OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol: 03 - Issue 03 (2026)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

URDU FAKE NEWS DETECTION USING LSTM AND HYBRID CNN-LSTM MODELS

Sabiha Anum[1], Sqlain Majeed[2], Muhammad Nabeel Sarwar[3]

[1]Administrative Sciences (Business Management), Boston University (Metropolitan College).
[2]Faculty of Computer Science & Information Technology, The Superior University, Lahore, Pakistan.
Lecturer, Department of Software Engineering, Islamia University of Bahawalpur, Pakistan.

Corresponding Author: (Sabihaa@bu.edu)
DOI:(*https://doi.org/10.71146/kjmr921*)

Article Info

Abstract

Fake news is an increasing point of interest among the research community as it can be transmitted due to numerous media in one of the shortest periods of time. False information, particularly on those languages that lack substantial resources, has turned into a large issue that is increasingly getting worse due to social media and internet crimes. It is difficult to find fake news in Urdu as it is a complex language and there are not many label datasets, which is why it is not a well-researched discipline. The accuracy of the pre-existing machine learning studies in the field of the Urdu fake news detection has been insufficient. In their present work, the authors make use of the deep learning-based methodology of Long Short-Term Memory (LSTM) and a composite method (CNN-LSTM), i.e., both TF-IDF and word embeddings employing LSTMs to address this problem. The LSTM-based word embedding new method was doing significantly better than the best research, having a 95.85% accuracy and a 94.65% F1 score in the D2 dataset. The model was very accurate with 93.87% and F1 score of 94.21% on D1 data. These findings are an important step toward investigating fake news in Urdu and a promising source to reduce the negative impact of information warping in the online society.

Keywords:

Fake news Detection, Urdu language, deep learning model, long-short-term memory. UFNDL detection.

1. INTRODUCTION

The fake news is characterized as misinformation that is both untrue and provable and disseminated intentionally in social media sites [1], and sometimes involves hoax, fabricated and financial frauds [2]. The quick development of social media has contributed to a number of cybercrimes, among which there is fake news. The process of identification of synthetic and original news has become a significant burden within the past few years [3]. The fake news detection has become a topic of interest both to the research community and by the general population in the recent past. The tables of fake news detection in under- resources such as Urdu are on the increase. The reason is that there has been already significant advancement in fake news detection in English using models that have already been established. One should apply extraordinary efforts in developing improved natural language processing [4] techniques of Urdu language. They may effectively be designed to manage realistically manipulated contents [5] to facilitate the ease of individuals online and other users [6]-[8]. The problem of fake news detection in Urdu is particularly important because of the linguistic complexity and the lack of the labeled sample.

Fake news detection by way of the application of natural language processing in the study has become a major concern as described in recent studies. Particularly, the initial experiment of fake news detection in Urdu developed a fake news detecting system by generating a preliminary dataset [9]. The primary weakness is that the data on news is very small in the dataset, and it might not suffice to use machine-learning methods. One more recent study explains the use of deep learning algorithms, including Text CNN and insertions of three words as well as TF-IDF attributes. The accuracy of this study was 0.716 [1]. Another work that is highlighted in the literature has suggested a generalized auto regressor based model to obtain an accuracy of 0.8400 [10] to perform classification is mentioned. Nevertheless, the problem of the accuracy improvement remains unsolved. This is why; effective development must be made to improve the quality of the Urdu language.

A simple solution is designed in this study particularly when resources and data are few in order to provide a better solution on the pursuit of fake news in the Urdu field. Using a deep learning model Long-short term memory word embedding based on LSTM algorithm is to be used to classify two classes. Various tests are conducted with TF-IDF and word embedding using LSTM classifier and this method revealed superior performance using word embedding by LSTM classifier. The proposed solution in detecting Urdu fake news demonstrates optimal behaviour on both datasets which is superior in comparison with the state of the art models. As long as any misinformation passes through social media, the UFNDL algorithm will be used systematically to evaluate its authenticity, and hence the society will be prepared to distinguish between authentic and fake news, therefore, resulting in creating a safer and more informed society.

2. Literature Review

The researchers addressed the problem of fake news by applying different feature-based methods [11]- [14] which are mainly, used to classify a text, and carried out by machine learning (ML) models Naive Bayes, Support Vector machine (SVM), Logistic Regression, (LR), Decision Tree, Deep Learning (DL) models, Convolution Neural Network (CNN), Recurrence Neural Network (RNN) and Deep Neural Network (DNN) models [15]. The hierarchy of fake news detection is shown in Figure 1 and was divided into machine learning and deep learning model and subdivided into supervised and unsupervised learning models.

2.1 Machine Learning

There are several machine-learning methods that have been developed and applied to the Urdu fake news detection. Here is the most obvious contribution made by the leading research works in the finding of Urdu

language fake news [1], [9], [16]. In [1], the benchmark data of three samples and four hundred anchors, which we refer to as Bend the truth, has 900 total samples including 500 real and 400 Urdu false news. In this paper, a supervised machine learning classifier AdaBoost with different weight schemes TF-IDF, norm and binary was used. The top score was attained in Ada boost and weight scheme TF-IDF with 87% F1 Fake and 90% F1 Real. Nevertheless, the dataset was also small in regards to domains and samples, and thus, the results were not significant. At the same time, researchers in the study [16] carried out another research to augment the dataset to yield higher accuracy by integrating real corpus with the augmented dataset. These three augmented, machine-translated and augmented-downsized techniques are used in improving the datasets. A mere comparison of the methods has already been made minus new contributions.

More recently, [9] used a baseline technique with the F1-macro score of 0.679 on a test set of 1600 news articles. The researchers developed a fake representation of the actual articles that is not necessarily a proper procedure [1], the small size of the dataset is another weakness of this research. News can be generated in real-time in realms with a low number of samples where the classification model might not have the capability of identifying fake news in the Urdu dialect. The cutting-edge research has added a novel dataset and has executed analysis of standard dataset that is produced by [16]. In addition to it, [17] investigated three ensemble strategies, which include voting, grading, and stacking, and five machine learning strategies to detect fake news. Compared to all the remaining models, voting ensemble method spent a small amount of time to alleviate the model. In the same vein, SVM was the relative champion in all the monitored classifiers with an 87.3 percentage BA score. Nonetheless, this experiment cannot generate an accurate technique of detecting fake news in Urdu [18]. Besides, this study used an English-to-Urdu dataset that was translated using Google Translate and did not pass through human translation. The information provided by five domains was used; Thereby, the breadth is small.

A recent research with significant results [18] introduced a data set of 4097 articles that comprised nine different domains. SVM, k-NN, and group models such as Stacked Feature, Random Forest and Extra Tree with TF-IDF and a Bag of Words were suggested as machine learning models. Comparatively, the most accurate (93.3 per cent) on an independent TFIDF configuration was on stacked approach. This study does not make use of deep learning techniques. Therefore, it is listed as one of the significant constraints. As the data volume is growing and conventional machine learning models fail to do well in demand, it is better to use deep learning models. Table 1 represents the analytical representation of the advanced machine learning methods.

2.2 Deep Learning

Comparing the results of various machine learning and deep learning algorithms research has been done by [19] that suggested the Char, CNN-RoBERT to detect Urdu fake news. The study took the dataset that was developed by [16] of 900 experimental samples. In this work, the combination of pre-trained models, RoBERTa, char CNN, as well as label smoothing, and 90 percent accuracy was proposed. Another research also does the same and applies n-grams methods [20]. The ensemble learning method is a machine method of learning that post ultimately attains 78% F1 score. Therefore, the deep learning method was more effective than machine learning, yet it is necessary to enhance the accuracy of the model. Therefore, a study [2] investigated three DL methods to classify and detect Urdu news such as CNN, RNN and Text CNN and in combination with classical feature extraction method, Term Frequency Inverse Document Frequency (TF-IDF), word embedding methods Word2Vec, fast Text, GlovE among all the models Text CNN with TF-IDF features yielded the best values of accuracy 0.716 and macro average F1 values 0.663. The given model was however not as accurate. Three different models were proposed in this study and they include DNN based

model and second one is the Ensemble Model which is founded on majority voting, and finally third founded on probability averaging. Upon comparison of the results, the Dense Neural Network-based model resulted in better performance with a macro F1-score of 59% and an accuracy of 72%. The model was trained using a smaller dataset. That is why this method cannot be regarded as dependable. Table 1 explains the constraints of being studied to detect fake Urdu news.

Table 1 Comparative literature on machine learning and deep learning

Study	**Model**	**Features**	**Limitation**
(Alhindi, Petridis, & Muresan, 2018)	LR SVM (linear) BiLSTM	"Unigram features Word embeddings (Glove)"	61.00 59.00 60.00
Khan, Khondaker, Afroz, Uddin, & Iqbal, 2021)	AdaBoost Decision Tree	"Lexical (word count, average word length, number count, parts of speech count, exclamation point count) & Sentiment (positive/negative polarity)"	56.00 51.00
	Naïve Bayes	"n-grams (TF-IDF of word-based unigram and bi-gram)"	60.00
	K-NN	"Lexical categories Empath (Empath tool - e.g.: violence, crime, war)"	54.00
	CNN LSTM C-LSTM HAN	"Word embeddings, GloVe"	54.00 58.00 54.00 57.00
	Feed-Forward	"Word embeddings, RoBERTa, Fast Text"	62.00
(Aslam, Khan, Alotaibi, Aldaej, & Aldubaikil, 2021)	BiLSTM – GRU?		89.90

Challenges

Because of the syntactic nature of the language Urdu and the absence of labeled datasets, training effective models to be used in fake news detection is challenging. Due to the need to ensure consistency and quality of the dataset, the need to do a large amount of data augmentation and preprocessing in the form of dealing with special characters as well as noise removal increased the complexity and time required to do the research as traditional models could not produce the desired outcomes in the task.

3. Material and Methods:

In this section, the information and procedures include data collection, preprocessing, feature extraction, model selection and evaluation are expounded upon. As seen in Figure 2, we start with the dataset by acquiring the appropriate text data of various recent research papers [1], [9], [18] to detect the fake news in constrained resource of Urdu language accurately. One of the major steps is preprocessing, which entails four major tasks which are: Data cleaning encompasses removal of stop words, removal of punctuation and scrubbing alphanumeric characters which will guarantee high quality analysis of text based data so that there are no inconsistencies, noise, and irrelevant data in the dataset. The second step is featuring construction which has two approaches used to acquire the feature vectors. Moreover, the process of tokenization is implemented to divide the text into separate words or tokens and LSTM or Hybrid CNN-LSTM deep learning architecture is chosen. In addition, models are also trained and their performances are assessed using different performance metrics.

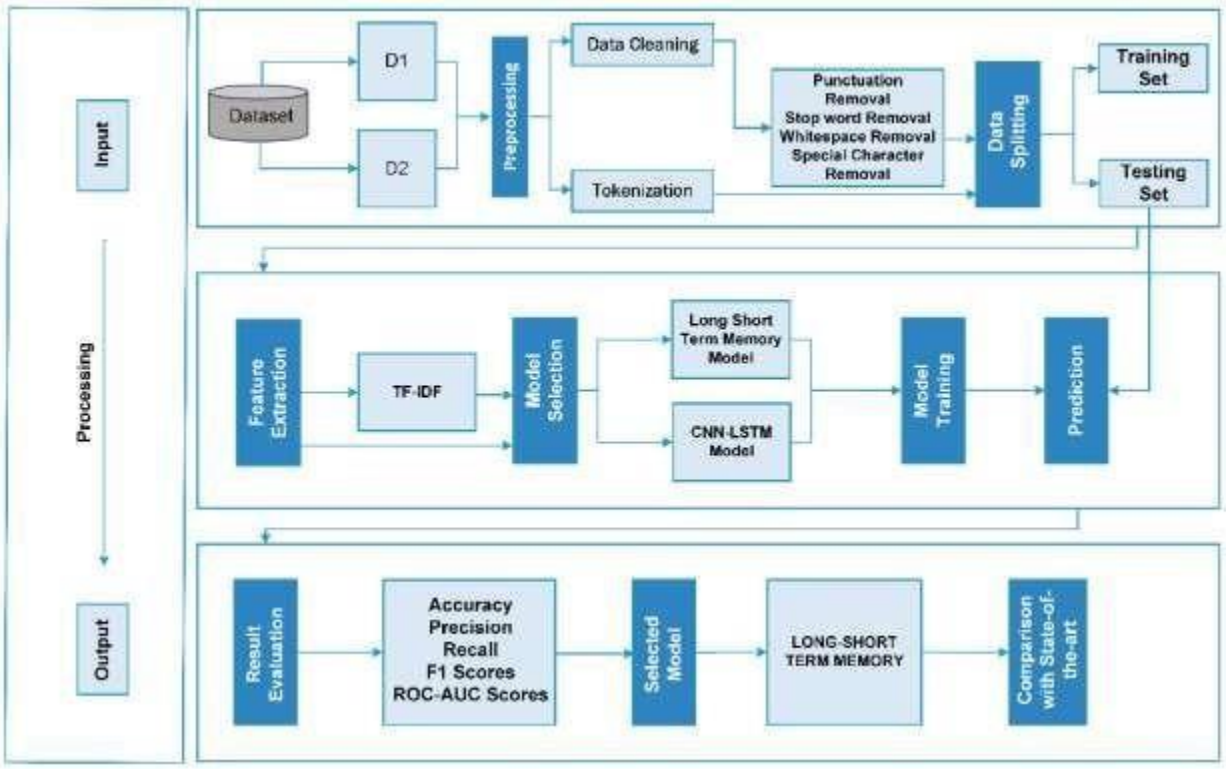


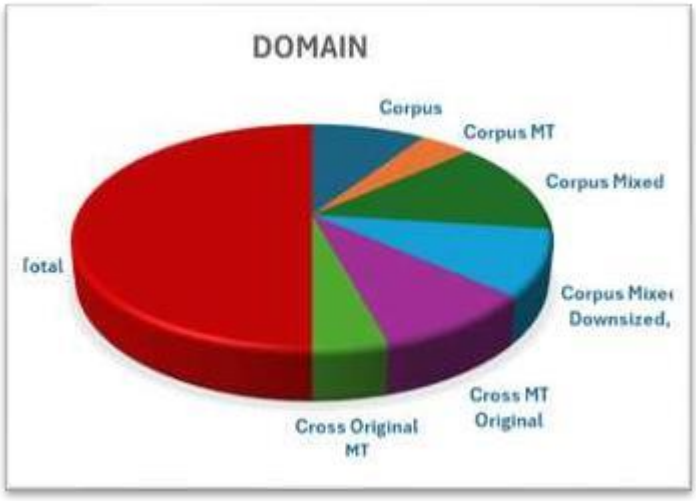
Figure 1 Workflow of Adopted Methodology

3.1 DATASET

This paragraph briefly explains two of the dataset to evaluate the performance of the model named D1 and D2 [10]. Table 3 demonstrates the D1 data that consisted of 4809 news articles of which 1413 are true and 3396 are fake which are sampled in 6 different techniques: i) Corpus, ii) Corpus MT, iii) Corpus Mixed, iv) Corpus Mixed Downsized, v) Cross MT original and vi) Cross original MT [1], [9]. In order to research the effectiveness of model on augmented dataset. As far as I know, this D1 dataset is the most complete augmented corpus of Urdu news in terms of detection of fake news on Urdu language because professional journalists have been employed to produce fake news [1]. The news items in truthful subset were checked and

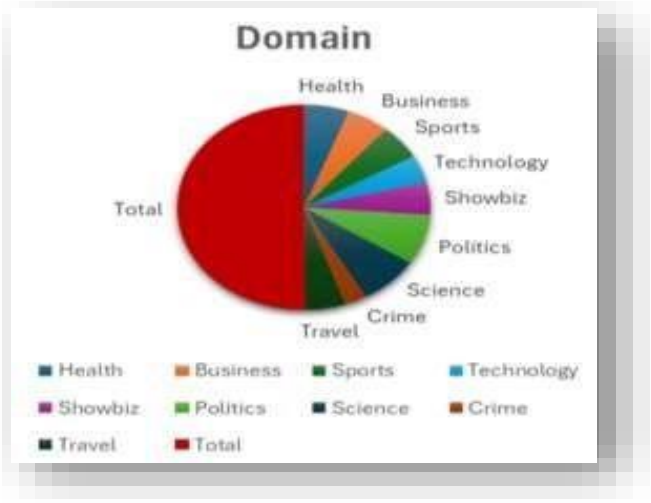
selected by hand and were based on valid news sources in information like BBC news, CNN Urdu, Daily Pakistan, and Etmad news.

Table 2: In D1 dataset class wise distribution of fake and real news



Furthermore, in this research study another manually-verified, slightly biased fake news dataset under a gold standard was released in 2023 [18] named D2. Table 4 presents the D2 data which will comprise nine distinct areas of news items i) Health, ii) Business, iii) Sports, iv) Technology, v) Showbiz, vi) Politics, vii) Science, viii) Crime, and ix) Travel [18]. D2 dataset was also professionally compiled based on the various sources on media platforms like sites and crowdsourcing and third-party build news. Moreover, the fake news data is obtained as well through the news of Vishvas [1]. The most tolerable method of gathering fake news and regarded as the most diverse and comprehensive dataset with 92% of annotation effectiveness [18]. Thus, the D1 and D2 represent the "Gold standard and "Benchmark Datasets" of Urdu fake news detection.

Table 3: In D2 dataset class wise distribution of fake and real news



3.2 Pre-Processing:

We clean up data by using methods used to ensure that it is good and easy to use. This improves the data set by eliminating the stuff, replicas and punctuations that are unnecessary. We eliminated the objects such as those which are irrelevant or the ones which distract the text. And we did it by locating and deleting characters HTML tags and additional spaces which are unnecessary in order to have the text clean and intelligible. All this is done to the data by the use of data cleaning methods.

1. Data Cleaning: There were four tasks in this step. These activities were to eliminate characters eliminate punctuations eliminating additional spaces and eliminate common words, such as the, and, which add negligible value to the Data Cleaning process. The primary objective of Data Cleaning was to make the data more practical. Another significant step is the Data Cleaning.

i. I will address the elimination of punctuation and other special characters. The process makes the text more straightforward. During the Special Character and Punctuation Removal operation we remove non-textual elements such as numbers and symbols. To remove the special characters and punctuations in the Urdu text, we compose rules called regular expressions. It is this way we will be able to guarantee that the Special Character and Removal of Punctuation process will only leave behind the most important words and meanings.

ii. Eradicating Whitespaces and Stop Words: Noise was eradicated in the text. This sound consisted of insignificant characters and HTML tags. This was done by removing these characters and HTML tags. We also removed unnecessary spaces. This was carried out in order to have a clear and relevant text. This we could do by removing stop words and white spaces.

2. Text data The text data was tokenized with the help of the tokenizer class of Kera's API of TensorFlow and it was converted to integer sequences. The resulting sequences are then padded with the pad sequence function to have a predetermined length. This step involved subdivision of the text into separate tokens, consideration was given to punctuations and consideration given to suitable segmentation of text based on context to represent the text accurately.

3.4 Data Splitting

Our experiments involved an attempt to use varying proportions of the training and testing set; 60:40, 70:30 and 80:20. Of these, the ratio of data splitting with 80 percent as the training data and 20 percent as the testing data gave the best results, as it is in low- resource language and higher share of the data to be used during training results can be advantageous. It enables the model to learn better with the available data, which may model more representative trends in the language.

3.5 Feature Extraction

In order to identify the relevance of words in the corpus, we used two extraction features, i.e., TF-IDF(Term Frequency Inverse Document Frequency) and LSTM word embedding, as evidenced by equation 1. In order to obtain the values of features using the statistical approach TF-IDF since it is able to weight the frequency of words as well as the inverse frequency of words in a corpus. In this way, one can emphasize the most significant words that should be used to classify documents in order to learn spatial characteristics. TF-IDF vectorization is among the feature engineering methods that are employed to convert textual data to numerical vectors depending on the significance and prevalence of words [2].

Where:

i is weight of term

j is the document and corpus shows N documents

(tf_{ij}) is the number of terms I in j documents

$$\text{Weightij} = \text{tf}_{ij} \times \log\left(\frac{N}{d_{fi}}\right) \quad (1)$$

Compared to TF-IDF, word embedding through LSTM is utilized, as it is known to be effective in gaining a semantic association and contextual dependence in language. The successfully LSTM-based word embedding is sensitive to both time and context of words such as the order of words and their different meanings depending on the context. Proper feature sets, which can be utilized in deep learning model, are produced by use of vocabulary reduction approaches like eliminating stop-words. After drawing up the values of the features, a deep learning classifier must identify the fake and real news articles.

3.6 Proposed Model

The problem of fake news classification in the textual field is rather challenging to cope with and the classification of fake news is proposed in several different approaches. Multilayer neural network architectures, or deep learning architectures, are necessary in sequential text analysis applications. The Long Short-Term Memory model, hybrid (CNN-LSTM) are compared with each other to capture textual dependencies with deep learning layers. Hyperparameter tuning of every model is a major part of a training.

3.7 Model Training

To facilitate training, validation and test sets, data splitting has been performed and easier ways of checking the performance of the model. In addition, the hyper parameters tuning technology is essential concerning the enhanced performance of the models.

Hyperparameter:

The learning rates, batch quantities and regularization conditions on the validation sets as well as epochs are cautiously adjusted every time to ensure that the pattern of model convergence in a dependable and successful manner. Activation functions introduce non-linearity to the model by identifying complicated data patterns. There is also the use of early stopping in order to prevent overfitting. Dropout strategies and regularization also reduce overfitting to come up with a reliable and efficient model that can be used to find fake news in Urdu. The UFNDL algorithm categorizes text into two categories: "Fake" or "Real." The initial one is preprocessing that purifies the text removing stop words, spaces, punctuations, and special characters. The data will be tokenized with maximum word count and length after which it will be inputted to a neural network with dense layers in which word embedding based on LSTM will occur. It also divides the frequency-inverse document frequency of each term and identifies the five most key keywords. The likelihood was estimated by the model using the sigmoid activation function. Lastly, we employ accuracy, precision, and recall, F1 score, and ROC-AUC measures of the performance of the model. Ultimately, there will be the division of the real news stories and fake news stories.

3.4 Result Evaluation

Different evaluation metrics are used to evaluate the performance of the model. This paper will apply such parameters as accuracy, precision, recall, F1 score and Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) Score. These measures are calculable as indicated in Equations 1-4.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Precision} = \frac{|TP|}{|TP| + |FP|} \quad (3)$$

$$\text{Recall} = \frac{|TP|}{|TP| + |FN|} \quad (4)$$

Accuracy is a basic measure to the performance of a model. The formula of accuracy is as below. Precision The reliability of the predictions which are true of all the positive predictions. The formula of the precision is as follows. Recall The rate of correctly anticipated positive predictions among all the real positive cases.

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

Equation 5 considers the F1 scores. F1 score represents the tradeoff between the accuracy of a model in terms of recalling all the objects it is supposed to remember. In case a model receives F1 score, which means that it is indeed very good at being precise and remembering everything. When it scores a low F1 that indicates that it is not performing a commendable task. ROC AUC score or Receiver Operating Characteristic Area Under Curve score in short is used to assess the ability of a model to distinguish between bad things. The F1 score and the ROC AUC score are essential each time we desire to be aware of how a model is performing. The computation is based on the area beneath the ROC (represents operating characteristics) curve which is graphically used to illustrate the dependence of the TPR (true positive rate) with the FPR (false positive rate). A classifier that has a large Receiver Operating Characteristic Area Under the Curve (ROC AUC) value is a source of a high rate of success of both positive and negative instances successfully found.

4 RESULTS AND DISCUSSION

In this part, there are outcomes that are measured by the means of two datasets. There are different experiments that are followed with the problem of fake news classification. The dataset was cleaned using feature extraction and data cleaning methods so as to be ready to bring better results as seen in the previous section of this paper.

Table 4 The number of words and feature vectors in the corpus

Data	**No of words**
Clean corpus	316,355
Feature Vector in D1	3219
Feature Vector in D2	917

Table 4 shows the amount of the top five feature vectors presented by TF-IDF of the two datasets. It implied that the TFIDF used in feature extractions has been proposed to be more effective in comparison with traditional n-gram of characters and words [2]. However, embeddings of words by the LSTM model form dynamic, context-dependent representation of every word that includes sequential dependencies. In order to check the model performance and comparative results, two sets of experiments are carried out and the results are presented in the table below.

Table 5 Model Performance on the D2 Dataset

Approach	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
LSTM Model on D2	95.85	96.17	93.19	94.65	99
Hybrid Model on D2	93.17	94.06	88.24	91.05	97
LSTM with TF-IDF on D2	80.73	71.1	86.07	77.87	88
Hybrid with TF-IDF on D2	92.93	96.82	84.83	90.43	98

Table 5 depicts that the long-term short term memory is a good performer. It achieves the accuracy of 95.85 on dataset D2 in fake news detection in Urdu. The Hybrid CNN-LSTM model also performs great with an accuracy score of 93.17 on the data. The long short-term memory is even better as far as the precision, recall, F1-Score, and ROC-AUC are concerned. This implies that the long short-term memory is great at the detection of the hard to observe patterns which is essential in correctly categorizing the objects. The long short term memory is good at this point. Conversely, the accuracy of the LSTM with TF-IDF was lower than the LSTM, which was 80.73% indicating potential challenges in using LSTM with TF-IDF to classify Urdu fake news.

Approach	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
LSTM Model on D1	94.07	93.19	95.91	94.53	98
Hybrid Model on D1	92.88	90.93	95.53	93.17	97
LSTM with TF-IDF on D1	92.52	90.93	95.53	93.17	97
Hybrid with TF-IDF on D1	88.88	87.48	89.58	88.49	96

Table 6 Model Performance on the D1 Dataset

The table 6 reveals the performance of LSTM model on the D1 dataset. The accuracy was 94.07%. Also, the LSTM model was also performing quite well with the accuracy of 92.88. This shows the LSTM model levels of understanding the complexity of language and content. The LSTM model possesses a sense of truth and falsity besides being capable of working with numbers. Due to this reason, LSTM model can be employed to identify whether the Urdu news is fake or not. Here the LSTM model merits. Since the LSTM word embedding method possesses a superior tendency to incorporate semantic relations and contextual peculiarities in the language, the extraction with the help of the Term Frequency-Inverse Document Frequency (TF-IDF) did not differ in order of success.

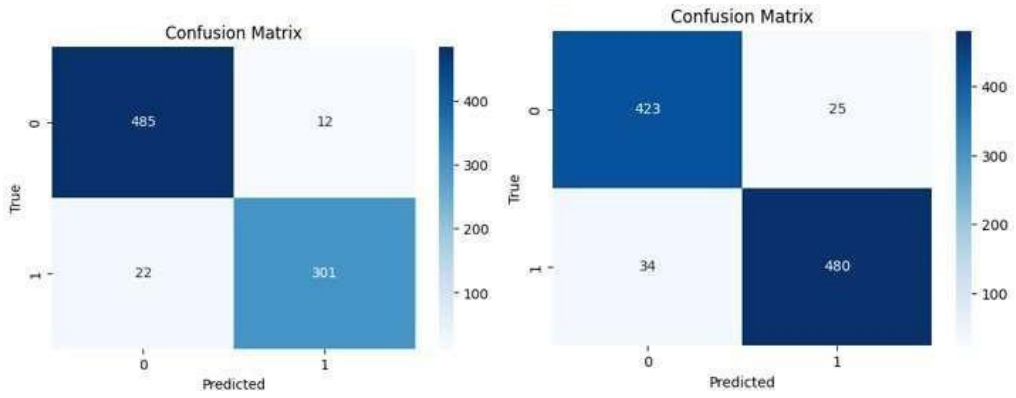


Figure 3 Confusion matrix of D1 and D2

Figure 3 presents the confusion table to be used in the best model and independent set testing. It was noted that the LSTM model which did not have the TF-IDF feature vectors was found to be more effective than

the hybrid approach. Figure 4 shows the best results obtained on independent set testing on D1. The best results have been demonstrated by the confusion matrix of the LSTM approach.

5 Comparative Analysis

In comparison to the state-of-art methods of the Urdu fake news classification, the analysis in question is also successful when it comes to the assessing the outcomes of the proposed framework. Training Model Training was carried out on the existing methods in order to compare the model performance against existing datasets. In figure 5, the effectiveness of the offered model is demonstrated.

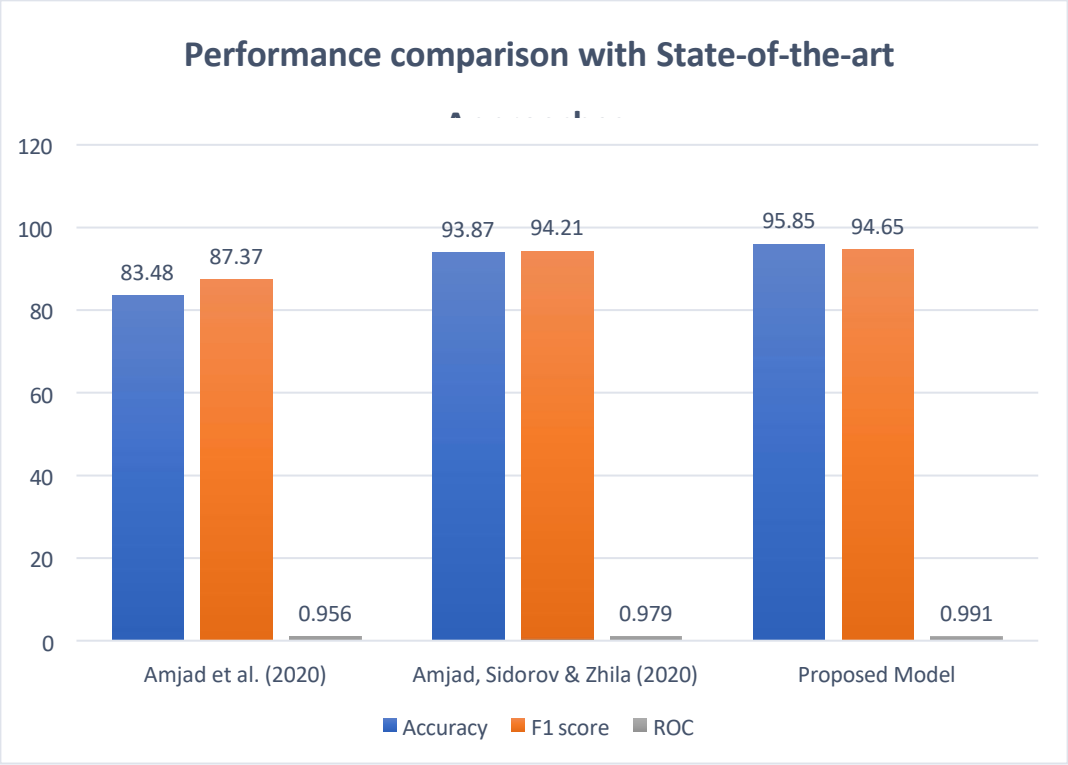


Figure 4 Comparative Analysis

The state-of-the-art studies [3], [9] are also applied to the present study using the two datasets to consider the fair comparative analysis. The verification and comparison of the results are decided. Figure 5 demonstrates findings that imply that the suggested methodology demonstrates desirable outcomes against the recent discourse. The robust performance measures used to test the model include F1 score. This research offers a superior ROC-ACU that is 99% relative to research study.

6. Discussion

This research paper is concerning identifying a method to identify the news in the Urdu language. These individuals who had done this This research is, regarding the identification of a solution to detect the news that were in Urdu language. The individuals, who conducted this study, had the desire to determine how computer models could be used to address the issues with the Urdu language. Urdu is not an easy language. It does not contain sufficient information. The major aim of the research was to enhance precision and ease with which news in Urdu can be spotted. The researchers did this using Long Short-Term Memory. They

also used the Long Short Memory with the Convolutional Neural Networks. There were also approaches used by them, e.g. word embedding and TF-IDF. They are made use of by using recycling Long Short-Term Memory, or LSTM in short. The aim of the study was to establish whether these computer models were actually able to assist in the process of identifying news pieces in Urdu.

The findings of the researchers are very good. The results prove that the models offered by the researchers can be effective in finding news in Urdu by Long Short-Term Memory. Language models are very useful to people who are dealing with language. They can also be helpful to people who desire to know what news is actual and what is not. Language models can be useful to both the language researchers and those who would want to avoid the spread of news. Language models can immensely benefit people that are concerned with this phenomenon, as well as those that are interested in differentiating news and non-news.

The complexity of the language, and the limited number of labeled datasets have led to the problem statement as it is very challenging to detect fake news in the Urdu language. The fact that previous approaches based on conventional machine learning models had accuracy problems highlighted the necessity of more advanced approaches. To mitigate these concerns, the objectives of the study were clear, and the study sought to employ advanced deep learning models in improving the detection efficiency and accuracy.

7. Conclusion

The Urdu language is one that is not rich in resources and, one that does not have specific repositories on where to dichotomize fake and real news. The literature demonstrates that research here has been limited due to the small scope of datasets applied and lacks in the multi-domain news. The greatest contribution is to provide a better accuracy in Urdu that will detect fake news. In this paper, a deep learning long-short-term memory classifier is introduced. TF-IDF and word embedding are conducted with the LSTM classifier and proposed model performs better when using word embedding through LSTM classifier. The offered detection system of Urdu fake news demonstrates the highest accuracy on the D2 dataset (95.85%), and D1 dataset (93.87%), it is higher than with the other models used in literature.

This study is based on TF-IDF and LSTM word embedding feature Extraction. We shall in the future develop a real time Urdu fake news detector system that will detect live Urdu fake news. We also attempt to add further dataset, which assists in the improved training of deep learning models.

- developed a high-precision LSTM-based model, which is better in addressing Urdu fake news detection compared to existing methods.
- better datasets of fake and real news expertly verified, getting a ROC-AUC score of 0.99 and an F1 score of 94.65.
- Future efforts will be to expand the data set and to create a real time detection system.

References:

[1] 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', *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 2, pp. 2457–2469, 2020, Doi: 10.3233/JIFS-179905.

[2] M. Abdullah Ilyas, M. A. Ilyas, and K. Shahzad, 'Urdu fake news detection using TF-IDF features and Text CNN', 2021. [Online]. Available: http://ceur-ws.org

[3] M. S. Farooq, A. Naseem, F. Rustam, and I. Ashraf, 'Fake news detection in Urdu language using machine learning', *Peer J Comput Sci*, vol. 9, 2023, Doi: 10.7717/peerj-cs.1353.

[4] W. H. Bangyal *et al.*, 'Detection of Fake News Text Classification on COVID-19 Using Deep Learning Approaches', *Comput Math Methods Med*, vol. 2021, pp. 1–14, Nov. 2021, Doi: 10.1155/2021/5514220.

[5] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, 'Deepfakes and beyond: A Survey of face manipulation and fake detection', *Information Fusion*, vol. 64, pp. 131–148, Dec. 2020, Doi: 10.1016/j.inffus.2020.06.014.

[6] S. Butt, N. Ashraf, G. Sidorov, and A. Gelbukh, 'Sexism Identification using BERT and Data Augmentation-EXIST2021', 2021.

[7] N. Ashraf, R. Mustafa, G. Sidorov, and A. Gelbukh, 'Individual vs. Group Violent Threats Classification in Online Discussions', in *Companion Proceedings of the Web Conference 2020*, New York, NY, USA: ACM, Apr. 2020, pp. 629–633. Doi: 10.1145/3366424.3385778.

[8] S. Latifi, Ed., *17th International Conference on Information Technology–New Generations (ITNG 2020)*, vol. 1134. Cham: Springer International Publishing, 2020. Doi: 10.1007/978-3- 030-43020-7.

[9] M. Amjad, S. Butt, H. I. Amjad, A. Zhila, G. Sidorov, and A. Gelbukh, 'Overview of the Shared Task on Fake News Detection in Urdu at FIRE 2021', Jul. 2022, [Online]. Available: http://arxiv.org/abs/2207.05133

[10] A. Faiz Ur, R. Khilji, S. Rahman Laskar, P. Pakray, and S. Bandyopadhyay, 'Urdu Fake News Detection using Generalized Auto regressors', 2020. [Online]. Available: https://github.com/urduhack

[11] M. Zuckerman, B. M. DePaulo, and R. Rosenthal, 'Verbal and Nonverbal Communication of Deception', 1981, pp. 1–59. Doi: 10.1016/S0065-2601(08)60369-X.

[12] G. Loewenstein, 'The psychology of curiosity: A review and reinterpretation.', *Psychol Bull*, vol. 116, no. 1, pp. 75–98, Jul. 1994, Doi: 10.1037/0033-2909.116.1.75.

[13] B. E. Ashforth and F. Mael, 'Social Identity Theory and the Organization', *Academy of Management Review*, vol. 14, no. 1, pp. 20–39, Jan. 1989, Doi: 10.5465/amr.1989.4278999.

[14] M. Deutsch and H. B. Gerard, 'A study of normative and informational social influences upon individual judgment.', *The Journal of Abnormal and Social Psychology*, vol. 51, no. 3, pp. 629–636, Nov. 1955, Doi: 10.1037/h0046408.

[15] D. Ali, M. M. S. Missen, and M. Husnain, 'Multiclass Event Classification from Text', *Sci Program*,

vol. 2021, pp. 1–15, Jan. 2021, Doi: 10.1155/2021/6660651.

[16] M. Amjad, G. Sidorov, and A. Zhila, 'Data Augmentation using Machine Translation for Fake News Detection in the Urdu Language', 2020. [Online]. Available: https://translate.google.com

[17] M. P. Akhter, J. Zheng, F. Afzal, H. Lin, S. Riaz, and A. Mehmood, 'Supervised ensemble learning methods towards automatically filtering Urdu fake news within social media', *Peer J Comput Sci*, vol. 7, p. e425, Mar. 2021, Doi: 10.7717/peerj-cs.425.

[18] M. S. Farooq, A. Naseem, F. Rustam, and I. Ashraf, 'Fake news detection in Urdu language using machine learning', *Peer J Comput Sci*, vol. 9, p. e1353, May 2023, Doi: 10.7717/peer j- cs.1353.

[19] N. Lin, S. Fu, and S. Jiang, 'Fake News Detection in the Urdu Language using Char CNN-RoBERTa'. [Online]. Available: https://huggingface.co/urduhack/urdu-roberta

[20] F. Balouchzahi and H. L. Shashirekha, 'Learning Models for Urdu Fake News Detection', 2020. [Online]. Available: https://mangaloreuniversity.ac.in/dr-h-l-shashirekha.