Detecting Stance in Urdu Content on Social Media and Websites for Fake News and Propaganda Identification

Shan Muhammad Khan[1], Babar Jehangir[2], Dr. Muhammad Imran[1]
[1]Department of Robotics & Artificial Intelligence (SZABIST University Islamabad).
[2]Department of Computer Science (SZABIST University Islamabad).
***Correspondence:** Babar Jehangir Email: babar.jehangir@szabist-isb.edu.pk.

Citation | Khan. S. M, Jehangir. B, Imran. D. M, "Detecting Stance in Urdu Content on Social Media and Websites for Fake News and Propaganda Identification", IJIST, Vol. 07 Issue. 02 pp 1396-1408, July 2025
DOI| https://doi.org/10.33411/ijist/20257313961408
Received| June 08, 2025 **Revised|** June 30, 2025 **Accepted|** July 05, 2025 **Published|** July 11, 2025.

The extensive spread of fake information has rendered various news types questionable, leading to a significant decline in trust in news. Social media is the primary channel by which fake news is disseminated widely. Worldwide, several deep learning methods have been created to identify fake news, achieving significant success with content in the English language. However, to our knowledge, there is no deep learning method available for detecting fake news or stance detection in content written in Urdu. Therefore, it is crucial to create a method that can detect fake news within Urdu language content. This study seeks to identify a method for detecting fake news in the Urdu language by proposing a framework that employs advanced Bidirectional Encoder Representations from Transformers (BERT), Embeddings from Language Models (ELMO), and various deep learning models (CNN, LSTM, Bi-LSTM) to evaluate performance accuracy on Urdu datasets (Liar-ProSOUL and Bend the Truth-Benchmark). We utilized Embeddings from Language Models (ELMO) for feature extraction and a convolutional neural network (CNN) for the classification task. The findings from the suggested framework indicate that ELMO excels with extensive datasets

Keywords: Deep Neural Networks, Urdu Fake News Detection, Natural Language Processing, ELMO, Deep Machine Learning

Introduction:

The world has fully transitioned into the Digital Age, with social media emerging as one of its most influential innovations. These platforms have united people around the world like never before, enabling constant and easy interaction. Facebook claims to have over 2.7 billion active users worldwide. YouTube boasts 1.9 billion monthly active users, WhatsApp has 1.5 billion, Messenger records 1.3 billion, and Instagram has around 1 billion monthly active users. This list only scratches the surface of the many social media apps available today. This level of interconnectedness has opened up endless possibilities across business, economic, and cultural spheres.

These apps handle hundreds of billions of dollars in transactions. However, it has a number of drawbacks [1]. Some of the disadvantages are minor, such as catfishing, but others are serious. Since 2016, the spread of false information, ranging from misleading news to intentional disinformation has grown noticeably, tearing communities apart, increasing polarization, and fueling violent behavior. Without a doubt, political systems in many countries have been the greatest victims of fake news.

The journalism industry has suffered the most significant blow to its credibility and its capacity to deliver accurate information essential for an informed social and political life. Because this issue is relatively new, no reliable, scalable systems have yet been developed to effectively address and counter it. Facebook and Twitter have made some progress in developing artificial intelligence as well as deep learning systems to detect, track, and remove fake news but these are evolving technologies and provide solutions to fake news spread in the English language. Commercially available and reliable systems to detect fake news are still non-existent. Theory-driven models of the Detection of fake news are being evaluated using real-world datasets. Pakistan faces a particularly daunting challenge in detecting, tracking, and eliminating fake news because most of the content is in Urdu [2]. Despite the widespread presence of Urdu content across the Indian subcontinent, the Middle East, and among the Urdu-speaking diaspora in Europe and the United States, no advanced deep learning systems tailored for the Urdu language have been developed to date. In Pakistan alone, Facebook reports approximately 39 million monthly active users. In comparison, research on this problem remains virtually non-existent. While several stance detection models have been developed, including both neural network-based approaches and traditional classifier-based methods, none have been specifically tailored to the Urdu language. This research explores reproducibility with a special interest in transfer learning because of its recent breakthroughs in Natural Language processing. Initial investigations in previous studies have identified the use of various artificial intelligence algorithms and components, including Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) models, Multi-Layer Perceptrons (MLPs), and Artificial Neural Network (ANN) architectures [3]. Nevertheless, these have achieved restricted success in instance detection. In this new approach, I will emphasize word embedding and deep context through ELMO (utilized in advanced language models like GPT-3 and MEA), a deep learning algorithm that has gained popularity in the business industry in recent years

Given its novelty and ambiguity, there is still no universally accepted definition of the term "fake news." However, some evolving definitions are gradually gaining traction. For instance, Cornell University defines fake news as “false information that imitates news media content in structure, but not in its organizational process or purpose. In contrast, fake news websites do not adhere to the editorial standards and procedures that the news media uses to verify the accuracy and authenticity of information [4]. For human and democratic societies, Fake news presents serious problems. According to Matthew Baum at Harvard University, “Recent changes in the media landscape heighten worries about democratic societies' susceptibility to false information and the public's restricted capability to manage it” [5]

Another recent event that highlighted the dangers of social media was the Coronavirus pandemic. Similar to the disinformation observed during the U.S. election, deliberately misleading content was disseminated through social media platforms, promoting fake medicines, therapies, and cures for the virus. Recently, a large-scale Indian disinformation campaign was unearthed by EUDisinfoLab that targeted European leaders and public opinion on the Kashmir issue. [6]

In Pakistan, polio immunization has suffered a huge setback because of fake news online. In the last ten years, 70 [7] polio workers have been killed by extremists inspired by these kinds of views. Fake news of blasphemy on social media has incited violence against minorities in Pakistan and the world over. To stop such wanton actions, policymakers are looking for solutions to curb the spread of fake news. Advanced algorithms that learn from deep learning, machine learning, and neural networks are being employed to build autonomous systems that can detect and eliminate fake news online [8]. Some of the ways in which this is done are by utilizing neural models with bidirectional encoder portrayal from transformers embedding. Sentences from fake news content are recognized by fine-grained examination of the text. In some cases, the text style and intelligibility have been to identify fake news. However, various algorithms are available to detect fake news in the English language content. There are no Urdu language processing fake news detecting solutions available. Some deep learning models used for fake news detection in the literature so far. The purpose of this study is to design a framework using a deep learning approach to detect fake news from Urdu language content.

Related Work:

The rapid proliferation of fake news across digital platforms has prompted extensive research into automated detection techniques. Numerous algorithms and applications have been proposed globally, continually evolving to enhance accuracy and reliability in identifying misinformation. In the first research the researcher used a label data set of fake news and detected using an advanced deep learning approach and traditional natural language processing techniques. The study revealed that the adoption of complex techniques does not always guarantee better classification performance. [9] In another research, researchers have identified propaganda elements within online Urdu literature [10]. They compiled and labeled a dataset of over 11,574 Urdu news articles to train machine learning classifiers. Then they formed a linguistic examination and word calculation dictionary to get the psycho-linguistic characteristics of the Urdu manuscript. They evaluated different classifiers and found that the Pro-SOUL framework performed most useful in identifying propaganda in the online Urdu-based news content compared to the overall internet content. In another case, the researchers used 5 data sets of different fields and crafted various text representation features and their combinations. The results reveal a sizeable performance gained by the Ada-Boost classifier with 0.870 F1 for Fake news and 0.90 F1 score for Real news. [11]. A three-part method with the help of a Naive Baye Classifier, semantic analysis, and support vector machines has also been proposed, which is an exact way to identify false news on different online media platforms [11].

Another study proposes a Multi-Kernel Optimized Convolutional Neural Network (MOCNN) optimized via grid search for fake news detection [12]. MOCNN outperforms ten deep learning models, achieving 85.8% (UFN) and 68.2% (BET) accuracy, confirming its superior performance. In another study, MuRIL and T5 models were used for Urdu fake news detection, with a manually verified dataset covering diverse topics [13]. Experimental results show that MuRIL outperforms T5, achieving a 0.96 F1 score and 0.83 validation accuracy. Other researchers have identified propaganda elements within online Urdu literature. They compiled and labeled a dataset of over 11,574 Urdu news articles to train machine learning classifiers. Qualitative and quantitative data analysis has been applied to two different data sets.

Another study applied qualitative and quantitative data analysis to two different data sets. The first is from authentic texts downloaded from a website and the other contains twenty news reports extracted from different Facebook pages. The study found that linguistic features help determine unreliable texts and show that most of the test news data sets tend to be untrustworthy articles [14].

Test results were obtained using two datasets constructed from Sina Weibo and Twitter, where Convolutional Neural Networks (CNN) and other neural network techniques were applied. News classification showed that the proposed novel deep neural network model could detect fake news with a significant accuracy of more than 90 percent accuracy within five minutes after it starts to spread and before it is shared fifty times. This is substantially faster than state-of-the-art baselines [15]. An investigation was carried out to detect fakes in an unsupervised way. The experiments conducted on two data sets showed that the collapsed Gibbs sampling approach outperformed the unsupervised method [16].

An attempt was made to enable the users to come up with a solution to find and filter out those sites that contain misleading and false information. To detect fake posts, key features extracted from titles and content were utilized. The experiments demonstrated that the logistic classifier reached an impressive accuracy of 99.4% [17]. A study conducted has raised awareness about opportunities and concerns for different businesses that are trying to automatically detect fake news. This study described two contradictory approaches and proposed algorithmic solutions [18]. To aware people of the right information, a competitive independent Cascade Model with User Biasand K- Truth score has been proposed. The researchers presented effects on a real-world network in the study, and the products depicted that the K-truth score outpaced the original methods [19].

Similarly, topology and interaction-based trust properties of nodes in real-world Twitter networks have been used to identify false information spreaders, achieving an accuracy of over 90%. [20]. The MisInfo Text Repository was introduced to support the research community after a comprehensive review of existing datasets. As part of this study, a topic modeling experiment was conducted. The research also emphasized the need for collecting additional data to facilitate future research initiatives. [21]. Based on Bi-directional LSTM-recurrent neural network a fake news detection model was presented. The data sets included unstructured news articles from two publicly available sources. It was found that the Bi-directional LSTM model is superior to CNN, unidirectional LSTM, and vanilla RNN for the detection of fake news [22]. Using dynamic relational networks, a bottom-up approach with mutual, relative, and changing reliability evaluation has been proposed by another study. However, this leads to the problem of fake assessment [23].

In another study, agents were characterized and the platform inspection problem was examined, revealing that optimal policy adaptation to user-generated inspections can exhibit behaviors that may not be immediately intuitive, particularly in environments with a low incidence of fake news [24]. A systematic literature review identified key drivers of fake news dissemination, including social conformity and peer influence, social segregation, political and financial motivations, ignorance, and the intentional spread of malicious content [25]. Furthermore, a deep learning-based method was proposed to classify the authenticity of claims using a modular architecture built on Deep Neural Networks. Experiments conducted on benchmark datasets demonstrated classification accuracies of 72% and 81% for two distinct models, with an overall certainty of 82.4% in detecting fake news [26].

Objectives:

The primary objective of this study is to develop and evaluate a deep learning–based framework for detecting fake news in Urdu language content. Specifically, the research aims to:

- Identify and apply suitable deep learning models including CNN, LSTM, Bi-LSTM, and ELMo for Urdu stance detection.
- Compare performance across traditional and pre-trained models using real-world datasets.
- Address the technological gap in Urdu fake news detection by proposing a high-performing, language-specific solution.

Material and Methods:

In this study, we proposed a framework that utilized the ELMo model to detect fake news in the Urdu language. The performance of the proposed framework was evaluated on two datasets of different sample sizes, along with their respective baseline results. We utilized the deep learning (CNN) model to classify the Urdu fake news detection process which achieve high accuracy as compared to existing models. The proposed model to detect fake news is presented in the Figure. 1.

Data Preprocessing:

In the pre-processing phase, we used various text pre-processing methods to clean the data and extracted features using the ELMo feature extraction technique. After that, we used different deep-learning models to train our model based on the given data. The phases of our framework model are defined in the next sections with the section flow respectively.

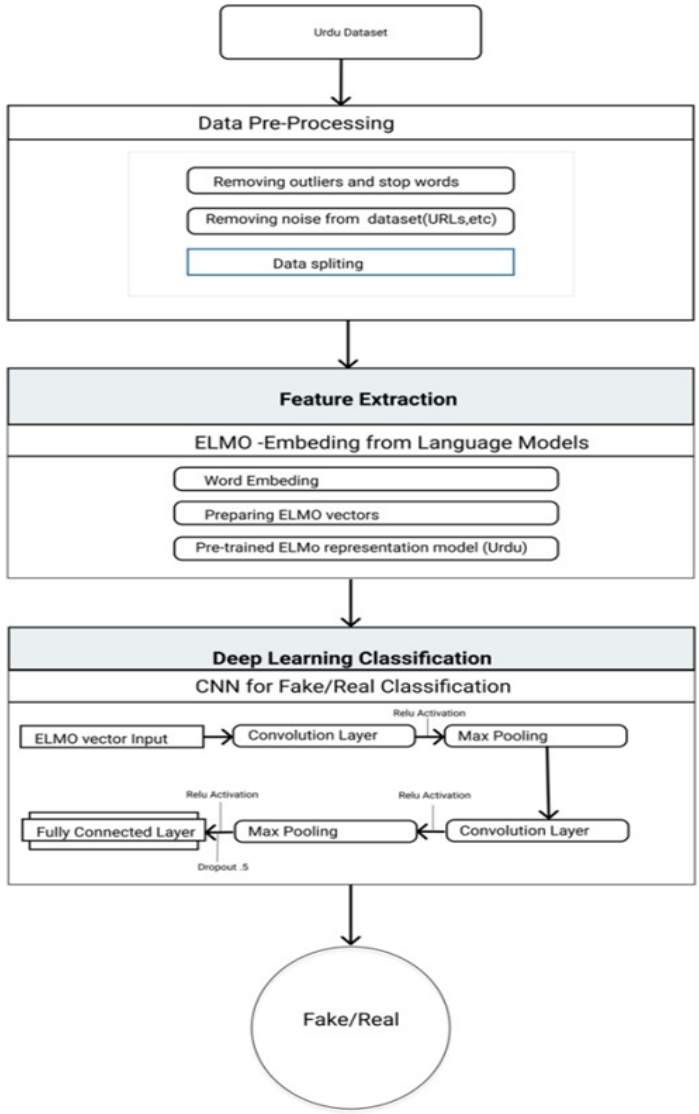

Figure 1. Architecture of Fake News Detection Framework

Word Embedding:

Word embedding is a popular way to represent document vocabulary or vector representations of specific words [27]. It is capable of detecting a word's context in a document, semantic and syntactic resemblance, and the relationship with other words. Usually, as a real-valued vector that represents the word's meaning, words located near each other in vector space are anticipated to have similar meanings. There are two categories of embeddings: traditional and contextual, with ELMO utilizing contextual embeddings. Contextual embedding is a representation to allocate each word based on its context, capturing word usage across contexts and encoding the knowledge. To compare the news length between the two classes (fake and real), we implemented a function and created a visualization that displays the distribution of news lengths for both fake and real news. The visual form of this data is shown in the graph Figure. 2 below:

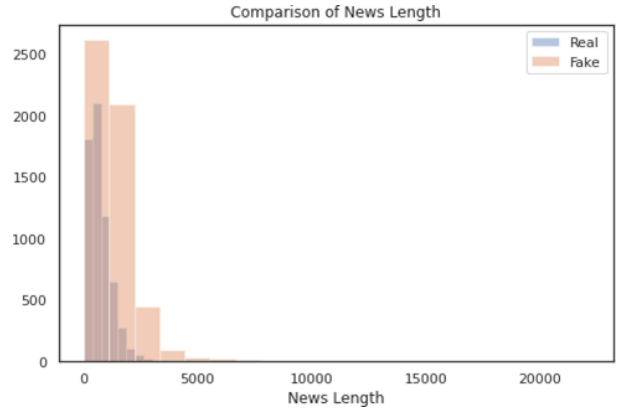

Figure 2. Comparison of news length (Fake and Real)

Figure. 3 and Figure. 4 show the total word count in the news article that considered fake words are shown. The most to the least fake words are listed in the visualization below.

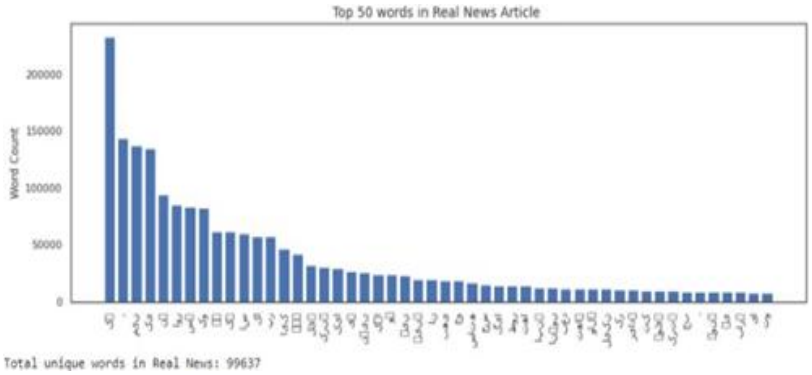

Figure 3. Total word counts

Figure 4. Word Cloud

In the pre-processing phase, we have used different text pre-processing methods to clean the data and extract the features through feature extraction techniques ELMO. After

that, we trained our model using various deep-learning architectures based on the provided data. The phases of our framework are detailed in the following sections, presented in sequential order.

ELMO-Embeddings from Language Models:

It is well known that an effective language embedding model should provide efficient representations, be sensitive to contextual information, and employ vector mathematics to capture semantic relationships. Techniques like Word2Vec and GloVe have been commonly used for this purpose; however, they do not account for context, as they generate static embeddings. Whereas the ELMO solves many of the problems presented by these. It is context-aware and understands the words and their distinct meanings in the context of the sentence. It is deep and character-based too, meaning the words learn from pre-trained neural networks and allow clues for representation even when the word is new and was not in the training [30].

ELMO is more efficient the traditional word embeddings. It is, however, not tested for Urdu fake news detection. Fig. 5 (a) below shows the architectural representation of the mathematical form. Fig. 5 (b) shows vanilla language model flow from the previous related work (bi-directional) in which x shows the kind of input (word) that processes it further with a layer of LSTM (the arrows show forward and backward responses) and that produces the sequence of output token. The Softmax output function section transforms a previous layer's output into a vector of probabilities. It is commonly used for multiclass classification. This is the stage to give all possible combinations of the semantic units. Given an input vector x and a weighting vector w, we have Equation 1:

$$P(y = j \mid x) = \frac{e^{x^T w_j}}{\sum_{k=1}^{K} e^{x^T wk}} \quad (1)$$

The Fig. 5 (b) shows the flow of biLM + ELMo. So, it shows that ELMo enhanced the percentage of the overall results with the previous work but their contribution architecture is quite similar to the previous model (like-vanilla) and the explanation of the linear combination of the vector length is also not written in the paper which is unsatisfied. We have used this model feature with our framework to check if it increases the accuracy of the identification of fake news from the dataset or not. In the equation below, softmax-normalized weights, which are a scalar parameter that enables the task model to scale the ELMo vector in its entirety. It is essential to assist the optimization process. Applying layer normalization to each biLM layer before weighting, improved in some situations. The following equation Equation 2 is the mathematical form of ELMo.

$$ELMo_k^{task} = E(R_k; \theta^{task}) = \gamma^{task} \sum_{j=0}^{L} S_j^{task} h_{k,j}^{LM} \quad (2)$$

ELMo utilizes multiple layers of recurrent units and retains the representations from all internal layers, including the final recurrent layer. This enables the generation of task-specific embeddings by allowing combinations of representations from different layers through linear weighting.

Experiments and Results:

To evaluate our framework on this dataset, we used TensorFlow version 1.5 on a non-GPU Google Colab environment, implementing a two-layered CNN model. L2 regularization was applied during the testing phase. The deep learning classifier for the weighting scheme was also arranged and formulated. The tests were run on the real and fake datasets separately and also on the combined dataset to get overall result accuracy.

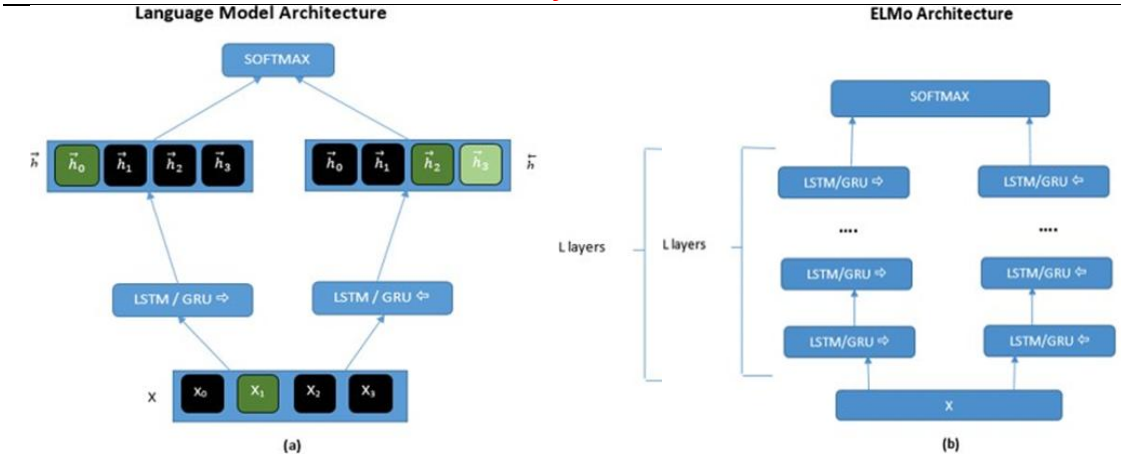

Figure 5 (a) and (b). Conversation Flow of Language Model Architecture into ELMo Architecture

Dataset: Two different datasets were used in our research and we set up two different experimental setups for each dataset. The Urdu Benchmark Dataset "Bend the Truth" is manually curated and labeled as the Urdu dataset available publicly. And the LIAR dataset is a translated dataset from English. The brief statics of the datasets are shown in Fig. 6. The Benchmark Urdu dataset used a total of 900 news including domain multiple topics; business, health, sports, technology, and showbiz. Whereas the other labeled dataset includes the propaganda and non-propaganda of 11, 574 fake news.

On our large dataset, we found that n-gram features performed the best among traditional deep learning models, achieving 93 percent accuracy. Our experiments on the large dataset showed that n-gram features yielded the highest accuracy among traditional deep learning models, reaching 93%. Empath-generated features did not show promising results for fake news detection, primarily due to their lack of context awareness, despite their earlier use in identifying deception in review systems. In Table 2 and Table 3, we presented the results of various deep-learning models. While the basic CNN model was widely regarded as the best for the LIAR-ProSoul dataset, our findings showed that it ranked as the second-best among all evaluated models. For this dataset, LSTM-based models were the most susceptible to overfitting, as evidenced by their performance. Despite the fact that Bi-LSTM suffers from overloading on the Labeled dataset, we rank it as the third-best neural network-based model based on its performance on the Urdu dataset.

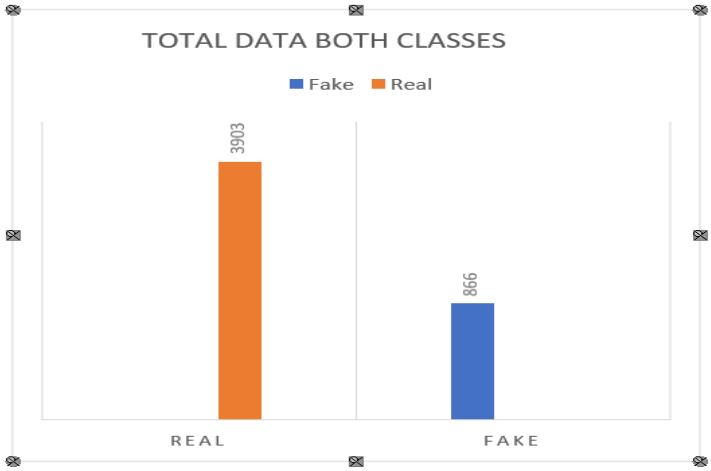

Figure 6. Total news items in the dataset

Table 1. Selected Datasets

		Bend the Truth Dataset [9]	LIAR Urdu Dataset [8]
Train	Fake	288	8,101
	Real	350	
Test	Fake	150	3,473
	Real	112	
Total		**900**	**11,574**

LSTM-based models performed best on the dataset, with both Bi-LSTM achieving 0.95 accuracy and 0.95 F1-score. On the other hand, CNN and all hierarchical attention models perform admirably, with accuracy and an F1- score of over 0.90. Although neural network-based models may suffer from overfitting on a small dataset (Benchmark Dataset), they demonstrate good accuracy and F1-score on a fairly big dataset, according to this finding (Labeled Dataset). Since the performance of the Bi-LSTM model improves with an increase in the amount of training data, it can be concluded that sufficient training samples are essential for optimal results. Based on the above analysis, we found that deep learning models perform effectively on Urdu datasets, comparable to their performance on datasets in other languages. In the comparison, we found that Bi-LSTM was the most recommended among the deep learning models we used, while CNN was considered the second-best for the Urdu datasets.

Table 2. Traditional Deep Learning Model Performance Evaluation

Models	Datasets							
	Benchmark – BendTruth				Liar – Prosoul Dataset			
	A	P	R	F1	A	P	R	F1
Ada Boost [9]	.87	.88	.86	.86	.82	.82	.58	.57
CNN [8]	.90	.89	.95	.90	.90	.89	.89	.90

Table 3. Pre-trained Advanced Deep Learning Language Model Performance Evaluation

Models	**Feature**	**Datasets**							
		Benchmark – Bend Truth Urdu Dataset				Liar – Prosoul Labeled Urdu Dataset			
		A	P	R	F1	A	P	R	F1
ELMo		.91	.90	.92	.91	.93	.91	.91	.93

We evaluated the performances of two pre-trained language models on two Urdu datasets. While these models have more sophisticated architectures, they do not suffer from overloading as much as deep learning models do on smaller datasets. It was because of the exception of the final classification layers, that all of the layers in these models use pre-trained weights. Consequently, fine-tuning their compound architectures did not require large datasets, and all the pre-trained models we tested performed well on both traditional language models and deep learning-based models, achieving F1-scores of 0.93 on the LIar-labeled dataset (Table 3) and 0.91 on the Benchmark dataset (Table 2). These pre-trained models perform better in the Urdu fake news detection job on the large dataset (i.e., Liar – Prosoul Labeled Dataset). We noticed that comparing the pre-trained language models, the ELMo is generally better than the previously used model for Urdu context without ELMO on a large dataset. For example, [9] achieve 0.91 accuracy, on the large dataset while ELMo achieves 0.92. We also found that the performance of transformer-based models was proportional to the number of their parameters and the extent of pre-training. As shown in Table 4, we also report the performance of traditional machine learning models such as AdaBoost, CNN, Bi-LSTM, and the pre-trained ELMo models. The results confirm that ELMo consistently outperforms other models on both datasets, especially on the larger dataset, where ELMo achieved a higher accuracy and F1 score than traditional deep learning models.

Overall, our findings suggest that pre-trained language models, particularly ELMo, are more suitable for fake news detection in Urdu, especially when working with large datasets. The performance gains observed with ELMo are significant, making it a highly recommended model for future applications in Urdu fake news detection.

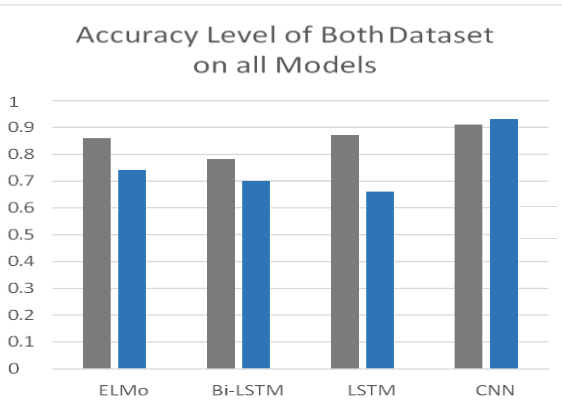

Figure 7. Performance Accuracy Result of both Datasets on all Deep Learning Models

ELMo performance accuracy was noted as 0.93 during the experiment on dataset_2. We noticed our model performed 2 to 3 % better than our baseline papers ProSOUL and Bend the Truth there are two reasons for this better performance.

Elmo can uniquely account for a word's context. Previous language models such as n-gram, GloVe, Bag of Words, and Word2Vec simply produce an embedding based on the literal spelling of a word. They do not factor in how the word is being used. Secondly, the original ELMo model was trained on a corpus of 5.5 billion words, and even the "small" version had a training set of 1 billion words. That's a lot of data! Being trained on that much data means that ELMo has learned a lot of linguistic knowledge and will perform well on a wide scope of datasets.

Table 4. Summary results of all deep learning models on Dataset 1 (Bend the Truth) and Dataset 2 (LIAR Translated).

Model	Dataset	Accuracy (A)	Precision (P)	Recall (R)	F1-Score(F1)
AdaBoost	Benchmark – BendTruth	0.87	0.88	0.86	0.86
	Liar – Prosoul Dataset	0.82	0.82	0.58	0.57
CNN	Benchmark – BendTruth	0.90	0.89	0.95	0.90
	Liar – Prosoul Dataset	0.90	0.89	0.89	0.90
Bi-LSTM	Benchmark – BendTruth	0.89	0.88	0.90	0.89
	Liar – Prosoul Dataset	0.90	0.94	0.96	0.95
ELMo	Benchmark – BendTruth	0.91	0.90	0.92	0.91
	Liar – Prosoul Dataset	0.93	0.91	0.91	0.93

As we know, ELMo works independently to generate pre-trained LSTM features from left-to-right and right-to-left for the downstream task. Where ELMo accuracy performance was high in the large Urdu dataset. As it outperforms independently on the downstream task from both sides of the words. Fig. 7 below shows the visual representation of the summarized result of both Urdu datasets on all deep learning models we used and evaluated in our experiment with the performance accuracy rate on each.

Discussion:

The findings of our study demonstrate that the proposed ELMo-based deep learning framework performs significantly better than traditional models for fake news detection in Urdu language content. On the LIAR-translated dataset, ELMo achieved an accuracy of 93%, outperforming CNN (90%) and Bi-LSTM (90%) (Table 4). These results are consistent with

previous studies that highlight the effectiveness of contextual embeddings in fake news detection tasks.

For example, author[12] reported an accuracy of 85.8% using a Multi-Kernel Optimized CNN on Urdu datasets, which is lower than the 93% accuracy achieved by our ELMo-based model. Similarly, author[13] compared MuRIL and T5 models and found MuRIL to perform better with a 0.96 F1-score, though this was achieved using a manually verified dataset not publicly available, which limits reproducibility. Compared to these studies, our framework benefits from the use of pre-trained ELMo embeddings, which significantly reduce dependency on large labeled datasets and enhance contextual understanding.

ProSOUL, as described by author[10], utilized a machine learning pipeline with a labeled Urdu dataset, achieving good performance in propaganda detection. However, our approach improves upon their framework by incorporating deep contextual embeddings, resulting in more robust performance across both propaganda and general fake news categories.

Despite the superior results, our framework has some limitations. First, it depends heavily on the quality of preprocessing and tokenization for Urdu, which lacks standardization compared to English. Second, while ELMo performs well with relatively smaller datasets, transformer-based models like BERT may offer better scalability and contextual depth, especially when trained on larger corpora.

Conclusion:

In this study, we proposed a deep learning framework for fake news detection in the Urdu language, leveraging both traditional architectures and pre-trained language models. Our experimental results demonstrate that ELMo-based models significantly outperform conventional approaches such as CNN and Bi-LSTM, particularly on larger datasets. These findings underscore the importance of contextual embeddings and transfer learning in enhancing performance for under-resourced languages like Urdu. This research addresses a critical gap by introducing a scalable and effective solution for fake news detection in regional languages.

Future Work:

In this study, we evaluated five deep learning models within our proposed framework over two Urdu datasets for fake news detection. Firstly, we critically reviewed popular deep learning models used for fake news detection in the literature. After that, we developed our framework using both deep learning models and pre-trained deep learning models. The overall results show that each model has its strengths and limitations depending on the size of the dataset. For instance, the Bi-LSTM achieved higher accuracy on the large dataset, while the CNN model performed more accurately on the small dataset. Furthermore, evaluation of pre-trained deep learning models revealed that ELMo achieved the best performance on the large Urdu dataset. In future research, we aim to expand the scope of our framework by incorporating additional Urdu fake news datasets from diverse domains, such as politics, health, and finance, to further validate model robustness. We also plan to explore transformer-based architectures, such as BERT and its multilingual variants, to assess their effectiveness in capturing contextual nuances in the Urdu language. Moreover, integrating explainable AI techniques will be a key focus to enhance model interpretability and provide deeper insights into fake news detection decisions.

References:

[1] I. A. Muhammad Shoaib Farooq, Ansar Naseem, Furqan Rustam, "Fake news detection in Urdu language using machine learning," *PeerJ Comput Sci*, vol. 9, p. e1353, 2023, doi: 10.7717/peerj-cs.1353.

[2] N. A. & M. A. A. Sheetal Harris, Hassan Jalil Hadi, "Multi-domain Urdu fake news detection using pre-trained ensemble model," *Sci. Rep.*, vol. 15, no. 8705, 2025, doi:

https://doi.org/10.1038/s41598-025-91054-4.

[3] M. G. & J. K. Muhammad Rizwan Rashid Rana, Asif Nawaz, Saif Ur Rehman, Muhammad Ali Abid, "BERT-BiGRU-Senti-GCN: An Advanced NLP Framework for Analyzing Customer Sentiments in E-Commerce," *Int. J. Comput. Intell. Syst.*, vol. 18, no. 21, 2025, [Online]. Available: https://link.springer.com/article/10.1007/s44196-025-00747-1

[4] S. Munir and M. Asif Naeem, "BiL-FaND: leveraging ensemble technique for efficient bilingual fake news detection," *Int. J. Mach. Learn. Cybern.*, vol. 15, no. 9, pp. 3927–3949, Sep. 2024, doi: 10.1007/S13042-024-02128-0/METRICS.

[5] D. I. Mahmood, "Disinformation and Democracies: Understanding the Weaponization of Information in the Digital Era," *Policy J. Soc. Sci. Rev.*, vol. 2, no. 2, Aug. 2024.

[6] R. Z. Xinyi Zhou, Atishay Jain, Vir V. Phoha, "Fake News Early Detection: A Theory-driven Model," *Digit. Threat. Res. Pract.*, vol. 1, no. 2, pp. 1–25, 2020, doi: https://doi.org/10.1145/3377478.

[7] M. N. A. and A. D. F. Gulzar Hussain, M. Wasim, S. Hameed, A. Rehman, "Fake News Detection Landscape: Datasets, Data Modalities, AI Approaches, Their Challenges, and Future Perspectives," *IEEE Access*, vol. 13, pp. 54757–54778, 2025, doi: 10.1109/ACCESS.2025.3553909.

[8] P. H. A. Faustini and T. F. Covões, "Fake news detection in multiple platforms and languages," *Expert Syst. Appl.*, vol. 158, p. 113503, Nov. 2020, doi: 10.1016/J.ESWA.2020.113503.

[9] I. Palacio Marín and D. Arroyo, "Fake News Detection: Do Complex Problems Need Complex Solutions?," *SSRN Electron. J.*, Sep. 2020, doi: 10.2139/SSRN.3721249.

[10] B. T. and M. A. M. S. Kausar, "ProSOUL: A Framework to Identify Propaganda From Online Urdu Content," *IEEE Access*, vol. 8, pp. 186039–186054, 2020, doi: 10.1109/ACCESS.2020.3028131.

[11] M. Amjad, G. Sidorov, A. Zhila, H. Gómez-Adorno, I. Voronkov, and A. Gelbukh, "'Bend the truth': Benchmark dataset for fake news detection in Urdu language and its evaluation," *J. Intell. Fuzzy Syst.*, vol. 39, no. 2, pp. 2457–2469, Jun. 2020, doi: 10.3233/JIFS-179905;JOURNAL:JOURNAL:IFSA;REQUESTEDJOURNAL:JOURNAL:IFSA;PAGE:STRING:ARTICLE/CHAPTER.

[12] M. K. H. and M. U. S. K. Zaheer, M. R. Talib, "A Multi-Kernel Optimized Convolutional Neural Network With Urdu Word Embedding to Detect Fake News," *IEEE Access*, vol. 11, pp. 142371–142382, 2023, doi: 10.1109/ACCESS.2023.3341870.

[13] M. A. Farah Rauf, Roha Irfan, Lyba Mushtaq, "Fake News Detection in Urdu using Deep Learning," *VFAST Trans. Softw. Eng.*, vol. 10, no. 4, pp. 151–167, 2022, doi: https://doi.org/10.21015/vtse.v10i4.1290.

[14] Z. Nasim and S. Ghani, "Sentiment Analysis on Urdu Tweets Using Markov Chains," *SN Comput. Sci.*, vol. 1, no. 5, pp. 1–13, Sep. 2020, doi: 10.1007/S42979-020-00279-9/METRICS.

[15] M. Q. Alnabhan and P. Branco, "Fake News Detection Using Deep Learning: A Systematic Literature Review," *IEEE Access*, vol. 12, pp. 114435–114459, 2024, doi: 10.1109/ACCESS.2024.3435497.

[16] M. A. Mohammad Mahyoob , Jeehaan Algaraady, "Linguistic-Based Detection of Fake News in Social Media," *Int. J. English Linguist.*, vol. 11, no. 1, 2021, doi: 10.5539/ijel.v11n1p99.

[17] Y. Liu and Y. F. B. Wu, "FNED: A Deep Network for Fake News Early Detection

on Social Media," *ACM Trans. Inf. Syst.*, vol. 38, no. 3, Jun. 2020, doi: 10.1145/3386253;REQUESTEDJOURNAL:JOURNAL:TOIS;TAXONOMY:TAXONOMY:ACM-PUBTYPE;PAGEGROUP:STRING:PUBLICATION.

[18] H. L. Shuo Yang, Kai Shu, Suhang Wang, Renjie Gu, Fan Wu, "Unsupervised Fake News Detection on Social Media: A Generative Approach," *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, pp. 5644–5651, 2019, doi: https://doi.org/10.1609/aaai.v33i01.33015644.

[19] A. A. Aldwairi, Monther, "Detecting Fake News in Social Media Networks," *Procedia Comput. Sci.*, vol. 141, pp. 215–222, 2018, doi: https://doi.org/10.1016/j.procs.2018.10.171.

[20] L. O. Álvaro Figueira, "The current state of fake news: challenges and opportunities," *Procedia Comput. Sci.*, vol. 121, pp. 817–825, 2017, doi: https://doi.org/10.1016/j.procs.2017.11.106.

[21] A. Saxena, H. Saxena, and R. Gera, "k-TruthScore: Fake News Mitigation in the Presence of Strong User Bias," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12575 LNCS, pp. 113–126, 2020, doi: 10.1007/978-3-030-66046-8_10.

[22] B. Rath, A. Salecha, and J. Srivastava, "Detecting Fake News Spreaders in Social Networks using Inductive Representation Learning," *Proc. 2020 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2020*, pp. 182–189, Dec. 2020, doi: 10.1109/ASONAM49781.2020.9381466.

[23] Fatemeh Torabi Asr and Maite Taboada, "Big Data and quality data for fake news and misinformation detection," *Sage Journals*, 2019, [Online]. Available: https://journals.sagepub.com/doi/10.1177/2053951719843310

[24] R. K. Pritika Bahad, Preeti Saxena, "Fake News Detection using Bi-directional LSTM-Recurrent Neural Network," *Procedia Comput. Sci.*, vol. 165, pp. 74–82, 2019, doi: https://doi.org/10.1016/j.procs.2020.01.072.

[25] S. K. Ishida, Yoshiteru, "Fake News and its Credibility Evaluation by Dynamic Relational Networks: A Bottom up Approach," *Procedia Comput. Sci.*, vol. 126, pp. 2228–2237, 2018, doi: https://doi.org/10.1016/j.procs.2018.07.226.

[26] A. A. Tahani Alsaedi, Muhammad Rizwan Rashid Rana, Asif Nawaz, Ammar Raza, "Sentiment Mining in E-Commerce: The Transformer-based Deep Learning Model," *Int. J. Electr. Comput. Eng. Syst.*, vol. 15, no. 8, 2024, doi: https://doi.org/10.32985/ijeces.15.8.2.

[27] A. Trivedi and S. Sangeetha, "Enhancing neural network predictions with finetuned numeric embeddings for stock trend forecasting," *Soft Comput.*, vol. 29, no. 3, pp. 1829–1844, Feb. 2025, doi: 10.1007/S00500-025-10483-5/METRICS.