International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Review Paper on Multilingual Deep Learning Framework for Fake News Detection using Capsule Neural Network

Sachin Dubey[1], Anshul Khurana[2], Vivek Jain[3]
[1]M. Tech Scholar(CSE), Department of Computer Science & Engineering, Shri Ram Institute of Technology, Jabalpur, MP.
[2]Head of Department, Department of Computer Science & Engineering, Shri Ram Institute of Technology, Jabalpur, MP.
[3]Assistant Professor, Department of Commerce, Shri Ram College of Commerce, Jabalpur, MP.

ABSTRACT

Fake news detection is an essential task; however, the complexity of several languages makes fake news detection challenging. It requires drawing many conclusions about the numerous people involved to comprehend the logic behind some fake stories. Existing works cannot collect more semantic and contextual characteristics from documents in a particular multilingual text corpus. To bridge these challenges and deal with multilingual fake news detection, we present a semantic approach to the identification of fake news based on relational variables like sentiment, entities, or facts that may be directly derived from the text. Our model outperformed the state-of-the-art methods by approximately 3.97% for English to English, 1.41% for English to Hindi, 5.47% for English to Indonesian, 2.18% for English to Swahili, and 2.88% for English to Vietnamese language reviews on TALLIP fake news dataset. To the best of our knowledge, our paper is the first study that uses a capsule neural network for multilingual fake news detection.
Keywords Fake news · Cross-language fake news · Machine learning · Deep learning

INTRODUCTION

Currently, fake news is more prevalent on social media channels as compared to traditional media . As a result of this problem, many researchers are focusing on developing several fake news detection frameworks, which is a crucial and challenging task. A model to detect fake news aims to spot or identify misleading news by analyzing previously reviewed real and fake news. As a result, a high-quality and large-scale dataset is required to perform this task accurately and efficiently. Fake news and natural language processing (NLP) researchers face the difficult task of creating multilanguage models that can be used in any of the world's 7,000 + languages. Multimodal data from multiple language can be challenging to collect and analyze simultaneously for a given task, making it essential to employ a framework that allows for the manageable generalization of a visual language model (Nelson et al., 2020).

In late 2019, the biggest pandemic around the globe Coronavirus Disease, known as COVID-19, generated a massive amount of informative data about COVID-19 (Nelson et al., 2020). Platforms for spreading such information, like mass media and social media, make it possible for the information to reach a large number of the audience (Cinelli et al., 2020). Unfortunately, not all of the information is accurate or trustworthy. Some of the information spreading around those platforms can be categorized as misleading or even identified as false news. Notably, various countries may have different situations and strategies to control the spread of COVID-19, which also leads to a considerable amount of inappropriate news sharing (Apuke & Omar, 2021). For example, "The spread of COVID- 19 is linked to 5G mobile networks", "Sunny weather protects you from COVID-19″, and "Place a halved onion in the corner of your room to catch the COVID-19 germs". Social media facilitated the rapid dissemination of this and similar false news stories. During the early stages of the pandemic. The wave of misinformation was so massive that the authorities had coined a word for it: "infodemic" (Shapley, 2016). Meanwhile, a lot of fake news was produced in various languages to spread more easily to particular ethnic groups. Thus, it is very challenging for authoritative organizations to respond promptly to the spread of fake news. Failing to detect and stop multilingual COVID-19 misinformation can lead to mistrust, panic buying, social distancing, and refusal to get tested and vaccinated.

Researchers have been focusing on machine learning-based NLP (Natural Language Processing) strategies to prevent the spread of misinformation. To identify misinformation, one study (Ozbay & Alatas, 2020) employed twenty-three supervised machine learning models. Increasingly successful models based on deep learning approaches have recently been adopted to detect misinformation. For example, (Aggarwal et al., 2020) used BERT to identify misinformation for short text and had excellent results. Similarly, (Patwa et al., 2021) propose a model for detecting fake COVID-19 news. Research studies have demonstrated outstanding results in classifying COVID-19 news using machine learning methods. Furthermore, NLP models trained in one language can be transferred to another without the need for extra annotated training data (Chu et al., 2021). In the process of adapting natural language processing (NLP) models to other languages, the zero-shot cross-lingual transfer is one method that shows promising results. Recent strategies for cross-lingual transfer have demonstrated that models can still generalize to a non-English language through the translation models for multiple languages or unsupervised language learning, followed by task-specific fine-tuning using only English annotation. This achievement can be traced back because several languages have significant common words in their vocabulary or structural in another way. For instance, several terms in the English and German vocabularies share a Latin root, such as the words “desk” and “tisch”. All languages have a repetitive structure, and many of their morphologies or word orders are the same in many languages.

Visual information is definitely necessary for the cross-lingual transmission of vision-language models. Detecting fake news in news statements of varying lengths using multilingual models is a crucial but understudied step in this direction (Vo, 2022). Many visual concepts are universally recognized, despite the fact that human languages may differ. Even though in English it is called “cat”, and in French, it is “chat”, both terms mean the same thing. In order to facilitate the transfer of vision-language models between languages, we learn to correlate phrases in various languages with perceptual notions by using this observation as a basis for our learning.

Galli et al. (2022) presented a detailed analysis of recent machine learning algorithms and their findings for fake news detection. This study provided the detail of datasets used for the experiments in terms of the efficiency and accuracy of the proposed models. Kanfoud and Bouramoul (2022) used standard language features for multilanguage sentiment analysis with high-accuracy results. The authors suggested that this is incredibly convenient for real-time applications. Zhang et al. (2022) proposed a technique for fake news detection based on the statement conflict to be classified as agree or disagree. The results showed the significance of the proposed technique and outperformed state- of-the-art methods using one dataset.

The fake news detection models based on textual contents (Yu et al., 2017, Shu, et al., 2019, Asghar et al., 2021) and visual features (Gupta et al., 2012, Qi et al., 2019) separately are ineffective with poor classification results. Thus, a multimodal approach has been introduced based on visual and textual features to improve the performance of fake news detection models. The convolutional neural network has been used in many studies to achieve the visual features for multimodal fake news detection (Wang et al., 2018, Yang et al., 2018, Singhal et al., 2019). However, CNN requires more training time and data to generalize the model and loses some crucial image features due to the pooling process that resultant in poor performance with combined with textual features. To deal with this issue, a capsule neural network has been proposed, which is faster than CNN and able to capture the most informative features from the image. In addition to this, we used LSTM to capture the sequence-to-sequence analysis to extract the textual features.

The goal of this study is to develop a multilingual framework based on capsule neural networks that can be used to improve the contextual representations of news text. How- ever, with the zero-shot default, a significant efficiency gap for queries exists between English and languages other than English. Since large-scale language models and train- ing have recently advanced, we propose a Multilingual technique to take advantage of the inadequate supervision provided by massive amounts of multilingual data in order to tackle this issue. Our technique has two significant advantages. Prior training on solely English data does not take advantage of the additional supervision provided by multi- lingual data and consequently does not improve the performance in any other language. To detect fake news from multilingual COVID-19, we propose a new framework for the detection of multilingual fake news using transfer learning and capsule neural net- works. Word embedding and n-gram feature extractors are used as feature extractors to develop models for detecting false news in various duration of press announcements.

CROSS-LINGUAL TRANSFER LEARNING

A survey of fake news detection methods, including the primary categories and the end aims of these methods, has been published (Chu et al., 2021). These researchers employed both linguistic and social approaches in their research: linguistic methods, where they extracted relevant language patterns from the news content, and network approaches, where they used criteria such as message metadata to make decisions about fresh incoming news.

Fake news detection is made more accessible with the help of Rubagotti and colleagues (Basile and Rubagotti (2018)) invention of an automated system known as CSI. The word- ing, response, and source of the incoming news are all factors taken into consideration by this detector when predicting the spread of false information. Three modules make up the model, the first of which extracts the articles' temporal representation. For classification purposes, the third and final module takes data from the first two modules, which are based on user behavior. The results of their trials show that CSI improves accuracy. Known as ISOT, a novel dataset for detecting fake news, Ahmed et al. The data in this set was gathered only from publicly available records. N-gram models and six machine learning approaches were employed to identify bogus news in the ISOT dataset. This combination of TF–IDF extraction and linear support vector machine classification yielded the best results.

New research by (Faruqui et al., 2014) attempts to determine whether or not a news article is bogus based on how many people have interacted with and/or liked it. They out- lined two distinct approaches to categorization. Using user interactions as characteristics, a logistic regression model is used in the first method. The subsequent is an innovative use of the crowdsourcing methods of the Boolean label. Both methods had impressive accuracy, as evidenced by the results. In addition, the trials demonstrated that taking into account the related informational data of people that engage along with the news is a significant feature in order to make a judgement regarding that news.

Rather than short statements containing bogus news material, (Pérez-Rosas et al., 2017) offered two new datasets connected to seven distinct categories, and their datasets contain authentic news extracts. Linguist-tic variables, and language features like syntactic, semantic, and lexical, can help identify between fake and real news using a linear support vector machine classifier. Human performance in this area was compared to the performance of the produced system.

TRANSFORMER-BASED MODELS

Later, multilingual mBERT was developed by the inventors of BERT by pre-training the BERT model using Wikipedia material in 104 languages (Wang et al., 2021). The multi- lingual mBERT to support several languages is mBERT. Later, many multilingual models were proposed following mBERT. In recent studies, large amounts of text were found to boost the model's performance.

Because of this, XLMR model was proposed (Hu et al., 2021), which is pre-trained on CC-100, which is a much larger corpus of text than Wikipedia's for low-resource languages. Significantly larger models like XLM-RXL and XLM- RXXL (Goyal et al., 2021) are pre-trained on CC-100 and obtain much better results. mBERT, XLM-R, and its variations are trained on non-parallel data. As a result, it is considerably more manageable for a model to acquire cross-lingual representations if it is trained on parallel and non-parallel data.

Sabour et al. (2017) introduced a capsule neural network, and it performed better than CNNs using the MINIST dataset. In the field of computer science, neural networks include capsule networks that make an effort to perform converse graphics processing. This method attempts, in order to figure out how the objective process that creates an image works back-wards. The capsule neural network consists of numerous capsules to perform functions and attempt to forecast parameters for instantiation and a specific object located in a specific position (Sabour et al., 2017).

Goldani et al. (2021) proposed a fake news detection model using a capsule network using two datasets, ISOT and LIAR. They used different embedding models to analyze the distance length of new items. The results demonstrate a 7.8% improvement in ISOT, a 3.1% improvement in the validation set, and a 1% improvement in the test set of the LIAR dataset when compared to the state-of-the-art approaches. The bidirectional encoder representations from transformers (BERT) model is used by Palani et al. (2022) to extract textual information, preserving word semantic links. Unlike CNN, the CapsNet model collects an image's most informative visual elements. Combining these variables creates a richer data representation that helps identify fake news. The proposed model outperformed the SpotFake + model using Politifact and Gossipcop datasets.

Sun et al. (2022) proposed a fake news detection model using bi-GRU and capsule net- works. This approach considered the contextual relationship in its entirety without sacrificing any of the text's spatial structural data. The suggested model employs a two-way GRU (BiGRU) network to enhance the capsule network, which can enhance the accuracy of rumor detection. The proposed model outperforms benchmark rumor detection algo- rithms on the Twitter dataset by 6.1% under static routing and 6.7% under dynamic routing.

Braşoveanu and Andonie (2019) proposed a hybrid approach incorporating machine learning, semantics, and NLP. They provided a semantic false news detection system based on text-extracted sentiment, entities, and facts. The findings of this study revealed that semantic characteristics increase false news categorization accuracy.

Word embedding models and their variants are discussed in detail in this section. For the second part of our study, we suggested two capsule neural network models for fake news identification based on the duration of the news statements.

Syntactic and semantic information can be extracted from words using dense word representation. Word embedding refers to the mapping of word representations to a low- dimensional space.
Closely related words are clustered together in the vector space in these representations. (Mikolov et al., 2017) proposed word2vec in 2013, a set of mathematical methods for generating word embeddings from raw text that has shown to be highly efficient. In order to build these models, enormous amounts of text are used to train neural networks with two layers of input. Using these methods, each word can be represented as a vector in a vector space with hundreds of dimensions. Similar-sounding words are clustered together in this area.
Word2vec vectors like "Google News", which is trained on 100 billion Google news words, are now available.

Using these pre-trained vectors for initializing word vectors is a popular way to increase text processing performance, especially in the absence of an extensive, supervised training dataset. Deep neural networks can employ these distributed vectors for text classification (Mobiny & Van Nguyen, 2018). However, further refinement of these pre- trained embeddings is possible. The authors (Mobiny & Van Nguyen, 2018) were the first who used multiple learning set- tings for vector representation of words using word2vec and demonstrated their superiority over standard pre-trained embeddings in a CNN model. These are the options: pre-trained vectors are utilized as inputs for the framework of the neural network in the static word embed- ding model. There are no changes made to these vectors while they are being trained.
This model uses the pre-trained vectors at the beginning of the learning process. Although the vectors are adjusted for optimal performance for each job during training, they are based on training data for that specific target task. It uses two sets of word2vec vectors, one non- static and the other during training fine-tuned .

THE PROPOSED FRAMEWORK – MULTILINGUAL-FAKE MODEL

We proposed a multilingual capsule network learning-based model with multilingual embed- dings integrated with semantics infusion. There are several layers in the proposed architecture for effective multilanguage fake news detection. The semantic infusion is used for extra lexical semantics (Fig. 1).
Let there be two input layers w_1, w_2, …… , w_n for O_s input, which is the source language and w_1, w_2, …… , w_n for T_s inputs for translated target language provided to the embedding layer. The embedding layer receives the input sequence, which contains a sequence of words as text. This input sequence is a matrix $X \in \mathbb{R}^{e \times V}$ where V is vocabulary size, and e represents a vector of words. This input also includes extra semantic infusion for fake news detection. The feature extraction layer contains the forward hidden units and backward hidden units aligned with the bidirectional Long Short-Term Memory (Bi-LSTM) network. The concept of using LSTM is to extract the contextual feature relationship with the local features.

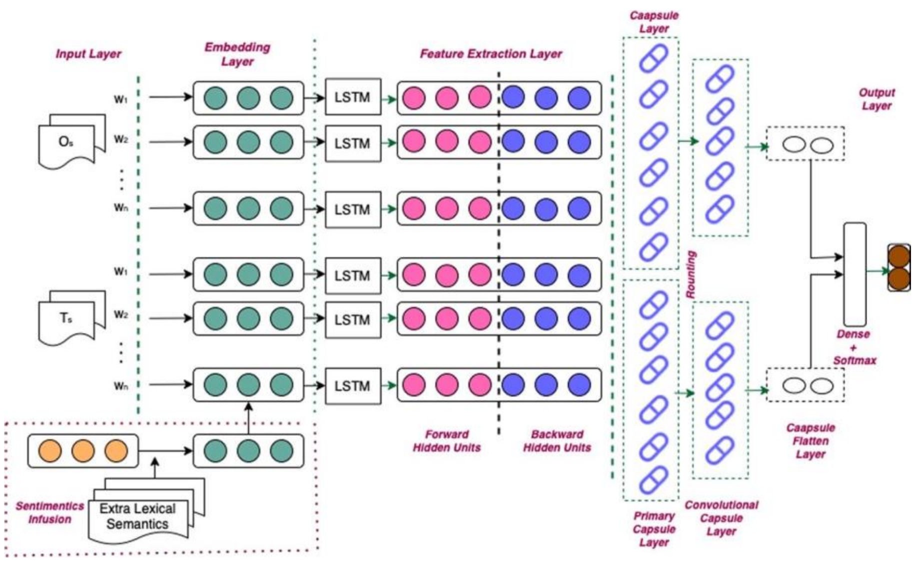

Fig. 1 Proposed multilingual deep learning architecture for Fake News Detection using capsule neural net- work.

REFERENCES

Aggarwal, A., et al. (2020). Classification of fake news by fine-tuning deep bidirectional transformers-based language model. *EAI Endorsed Transactions on Scalable Information Systems, 7*(27), e10–e10.

Aghakhani, H., et al. (2018). Detecting deceptive reviews using generative adversarial networks. In 2018 IEEE Security and Privacy Workshops (SPW), IEEE.

Apuke, O.D., & Omar, B. (2021). Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, *56*: 101475.

Asghar, M. Z., et al. (2021). Exploring deep neural networks for rumor detection. *Journal of Ambient Intel- ligence and Humanized Computing, 12*, 4315–4333.

Basile and Rubagotti (2018).Basile, A., & Rubagotti, C. (2018). CrotoneMilano for AMI at Evalita2018. A performant, cross-lingual misogyny detection system. EVALITA Evaluation of NLP and Speech Tools .u, A.M., & Andonie, R. (2019). Semantic fake news detection: a machine learning perspective. In Advances in Computational Intelligence: 15th International Work-Conference on Artificial Neural Net- works, IWANN 2019, Gran Canaria, Spain, June 12–14, 2019, Proceedings, Part I 15, Springer.

Chen, X., et al. (2018). Adversarial deep averaging networks for cross-lingual sentiment classification. *Transactions of the Association for Computational Linguistics, 6*, 557–570.

Chu, S. K. W., et al. (2021). Cross-Language fake news detection. *Data and Information Management,5*(1), 100–109.
Cinelli, M., et al. (2020). The COVID-19 social media infodemic. *Scientific Reports*, *10*(1), 1–10.
Conneau, A., et al. (2020). Unsupervised cross-lingual representation learning at scale. *ACL*.
De, A., et al. (2021). A transformer-based approach to multilingual fake news detection in low-resource languages. *Transactions on Asian and Low-Resource Language Information Processing, 21*(1), 1–20.

Devlin, J., et al. (2019). Pre-training of deep bidirectional transformers for language understanding In: Pro- ceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Minneapolis, MN: Association for Computational Linguistics: 4171–4186.
Faruqui, M., et al. (2014). Retrofitting word vectors to semantic lexicons. arXiv preprint arXiv:1411.4166.

Galli, A., et al. (2022). A comprehensive Benchmark for fake news detection. *Journal of Intelligent Infor- mation Systems, 59*(1), 237–261.
Goldani, M.H., et al. (2021). Detecting fake news with capsule neural networks. *Applied Soft Computing, 101*: 106991.
Goyal, N., et al. (2021). Larger-scale transformers for multilingual masked language modeling. REPL4NLP.

Gupta, M., et al. (2012). Evaluating event credibility on twitter. Proceedings of the 2012 SIAM international conference on data mining, SIAM.

Hu, J., et al. (2021). Explicit alignment objectives for multilingual bidirectional encoders. NAACL.

Kanfoud, M. R., & Bouramoul, A. (2022). SentiCode: A new paradigm for one-time training and global pre- diction in multilingual sentiment analysis. *Journal of Intelligent Information Systems, 59*(2), 501–522.

Kingma, D.P., & Ba, J. (2015). Adam: A Method for Stochastic Optimisation. CoRR abs/1412.6980.

Li, Z., et al. (2020). Learn to cross-lingual transfer with meta graph learning across heterogeneous lan- guages. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Manning, C.D., et al. (2014). The Stanford CoreNLP natural language processing toolkit. Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations.

Mikolov, T., et al. (2017). Advances in pre-training distributed word representations. arXiv preprint arXiv:1712.09405.