Fake News Detection in Urdu

Muhammad Zeeshan Nazar[1], Shakir Ullah Shah[2], Muhammad Taimoor Khan[3*]

[1] National University of Computer and Emerging Sciences, Pakistan
[2] School of Architecture, Computing and Engineering University of East London, United Kingdom
[3] GESIS - Leibniz Institute for the Social Sciences, Germany
taimoor.khan@gesis.org

Abstract

Fake news presents misleading information as legitimate news to influence public opinion and deceive readers. Fake news detection techniques distinguish between fake news and real news, having credible information. These techniques analyze the linguistic patterns in the text, contextual inconsistencies in user responses, and propagation behavior on social networks. Unlike high-resource languages, Urdu has limited basic tools that restrict the application of state-of-the-art machine learning models for Urdu-based challenges. Therefore, the available approaches for fake news detection in Urdu do not perform well on benchmark datasets. Bag-of-words approaches consisting of frequency-based sparse vectors are often used to represent features as n-grams, which are inadequate for detecting linguistic indicators related to legitimacy in news. In this paper, we propose a methodology that uses Urduhack text preprocessing tools to prepare the data, Urdu embeddings to represent the news text as dense vectors, and finally, a long short-term memory (LSTM) based deep sequence model to classify fake and real news. The proposed methodology outperforms traditional machine learning approaches in identifying linguistic characteristics and utilizing them for decision-making, achieving considerable performance gains with an accuracy of 85% and 83% on the Bend the Truth (BET) and Urdu fake news (UFN) datasets.

Introduction

The rise of social media and the widespread use of the internet have brought about a notable change in news spread (Zhou and Zafarani 2020). However, this accessibility has also given rise to the problem of fake news. This refers to news stories that are intentionally false and are meant to mislead readers. Fake news can cause harm, from damaging people's trust in reliable information to causing confusion and even social unrest. Fake news may be used to promote or demote individuals, groups, organizations, governments, etc. Various sociology studies review the propagation and impact of fake news through responses on social media (Shu et al. 2017). Fake news can be any incorrect material intended to convince its users to accept it. It may target the promotion or demotion of individuals, groups, or organizations. For instance, *"Pope Francis staggers the world, accepts Donald Trump for president, releases verbalization"*. The news was exceptionally mainstream and accumulated more than 960K user commitments on Facebook. Figure 1 has a few examples of Urdu fake and real news from BET dataset. It can be observed that the claims in fake news are stronger as compared to true news, given the real-world situation. For example, people following Pakistan politics would deny the startling claim in the first fake news, given the relationship between the two politicians.

Fake news has profound effects on societies on financial, political, religious, moral, cultural, and social grounds. Social media platforms are increasingly becoming the main source of information. They have made access to information faster, simpler, and more engaging than ever before. As per the Pew Research Center, around 66% of American grown-ups in 2013 got news from online media (Gottfried and Shearer 2016). It also exposes them to the noise and fake news that is also spread through social media. Distinguishing fake news from genuine information and minimizing its propagation and impact is a critical concern for online platforms (Yang et al. 2019). According to a report, 50% of Facebook postal links are fake, and only 20% of referral links are from well-known websites (Pérez-Rosas et al. 2017). Fake news articles and their sources are unverified, often wrong, and lead viewers astray (Raza et al. 2020). Fake news lacks genuineness, and has falsehood about the substance, excluding paranoid ideas that are hard to decide as true or false. They are written to deceive the reader, therefore, indicating specific patterns, vocabulary, and writing styles that appeal to a wider audience (Neal et al. 2017).

Automated detection of fake news intends to reduce human effort and time to identify fake news (Riedel et al. 2017; Jain and Kasbe 2018). Several types of approaches are used for content-based fake news detection in literature (Amjad et al. 2020). The knowledge-based approaches that rely on external resources to verify the legitimacy of a news item (Zhou and Zafarani 2020). Context-specific vocabulary has been used to distinguish between fake and legitimate news (Gómez-Adorno et al. 2018), and patterns in writing style may also serve as indicators of fake news (Amjad, Sidorov, and Zhila 2020; Akhter et al. 2021). Additionally, network propagation-based approaches, depending on news propagation and response from social media users are also indicative of detecting fake news (Shu et al. 2017). However, the lack of basic NLP tools to preprocess the data

Label	News Text
Fake	عمران خان نے یہ اعلان کردیا کہ اگلی حکومت نواز شریف کی ہی ہوگی
True	پاکستان مسلم لیگ نواز (پی ایم ایل این) کی صدر مریم نواز نے بدھ کو حکومت پر آڑے ہاتھوں لیتے ہوئے سوال کیا کہ وزیراعظم عمران خان نے کیوں عہدہ سنبھالا جب وہ چیلنجز کے لیے تیار نہیں تھے
Fake	کرونا وائرس ایک سازش ہے۔
True	کوویڈ -19 استثنیٰ کم از کم آٹھ ماہ تک جاری رہتا ہے ، ویکسین کی لمبی عمر کی امید ہے
Fake	شاہ رخ خان کا ڈوپل گینگر کشمیری لڑکا ہے۔
True	سنجے دت کا کہنا ہے کہ چھوٹے مسائل طویل مدت میں حل ہو جائیں گے۔
Fake	پاکستان نے یوٹیوب کو ہمیشہ کے لئے بین کر دیا
True	ایک وقت تھا جب ایک ہفتے میں 100 ملین یوٹیوب ویوز حاصل کرنا ناقابل تصور تھا، لیکن 2021 میں، اس قسم کی ٹریفک حاصل کرنے والا چینل دنیا بھر میں سب سے زیادہ دیکھے جانے والے 50 سب سے زیادہ دیکھے جانے والے یوٹیوب چینلز کی ہماری درجہ بندی میں بھی کمی نہیں کرے گا۔

Figure 1: Fake and real news examples from Bend the Truth (BET) dataset, where subsequent examples are from the same domain.

with higher accuracy and to represent the data as dense vectors that preserve the semantic association among words has been a limitation. The use of n-grams and frequency-based bag-of-words approaches plateaus before achieving acceptable accuracy (Amjad, Sidorov, and Zhila 2020). These approaches exhibited limited ability to capture the contextual semantics and long-distance associations among words. The Urduhack library[1] offers NLP tools for preprocessing Urdu text, and gensim library[2] contains pre-trained Urdu word embeddings. These two resources provide the means to apply modern NLP models to Urdu text. It leads us to our research question i.e., can these resources, integrated into a deep learning architecture, improve the accuracy of fake news detection in Urdu? We hypothesize that since the word embeddings have proven much better than frequency-based approaches for NLP tasks, the Urdu embeddings will improve the accuracy of the fake news detection task. The embeddings are trained on Wikipedia articles and web news that have relevance to the task at hand.

Deep learning models such as convolutional neural networks, recurrent neural networks, and their variants, i.e., bidirectional recurrent neural networks and long short-term memory (LSTM) models, are frequently used for similar tasks in other languages (Deepak and Chitturi 2020). Benefiting from the literature on high-resource languages, we propose to address fake news detection in Urdu by using a deep sequence classification methodology that utilizes Urduhack for preprocessing tools and the gensim Urdu Word2Vec embeddings for dense vector representation. The proposed approach performs better in identifying patterns of fake and real news and better detection of semantic associations and long-distance word dependencies. We use the Benchmark Urdu fake news datasets i.e., Bend the truth (BET) dataset[3], having 900 instances, and the Urdu fake news (UFN) dataset[4], having 10,083 instances, to train and evaluate our approach. These datasets have fake news either written or labeled by expert journalists and cover multiple categories, including health, technology, politics, sports, etc. Our proposed approach achieves 85% accuracy on the BET and 83% accuracy on the UFN datasets. In the UFN dataset with domains explicitly mentioned, the model faced difficulty in correctly classifying news in the Entertainment and Technology domains.

Literature Review

Automated investigation of fake news is a subject of higher consideration. The evidence for fake news started in 1439 with the creation of the printing press (Amjad, Sidorov, and Zhila 2020). Fake news is a piece of information that is false. Various sociology studies have analyzed the impacts of fake news and how people have responded to it (Shu et al. 2017). Social media platforms provide all sorts of information to their users that has fake news disguised as genuine information. It has affected various platforms where a higher number of user activities are expected, such as social media platforms, emails, broadcasting platforms, and even standalone news sites. Fake news may contain serious fabrication, deception, and satire. serious fabrications are news

[1] https://github.com/urduhack/urduhack

[2] https://pypi.org/project/gensim/

[3] https://github.com/MaazAmjad/Datasets-for-Urdu-news

[4] https://github.com/pervezbcs/Urdu-Fake-News

items about bogus and non-existent occasions or data. Deceptions are giving incorrect information from customary news sites. Satire is diverting news that imitates authentic news while containing disjointedness and craziness (Amjad et al. 2020).

The trustworthiness of the news item is determined by using different approaches to offer ground truth. The expert-oriented approaches depend on specialists such as journalists and fake news researchers to survey, review, and label fake news items. Crowd-sourcing-oriented approaches benefit from the collective intelligence of adequately large groups to annotate data. The lack of expertise of the annotators is compensated through higher annotator agreement among multiple annotators annotating the same news item. Computation-oriented approaches depend on external resources, e.g., DBPedia, to determine fake news. Machine learning models require high-quality training data to build reliable models, more often treating fake news detection as a binary classification problem (Jain and Kasbe 2018). Since the vocabulary of fake and legitimate news on the same topic has a high overlapping vocabulary, the n-grams and frequency-based approaches that are invested in the skewed use of vocabulary in a specific category are found to struggle. Fake news detection requires deeper analysis of the semantic association and writing style patterns of the two types of news items. There are also news items that can certainly be labeled as true or false, and stay close to the fence. It also hints towards either increasing the number of classes to treat it as a multi-class classification problem or projecting a regression model on it that assigns a validity score to a news item, indicating how close it is to reality (Pérez-Rosas et al. 2017). Typically, the difference between predicted scores and label scores may be computed through Pearson or Spearman corrections. However, since the available datasets have discrete ground truth scores, the challenge here is how to convert single labels to numeric scores. If the dataset contains the full length of the article, the rhetorical approach can be used as one of the craft features (Oshikawa, Qian, and Wang 2018; Parikh and Atrey 2018).

Machine learning models such as naive bayes classifier, logistic regression, random forest, support vector machines, and decision trees are utilized (Amjad, Sidorov, and Zhila 2020). Deep learning models such as convolutional neural networks, recurrent neural networks, and generative adversarial networks are also used (Deepak and Chitturi 2020; Manna et al. 2020). Although left and right-handed handwriting languages are similar for detecting style patterns, with reported difficulties in detecting fake news (Potthast et al. 2017). Naive Bayes classifier, focusing on the presence or absence of vocabulary across different labels, is used for text classification tasks, e.g., spam detection (Granik and Mesyura 2017). Satirical cues may indicate fake news, where Support vector machines are used to classify fake news (Rubin et al. 2016). Support vector machines and AdaBoost have shown better results in classifying Urdu fake news. However, the machine learning approaches are often used with frequency-based bag-of-words representations, i.e., term frequency-inverse document frequency, log entropy of frequency, and n-grams that cannot capture the semantic associations among words. There are also fake news-specific evaluation techniques that closely focus on the performance of fake and real news items detection (Burzo et al. 2018). However, the fake news detection task has relatively smaller annotated datasets as compared to other NLP tasks.

Fake news detection approaches analyze the content of the news item to determine fake news. These approaches may benefit from external resources that provide direct information on the content or writing styles for fake and legitimate news (Amjad et al. 2020; Zhou and Zafarani 2020). The context-based approaches that focus on the use of specific vocabulary to detect fake news (Gómez-Adorno et al. 2018). Unigrams of varying orders with different weighting schemes to improve accuracy (Hossain et al. 2020). The style-based approaches that detect patterns in the writing of fake and real news items (Akhter et al. 2021). Increasing the order of n-gram beyond a limit decreases performance, as determined by the receiver operating characteristic curve. The function words were also found crucial with a suitable weighting scheme to improve results. AdaBoost has generally achieved higher performance for fake news detection across domains (Amjad, Sidorov, and Zhila 2020). Writing style of the content creator at lexical, i.e., the degree of characters, grammatical, semantic, primary, and subject levels, also improves classification accuracy (Lagutina et al. 2019). Boundaries at the character, syllable, or word level in text are regularly alluded to and are among the most crucial decisions. However, multiple levels of n-grams drastically increase the number of features, supplemented by recurrence of characters, lower and upper cases, numbers and the use of white spaces (Bühlmann 2011).

External fact-checking sites such as PolitiFact and Snopes can help to separate real news from fake news (Asr and Taboada 2019). However, exploring the World Wide Web for information on the veracity of a news item is too complicated due to multi-modal data, i.e., database records, videos, audios, and articles in different text formats and scattered across different organizations' websites (Shu et al. 2017). There are also efforts in collecting, annotating, and enriching good-quality data on the topic (Ahmad et al. 2020). However, low-resource languages are not keeping up with these developments and therefore, have only a limited supply of annotated datasets, which restricts the model to analyze content on multiple levels of n-gram features only (Pérez-Rosas et al. 2017). These styles are often attempts to entice customers, impact their behavior, push towards quick action, and manipulate their feelings (Amjad, Sidorov, and Zhila 2020). Attempting to catch the objectivity of fake news heading also indicates similar behavior, e.g., *"You will never accept what he did?"*. However, these approaches suffer from the unavailability of sufficient tools in the case of low-resource languages. Diverse literary features with ensemble learning are used with ensemble learning to separate fake news from real news (Ruchansky, Seo, and Liu 2017).

Some datasets do not contain hard instances and have rather clearer separation between real and fake news. Although a model may achieve higher accuracy on such data, but is not effective for building real-world applications. LIAR[5], dataset and Buzz-feed information collection[6] also lack enough hard-to-classify instances as compared to the TruthFinder[7] and CRH datasets (Yang et al. 2019). In datasets with more hard-to-classify instances, the accuracy of the traditional machine learning models does not improve beyond a point despite parameter optimization. Deep learning models such as convolutional neural networks, recurrent neural networks, and improved variants, Bi-recurrent neural networks, and LSTM better adapt to the detection of complex patterns for better detection of fake news. A deeper analysis of the news content and its headings offers additional features to investigate fake news, such as geographical locations, writer subtleties, news domain, etc. This information, when concatenated with the content embeddings, has been shown to improve the classification accuracy of fake news detection (Nguyen et al. 2019).

The propagation features offered by social media platforms, e.g., retweets, shares, likes, comments, etc., are also used to detect user behavior in responding to fake news. It involves the behavior patterns of news consumers in responding to fake or legitimate news. Intentionally or unintentionally, the consumers may also contribute to spreading fake news in their network. The users' behavior in interacting with news on social media is checkers who verify information and spreaders who do not verify and contribute to spreading fake news. The users' behavior patterns, personality traits, and social network indicate checkers and spreaders (Manna et al. 2020). With the recent advancements in machine learning and their application in NLP, the models are getting bigger, better, and therefore, more effective in capturing complex patterns in the text. The algorithmic enhancements are also complemented by annotated datasets with more diverse samples. The NLP tools have also improved in text preprocessing, information extraction, and enrichment offered to the model. As a low-resourced language, the fake news detection approaches in Urdu are limited to n-grams and frequency-based approaches with traditional classification models. With the recent availability of better NLP tools for Urdu (Urduhack library) and Urdu word embeddings (gensim Word2Vec), we investigate their utility for the fake news detection task.

Proposed methodology

The proposed methodology is designed to identify distinguishing linguistic patterns in Urdu fake and real news by leveraging a deep learning approach that combines NLP preprocessing tools, word embeddings, and sequence modeling via LSTM networks. The complete methodology comprises three main components:

Data Preprocessing

We utilize the Urduhack library for customized preprocessing tailored to the Urdu language. This step includes:

Segmentation and Tokenization: Handling the tokenization of Urdu sentences with attention to compound and inflected forms common in Urdu script.

Normalization and Stemming: Reducing words to their base or root forms while preserving semantic context.

Noise Removal: Eliminating stopwords, domain-specific noise words, numeric tokens, and irrelevant symbols such as currency marks, punctuation, whitespace anomalies, and diacritics (e.g., zabar, zer, pesh).

This stage ensures that the input text is clean, standardized, and linguistically relevant for downstream embedding and modeling.

Word Embedding using Gensim

The cleaned corpus is then transformed into vector representations using pre-trained Urdu word embeddings from Gensim. Each token is replaced with its corresponding 300-dimensional vector, derived from embedding models trained on Wikipedia and news-related corpora. These embeddings offer two main advantages. They capture contextual relationships between the words. They enhanced the performance due to the similarity between the pretraining corpus and the task domain i.e., news classification. The word embeddings serve as the input layer to the neural architecture, allowing the model to ingest rich semantic representations of the text.

LSTM-Based Deep Learning Model

The core of the classification model is a Long Short-Term Memory (LSTM) network, which is particularly well-suited for processing sequential data like text. LSTM networks address the limitations of traditional RNNs, particularly the vanishing gradient problem, by introducing memory cells and three types of gates i.e., input gate Controlling the flow of input into the memory cell, forget gate Regulating parts of the previous memory to discard, and output gate determining memory passed to the next layer (Deepak and Chitturi 2020). This approach allows LSTM to learn both short-term and long-term dependencies, making them ideal for capturing the temporal and contextual cues that differentiate fake news from real news. The model processes sequences of embedded tokens and outputs a classification label indicating the veracity of the input news item. We train and evaluate our model on publicly available Urdu fake news datasets. The model uses a single-node output layer with a sigmoid as a binary classifier.

Results

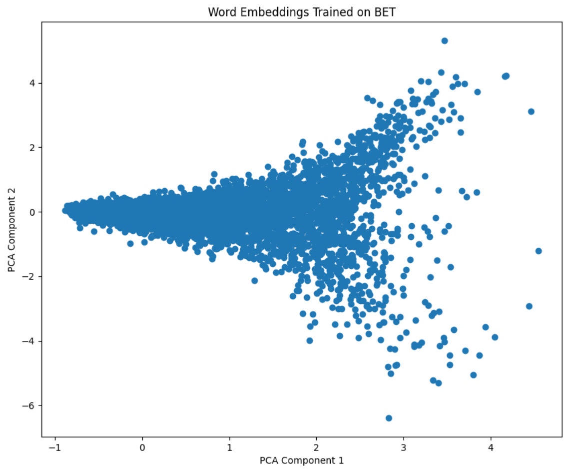
Datasets: We used two benchmark datasets for fake news detection in Urdu. Urdu Fake News (UFN)[8] has 10,083 instances with 5030 real and 5053 fake news. The news items are collected from leading newspapers and news channels in Pakistan and India from 2017 to 2023 and are annotated by expert journalists. It covers 15 domains, i.e., politics, health, sports, entertainment, technology, weather, agriculture, economy, showbiz, social media, education, women's

[5]https://huggingface.co/datasets/ucsbnlp/liar

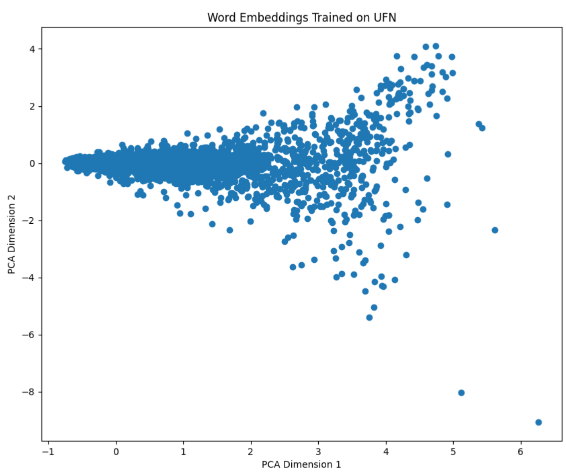
[6]https://webis.de/data/buzzfeed-webis-fake-news-16.html

[7]https://github.com/IshitaTakeshi/TruthFinder

[8]https://github.com/Sheetal83/Ax-to-Grind-Urdu-Dataset/tree/main



(a) Word embeddings (128-dimensional vectors) trained from scratch on the BET dataset



(b) Word embeddings (128-dimensional vectors) trained from scratch on UFN dataset

Figure 2: Embeddings trained on BET and UFN datasets from scratch.

rights, religion, foreign affairs, and international news. However, the dataset does not contain domain-specific labeling for individual news items. Bend the Truth (BET)[9] has 900 items with 500 real and 400 fake news. It covers only 5 domains, i.e., technology, business, sports, health, and entertainment. The fake news items are written by professional journalists for the dataset. Table 1 provides statistics on the preprocessed BET and UFN datasets. Interestingly, there is very little overlap in the vocabulary of fake and real news in both datasets. In comparison to BET, UFN not only has more documents but also much longer documents providing more context and relevant information about the news items. UFN also uses a much more diverse vocabulary as compared to BET. The overlap in the vocabulary of fake and real news makes the fake news detection task even more challenging.

[9]https://github.com/MaazAmjad/Datasets-for-Urdu-news

Table 1: Dataset features i.e., vocabulary, longest document, shortest document, and total documents in the BET and UFN datasets.

Bend the Truth (BET)		
Feature	Fake	Real
Vocabulary	184K	120K
Longest document	6516	1096
Shortest document	2	24
Total documents	500	400
Urdu Fake News (UFN)		
Vocabulary	1147K	954K
Longest document	6976	14,123
Shortest document	432	407
Total documents	5053	5030

Table 2: Training parameters of the proposed architecture

Stage	Hyperparameter	Instance
Initialization	Weights	Xavier (Default)
Activation	Function ReLU, Sigmoid	Alpha=0.0 (Default)
Training	Epochs	10
	Batch Size	32
	Loss Function	binary crossentropy
	Optimizer	Adam

Figure 2 shows 2D projection of the word embeddings trained on the BET and UFN datasets, with word dimensions reduced from 128 to 2 using principal component analysis (PCA). It broadly exhibit similar spatial structures, both characterized by a skewed distribution along the horizontal axis. In each case, the embeddings show a dense cluster concentrated on the left side, gradually spreading out toward the right. However, the BET embeddings are more compact suggesting stronger clustering or semantic similarity among frequently occurring terms. In contrast, the UFN dataset exhibits a slightly more uniform spread, with less density on the left and a broader dispersion across the horizontal plane, indicating higher variability in word usage or a wider semantic range. Despite these nuances, the overall shape and directional skew remain consistent, highlighting shared underlying linguistic structures.

We used early stopping, dropout, and regularization techniques to control over-fitting and improve generalization. The model uses the ReLU activation function for dense layers in the architecture, while the sigmoid for the single-node output layer performs binary classification. The proposed architecture is trained for 10 epochs, where the performance of the model stabilizes. We used 32 batch size, while also trying larger batches that did not help improve the classification accuracy. We used binary cross-entropy loss with the Adam optimizer while experimenting with the stochastic gradient descent optimizer as well. The model's parameter settings are provided in Table 2.

Next, we present the results of our proposed approach for both the BET and UFN datasets. To estimate the improvement in accuracy with the help of Urdu embeddings, we performed three experiments on both BET and UFN datasets and compared the results with both deep learning and traditional machine learning architectures. Among traditional machine learning approaches, we used the most frequently used ones i.e., support vector machines (SVM), logistic regression (LR), decision trees (DT), XGBoost, and random forest (RF). Among deep learning approaches, we compared our proposed LSTM-based architecture with convolutional neural networks (CNN))-based architecture. The deep learning architectures have experimented with different available embeddings represented as "_" i.e., embeddings trained on the datasets from scratch, "wbn" i.e., embeddings trained on web news, and "wkn" i.e., embeddings trained on Wikipedia and web news.

Table 3 shows the results of the proposed approach in comparison to traditional machine learning and deep learning approaches. It can be observed that BET being a relatively smaller dataset has the traditional machine learning approaches performing well too. For example, SVM and logistic regression are just next to the proposed approach, while some of the deep learning architectures e.g., $\text{CNN}_{\_}$, $\text{LSTM}_{\_}$ have the lowest accuracy. Since the instances are not enough, the embeddings trained from within this dataset also do not help to improve accuracy. However, the results are relatively better with pre-trained embeddings without allowing them to train on the given data. Thus, on BET when the embeddings are loaded while the architecture is trained on the available data, the model, the proposed approach performs well. However, the CNN architecture could not perform well despite using pre-trained embeddings. On the UFN dataset, having many more instances, there is a wider gap between the performance of traditional machine learning and deep learning-based approaches. SVM and logistic regression which were very competitive on the BET dataset have under-performed for the UFN dataset. While CNN, being a deep learning approach has improved accuracy on the UFN dataset as compared to its accuracy on the BET dataset. In general LSTM with "wkn" embeddings has the highest accuracy with competition from the traditional machine learning approaches in the case of BET, while a clear winner for the UFN dataset. We also experimented with different embedding vector lengths, starting from the default 300 dimensions. The proposed approach uses 128-dimension embedding vectors for improved computational efficiency with a minimal drop in accuracy. Figure 3 plots the accuracies of the deep learning approaches using different embeddings. It shows an upward trend from left to right on both datasets, where the LSTM versions are better than the CNN. Similarly, the no embeddings, Urduhack web news, and Urduhack Wikipedia news also show an upward trend from left to right. The only difference is in CNN_{wpn} outperformed by both LSTM_{wbn} and LSTM_{wpn}.

The proposed approach performed well in sports, health, politics, and other domains, however, it particularly struggled in distinguishing between fake and real news in the entertainment and technology domains. It indicates interesting directions about the nature of fake and real news across domains that need further investigation.

Table 3: Results of the proposed approach in comparison to traditional machine learning and deep learning approaches using different variants of Urdu embeddings on BET and UFN datasets.

		Accuracy	Precision	Recall	F1-score
BET	SVM	0.82	0.81	0.85	0.82
	LR	0.0.80	0.78	0.86	0.81
	DT	0.73	0.74	0.76	0.73
	XGBoost	0.77	0.78	0.81	0.77
	RF	0.72	0.73	0.81	0.71
	$\text{CNN}_{\_}$	0.56	0.58	0.86	0.69
	$\text{LSTM}_{\_}$	0.57	0.57	0.98	0.73
	CNN_{wbn}	0.69	0.71	0.68	0.69
	LSTM_{wbn}	0.79	0.75	0.71	0.73
	CNN_{wkn}	0.70	0.71	0.64	0.67
	LSTM_{wkn}	**0.85**	**0.79**	**0.68**	**0.76**
UFN	SVM	0.68	0.66	0.68	0.66
	LR	0.66	0.62	0.71	0.60
	DT	0.56	0.55	0.55	0.54
	XGBoost	0.68	0.65	0.70	0.64
	RF	0.67	0.63	0.73	0.61
	$\text{CNN}_{\_}$	0.66	0.71	0.73	0.72
	$\text{LSTM}_{\_}$	0.70	0.74	0.69	0.72
	CNN_{wbn}	0.77	0.74	0.77	0.75
	LSTM_{wbn}	0.82	0.78	0.83	0.80
	CNN_{wkn}	0.78	0.80	0.76	0.78
	LSTM_{wkn}	**0.83**	**0.81**	**0.79**	**0.80**

Conclusion

Fake news harms society as more and more people use online sources for information. Fake news detection approaches identify and intervene in fake news propagation. Urdu fake news detection approaches are restricted by a lack of efficient NLP tools and pre-trained models. The proposed approach uses Urduhack to preprocess Urdu text using multiple functionalities available. Urdu embeddings are used for presenting text as dense vectors. The proposed methodology uses the embeddings of preprocessed text with LSTM-based deep sequence model. The proposed approach shows improvement by benefiting from the latest technologies with higher accuracies on the standard Urdu fake news datasets. Although the traditional machine learning approaches were competitive on BET (the smaller dataset), however, as the size, diversity in vocabulary, and size of the documents increased (in the UFN dataset), the deep learning approaches showed dominance. It shows that by developing the basic NLP tools for low-resourced languages, they can be included to benefit from the recent improvements in machine learning.

References

Ahmad, I.; Yousaf, M.; Yousaf, S.; and Ahmad, M. O. 2020. Fake news detection using machine learning ensemble methods. *Complexity*, 2020(1): 8885861.

Akhter, M. P.; Zheng, J.; Afzal, F.; Lin, H.; Riaz, S.; and Mehmood, A. 2021. Supervised ensemble learning methods

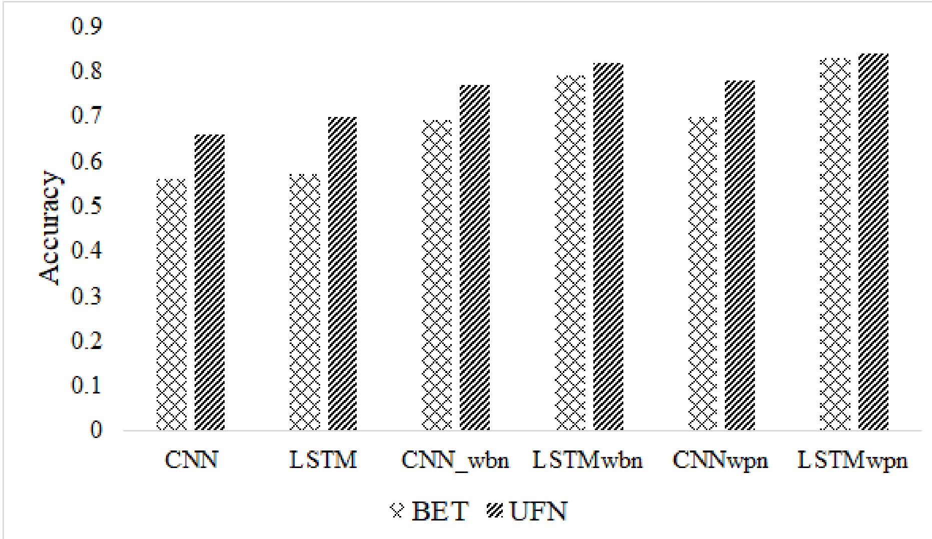


Figure 3: Results of CNN and LSTM architectures different embeddings variants.

towards automatically filtering Urdu fake news within social media. *PeerJ Computer Science*, 7: e425.

Amjad, M.; Sidorov, G.; and Zhila, A. 2020. Data augmentation using machine translation for fake news detection in the Urdu language. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, 2537–2542.

Amjad, M.; Sidorov, G.; Zhila, A.; Gómez-Adorno, H.; Voronkov, I.; and Gelbukh, A. 2020. "Bend the truth": Benchmark dataset for fake news detection in Urdu language and its evaluation. *Journal of Intelligent & Fuzzy Systems*, 39(2): 2457–2469.

Asr, F.; and Taboada, M. 2019. Misinfotext: a collection of news articles, with false and true labels. *Dataset. URL: http://fakenews. ngrok. io/# parseWebs [08/10/19]*.

Bühlmann, P. 2011. Bagging, boosting and ensemble methods. In *Handbook of computational statistics: Concepts and methods*, 985–1022. Springer.

Burzo, M.; Abouelenien, M.; Perez-Rosas, V.; and Mihalcea, R. 2018. Multimodal deception detection. In *The Handbook of Multimodal-Multisensor Interfaces: Signal Processing, Architectures, and Detection of Emotion and Cognition-Volume 2*, 419–453.

Deepak, S.; and Chitturi, B. 2020. Deep neural approach to Fake-News identification. *Procedia Computer Science*, 167: 2236–2243.

Gómez-Adorno, H.; Posadas-Duran, J.-P.; Ríos-Toledo, G.; Sidorov, G.; and Sierra, G. 2018. Stylometry-based approach for detecting writing style changes in literary texts. *Computación y Sistemas*, 22(1): 47–53.

Gottfried, J.; and Shearer, E. 2016. News use across social media platforms 2016.

Granik, M.; and Mesyura, V. 2017. Fake news detection using naive Bayes classifier. In *2017 IEEE first Ukraine conference on electrical and computer engineering (UKRCON)*, 900–903. IEEE.

Hossain, M. Z.; Rahman, M. A.; Islam, M. S.; and Kar, S. 2020. Banfakenews: A dataset for detecting fake news in bangla. *arXiv preprint arXiv:2004.08789*.

Jain, A.; and Kasbe, A. 2018. Fake news detection. In *2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 1–5. IEEE.

Lagutina, K.; Lagutina, N.; Boychuk, E.; Vorontsova, I.; Shliakhtina, E.; Belyaeva, O.; Paramonov, I.; and Demidov, P. 2019. A survey on stylometric text features. In *2019 25th Conference of Open Innovations Association (FRUCT)*, 184–195. IEEE.

Manna, R.; Pascucci, A.; Monti, J.; et al. 2020. Profiling fake news spreaders through stylometry and lexical features. UniOR NLP@ PAN2020. In *CEUR WORKSHOP PROCEEDINGS*, volume 2696. CEUR.

Neal, T.; Sundararajan, K.; Fatima, A.; Yan, Y.; Xiang, Y.; and Woodard, D. 2017. Surveying stylometry techniques and applications. *ACM Computing Surveys (CSuR)*, 50(6): 1–36.

Nguyen, D. M.; Do, T. H.; Calderbank, R.; and Deligiannis, N. 2019. Fake news detection using deep markov random fields. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 1391–1400.

Oshikawa, R.; Qian, J.; and Wang, W. Y. 2018. A survey on natural language processing for fake news detection. *arXiv preprint arXiv:1811.00770*.

Parikh, S. B.; and Atrey, P. K. 2018. Media-rich fake news detection: A survey. In *2018 IEEE conference on multimedia information processing and retrieval (MIPR)*, 436–441. IEEE.

Pérez-Rosas, V.; Kleinberg, B.; Lefevre, A.; and Mihalcea,

R. 2017. Automatic detection of fake news. *arXiv preprint arXiv:1708.07104*.

Potthast, M.; Kiesel, J.; Reinartz, K.; Bevendorff, J.; and Stein, B. 2017. A stylometric inquiry into hyperpartisan and fake news. *arXiv preprint arXiv:1702.05638*.

Raza, A.; Ahmed, S.; Bibi, M.; Rehan, M.; and Anwar, H. 2020. Automatic fake news detection: issues and solutions. *International Journal of Computing and Communication Networks*, 2(1): 21–28.

Riedel, B.; Augenstein, I.; Spithourakis, G. P.; and Riedel, S. 2017. A simple but tough-to-beat baseline for the fake news challenge stance detection task. *arXiv preprint arXiv:1707.03264*.

Rubin, V. L.; Conroy, N.; Chen, Y.; and Cornwell, S. 2016. Fake news or truth? using satirical cues to detect potentially misleading news. In *Proceedings of the second workshop on computational approaches to deception detection*, 7–17.

Ruchansky, N.; Seo, S.; and Liu, Y. 2017. Csi: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 797–806.

Shu, K.; Sliva, A.; Wang, S.; Tang, J.; and Liu, H. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19(1): 22–36.

Yang, S.; Shu, K.; Wang, S.; Gu, R.; Wu, F.; and Liu, H. 2019. Unsupervised fake news detection on social media: A generative approach. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, 5644–5651.

Zhou, X.; and Zafarani, R. 2020. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)*, 53(5): 1–40.