Normalized effect size (NES): a novel feature selection model for Urdu fake news classification

Muhammad Wasim[1], Sehrish Munawar Cheema[2] and Ivan Miguel Pires[3,4]

[1] Department of Computer Science, University of Management & Technology, Sialkot Campus, Sialkot, Pakistan
[2] Department of Computer Science, University of Management and Technology, Lahore, Pakistan
[3] Instituto de Telecomunicações, Covilhã, Portugal
[4] Escola Superior de Tecnologia e Gestão de Águeda, Universidade de Aveiro, Águeda, Portugal

ABSTRACT

Social media has become an essential source of news for everyday users. However, the rise of fake news on social media has made it more difficult for users to trust the information on these platforms. Most research studies focus on fake news detection in the English language, and only a limited number of studies deal with fake news in resource-poor languages such as Urdu. This article proposes a globally weighted term selection approach named normalized effect size (NES) to select highly discriminative features for Urdu fake news classification. The proposed model is based on the traditional inverse document frequency (TF-IDF) weighting measure. TF-IDF transforms the textual data into a weighted term-document matrix and is usually prone to the curse of dimensionality. Our novel statistical model filters the most discriminative terms to reduce the data's dimensionality and improve classification accuracy. We compare the proposed approach with the seven well-known feature selection and ranking techniques, namely normalized difference measure (NDM), bi-normal separation (BNS), odds ratio (OR), GINI, distinguished feature selector (DFS), information gain (IG), and Chi square (Chi). Our ensemble-based approach achieves high performance on two benchmark datasets, BET and UFN, achieving an accuracy of 88% and 90%, respectively.

INTRODUCTION

The term "fake news" represents the stories that are intentionally and undeniably bogus and intended to control individuals' views of genuine realities, events, explanations, and occasions (*Miro-Llinares & Aguerri, 2023*; *Choudhury & Acharjee, 2023*). Fake news covers deception and false data or disinformation deliberately spread to delude individuals (*Ruffo et al., 2023*). It is all about the information projected as news being misleading as it is based on demonstrably incorrect facts and events that never occurred (*Monsees, 2023*; *Cantarella, Fraccaroli & Volpe, 2023*). The internet, social media, and news media provide

a platform for millions of users to access up-to-date information and perform social interactions (*Cheng, Ge & Cosco, 2023*; *Longo, 2023*; *Lytos et al., 2019*; *Rodríguez-Ferrándiz, 2023*). The data significantly impacts their opinions and choices for different aspects of their lives (*Xing et al., 2022*; *Kozitsin, 2023*; *Vuong et al., 2019*). Studies have shown that a viral story has an echo chamber effect, and the user is more inclined to have a favorable opinion about it (*Robertson et al., 2023*; *Scheibenzuber et al., 2023*; *González-Bailón & Lelkes, 2023*). Manually differentiating between a real and fake viral story is becoming challenging with the ever-increasing amount of online data (*Khan, Michalas & Akhunzada, 2021*). There are a lot of examples of fake news that we can see throughout history impacting societal values and norms, changing opinions on critical issues, and redefining truths, facts, and beliefs (*Farago, Kreko & Orosz, 2023*; *Olan et al., 2022*). As unreliable and fake information has significant social and economic consequences for society (*Aïmeur, Amri & Brassard, 2023*), it is essential to automatically distinguish between real and fake news (*Buzea, Trausan-Matu & Rebedea, 2022*).

There has been significant research on fake news classification for the English language in the past few years (*Lillie & Middelboe, 2019*; *Shu et al., 2017*; *de Souza et al., 2020*; *Rohera et al., 2022*). However, the work for fake news classification in the Urdu language remains minimal (*Amjad et al., 2020b*; *Amjad, Sidorov & Zhila, 2020*; *Khiljia et al., 2020*) due to the unavailability of a sufficient annotated corpus. Therefore, many challenges must be addressed to solve this problem for the Urdu language. Urdu belongs to the Indo-Aryan group, is written in Arabic Person script, and is the official language of Pakistan. Urdu is spoken by more than 100 million speakers worldwide, but it remains a resource-poor language (*Nazir et al., 2021*; *Ullah et al., 2022*). The datasets for the Urdu language are also small, and deep learning approaches do not perform well on downstream natural language processing tasks such as sentiment analysis and fake news classification (*Rana et al., 2021*). In this regard, one of the foremost concerns is understanding the basis on which an Urdu fake news piece can be classified accurately. It includes extracting appropriate features and ranking discriminative features for Urdu fake news classification.

There are three different paradigms for fake news classification, as shown in Fig. 1: style-based, context-based, and knowledge-based (*Raza & Ding, 2022*). Style-based approaches mainly classify fake news based on deception detection and text categorization (*Kasseropoulos & Tjortjis, 2021*; *Hangloo & Arora, 2021*). Such methods use the news content and extract lexical features to discriminate between real and fake news. Such approaches also require effective feature selection techniques to classify fake news. The context-based paradigm exploits social network analysis to classify fake news (*Donabauer & Kruschwitz, 2023*; *Sivasankari & Vadivu, 2021*). Such approaches use the user's social engagement with the news content and the network of users to identify fake news content. Finally, knowledge-based classification (fact-checking) uses information retrieval techniques, or the semantic web, to detect fake news (*Seddari et al., 2022*; *Pathak & Srihari, 2019*; *Ceron, de Lima-Santos & Quiles, 2021*).

Style-based approaches for fake news classification are based on machine learning and deep learning models. Deep learning models do not perform well on small datasets, and because Urdu is a resource-poor language, no pre-trained models can be fine-tuned for such
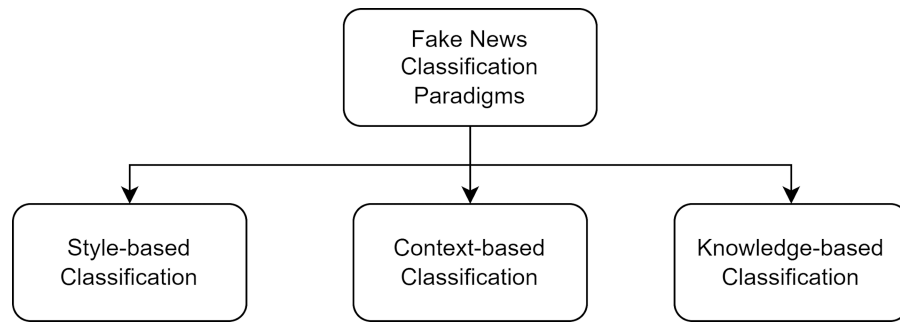
Figure 1 Different paradigms for fake news classification (*Potthast et al., 2017*).
Full-size 🖼 DOI: 10.7717/peerjcs.1612/fig-1

small datasets. Machine learning models are more suitable for resource-poor languages. These models rely on selecting the most relevant features to discriminate between fake and real news (*Rafique et al., 2022*; *Pal, Pranav & Pradhan, 2023*). The first step is to extract the features from the raw text. Secondly, the terms present in the corpus are weighted based on weighting schemes such as the binary vectorizer or TFIDF vectorizer. Finally, different statistical approaches are used for feature vectors and their class labels. After assigning weights to each term, some feature selection measure is used to find and rank the most discriminating terms. In recent studies, ensemble learning approaches outperformed for fake news detection (*Akhter et al., 2021*; *Mahabub, 2020*; *Hakak et al., 2021*; *Fayaz et al., 2022*; *Al-Ash et al., 2019*). Ensemble learning aims to exploit the diversity of base models to increase overall performance by handling multiple error types.

This article proposes a novel feature weighting model named normalized effect size (NES) that is based on a global weighting scheme (TFIDF). We extracted three features from the Urdu news articles: word n-grams, character n-grams, and function n-grams. The top k terms ranked by NES are used to train an ensemble model. The performance of the proposed approach is compared with seven feature selection measures, namely Normalized Difference Measure (NDM), Bi-normal Separation (BNS), Odds Ratio (OR), GINI, Distinguished Feature Selector (DFS), Information Gain (IG) and Chi Square (Chi). Our significant contributions to this research are as follows:

1. Proposal of a robust feature selection approach to discriminate between fake and real news;
2. Comparison of the performance of the proposed feature metric with seven well-known feature selection methods, showing the high performance of our feature metric;
3. Analysis of the performance of the proposed feature selection metric on two benchmark Urdu fake news datasets.

All example articles in the benchmark datasets have two classes assigned to them, *i.e.,* real or fake news. Therefore, our problem is a binary classification problem compared to the multi-class classification problem, in which more than two classes are used for labeling a dataset's examples. The following section briefly reviews the literature on fake news classification.

LITERATURE REVIEW

Previously, researchers have used feature selection for machine learning on different genres of textual data (*Katakis, Tsoumakas & Vlahavas, 2005*; *Rehman, Javed & Babri, 2017*; *Khan, Alam & Lee, 2021*; *Ramasamy & Meena Kowshalya, 2022*). In this section, we shall focus on feature selection for fake news classification, specifically for Urdu fake news classification. We shall review the feature selection techniques that have been used previously for the same or related tasks and summarize their findings.

Urdu is the national language of Pakistan and the 8th most spoken language globally, with more than 100 million speakers (*Akhter et al., 2020*). It is a South Asian language with limited resources (*Nazir et al., 2021*). A few annotated corpora in a few domains are available for research purposes, compared to English, which is a resource-rich language (*D'Ulizia et al., 2021*). The availability of insufficient linguistic resources like stemmers and annotated corpora makes the research on Urdu fake news classification more challenging and inspiring. Labeling a news article as "fake" or "legitimate" requires experts' opinions and is time-consuming. Also, hiring experts for each related domain is costly. In *Amjad et al. (2020b)*, *Amjad, Sidorov & Zhila (2020)*, *Amjad et al. (2020a)* and *Amjad et al. (2022)*, the authors proposed an annotated fake news corpus with a few hundred news articles. Their experimental results reveal the poor performance of machine learning models. Deep Learning models perform poorly due to the small corpus available for Urdu fake news.

Ensemble learning techniques boost the efficiency of individual machine learning models by aggregating the predictions in a way, also called base learners, base models, and base predictors (*Sagi & Rokach, 2018*). *Mahabub (2020)* applied eleven machine learning classifiers on a fake news corpus, including neural network-based models. Three of eleven machine learning models were selected to ensemble a voting model. Ensemble soft voting results reflect better performance than other models. To detect fake reviews, two ensemble learning approaches, bagging and boosting, were applied with SVM and MLP-based learners, and their research findings reflect that boosting with MLP outperforms the others (*Singh & Selva, 2023b*; *Singh & Selva, 2023a*; *Gutierrez-Espinoza et al., 2020*). Numerous ways are used to achieve this, such as re-sampling the corpus, heterogeneous models, homogeneous models with diverse parameters, and using various methods to combine the predictions of base models (*Kunapuli, 2023*).

Machine learning models are applied for fake news detection and classification tasks for languages such as English, Portuguese, Urdu, Arabic, Spanish, and Slavic (*Lahby et al., 2022*; *Nirav Shah & Ganatra, 2022*; *Ahmed et al., 2021*). Less effort has been made to explore ensemble learning techniques to fake news classification compared to machine learning methods (*Capuano et al., 2023*; *Chiche & Yitagesu, 2022*).

Posadas-Durán et al. (2019) defined three categories to detect fake news: knowledge-based, context-based, and style-based. They used bag-of-words (BOW), POS tags, and n-gram features. In normalization, they removed tags such as the editor's emails or phone numbers. Their experiments showed that the random forest outperformed with the highest accuracy of 76.94% using BOW, POS, and n-grams. Their work was primarily focused on the Spanish language.

Reis et al. (2019) used PolitiFact, Channel 4, and Snoops to gather fake news articles. They used linguistic inquiry word count (LIWC) and could attain an accuracy of 60%. Similarly, in another similar study (*Ahmed, Traore & Saad, 2018*), the authors used two feature weighting techniques, TF and TF-IDF, along with N-gram features. Their extracted features with the linear SVM classification model outperformed the baseline with an accuracy of 90% by using Bigram and 10,000 features.

Krešňáková, Sarnovský & Butka (2019) used preprocessing techniques such as stopwords, punctuation removal, and word2vec word embedding to represent the features. They conducted experiments with four classification models: feedforward neural networks, CNNs with one convolutional layer, CNNs with more convolutional layers, and LSTMs. They used only textual data to train the model and achieved the highest F1 score of 97.52% by using CNN mode. Using the CNN model, they also used the title and text to train the model and achieved an F1-score of 93.32%.

Bajaj (2017) collected a fake news dataset from Kaggle. They used different classification models: logistic regression (LR), feedforward neural networks, recurrent neural networks (RNN), long short-term memory (LSTM), gated recurrent units (GRU), bidirectional RNN with LSTM, a convolutional neural network with max-pooling, and an attention-augmented convolutional neural network (CNN). In this model comparison, GRU achieved the highest F1-score of 84%, and the performance of CNN was deficient (6%). Similarly, *Saikh et al. (2020)* collected the dataset in six domains and analyzed it. They used hand-crafted linguistic features and support vector machines. They achieved accuracies of 74% and 76% in the AMT and Celebrity News datasets, respectively. Also, they solved this problem using deep learning approaches. The first model was Bi-directional Gated Recurrent Unit (BiGRU), and the second was embedded from the language model (ELMo) and got accuracies of 54% and 68%, respectively.

Granik & Mesyura (2017) worked on unsolicited messages (spam) and fake news. Both types of messages have grammatical mistakes and false content. Fake news and spam both have some similarities. The set of words in one spam article is also in fake news and other spam articles. In this paper, the source dataset was BuzzFeed News. This dataset was contained in Facebook news posts. They collected data from three news pages (ABC News, Politico, and CNN). The dataset contained 2,282 posts. After cleaning the data, they obtained 1,771 articles. This data was classified into three subsets (training, validation, and testing). For true probability, the threshold value was [0.5; 0.9]. The unconditional probability of the news article was 59%, and the true probability of the threshold was 80%. A total number of fake news articles in the dataset contained 46 news articles, and 333 news articles were successfully classified with 71.73% accuracy. From a total of 927 articles, 699 were successfully classified. The accuracy of true news articles was slightly better than that of fake news articles. They used 2,000 articles, and this dataset was minimal to improve performance. The performance would also improve by using stemming and removing stop words. The total number of true news articles in the dataset contained 881 and 666 news articles successfully classified with 75.99% accuracy.

Liu & Wu (2018) collected three datasets (Weibo, Twitter15, and Twitter16) from Chinese and US social media sites. In the Weibo dataset, stories have binary labels, *i.e.,* fake

and true. On the other hand, in the Twitter15 and Twitter16 datasets, stories contained four labels (fake, true, unverified, and debunked). The fake news detection modal had four components: propagation path construction and transformation, RNN-based propagation path representation, CNN-based propagation path representation, and propagation path classification. They used 75% of the collected dataset for training and 25% of the dataset for testing. They gathered user information from their profiles. Furthermore, they found eight common characteristics between Weibo and Twitter and used stochastic gradient descent to train the model over 200 epochs. They used three proposed models: propagation path classification (PPC) (PPC_RNN, PPC_CNN, PPC_RNN+CNN). They compared their model with the baseline comparison model (DTC, SVM-RBF, SVM-TS, DTR, GRU, RFC, and PTK). The performance of the PPC_RNN+CNN model was the best among all. The accuracy of Twitter15 was 84.2%, Twitter16 was 86.3%, and Weibo was 92.1%.

Vogel & Meghana (2020) worked in two languages (English and Spanish). They used the n-gram feature and support vector machine (SVM) for English and logistic regression for Spanish. They used the PAN 2020 dataset, which has 300 English and 300 Spanish Twitter accounts, and each user had 100 tweets. Likewise, they removed stop words by using the NLTK library. They used 70% of the data for training and 30% for testing. The accuracy using SVM and the TF-IDF char n-gram feature in English was 73%. Also, the accuracy using logistic regression and the TF-IDF char n-gram feature in Spanish was 79%.

Ahmed, Traore & Saad (2018) used n-gram and machine learning techniques. They used two feature extraction techniques (term frequency-inverted document (TF-IDF) and n-gram) and six classification techniques (Stochastic Gradient Descent (SGD), Decision Tree (DT), Linear Support Vector Machine (LSVM), Logistic Regression (LR), Support Vector Machine (SVM) and K-nearest neighbor (KNN)). The TF-IDF feature and LSVM classifier achieved the highest accuracy of 92%. The dataset contained 25,200 articles, with an equal number of real and fake documents. They used n-gram features ranging from 1 to 4. In all the non-linear classifiers, DT achieved the highest accuracy of 89%. The linear base classifiers (LR, LSVM, and SDG) performed better than the non-linear classifiers. The performance of the model decreased as the number of n-grams increased. KNN and SVM achieved the lowest accuracy of 47.2%. The performance of SVM with a linear kernel attained an accuracy of 71%.

Monti et al. (2019) used deep geometric learning for fake news detection. The author collected datasets from Snopes, Politifact, and BuzzFeed. They extracted features into four categories: user profile, activity, network and spreading, and content. They used four layers of CNN graphs and two convolutional layers to predict the probability of fake and real articles. Furthermore, they used scaled exponential linear units (SELU) in their network. The dataset was split into training (677), testing (226), and validation (226) for URL classification. They used the same pattern to split data, like URL classification for cascade classification. The ROC AUC was 92.70 ± 1.80% URL-wise classification and 88.30 ± 2.74% cascade-wise classification. For URL settings, they split the data into 80% of URLs for training and 20% for testing. In the ablation experiment for both settings, two features (user profile and network spreading) had high importance, with nearly 90% ROC AUC. There was a time t for the first tweet, and t was 0 to 24 h. Each value of the

t model was trained separately. They also used five cross-validations to reduce bias. The method improved performance with cascade duration. There was different behavior due to different characteristics of cascades and URLs. The highest performance was 92.7% ROC AUC.

Only a handful of researchers have worked on fake news in the Urdu language. *Amjad et al. (2020b)* manually collected and verified the datasets. Their dataset contained 500 real and 400 fake news articles. They used raw frequency, a binary weighting scheme, normalized frequency, log entropy, and TF-IDF. They got a good performance with the combinations of character-word 2-grams and 1-grams, reporting a F1-score of 87% for fake news and a F1-score of 90% for legit news using the AdaBoost classifier. In a different study by *Amjad, Sidorov & Zhila (2020)*, the researchers developed a new dataset using machine translation (MT). The model's performance on the new dataset was not satisfactory compared to the original. Binary weighting schemes were used for feature normalization (TF-IDF) and log-entropy (decrease classification performance). They used Support Vector Machines (SVM) and AdaBoost classifiers. Then they used the dataset of *Amjad et al. (2020b)*, using the Original Urdu dataset combination of character and word Unigrams, and achieved excellent results with an F1-score of 84% for fake and a ROC-AUC score of 94%.

Khiljia et al. (2020) utilized generalized autoregressive-based model techniques on a dataset of 350 real and 288 fake articles. The data was preprocessed using the UrduHack library, removing mobile numbers, email IDs, URLs, and extra spaces. The XLNet pre-trained model was used for training, but BERT was used for fine-tuning. The study achieved a F1 macro score of 83.70%.

The CharCNN-RoBETa model was proposed by *Lina, Fua & Jianga (2020)*, using a dataset of 500 real and 400 fake articles. The model represented sentences as word and character embeddings by concatenating vectors and applying softmax for prediction. Five-fold cross-validation was used to achieve high accuracy. The model achieved a F1-score of 99.99% with the RoBERTa+pretrain model, 91.18% with the RoBERTa+pretrain+label smoothing model, 91.25% with the RoBERTa+charcnn+pretrain model, and 91.41% with the RoBERTa+charcnn+pretrain model.

Table 1 presents a summary of the comparison of previous work and the features used for Fake News Classification, where it is verified that the word n-gram.

PROPOSED METHODOLOGY

Our proposed methodology for Urdu fake news classification follows four steps, where the first step involves feature extraction from the documents. Next, we apply different feature weighting schemes to the feature vectors. Once weights are assigned, we use state-of-the-art feature selection techniques to select highly discriminative features and compare them with our proposed feature selection metric. Finally, the classification model (AdaBoost) is trained based on these features. This methodology was implemented in *Wasim (2023)*, and Fig. 2 shows the complete process with the details of each component in the following subsections.

Table 1 Comparison of previous work and the features used for fake news classification.

| No.	Paper	Features	Models	Performance
1	*Amjad et al. (2020b)*	TF-IDF, log-entropy, character n-grams, word n-grams	AdaBoost	84% F1 Fake, 94% ROC-AUC scores, lower ROC-AUC of 93%.
2	*Amjad, Sidorov & Zhila (2020)*	Word n-gram, char n-grams, and functional n-grams	AdaBoost Classifier's	Accuracy 87%, F1 Fake 91%
3	*Monti et al. (2019)*	User profile, User activity, Network and spreading, Content	Convolutional Neural Network (CNN)	92.7% ROC AUC
4	*Humayoun (2022)*	Word n-gram, character n-gram	Support Vector Machine, CNN Embeddings	F1 macro 66%, Accuracy 72%
5	*Rafique et al. (2022)*	TF-IDF, BoW, Character N-gram, word N-gram	NF, LR, SVC, GB, PA, Multinomial NB	Accuracy 95%
6	*Amjad et al. (2020a)*	Character bi-gram, MUCS, BoW, Random	BERT 4EVER, Logistic Regression	Accuracy 90%
7	*Amjad et al. (2022)*	TF-IDF, count-based BoW, word vector embeddings	SVM, BERT, RoBERta	F1-macro 67%, Accuracy 75%
8	*Kalra et al. (2022)*	N/A	Ensemble Learning, ROBERTA, ALBERT, Multilingual Bert, xlm-RoBERTa	Accuracy 59%
9	*Salahuddin & Wasim (2022)*	TF-IDF	Logistic Regression	F1 Score 72%
10	*Akhter et al. (2021)*	BoW, IG	SVM, Decision Tree, Naive Bayes	BA 81.6%, AUC 81.5%, MAE 23.5%
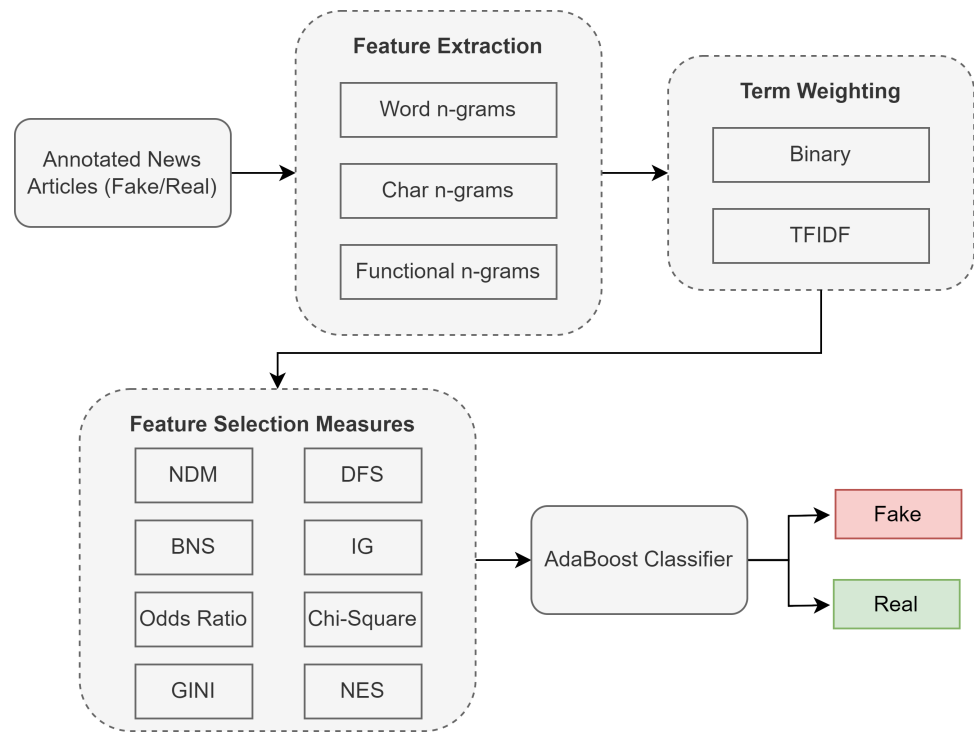
Figure 2 Proposed methodology with different feature selection techniques for Urdu fake news classification.

Full-size ☑ DOI: 10.7717/peerjcs.1612/fig-2

Feature extraction

We extracted three types of features from our datasets: word n-grams, character n-grams, and functional n-grams.

Word n-grams: The first feature we extracted from the document is word unigrams. We use the unigram feature, as we observed that these features have superior performance and low sparsity compared to higher-order n-grams for fake news classification.

Character n-grams: The second feature extracted from the document is character n-grams. The purpose of this feature is to capture the syntactic and morphological elements present in the document. We extracted 2-gram sequences of characters for the character n-gram feature.

Functional word n-grams: The third feature is functional word n-grams. We extract these features as previous studies on fake news classification have shown improved performance using functional words. The functional words include determiners, prepositions, articles, and auxiliary verbs. The functional word n-grams is a sequence of these words, omitting the content words, and we use a sequence of 2-gram functional words.

Term weighting schemes

Term weighting for the extracted features is vital to improving the performance of fake news classification. We perform experiments with two different types of term weighting schemes:

Binary weighting: In this weighting scheme, we use binary feature values. If a feature is present one or more times in the document, it is assigned a value of 1. Otherwise, it is assigned the value of zero.

TFIDF weighting: TFIDF is a well-known weighting measure used in information retrieval and classification. It is calculated as:

$$TFIDF = TF(t,d) \times log(\frac{N}{df_i}) \tag{1}$$

where TF is the value of a term *t's* presence in document *d*, df_i is the document frequency of term t_i, and N is the total number of documents in the corpus.

Feature selection (FS) methods

Feature selection methods play a vital role in machine classification tasks. We provide the necessary details of the state-of-the-art feature selection methods for comparison. We start by defining the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) of a term that forms the basis of all the feature selection measures, where TP is the number of documents in the positive class with the term in them, TN is the number of documents that are not in the positive class and have no term, FP is the number of documents in the negative class with the term in them, and FN is the number of documents that are not in the negative class and have no term.

Normalized difference measure (NDM)

Normalized difference measure (NDM) is a new feature selection technique proposed by *Rehman, Javed & Babri (2017)*. This is based on *tpr* and *fpr* where:

$$tpr = \frac{TP}{N_{pos}} \tag{2}$$

$$fpr = \frac{FP}{N_{neg}} \tag{3}$$

where N_{pos} is the number of positive documents and N_{neg} is the number of negative documents in the corpus. NDM is based on the following principles:

- An important term should have high $|tpr - fpr|$ value.
- One of the tpr or fpr values should be closer to zero.
- If two terms have equal $|tpr - fpr|$ values, then the term having a lower min (tpr, fpr) value should be assigned a higher rank where min is the function to find a minimum of the two values.

Mathematically, NDM is defined as:

$$NDM = \frac{|tpr - fpr|}{min(tpr, fpr)} \tag{4}$$

Note that if the min of *tpr* and *fpr* is zero, a smaller value like 0.01 is used.

Bi-normal separation (BNS)

Forman (2003) proposed bi-normal separation and is defined as:

$$BNS = |F^{-1}(tpr) - F^{-1}(fpr)| \tag{5}$$

where F^{-1} is the standard Normal distribution's inverse cumulative probability function or z-score. To avoid the undefined value $F^{-1}(0)$, zero is substituted by 0.0005.

Odds ratio (OR)

The odds ratio depends on the probability of a term's occurrence or whether that term is not present in a document. This feature extraction technique only relies on the probability of the occurrence of terms.

$$OddsRatio = \frac{TPR(1 - FPR)}{(1 - TPR)FPR} = \frac{TP \times TN}{FP \times FN}. \tag{6}$$

If fp or fn is zero, the denominator will be zero values ($fp * fn = 0$), which is replaced with a small value.

Gini

The Gini index was initially used to estimate income distribution across a population. It is also used as a feature ranking metric, where it is used to estimate the distribution of an attribute over different classes.

$$GINI = tpr^2 \times \left(\frac{tp}{tp \times fp} \right)^2 + fpr^2 \times \left(\frac{fp}{tp \times fp} \right)^2 \tag{7}$$

Distinguished feature selector (DFS)

Uysal & Gunal (2012) proposed a distinguished feature selector based on the idea that the terms present in one class only play an important role in discriminating between the two classes.

$$DFS(t) = \sum_{i=1}^{m} \frac{P(c_i|t)}{P(\bar{t}|c_i) + P(t|\bar{c}_i) + 1} \tag{8}$$

where $P(c_i)$ is probability of ith class and $P(\bar{t}|c_i)$ is probability of absence of term t when class c_i is given.

Information Gain (IG)

Information gain is widely used to assess the usefulness of features for machine learning. It measures the decrease in entropy when the feature is given *vs.* when the feature is absent. It is defined as:

$$IG(t) = -\sum_{i=1}^{m} P(c_i)logP(c_i) - \left(P(t)\sum_{i=1}^{m}P(c_i|t)logP(c_i|t) + P(\bar{t})\sum_{i=1}^{m}P(c_i|\bar{t})logP(c_i|\bar{t}) \right) \tag{9}$$

where c_i is set of classes in the dataset, $P(c_i)$ is probability of *ith* class and $P(c_i|t)$ is probability of *ith* class when term t is present and $P(c_i|\bar{t})$ is probability of class c_i when term t is absent.

Chi square (Chi)

Chi-square (Chi) measures the divergence from the expected distribution, assuming that the presence or absence of a term is independent of the class label (*Forman, 2003*). It is defined as:

$$CHI = t(tp, (tp+fp)P_{pos}) + t(fn, (fn+tn)P_{pos}) + t(fp, (tp+fp)P_{neg}) + t(tn, (fn+tn)P_{neg}) \quad (10)$$

where $t(count, expect) = (count - expect)^2/expect$, P_{pos} is the probability of the positive class and P_{neg} is the probability of the negative class.

Propose feature selection measure: normalized effect size (NES)

According to *Sullivan & Feinn (2012)*, the effect size is used in medical research to determine the magnitude of the difference between two groups. Inspired by their idea, we define normalized effect size (NES) as the absolute mean difference between a term's TF-IDF score distribution in positive and negative classes normalized by the sum of their standard deviation. NES is a global weighting technique for discriminative feature selection, mathematically defined as:

$$NES = \frac{|\mu_+ - \mu_-|}{\sigma_+ + \sigma_-} \quad (11)$$

where:

- μ_+ is the mean of the TFIDF score of a term across all the documents labeled as positive
- μ_- is the mean of the TFIDF score of a term across all the documents labeled as negative
- σ_+ is the standard deviation of the TFIDF score of a term across all documents labeled as positive
- σ_- is the standard deviation of a TFIDF score of a term across all documents labeled as negative

Classification model

Previous studies show that ensemble learning approaches, such as bagging and boosting, outperform the other algorithms for Urdu fake news classification. Therefore, we use the adaptive boosting technique, which boosts multiple serial estimators, to improve the classification performance. We experimentally set the hyperparameter values for alpha (learning rate) and the number of estimators. The value of the learning rate hyperparameter was set to 0.1, and the number of estimators for AdaBoost was set to 300.

EXPERIMENTAL EVALUATION

This section discusses the corpus used to evaluate the proposed methodology for evaluating different feature selection metrics. We use a well-known accuracy measure to analyze and compare feature selection measures. The results and comparison are presented in the last subsection.

Evaluation corpus

We use two datasets to evaluate the performance of our proposed feature selection measure, namely the Bend the Truth (BET) dataset (*Amjad et al., 2023*) and the Urdu Fake News (UFN) dataset (*Akhter, 2023*).

Table 2 The number of real and fake news documents present in each category for the BET dataset.

Category	Real	Fake
Business	100	50
Health	100	100
Showbiz	100	100
Sports	100	50
Technology	100	100
Total	500	400

Table 3 Classification performance on feature ranking metrics with varying number of terms on UFN dataset.

No. of terms	Feature selection measures							
	NDM	BNS	OR	Gini	DFS	IG	CS	NES
100	0.78	0.77	0.63	0.55	0.85	0.84	0.84	0.88
200	0.78	0.8	0.65	0.63	0.87	0.86	0.85	0.89
300	0.8	0.83	0.64	0.72	0.87	0.87	0.86	0.90
400	0.8	0.84	0.67	0.76	0.86	0.86	0.87	0.89
500	0.81	0.85	0.67	0.81	0.88	0.87	0.87	0.89
1,000	0.85	0.86	0.73	0.84	0.89	0.89	0.88	0.89
1,500	0.86	0.87	0.74	0.87	0.89	0.88	0.89	0.89
2,000	0.85	0.88	0.75	0.88	0.88	0.88	0.88	0.89

The BET dataset comprises five categories: business, health, showbiz, sports, and technology. The corpus contains 900 documents (500 real and 400 fake new documents). The statistics of each category with the number of real and fake news are shown in Table 2.

The Urdu Fake News (UFN) dataset contains 1,032 real and 968 fake news stories, representing 2,000 stories. The dataset does not provide any categories and is just a translation of the fake news English corpus. We use 65% of the data for training and 35% for testing purposes for both datasets.

Results and comparison

This section presents the performance of different feature selection metrics on the BET and UFN datasets. We compare the proposed method with seven well-known feature selection measures. These feature selection measures are evaluated on different numbers of terms selected, from 50 to 2,000 top terms. The section also compares the results of the proposed method with previous results on the same datasets.

Table 3 shows the results of the performance of the AdaBoost classifier for a different number of terms selected by the feature selection measures in our study compared with the proposed feature selection measure. The proposed measure performs better than all other feature metrics over all the different numbers of terms. The results are also depicted in Fig. 3.

Surprisingly, both binary and TF-IDF scores perform similarly on the UFN dataset, as shown in Fig. 3, attaining an accuracy of 89%. The odds ratio (OR) performance is worst

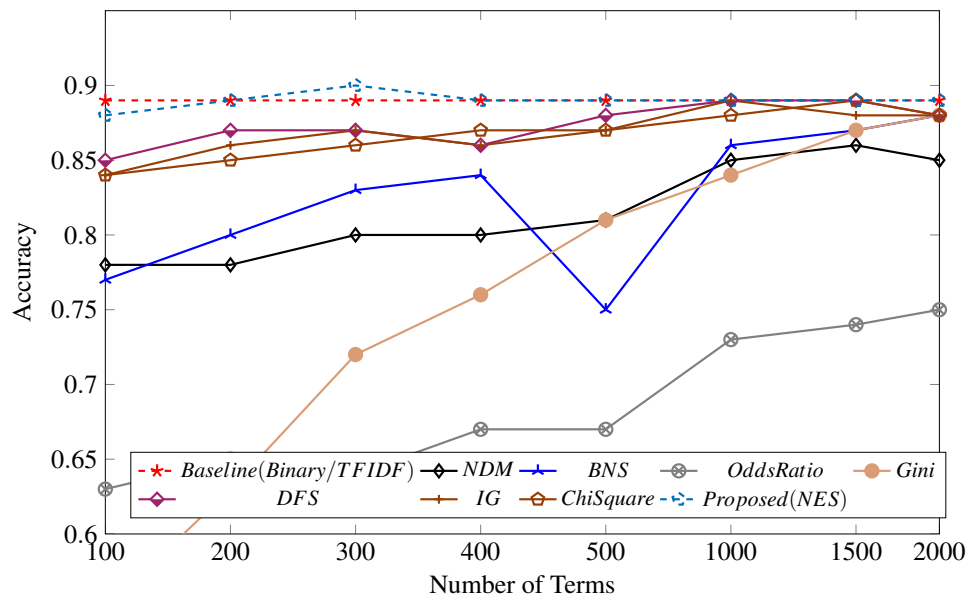**Figure 3** Feature selection measure applied to the UFN dataset.

Table 4 Classification performance on feature ranking metrics with varying number of terms on BET dataset.

No. of terms	Feature selection measures							
	NDM	BNS	OR	Gini	DFS	IG	CS	NES
100	0.69	0.85	0.74	0.51	0.88	0.88	0.88	0.87
200	0.81	0.78	0.8	0.8	0.86	0.84	0.85	0.88
300	0.82	0.76	0.76	0.81	0.85	0.85	0.85	0.86
400	0.83	0.73	0.78	0.84	0.87	0.83	0.86	0.86
500	0.85	0.77	0.79	0.86	0.87	0.87	0.87	0.87
1,000	0.84	0.85	0.82	0.86	0.87	0.87	0.87	0.85
1,500	0.85	0.85	0.83	0.87	0.85	0.87	0.86	0.85
2,000	0.84	0.84	0.84	0.85	0.85	0.87	0.87	0.85

across all different numbers of terms. All the other measures could not perform well in selecting discriminative terms for fake news classification.

Table 4 shows the feature selection metrics results compared with the proposed methodology. It shows that our proposed method performs well for 100, 200, and 300 features, achieving the highest accuracy of 88% with 200 top features. Similar to our observation for the UFN dataset, the binary and TF-IDF weightings achieved identical results, as shown in Fig. 4.

We compare the performance of feature weighting techniques without any feature selection metric and compare their performance when our proposed method is used, as shown in Fig. 5. It shows that performance significantly improves when our proposed feature selection metric is used for the classification task on both datasets. The performance
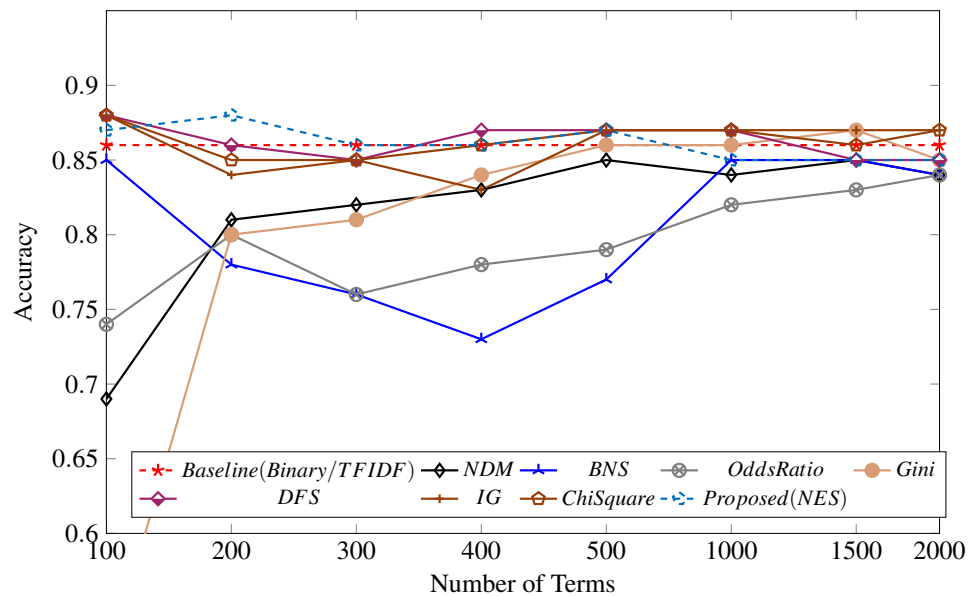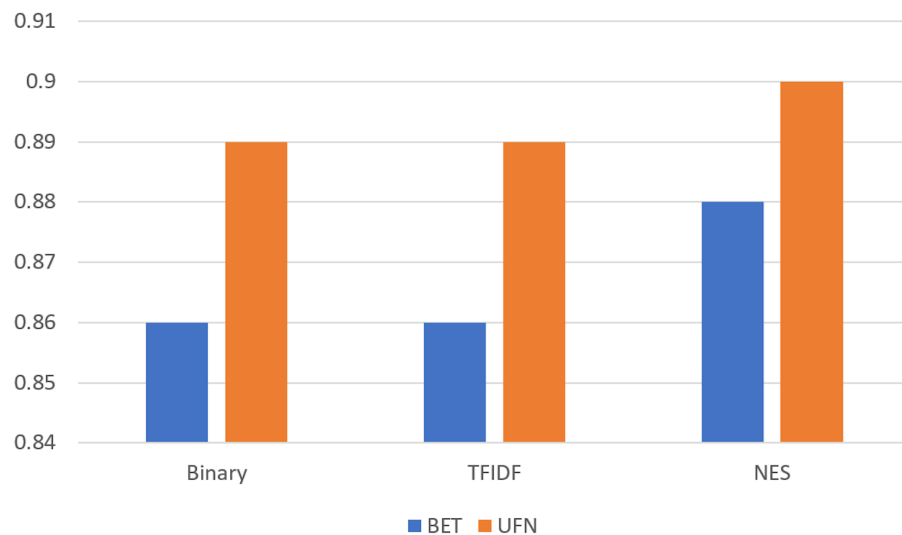
Figure 4 Feature selection measure applied to the BET dataset.

Figure 5 Comparison on BET and UFN dataset for simple weighting compared with the proposed feature selection metric (NES).

on the BET dataset is low for all weighting measures and the proposed feature selection measure. The reason for this is the small dataset size.

We also compare the performance of the proposed method and the final results with previous state-of-the-art results on the same dataset. Table 5 compares the achieved results to those of prior studies. The results show that the feature selection metric-based methodology performs better on both datasets.

Table 5 Comparison of the proposed method with the previous studies.

Study	BET dataset	UFN dataset
Salahuddin & Wasim (2022)	72%	NA
Amjad et al. (2020b)	83%	NA
Akhter et al. (2020)	83%	89%
Proposed approach (NES)	88%	90%

CONCLUSION

The findings of this study demonstrate the critical need for effective feature selection techniques for low-resource languages such as Urdu. The scarcity of language resources requires effective feature engineering and selection techniques, and machine learning approaches can be effectively applied to Urdu fake news classification. The study presented a novel feature selection measure (NES) to select discriminative features for the Urdu fake news classification. Our proposed method ranked the discriminating terms for filtering, decreasing the data dimensionality and improving the classification performance, as evident from the experimental results. The binary and TF-IDF weighting metrics performed similarly, and we used them as our baseline. To compare and analyze the performance of our feature selection model, we compared it with seven well-known feature selection methods.

Moreover, we evaluated the performance of our proposed model on the BET and UFN datasets. Our analysis of both datasets showed that our proposed approach works well in finding highly discriminative features to classify the Urdu news. Our proposed approach achieved an accuracy of 88% and 90% on the BET and UFN datasets, respectively. In the future, the proposed feature selection model can also be used for other text classification problems. Moreover, we plan to work on language resources to facilitate downstream tasks for the Urdu language.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

Ivan Miguel Pires is an Academic Editor for PeerJ Computer Science.

Author Contributions

- Muhammad Wasim conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Sehrish Munawar Cheema conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Ivan Miguel Pires conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The code is available at GitHub and Zenodo:

- https://github.com/dr-m-wasim/UrduFakeNewsFS.

- Muhammad Wasim, Sehrish Munawar Cheema, & Ivan Miguel Pires. (2023). Normalized Effect Size (NES): a novel feature selection model for Urdu fake news classification. https://doi.org/10.5281/zenodo.8320957.

The BET Dataset is available at GitHub: https://github.com/MaazAmjad/Datasets-for-Urdu-news.

Institution: Natural Language and Text Processing Laboratory, Center for Computing Research (CIC), Instituto Politécnico Nacional (IPN), Ciudad de México (Mexico City), Mexico

Contact: Maaz Amjad (maazamjad@phystech.edu)

The UFN Dataset is available at GitHub: https://github.com/pervezbcs/Urdu-Fake-News.

Institution: Department of Humanities and Basic Sciences, MCS, National University of Sciences and Technology, Islamabad, Pakistan

Contact: Farkhanda Afzal: (farkhanda@scm.edu.pk)

REFERENCES

Ahmed AAA, Aljabouh A, Donepudi PK, Choi MS. 2021. Detecting fake news using machine learning: a systematic literature review. ArXiv preprint. arXiv:2102.04458.

Ahmed H, Traore I, Saad S. 2018. Detecting opinion spams and fake news using text classification. *Security and Privacy* **1**(1):e9 DOI 10.1002/spy2.9.

Aïmeur E, Amri S, Brassard G. 2023. Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining* **13**(1):30 DOI 10.1007/s13278-023-01028-5.

Akhter MP. 2023. Urdu fake news dataset. *Available at https://github.com/pervezbcs/Urdu-Fake-News* (accessed on 07 June 2023).

Akhter MP, Jiangbin Z, Naqvi IR, Abdelmajeed M, Sadiq MT. 2020. Automatic detection of offensive language for Urdu and Roman Urdu. *IEEE Access* **8**:91213–91226 DOI 10.1109/ACCESS.2020.2994950.

Akhter MP, Zheng J, Afzal F, Lin H, Riaz S, Mehmood A. 2021. Supervised ensemble learning methods towards automatically filtering Urdu fake news within social media. *PeerJ Computer Science* **7**:e425 DOI 10.7717/peerj-cs.425.

Al-Ash HS, Putri MF, Mursanto P, Bustamam A. 2019. Ensemble learning approach on indonesian fake news classification. In: *2019 3rd international conference on informatics and computational sciences (ICICoS)*. Piscataway: IEEE, 1–6.

Amjad M, Butt S, Amjad HI, Zhila A, Sidorov G, Gelbukh A. 2022. Overview of the shared task on fake news detection in Urdu at Fire 2021. ArXiv preprint. arXiv:2207.05133.

Amjad M, Butt S, Amjad HI, Zhila A, Sidorov G, Gelbukh A. 2023. BET dataset. *Available at https://github.com/MaazAmjad/Datasets-for-Urdu-news* (accessed on 07 June 2023).

Amjad M, Sidorov G, Zhila A. 2020. Data augmentation using machine translation for fake news detection in the Urdu language. In: *Proceedings of the 12th language resources and evaluation conference*. 2537–2542.

Amjad M, Sidorov G, Zhila A, Gelbukh A, Rosso P. 2020a. UrduFake@ FIRE2020: shared track on fake news identification in Urdu. In: *Forum for information retrieval evaluation*. 37–40.

Amjad M, Sidorov G, Zhila A, Gómez-Adorno H, Voronkov I, Gelbukh A. 2020b. "Bend the truth": benchmark dataset for fake news detection in Urdu language and its evaluation. *Journal of Intelligent & Fuzzy Systems* **39**(2):2457–2469 DOI 10.3233/JIFS-179905.

Bajaj S. 2017. The pope has a new baby! fake news detection using deep learning. In: *CS 224N*. 1–8.

Buzea MC, Trausan-Matu S, Rebedea T. 2022. Automatic fake news detection for romanian online news. *Information* **13**(3):151 DOI 10.3390/info13030151.

Cantarella M, Fraccaroli N, Volpe R. 2023. Does fake news affect voting behaviour? *Research Policy* **52**(1):104628 DOI 10.1016/j.respol.2022.104628.

Capuano N, Fenza G, Loia V, Nota FD. 2023. Content based fake news detection with machine and deep learning: a systematic review. *Neurocomputing* **530**:91–103 DOI 10.1016/j.neucom.2023.02.005.

Ceron W, de Lima-Santos M-F, Quiles MG. 2021. Fake news agenda in the era of COVID-19: identifying trends through fact-checking content. *Online Social Networks and Media* **21**:100116 DOI 10.1016/j.osnem.2020.100116.

Cheng X, Ge T, Cosco TD. 2023. Internet use and life satisfaction among Chinese older adults: the mediating effects of social interaction. *Current Psychology* 1–8 DOI 10.1007/s12144-023-04303-y.

Chiche A, Yitagesu B. 2022. Part of speech tagging: a systematic review of deep learning and machine learning approaches. *Journal of Big Data* **9**(1):1–25 DOI 10.1186/s40537-021-00549-0.

Choudhury D, Acharjee T. 2023. A novel approach to fake news detection in social networks using genetic algorithm applying machine learning classifiers. *Multimedia Tools and Applications* **82**(6):9029–9045 DOI 10.1007/s11042-022-12788-1.

Donabauer G, Kruschwitz U. 2023. Exploring fake news detection with heterogeneous social media context graphs. In: *Advances in information retrieval: 45th european conference on information retrieval, ECIR 2023, Dublin, Ireland, April 2–6, 2023, Proceedings, Part II*. Cham: Springer, 396–405.

D'Ulizia A, Caschera MC, Ferri F, Grifoni P. 2021. Fake news detection: a survey of evaluation datasets. *PeerJ Computer Science* 7:e518 DOI 10.7717/peerj-cs.518.

de Souza JV, Gomes Jr J, Souza Filho FMd, Oliveira Julio AMd, de Souza JF. 2020. A systematic mapping on automatic classification of fake news in social media. *Social Network Analysis and Mining* 10:1–21 DOI 10.1007/s13278-019-0612-8.

Faragó L, Krekó P, Orosz G. 2023. Hungarian, lazy, and biased: the role of analytic thinking and partisanship in fake news discernment on a Hungarian representative sample. *Scientific Reports* 13(1):178 DOI 10.1038/s41598-022-26724-8.

Fayaz M, Khan A, Bilal M, Khan SU. 2022. Machine learning for fake news classification with optimal feature selection. *Soft Computing* 26(16):7763–7771 DOI 10.1007/s00500-022-06773-x.

Forman G. 2003. An extensive empirical study of feature selection metrics for text classification. *Journal of Machine Learning Research* 3(Mar):1289–1305.

González-Bailón S, Lelkes Y. 2023. Do social media undermine social cohesion? A critical review. *Social Issues and Policy Review* 17(1):155–180 DOI 10.1111/sipr.12091.

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

Gutierrez-Espinoza L, Abri F, Namin AS, Jones KS, Sears DR. 2020. Fake reviews detection through ensemble learning. ArXiv preprint. arXiv:2006.07912.

Hakak S, Alazab M, Khan S, Gadekallu TR, Maddikunta PKR, Khan WZ. 2021. An ensemble machine learning approach through effective feature extraction to classify fake news. *Future Generation Computer Systems* 117:47–58 DOI 10.1016/j.future.2020.11.022.

Hangloo S, Arora B. 2021. Fake news detection tools and methods–a review. ArXiv preprint. arXiv:2112.11185.

Humayoun M. 2022. The 2021 Urdu fake news detection task using supervised machine learning and feature combinations. ArXiv preprint. arXiv:2204.03064.

Kalra S, Verma P, Sharma Y, Chauhan G. 2022. Ensembling of various transformer based models for the fake news detection task in the Urdu language. In: *Forum for Information Retrieval Evaluation, December 13-17, 2021, India*.

Kasseropoulos DP, Tjortjis C. 2021. An approach utilizing linguistic features for fake news detection. In: *Artificial intelligence applications and innovations: 17th IFIP WG 12.5 international conference, AIAI 2021, Hersonissos, Crete, Greece, June 25–27, 2021, Proceedings 17*. Cham: Springer, 646–658.

Katakis I, Tsoumakas G, Vlahavas I. 2005. On the utility of incremental feature selection for the classification of textual data streams. In: *Panhellenic conference on informatics*. Cham: Springer, 338–348.

Khan J, Alam A, Lee Y. 2021. Intelligent hybrid feature selection for textual sentiment classification. *IEEE Access* **9**:140590–140608 DOI 10.1109/ACCESS.2021.3118982.

Khan T, Michalas A, Akhunzada A. 2021. Fake news outbreak 2021: can we stop the viral spread? *Journal of Network and Computer Applications* **190**:103112 DOI 10.1016/j.jnca.2021.103112.

Khiljia AFUR, Laskara SR, Pakraya P, Bandyopadhyaya S. 2020. Urdu fake news detection using generalized autoregressors. In: *Forum for information retrieval evaluation 2020, December 16-20, 2020, Hyderabad, India.*

Kozitsin IV. 2023. Opinion dynamics of online social network users: a micro-level analysis. *The Journal of Mathematical Sociology* **47(1)**:1–41 DOI 10.1080/0022250X.2021.1956917.

Krešňáková VM, Sarnovský M, Butka P. 2019. Deep learning methods for Fake News detection. In: *2019 IEEE 19th international symposium on computational intelligence and informatics and 7th IEEE international conference on recent achievements in mechatronics, automation, computer sciences and robotics (CINTI-MACRo)*. Piscataway: IEEE, 000143–000148.

Kunapuli G. 2023. *Ensemble methods for machine learning.* NY, USA: Simon and Schuster.

Lahby M, Aqil S, Yafooz WMS, Abakarim Y. 2022. Online fake news detection using machine learning techniques: a systematic mapping study. In: Lahby M, Pathan ASK, Maleh Y, Yafooz WMS, eds. *Combating fake news with computational intelligence techniques. Studies in computational intelligence*, Vol. 1001. Cham: Springer DOI 10.1007/978-3-030-90087-8_1.

Lillie AE, Middelboe ER. 2019. Fake news detection using stance classification: a survey. ArXiv preprint. arXiv:1907.00181.

Lina N, Fua S, Jianga S. 2020. Fake news detection in the urdu language using CharCNN-RoBERTa. *Health* **100**:100.

Liu Y, Wu Y-F. 2018. Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. In: *Proceedings of the AAAI conference on artificial intelligence, vol. 32.*

Longo GM. 2023. The internet as a social institution: rethinking concepts for family scholarship. *Family Relations* **72(2)**:621–636 DOI 10.1111/fare.12825.

Lytos A, Lagkas T, Sarigiannidis P, Bontcheva K. 2019. The evolution of argumentation mining: from models to social media and emerging tools. *Information Processing & Management* **56(6)**:102055 DOI 10.1016/j.ipm.2019.102055.

Mahabub A. 2020. A robust technique of fake news detection using ensemble voting classifier and comparison with other classifiers. *SN Applied Sciences* **2(4)**:525 DOI 10.1007/s42452-020-2326-y.

Miro-Llinares F, Aguerri JC. 2023. Misinformation about fake news: a systematic critical review of empirical studies on the phenomenon and its status as a 'threat'. *European Journal of Criminology* **20(1)**:356–374 DOI 10.1177/1477370821994059.

Monsees L. 2023. Information disorder, fake news and the future of democracy. *Globalizations* **20(1)**:153–168 DOI 10.1080/14747731.2021.1927470.

Monti F, Frasca F, Eynard D, Mannion D, Bronstein MM. 2019. Fake news detection on social media using geometric deep learning. ArXiv preprint. arXiv:1902.06673.

Nazir Z, Shahzad K, Malik MK, Anwar W, Bajwa IS, Mehmood K. 2021. Authorship attribution for a resource poor language—Urdu. *Transactions on Asian and Low-Resource Language Information Processing* **21(3)**:1–23.

Nirav Shah M, Ganatra A. 2022. A systematic literature review and existing challenges toward fake news detection models. *Social Network Analysis and Mining* **12(1)**:168 DOI 10.1007/s13278-022-00995-5.

Olan F, Jayawickrama U, Arakpogun EO, Suklan J, Liu S. 2022. Fake news on social media: the impact on society. *Information Systems Frontiers* 1–16 DOI 10.1007/s10796-022-10242-z.

Pal A, Pranav , Pradhan M. 2023. Survey of fake news detection using machine intelligence approach. *Data & Knowledge Engineering* **144**:102118 DOI 10.1016/j.datak.2022.102118.

Pathak A, Srihari RK. 2019. BREAKING! Presenting fake news corpus for automated fact checking. In: *Proceedings of the 57th annual meeting of the association for computational linguistics: student research workshop*. 357–362.

Posadas-Durán J-P, Gómez-Adorno H, Sidorov G, Escobar JJM. 2019. Detection of fake news in a new corpus for the Spanish language. *Journal of Intelligent & Fuzzy Systems* **36(5)**:4869–4876 DOI 10.3233/JIFS-179034.

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

Rafique A, Rustam F, Narra M, Mehmood A, Lee E, Ashraf I. 2022. Comparative analysis of machine learning methods to detect fake news in an Urdu language corpus. *PeerJ Computer Science* **8**:e1004 DOI 10.7717/peerj-cs.1004.

Ramasamy M, Meena Kowshalya A. 2022. Information gain based feature selection for improved textual sentiment analysis. *Wireless Personal Communications* **125**:1203–1219 DOI 10.1007/s11277-022-09597-y.

Rana TA, Shahzadi K, Rana T, Arshad A, Tubishat M. 2021. An unsupervised approach for sentiment analysis on social media short text classification in roman Urdu. *Transactions on Asian and Low-Resource Language Information Processing* **21(2)**:1–16.

Raza S, Ding C. 2022. Fake news detection based on news content and social contexts: a transformer-based approach. *International Journal of Data Science and Analytics* **13(4)**:335–362 DOI 10.1007/s41060-021-00302-z.

Rehman A, Javed K, Babri HA. 2017. Feature selection based on a normalized difference measure for text classification. *Information Processing & Management* **53(2)**:473–489 DOI 10.1016/j.ipm.2016.12.004.

Reis JC, Correia A, Murai F, Veloso A, Benevenuto F. 2019. Supervised learning for fake news detection. *IEEE Intelligent Systems* **34(2)**:76–81 DOI 10.1109/MIS.2019.2899143.

Robertson CE, Pröllochs N, Schwarzenegger K, Pärnamets P, Van Bavel JJ, Feuerriegel S. 2023. Negativity drives online news consumption. *Nature Human Behaviour* **7**:812–822 DOI 10.1038/s41562-023-01538-4.

Rodríguez-Ferrándiz R. 2023. An overview of the fake news phenomenon: from untruth-driven to post-truth-driven approaches. *Media and Communication* **11(2)**:15–29.

Rohera D, Shethna H, Patel K, Thakker U, Tanwar S, Gupta R, Hong W-C, Sharma R. 2022. A taxonomy of fake news classification techniques: survey and implementation aspects. *IEEE Access* **10**:30367–30394 DOI 10.1109/ACCESS.2022.3159651.

Ruffo G, Semeraro A, Giachanou A, Rosso P. 2023. Studying fake news spreading, polarisation dynamics, and manipulation by bots: a tale of networks and language. *Computer Science Review* **47**:100531 DOI 10.1016/j.cosrev.2022.100531.

Sagi O, Rokach L. 2018. Ensemble learning: a survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **8(4)**:e1249.

Saikh T, De A, Ekbal A, Bhattacharyya P. 2020. A deep learning approach for automatic detection of fake news. ArXiv preprint. arXiv:2005.04938.

Salahuddin R, Wasim M. 2022. Automatic identification of Urdu fake news using logistic regression model. In: *2022 16th international conference on open source systems and technologies (ICOSST)*. Piscataway: IEEE, 1–6.

Scheibenzuber C, Neagu L-M, Ruseti S, Artmann B, Bartsch C, Kubik M, Dascalu M, Trausan-Matu S, Nistor N. 2023. Dialog in the echo chamber: fake news framing predicts emotion, argumentation and dialogic social knowledge building in subsequent online discussions. *Computers in Human Behavior* **140**:107587 DOI 10.1016/j.chb.2022.107587.

Seddari N, Derhab A, Belaoued M, Halboob W, Al-Muhtadi J, Bouras A. 2022. A hybrid linguistic and knowledge-based analysis approach for fake news detection on social media. *IEEE Access* **10**:62097–62109 DOI 10.1109/ACCESS.2022.3181184.

Shu K, Sliva A, Wang S, Tang J, Liu H. 2017. Fake news detection on social media: a data mining perspective. *ACM SIGKDD Explorations Newsletter* **19(1)**:22–36 DOI 10.1145/3137597.3137600.

Singh G, Selva K. 2023a. A comparative study of hybrid machine learning approaches for fake news detection that combine multi-stage ensemble learning and NLP-based framework. *TechRxiv* DOI 10.36227/techrxiv.21856671.v3.

Singh G, Selva K. 2023b. Detection of fake news using NLP and various single and ensemble learning classifiers. *TechRxiv* DOI 10.36227/techrxiv.21856671.v1.

Sivasankari S, Vadivu G. 2022. Tracing the fake news propagation path using social network analysis. *Soft Computing* **26**:12883–12891 DOI 10.1007/s00500-021-06043-2.

Sullivan GM, Feinn R. 2012. Using effect size—or why the P value is not enough. *Journal of Graduate Medical Education* **4(3)**:279–282 DOI 10.4300/JGME-D-12-00156.1.

Ullah F, Chen X, Shah SBH, Mahfoudh S, Hassan MA, Saeed N. 2022. A novel approach for emotion detection and sentiment analysis for low resource Urdu language based on CNN-LSTM. *Electronics* **11(24)**:4096 DOI 10.3390/electronics11244096.

Uysal AK, Gunal S. 2012. A novel probabilistic feature selection method for text classification. *Knowledge-Based Systems* **36**:226–235 DOI 10.1016/j.knosys.2012.06.005.

Vogel I, Meghana M. 2020. Fake news spreader detection on Twitter using character N-grams. In: *CLEF (Working Notes)*.

Vuong T, Saastamoinen M, Jacucci G, Ruotsalo T. 2019. Understanding user behavior in naturalistic information search tasks. *Journal of the Association for Information Science and Technology* **70(11)**:1248–1261 DOI 10.1002/asi.24201.

Wasim M. 2023. dr-m-wasim/UrduFakeNewsFS. *Available at* *https://github.com/dr-m-wasim/UrduFakeNewsFS* (accessed on 07 June 2023).

Xing Y, Wang X, Qiu C, Li Y, He W. 2022. Research on opinion polarization by big data analytics capabilities in online social networks. *Technology in Society* **68**:101902 DOI 10.1016/j.techsoc.2022.101902.