one model [13].

The overall strategy is to discover awful news through human fact-checking and automated artificial intelligence (machine learning, natural language processing, and network analysis). Research communities have used two fundamental counter-strategies: lowering the importance of fake news and sending out warning messages [10, 14].

In the first strategy, objectionable content is pushed down by the search algorithm, such as the second or later pages of a search engine, making it less likely that visitors will see it (most users scan the first page of search results). However, two issues arise. One is that reality is not always black and white, and fact-checkers frequently differ on classifying the content included in computer training sets, risking false positives, and unjustifiable suppression [15]. Furthermore, because fake news evolves quickly, misleading identifiers may become obsolete in the future.

The second strategy entails attaching warnings to erroneous content by expert fact-checkers. Many studies show that corrections and cautions diminish misperceptions and sharing. Despite some early evidence that fact-checking could backfire, further research has revealed that these backfire effects are infrequent. However, the primary issue is that professional fact-checking is not scalable, i.e., it might take a significant amount of time and effort to research each claim. As a result, many (if not most) bogus claims go unchecked. Furthermore, the process is sluggish, and a warning may miss the peak period of viral propagation. Moreover, signs are usually connected to fake news rather than biased coverage of events.

A third strategy prioritizes trustworthy sources such as mainstream media and science communication publications. However, this technique has yielded inconsistent results, as many sites contain hype partisan commentary and confirmation bias, and specific community sections deny all scientific research.

The above, in combination with the fact that users play a key role in how fake news is spread, also users share information with other similar users (friends) and that the characteristics of the network are essential in categorizing or classifying fake news [3], suggesting that the exploring networks of user relationships could greatly facilitate the categorization of fake news [9, 16]. As a result, in this paper, we aim to take advantage of information derived from user friendships and to demonstrate their importance in the problem of detecting fake news. In most cases, standard tensor derivatization or decomposition methods are performed in an unsupervised environment [17]. The class information available for some of the data does not affect them. Instead, our goal is to construct class-aware factor tables using tensor factorization methods. We believe that using class information during and not after the factorization process can give more accurate representations for detecting false news with considerable reliability.

Specifically, using a pioneering DTCN scheme [18] assisted using the tensor Factorization non-negative RESCAL method [19], we can utilize factor-aware tables rather than after the factorization process to produce more accurate representations for localization of fake news with exceptionally high credibility [14].

2. Related Literature

Detecting misleading information ([20] p 19) is a constantly evolving problem, and recently, many survey papers have covered the topic of online fake news under numerous different approaches [7, 9].

Zhang and Ghorbani [3] presented a comprehensive overview of the findings relating to fake news. They assessed the harm caused by online fake news by looking at the various components that make up false news, including the source, audience, target, substance, and the broader social environment in which it circulates. The most up-to-date detection technologies are all about detecting the characteristics that signal it when fighting misinformation. Fake news detection was reviewed by comparing existing detection methods and examining existing datasets for supervised training models. Fake news was also classified using pre-existing databases, which the researchers studied. Finally, they outlined several intriguing directions for future research into online fake news.

To address the issue of missing ground-truth data, as well as its high dimensionality and possible redundancy, the authors of [21] present a unique unsupervised feature learning technique for extracting discriminative features from the original data. It learns the compressed representation of unlabeled input using recurrent neural network-based asymmetric autoencoders and can elaborate on spectral and spectral-spatial characteristics. Their extractors can be added to the unsupervised segmentation pipeline and followed by any clustering technique. The trials indicated that their techniques produce high-quality segmentation without prior class labels and are one order of magnitude faster than 3-D convolutional AEs. Their methods outperform or perform similarly to previous methodologies, allowing for significant data reduction.

Pérez-Rosas et al. [11] focused on the automatic identification of fake content in online news. While testing their models, the researchers found that their best models could detect false content with the same level of accuracy as humans can. As a result of their work, they offered two datasets to detect fake stories: one based on seven different types of news domains, six different types of news domains, and one that was gathered from the web that included celebrities. They also explained the collection, annotation, and
validation method and offered many exploratory analyses to discover counterfeit and authentic news content linguistic variations.

Faustini and Covões [4] suggested a method for detecting false news that relied only on text attributes that could be created independent of the source platform. They tested five datasets, including texts and social media postings, in some language groups and found their findings attractive to baselines. They analyzed the accuracy achieved using a particular set of characteristics to those acquired using other common natural language processing approaches. They then ranked the algorithms in order of highest performance. In terms of the feature set, the bag-of-words technique produced the best results but not being statistically more significant than the other attributes. Finally, they created datasets using various channels, such as websites and social media.

Yang et al. [22] developed a text and image information based Convolutional neural network model that could incorporate text and image information with the associated explicit and latent properties. More importantly, it was expandable. They performed trials on a dataset obtained before election procedures and real-world false news records to show that TI-CNN was successful in detecting fake news. The findings demonstrated that using explicit and latent information acquired from multilayer neurons; their method could correctly see bogus info with very high efficiency.

The propagation of fake news has become a global issue, undermining public trust. Fake news and actual news propagate differently on social media. Thus, propagation-based detection systems, which use graph neural networks to form graphs with users as nodes and news sharing chains as edges and simultaneously learn propagation patterns and user preferences, have received much attention. The authors of [23] offer a method for detecting false news using the graph transformer network, which can learn efficient node representations while recognizing functional connections between nodes in the original graph. The proposed method’s efficacy is proven by comparison studies utilizing a real-world dataset made of Twitter data.

Antoun et al. [7] proposed techniques for detecting fake news, identifying domains and bots in tweets. They used two distinct approaches to catching fake news. The first used breakthroughs in natural language understanding end-to-end deep learning models to recognize aesthetic distinctions between authentic and counterfeit news pieces. The second was built after the competition and surpassed the winner. The best approach for news domain recognition was a hybrid method that combined named entity characteristics with semantic embeddings generated from end-to-end models [4]. The technique for detecting bots on Twitter consisted of elements taken from news tweets mixed with information from the news tweets. The trials revealed the relevance of several aspects, and the findings suggested that the proposed models performed adequately. They want to enhance false news detection by using elements from fact-checking websites with Google queries.

3. Methodology

Sequence modeling problems have been addressed mainly through feedback neural networks. This work proposes to model the situation using TCN, which achieves top performance and reliability. Based on a discrete convergence operator that produces a map of output characteristics by dragging a kernel over the input, the suggested architecture is designed to create a map of output characteristics \( f \). Using the multiplication between the kernel and the input stride (i.e., the piece of the input with the same size as the kernel), the output characteristic map is constructed. The depth of the output volume is determined by the number of \( M \) cores (filters) utilized at each convergent level (i.e., the number of output feature maps). Stride and padding are hyper-parameters used to adjust the output feature maps’ remaining spatial dimensions [24]. The first can be set to any direction of motion and reflects the distance between two consecutive input strides. Specifically, padding refers to the ability to silently expand the inputs by adding to their limitations (typically zeros) zeros to regulate the size of the output. Without padding, the output dimension would decrease after each convergent level [3, 9, 22].

Considering a one-dimensional sequence \( x \in \mathbb{R}^{nT} \) and a one-dimensional nucleus \( w \in \mathbb{R}_k \), the \( i \)-th element of the convergence between \( x \) and \( w \) is [18, 22, 25]:

\[
f(i) = (x * w)(i) = \sum_{j=0}^{k-1} x(i - j)w(j).
\]

with \( f \in \mathbb{R}^{nT - k + 1} \) if padding is not used; otherwise, the padding has the dimension of the input, i.e., \( f \in \mathbb{R}^{nT} \). The previous equation applies to the case of one-dimensional inputs but can easily be extended to inputs of larger dimensions.

In the model we propose, in addition to the contemporary architecture, we also use a fully probabilistic and self-regression model, which can process arbitrary length sequences and extract a sequence of equal length. The network uses causal (expanded) events to achieve this, and the remaining connections are used to handle substantial historical values. In the proposed TCN, the estimated value at time \( t \) depends only on previous samples and not on future ones. To achieve this behavior, the causal convergence replaces the standard convergence operator/symbol.

In addition, zero-length padding (filter size \( \sim 1 \) ) is added to ensure that each level is the same length as the input level. Expanded causal events enhance the network’s capabilities further, allowing the network receptive field (i.e., the number of input neurons to which the filter is applied) to increase and its ability to learn long-term dependencies on time series. By this logic, considering a one-dimensional input \( x \in \mathbb{R}^{nT} \) and a nucleus \( w \in \mathbb{R}_k \), an
expanded convolution output using an expansion factor $d$ becomes [4, 22, 26]:

$$f(i) = (x * q) (i) = \sum_{j=0}^{k-1} x(i - dj)w(j).$$  

(2)

This is an essential advantage over simple causal convolutions, as in the latter case, the receptive field $r$ increases linearly with the lattice depth $r = k(L - 1)$. In contrast, with dilated convolutions, the dependence is exponential $r = 2L - 1k$, ensuring that the network uses a more extensive history size.

Despite the application of expanded convergence, the comprehensive system still needs many levels to learn the dynamics of the inputs. Besides, as it turns out experimentally, the performance often degrades with increasing network depth [27]. The degradation problem was addressed using a deep residual learning framework. Each layer in a typical neural network feeds into the next layer. Each layer in a network with leftover blocks provides the next layer and straight into the levels 2-3 hops away.

There are several interpretations of why residual blocks are lovely and how and why they are one of the essential principles that can enable a neural network to function at a high level on a wide range of tasks. We know, for example, that neural networks are universal function approximators with increasing accuracy as the number of layers grows. However, there is a limit to the number of layers that can be added to improve accuracy. If neural networks were universal function approximators, they should have learned any simplex or complex function. However, because of issues such as vanishing gradients and the curse of dimensionality, it turns out that even with suitably deep networks, it may be unable to learn simple functions such as the identity function. This is unwanted. Furthermore, as the number of layers increases, the accuracy will saturate at some point and subsequently decline. Moreover, in most cases, this is not due to overfitting. As a result, shallower networks can learn faster than deeper networks, which is counterintuitive. However, this is observed in practice and is commonly referred to as the degrading problem. We know that shorter networks outperform deeper equivalents with a few more layers in the degradation problem [19, 21].

In particular, the proposed methodology proposes that for a network of $L$ levels with training error $\epsilon$, introducing $k$ additional levels on it should either leave the error unchanged or improve it. In the worst case, the $k$ new added nonlinear levels should learn the identity mapping $y = H(x) = x$ where $x$ is the output of the $L$ level network and $y$ is the output of the $L + k$ level network. According to the proposed solution, these stacked levels are proposed to fit in a residual mapping $F(x) = H(x) - x$ instead of the desired $H(x)$. Thus, the initial mapping is reshaped to $F(x) + x$, implemented via a simple front-end neural network with shortcut connections. In this way, identity mapping is learned simply by driving stacked level weights to zero values.

The architecture used is shown in Figure 1 [18, 22, 28].

At the first level of the network, the information is processed separately from the external data (when available). Later, the individual results will be combined and processed by the deep residual network $L$ levels [29]. Each level consists of a residual block with one-dimensional expanded causal convolution, ReLU activation function, and dropout to avoid overfitting [30].

The process in the residual block is as shown in Figure 2 [26, 31, 32].

The output plane consists of a $1 \times 1$ convolution that allows the network to export a one-dimensional vector $y \in \mathbb{R}^nT$ with the same dimension as the input vector $x$.

To approach multistep prediction, which essentially means that multidimensional data representation is required, we adopt the RESCAL negative factorization method. RESCAL factorization is a new tensor factorization model that scales very well into large amounts of data and can produce state-of-the-art results in many machine learning problems. More specifically, given an $X$ tensor, each slice of $X_k$ is factorized as follows [33, 33, 34]:

$$X_k \approx AR_kA^T, \quad k = 1, \ldots, m,$$  

(3)

where $A$ is an array $n \times r$ containing the latent representations of the $n$ entities of the problem and $R_k$ is an array $r \times r$ with the latent interactions of the $r$ factors in the $k$ dimension of the tensor. Tables $A$ and $R_k$ are computed by solving the following minimization problem [30, 35, 36]:

$$\min_{A, R_k} \left\{ f(A, R_k) + g(A, R_k) \right\},$$  

(4)

$$f(A, R_k) = \frac{1}{2} \sum_k \| X_k - AR_kA^T \|^2_f,$$

is the problem of least squares of factorization and
isthenormalizationtermaddedtoavoidoverfittingthemodel.

Relationaldatasuchasthatgeneratedbysocialnetworkscangenerallyberegardednon-negative. Therefore, the data in our problem can be factorized using a non-negative version of RESCAL. First, we represent the data using a third-order tensor. Each slice of $T_k$ is factorized as follows:

$$T_k \approx AR_kA^T,$$

where $A$ is an array $n \times r$ containing the latent representations of $n$ entities and $R_k$ is an array $r \times r$ with the latent interactions of the $r$ factors in the $k$ dimension of the tensor. Parameters $A$ and $R_k$ are calculated by solving the following minimization problem [23, 26, 37]:

$$\min_{A, R_k} f(A, R_k) + g(A, R_k),$$

$$f(A, R_k) = \sum_k \|T_k - AR_kA^T\|_F^2,$$

is the problem of least squares of factorization and

$$g(A, R_k) = \frac{1}{2} \lambda \|A\|_F^2 + \sum_k \|R_k\|_F^2,$$

is the normalization term. The non-negative updates for Tables $A$ and $R_k$, respectively, are as follows[38, 39]:

$$A \leftarrow A * \frac{\sum_k T_kAR_k^T + T_k^TAR_k}{A^T(\sum_k R_kA^TAR_k^T + R_k^TAR_k^T) + \lambda A},$$

$$R_k \leftarrow R_k * \frac{A^T T_k A}{A^TAR_kA^T + \lambda R_k}.$$
them with class tags, and the remaining 2(n − p) without, where \( p < n \). There is an equal number of false and actual news [42]. First, we create two third-order tensors using binary representation to model user’s friendships: \( X_{\text{fake}} \in \mathbb{R} \times u \times p \) containing all news posts marked as false and \( X_{\text{real}} \in \mathbb{R} \times u \times p \) with news that has been flagged as valid.

Neighborhood tables represent news posts. By stacking these tables, one after the other, we configure the tensors in a binary way and with whether the user interacted with the fake post - if he is following the user.

It is noted that when we declare that a user has interacted with a post, it means that he has posted the related post/news on his social media profile, in our case, on Twitter. Next, we apply the non-negative RESCAL factorization to the tensors \( X_{\text{fake}} \) and \( X_{\text{real}} \). As a result of the factorization, we end up with tables \( A_{\text{fake}} \in \mathbb{R} \times u \times r \) and \( A_{\text{real}} \in \mathbb{R} \times u \times r \), respectively, through the following non-negative updates that result [39, 43, 44]:

\[
A_{\text{fake}} \leftarrow A_{\text{fake}} \ast \frac{\sum_k X_{\text{fake}}(k) A_{\text{fake}} R_k + X_{\text{fake}}^T(k) A_{\text{fake}} R_k}{A_{\text{fake}} \left( \sum_k R_k (A_{\text{fake}})^T R_k + R_k (A_{\text{fake}})^T A_{\text{fake}} R_k \right) + \lambda A_{\text{fake}} I} 
\]

(10)

where,

\[
R_k \leftarrow R_k \ast \frac{\sum_k A_{\text{fake}}^T X_{\text{fake}}(k) A_{\text{fake}}}{A_{\text{fake}} A_{\text{fake}} R_k (A_{\text{fake}})^T A_{\text{fake}} + \lambda R_k}
\]

(11)

\[
A_{\text{real}} \leftarrow A_{\text{real}} \ast \frac{\sum_k X_{\text{real}}(k) A_{\text{real}} R_k + X_{\text{real}}^T(k) A_{\text{real}} R_k}{A_{\text{real}} \left( \sum_k R_k (A_{\text{real}})^T R_k + R_k (A_{\text{real}})^T A_{\text{real}} R_k \right) + \lambda A_{\text{real}} I} 
\]

(12)

where:

\[
R_k \leftarrow R_k \ast \frac{\sum_k A_{\text{real}}^T X_{\text{real}}(k) A_{\text{real}}}{A_{\text{real}} A_{\text{real}} R_k (A_{\text{real}})^T A_{\text{real}} + \lambda R_k}
\]

(17)

The posts that were not used in the train set as slices of the tensors belong to the test set and do not have a tag yet. Suppose, we have a set of arrays \( M = \{P_1, P_2, \ldots, P_{2(n-p)}\} \), with each array corresponding to a different unlabeled post, where \( P_{idx} \in \mathbb{R} \times u \) and \(.idx \in [1, 2(n-p)] \). The \( P_{idx} \) table is created for each post \( idx \) in the known binary way described earlier[45].

Next, we perform the non-negative RESCAL factorization on the tensors to generate the new factor tables through the following non-negative updates that occur:

\[
A_{\text{fake}}' \leftarrow A_{\text{fake}}' \ast \frac{\sum_k X_{\text{fake}}(k) A_{\text{fake}} R_k + X_{\text{fake}}^T(k) A_{\text{fake}} R_k}{A_{\text{fake}}' \left( \sum_k R_k (A_{\text{fake}}')^T R_k + R_k (A_{\text{fake}}')^T A_{\text{fake}}' R_k \right) + \lambda A_{\text{fake}}' I} 
\]

(13)

where,

\[
A_{\text{real}}' \leftarrow A_{\text{real}}' \ast \frac{\sum_k X_{\text{real}}(k) A_{\text{real}} R_k + X_{\text{real}}^T(k) A_{\text{real}} R_k}{A_{\text{real}}' \left( \sum_k R_k (A_{\text{real}}')^T R_k + R_k (A_{\text{real}}')^T A_{\text{real}}' R_k \right) + \lambda A_{\text{real}}' I} 
\]

(14)

Since the Euclidean distance is a known metric for calculating the distance that is regularly used in similar difference calculation problems, we calculate the prediction as follows: Let \( l \) be the still unknown tag of each post, with \( l \in [0, 1] \), where 0 indicates an authentic post and 1 a false one. Finally, we remove the \( P_{idx} \) table from \( X'_{\text{fake}} \) and \( X'_{\text{real}} \), add the next \( P_{idx+1} \) table from the set of tables \( M \) to both tensors, and repeat the procedure for the remaining \( 2(n-p) \) posts [46, 47].

To evaluate the proposed method, we conducted experiments with two public English-language datasets from two platforms, BuzzFeed and PolitiFact, which include both news content and network information, along with class tags that indicate if each news item is false or true. Content provides information related to the news article’s content, such as the title, author, and text, while network information includes information such as user profiles, friendships, and activity. For our evaluation, we use only the knowledge of the network, and in particular, the friendship networks between the users who have posted the relevant news. To reduce the size and sporadicity of the data, we removed the users with node grade <3. Finally, we end up with two tensors of dimensions \( 182 \times 1449 \times 1449 \) and \( 240 \times 1697 \times 1697 \) for the BuzzFeed and PolitiFact datasets, respectively. The Colab environment with GPU was used for the experiments.

To measure the method’s performance, we choose the precision evaluation metrics, recall, F1-score, and accuracy, often used in similar problems. The first 70% of the news is
the train set, and the remaining 30% is the test set. The number of news items classified as accurate is equal to the number of items marked as false. We perform the experiments ten times, independently for each data set, and record the average results [48, 49].

The results of the process are presented in detail in Table 1. The abovementioned results show that by using only network data and some class tags, adding class information in the middle and not after the tensor factorization process, and with a small number of factors leading to short computation times, we can achieve outstanding performance, even in problems that require the combination of much heterogeneous information and complex calculations. With this in mind, we can confirm our original hypothesis that explore networks between users that can be helpful in the process of detecting fake news.

In conclusion, the use of the proposed method creates a highly efficient TCN which can exploit a sizeable historical size effectively. With this architecture, lower memory requirements are achieved during training, and predictions for later time steps are not made sequentially. Still, they can be calculated in parallel, taking advantage of parameter sharing. In addition, the training of the proposed system is much more stable than that which includes RNNs because it allows avoiding the problem of explosion/disappearance of inclination.

The success of the method is mainly due to the following three reasons [9, 28, 41, 50, 51]:

1. Local connectivity: One set of input neurons is connected to each hidden neuron (according to a specific space-time metric). Compared to a fully linked network, this feature significantly reduces the number of parameters that must be learned and facilitates calculations.

2. Parameter Sharing: The weights used to determine the output neurons in a feature map are the same for each location so that the same kernel is employed. There is less of a learning curve because there are fewer parameters to master.

3. Translation exchange rate: The network is resistant to a possible shift of its input.

### 5. Conclusions

In this work, using an innovative DTCN scheme assisted using the tensor Factorization non-negative RESCAL method, we manage to take advantage of class-aware factor tables rather than after the factorization process to produce more accurate representations in detecting fake news with exceptionally high reliability. Instead of applying factorization and classification separately, we proposed a method that combines them into a standard learning process. This approach uses user friendship networks that have interacted with the news and a set of class tags available for some of the news. We proposed a standard RESCAL negative factorization method to combine this information, which incorporates class tags into the factorization itself. In this way, we successfully arrive at a class-aware tensor-aware semi-supervised derivation. To evaluate the method, we conducted experiments with two public datasets. The results demonstrate integrating class information into the factorization phase as a single process. They also validated our original hypothesis: how individuals connect with news on social media directly affects the news’ legitimacy.

As a future extension, we would like to investigate how the proposed methodology is improved when more information is added to it, both from the network and from the news content. In addition, we plan to evaluate the performance of our approach in more databases and investigate the effect of data size on the scalability of the algorithm. It would be interesting to explore new ways of representing the available information with tensors to integrate it into the proposed method at the methodological level. [52–55].

### Data Availability

The data used in this study are available from the author upon request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This work was a project supported by Hunan Provincial Social Science Fund "Research on Translation of Miao Culture Classics in Western Hunan Area under the Perspective of Cultural Anthropology” (Grant no. 18ZDB005), also by Scientific Research Fund of Hunan Provincial Education Department “A study on the English Translation of Hmong Epics from the Perspective of Ethnographic Thick Translation” (Grant no. 19B130); by First-Class Undergraduate Major in Hunan Province ----Business English (Grant No. (2020)179).

### References

[1] P. Surendran, B. Navayasree, H. Kambham, and M. Anand Kumar, “Covid-19 fake news detector using hybrid convolutional and Bi-lstm model,” in Proceedings of the 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, July, 2021.

[2] F. Torgheh, M. R. Keyvanpour, and B. Masoumi, “A new method based on deep learning and time stabilization of the propagation path for fake news detection,” in Proceedings of the 2021 12th International Conference on Information and Knowledge Technology (IKT), pp. 57–61, Babol, Iran, December 2021.

[3] X. Zhang and A. A. Ghorbani, “An overview of online fake news: characterization, detection, and discussion,”...
C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," 2016, http://arxiv.org/abs/1611.05267.

D. Li, H. Guo, Z. Wang, and Z. Zheng, "Unsupervised fake news detection using graph neural networks with transformer," in Proceedings of the 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE), pp. 19-20, Kyoto, Japan, October 2021.

A. Dhillon and G. K. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection," Progress in Artificial Intelligence, vol. 9, no. 2, pp. 85–112, 2020.

Y. Zhou, X. Xiao, "An efficient residual-convolutional neural model for handwritten text recognition," in Proceedings of the 15th International Conference on Semantics, Knowledge and Grids (SKG), pp. 115–118, Guangzhou, China, September 2019.

G. Khan, Z. Tariq, J. Hussain, M. A. Farooq, and M. U. G. khan, "Segmentation of crowd into multiple constituents using modified mask R-CNN based on mutual positioning of human," in Proceedings of the International Conference on Communication Technologies (ComTech), pp. 19–25, Rawalpindi, Pakistan, March 2019.

A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," Artificial Intelligence Review, vol. 53, no. 8, pp. 5455–5516, 2020.

J. Du, L. Gui, R. Xu, and Y. He, "A convolutional attention model for text classification," Natural Language Processing and Chinese Computing, vol. 10619, pp. 183–195, 2018.

P. Bhattacharya, S. B. Patel, R. Gupta, S. Tanwar, and J. J. C. Rodrigues, "SaTyA: trusted Bi-LSTM-Based fake news classification scheme for smart community," IEEE Transactions on Computational Social Systems, no. –10, pp. 1–10, 2021.

Y. Wang, Y. Yang, W. Ding, and S. Li, "A residual-attention offline handwritten Chinese text recognition based on fully convolutional neural networks," IEEE Access, vol. 9, pp. 132301–132310, 2021.

A. Bala, I. Ismail, Ibrahim, and S. M. Sait, "Applications of metaheuristics in reservoir computing techniques: a review," IEEE Access, vol. 6, pp. 58012–58029, 2018.

K. Zhang, M. Sun, T. X. Han, X. Yuan, L. Guo, and T. Liu, "Residual networks of residual networks: multilevel residual networks," IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 6, pp. 1303–1314, 2018.

S. Y. Arafat and M. J. Iqbal, "Urdu-text detection and recognition in natural scene images using deep learning," IEEE Access, vol. 8, pp. 96787–96803, 2020.
[35] G. Peters, M. Lampart, and R. Weber, “Evolutionary rough k-medoid clustering,” in Transactions on Rough Sets VIII, J. F. Peters and A. Skowron, Eds., vol. 5084pp. 289–306, 2008.

[36] Q. Rao, B. Yu, K. He, and B. Feng, “Regularization and iterative initialization of softmax for fast training of convolutional neural networks,” in Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), pp. 1–8, Budapest, Hungary, July 2019.

[37] J. Gawlikiewska et al., “A Survey of Uncertainty in Deep Neural Networks,” 2021, http://arxiv.org/abs/2107.03342.

[38] M. Cai, Y. Shi, J. Kang, J. Liu, and T. Su, “Convolutional maxout neural networks for low-resource speech recognition,” in Proceedings of the 9th International Symposium on Chinese Spoken Language Processing, pp. 133–137, Singapore, September 2014.

[39] M. A. Hussain and T.-H. Tsai, “An efficient and fast softmax hardware architecture (EFSHA) for deep neural networks,” in Proceedings of the 2021 IEEE 3rd International Conference on Artificial Intelligence Circuits and Systems (AICAS), pp. 1–4, Washington DC, DC, USA, June 2021.

[40] C. C. Aggarwal, “Neighborhood-based collaborative filtering,” in Recommender Systems, C. C. Aggarwal, Ed., Springer, Cham, pp. 29–70, 2016.

[41] M. Srifi, A. Oussous, A. Ait Lahcen, and S. Mouline, “Recommender systems based on collaborative filtering using review texts-A survey,” Information, vol. 11, no. 6, pp. 317–6, 2020.

[42] K. Chatterjee and D. Mukherjee, “A new integrated likelihood for estimating population size in dependent dual-record system,” Canadian Journal of Statistics, vol. 46, no. 4, pp. 577–592, 2018.

[43] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, “Collaborative filtering recommender systems,” in The Adaptive Web, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds., vol. 4321, pp. 291–324, 2007.

[44] K. T. Ilayaraja, V. Vijayakumar, M. Sugadev, and T. Ravi, “Text recognition in moving vehicles using deep learning neural networks,” in Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIAS), pp. 279–283, Coimbatore, India, March 2021.

[45] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[46] Y. Yanagi, R. Orihara, Y. Sei, Y. Tahara, and A. Ohsga, “Fake news detection with generated comments for news articles,” in Proceedings of the 2020 IEEE 24th International Conference on Intelligent Engineering Systems (INES), pp. 85–90, Reykjavik, Iceland, July 2020.

[47] A. C. Rivera, M. Tapia-Leon, and S. Lujan-Mora, “Recommendation systems in education: a systematic mapping study,” Proceedings of the International Conference on Information Technology & Systems (ICITS 2018), pp. 937–947, Cham, January 2018.

[48] S. S. Haykin and S. S. Haykin, Neural Networks and Learning Machines, Prentice-Hall, New York, 3rd ed edition, 2009.

[49] L. E. B. Salasar, J. G. Leite, and F. Louzada, “Likelihood-based inference for population size in a capture-recapture experiment with varying probabilities from occasion to occasion,” Brazilian Journal of Probability and Statistics, vol. 30, no. 1, pp. 47–69, 2016.

[50] S. M. M. Seyednezhad, K. N. Cozart, J. A. Bowllan, and A. O. Smith, “A Review on Recommendation Systems: Context–Aware to Social-Based,” 2018, http://arxiv.org/abs/1811.11866.