

NLP-Based Approach to Multilingual Fake News Detection through Social Media in Low-Resource Languages: A Review

Sibgha Munir^{1*}, Haris Munir²

¹ Department of Artificial Intelligence, University of Engineering and Technology (UET), Pakistan.

² Department of Computer Science, COMSATS University Islamabad, Sahiwal Campus, Pakistan.

* Corresponding author. Email: 2024msaie14@student.uet.edu.pk (S.M.);
fa22-bcs-069@students.cuisahawal.edu.pk (H.M.)

Manuscript submitted October 17, 2025; accepted January 30, 2026; published April 28, 2026.

doi: 10.18178/JAAI.2026.4.2.76-93

Abstract: Especially in multilingual and impoverished language situations, the quick proliferation of false news on social media jeopardizes social stability. Emphasizing difficulties specific to low-resource environments, this paper offers a thorough examination of current Natural Language Processing (NLP) techniques for fake news detection across several languages. It examines prominent techniques, including named entity recognition, sentiment analysis, and text categorization, highlighting their uses, advantages, and drawbacks. Particularly focused on advanced methods like transfer learning, multilingual embeddings, and cross-lingual models all of which attempt to get around the scarcity of labeled data and the complexity of linguistic variety. The article also draws attention to deficiencies in present techniques and stresses the need for flexible models able to handle developing disinformation problems. The research provides ideas to help in the creation of strong, inclusive, and efficient tools for reducing the world-wide spread of false information by combining present progress with gaps.

Keywords: fake news detection, low-resource languages, multilingual Natural Language Processing (NLP), social media misinformation

1. Introduction

The ever-changing trend in sharing information worldwide that has resulted from the development of many social media platforms has changed into a potent activity for wild propagation. Equally, in the other light, this fast development has been marked by the wild spread of misinformation and fake news, which has dramatically become a significant challenge to societal well-being, democracy, and public safety [1]. Also, detecting and mitigating fake news have become a vital area of research in the multilingual social media setting. If significant progress has been achieved today for other major languages, such as English, the current situation is significantly different for low-resource languages where linguistic resources and advanced natural language processing tools are not expected [2].

Hundreds of millions worldwide speak languages with a relative lack of large, annotated datasets, fewer linguistic tools, and sparse research focus. This, in turn, leads to a gap between high-resource languages about fake news detection and fighting [3, 4]. The development of robust, scalable, and adaptable NLP-based approaches for the complexities of multilingual fake news detection is overnight in view, particularly for low-resource languages, as social media is used increasingly across different linguistic

groups.

The challenge that lies ahead is appropriately put into perspective by the following statistics in Table 1 on global language distribution and social media usage:

Table 1. Global Language Distribution and Social Media Statistics Applicable to the Detection of Fake News

| Statistics | Value |
|--|-----------------------------|
| Number of languages spoken worldwide | Approximately 7,000 |
| Percentage of languages classified as low-resource | Over 90% |
| Number of people using social media globally | 4.7 billion (as of 2023) |
| Percentage of social media users in low-resource language regions | 25% |
| Increase in fake news incidents on social media (2019–2023) | 78% increase |
| Languages with substantial NLP resources | Fewer than 100 |
| Number of research papers on fake news detection in low-resource languages (2018–2023) | Less than 5% of all studies |

The techniques of NLP have been very promising in application domains, ranging from sentiment analysis to machine translation, and hold immense promise for detecting fake news in several languages [5, 6]. However, developing such systems in low-resource languages is not without challenges. Those challenges include the rarity of big, annotated corpora, the complexity of linguistic diversity, and the requirement for models to be cross-linguistically adaptable. The handling of such topics calls for new methods based on recent advances in transfer learning, multilingual embeddings, and the exploitation of domain-specific knowledge.

This paper aims to fully review state-of-the-art NLP-based multilingual fake news detection methodologies, mainly focusing on low-resource languages [7]. We will present a few of the significant challenges and existing approaches in the area and study the potential of new emerging technologies in closing the gap across fake news detection on linguistic grounds. Finally, how social media acts as a facilitator and a battleground for fake news underlines the need to develop culturally and contextually aware detection systems.

It synthesizes present knowledge in this rapidly evolving field but also points out the gaps and opportunities for further research. The authors intend to contribute to developing more inclusive and practical tools to fight fake news, hence ensuring all linguistic communities can benefit from these protections as advanced NLP techniques have to offer [8].

Section 2 presents the review methodology of Natural Language Processing (NLP)-based approaches in multilingual fake news detection. Section 3 elaborates on the key NLP techniques, Text Classification, Sentiment Analysis, and Named Entity Recognition, that are basic in identifying fake news. Section 4 discusses state-of-the-art approaches designed explicitly for low-resource languages, examining the roles of Transfer Learning, Multilingual Embeddings, and Cross-linguistic Adaptation. Section 5: Critical analysis of the existing models with the presentation of strengths, limitations, and areas of potential improvements. Section 6 underlines the main challenges in multilingual fake news detection, outlining what the authors call a Scarcity of Annotated Data, Linguistic Diversity, and Model Adaptability problem. Section 7 outlines the future directions in the field. Section 8 concludes the review by outlining the main findings and implications for future research. The depiction of the topology of the review is shown in Fig. 1.



Fig. 1. Organization of the review.

2. Literature Review

Alnabhan *et al.* [9] conducted a systematic review on fake news detection using deep learning. After screening 1642 articles, 393 papers were analyzed, with 178 selected for detailed review. Commonly used models like CNN and Bi(LSTM) showed high accuracy, with 47% using BERT and hybrid models for practical purposes. Class imbalance was highlighted as a challenge, impacting detection accuracy. Transfer learning was recommended to enhance performance. The study identifies areas for future research, emphasizing the importance of advanced deep learning techniques and addressing class imbalance in fake news detection systems.

Alghamdi *et al.* [10] analyzed a framework's ability to model inter-domain relationships and intra-domain relationships within the same language. The study examined training and testing in the same language, as well as the transfer of domain-independent knowledge across different domains. The hybrid extractive-abstractive summarization method showed effectiveness in detecting fake news across languages and domains, improving model flexibility and accuracy in classification.

Fagundes *et al.* [11] conducted a study reviewing 20 papers to address specific research questions. Most excluded articles focused on syntax without delving into primary research. Shallow syntactic representations like parts of speech offer minimal improvements in fake news detection, while deeper representations like context-free grammar show promise. Combining syntactic information with other techniques enhances detection accuracy. Limitations include search engine precision and publication scope up to October 2021. This review highlights the value of syntactic information in enhancing automated fake news detection.

Krasadakis *et al.* [12] introduced a paper discussing challenges and developments in NLP within the legal domain. The focus is on tasks such as NER, EL, RelEx, and Coref, important for creating a legal citation network and law consolidation system. The study highlights the complexities of legal texts and the challenges in replicating results due to lack of standardization. There is a call for more research in multilingual NLP, particularly for low-resource languages, and the need for benchmarks like LEXTREME to measure progress.

Meesad *et al.* [13] developed a framework to identify fake news in Thai, with three main modules: Information Retrieval, Natural Language Processing, and Machine Learning. They collected data from Thai news websites using web crawlers, resulting in 41,448 samples. Various machine learning models were tested, with the Long Short-Term Memory model proving most effective. The study also highlighted challenges in Thai text processing, such as word segmentation issues.

Bahmanyar *et al.* [14] introduced a transformer-based deep learning pipeline for detecting fake news in multiple languages. By fine-tuning the model on a multilingual dataset with 16 languages, they achieved an accuracy of 97.38% and an F1-Score of 0.9734. Zero-shot testing demonstrated the model's ability for cross-lingual transfer and understanding, even for languages not seen in the training data. The study also developed a filtering model with 99.8% accuracy to remove impractical data during dataset enlargement, ensuring high-quality data for detecting fake news across various languages.

Schmidt *et al.* [15] conducted a study using the Snopes dataset to tune LLaMA and baseline models, with the best results seen with the LLaMA-2 model with 70 billion parameters. Performance differences were significant between SVM models relying on stylistic features and LLaMA-2 models leveraging pretrained knowledge. Experimentation on the FCTR1000 dataset showed high F1-macro scores of 0.89 and 0.828 from the LLaMA-13B model when both claim statements and summaries were input. However, LLaMA models performed poorly in Turkish, particularly when only claim statements were used, indicating the need for additional context for improved results.

Abboud *et al.* [16] developed a fake news detection system for low-resource languages using a knowledge base from high-resource languages. Their system overcomes the challenge of limited training data, with a significant accuracy improvement of 3.71% compared to baseline. By incorporating a cross-lingual source information feature, accuracy increased by an additional 3.03%, allowing for automated credibility checks on articles. This approach can be applied to various low-resource languages, making it a versatile solution for fake news detection.

Masis *et al.* [17] conducted a study using a dataset of 19,000 entries, splitting it 80/20 for training and testing. Logistic regression was employed, with results showing that news titles alone were sufficient for training the model effectively. The study highlighted the importance of news titles in identifying fake news, as demonstrated by the model's performance in classifying articles. However, the model's performance varied when tested on articles from different sources, dates, or agendas, indicating a need for fine-tuning or algorithm adjustments. The authors recommended using a hybrid dataset to improve performance and minimize data drift in machine learning models.

Ziyaden *et al.* [18] found that text augmentation methods, especially through translation approaches like Google Translate API, significantly improved model performance, with an accuracy rate of 87%. The baseline approach failed for certain classes, prompting the introduction of augmentation techniques that greatly increased F1-Scores by at least one-third for augmented classes. Comparing back-translation and neural network-based translations, Google Translate proved most effective with an accuracy of 0.87. Augmentation techniques were deemed crucial for addressing class imbalances and enhancing model generalization for low-resource languages, providing valuable gains for text classification tasks. The authors also made their Azerbaijani language dataset and pre-trained Roberta model publicly available for future NLP research.

3. Methodology

The steps adapted to perform this systematic review are as follows:

3.1. Collection of Articles

A structured search was performed across major databases on aspects of NLP-based [19] approaches to multilingual fake news detection [20, 21], focusing on low-resource languages. This search was conducted in June 2025 using the following databases: Google Scholar, Scopus, and Web of Science. These keywords were, therefore, targeted to capture relevant studies: fake news detection, multilingual NLP [22], low-resource languages [23], transfer learning, multilingual embeddings, cross-linguistic adaptation, and social media. This search returned 987 peer-reviewed articles. Further screening excluded those studies

which had a focus on detecting fake news in high-resource languages. Further, only the articles after the year 2014 were considered for this review since the development in NLP techniques for fake news detection happened at a very fast pace in the last decade.

3.2. Search Strategy

Clear inclusion and exclusion criteria were set to assess the quality of the selected literature. These criteria are crucial in ensuring the reliability of findings and consistency in the review process. The evaluation will consider the title of the article, the abstract, and the full text. The inclusion criteria that have been considered for the article review include the following:

- Only studies published between 2015 and 2025 were included to reflect recent research trends.
- It includes various research works on the methods of NLP-based multilingual fake news detection [24] focusing on low-resource languages [25, 26].
- More preferences were given to those papers which discussed technical aspects-feature extraction, model architecture, preprocessing of data, and transfer learning techniques for cross-linguistic adaptation.
- Those papers were more preferred which reported the results based on performance metrics like accuracy, precision, recall, F1-score, and AUC.
- Only peer-reviewed journal articles and conference papers were included to maintain the integrity of the review. In this context, editorials, posters, and technical reports were not considered to be peer-reviewed sources.

Each study was critically evaluated for inclusion, and only those studies that met the quality criteria were included in the final review. A schematic overview of the general plan for this review is summarized in Fig. 2 (hypothetical).

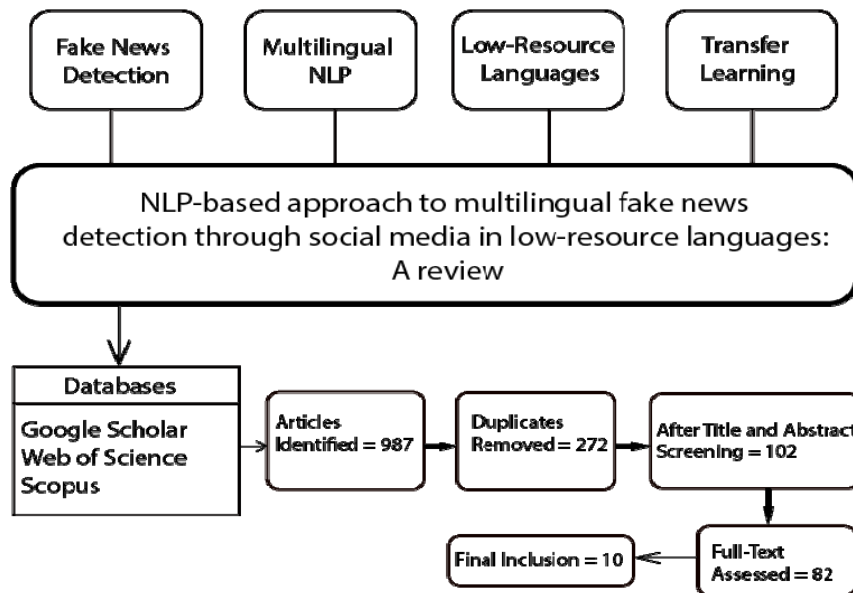


Fig. 2. Strategy for articles screening and inclusion.

4. Research Findings

The results of the research in this review are presented below:

4.1. NLP Techniques for Fake News Detection

4.1.1. Text classification

One of the most basic NLP techniques in fake news detection is text classification [27, 28], where the

given text is classified into pre-defined classes like “fake” or “real”. This is vital in filtering and identifying misleading information across various languages. In low-resource languages, there is an acute scarcity of annotated data. This problem has, however, been circumvented by techniques like transfer learning [27] and data augmentation [28], as well as the utilization of synthetic data generation among others. For example, pre-trained models in high-resource languages may enhance their performance on the classification task by fine-tuning them on small datasets in low-resource languages.

Other robust features are multilingual embeddings, which project words from different languages on a shared vector space. As the model can make better generalizations, this improves text classification across languages even when trained on small amounts of data from low-resource languages [29]. While there have been considerable gains in handling code-switching or dialectal variations quite common on social media there are still challenges to overcome.

4.1.2. Sentiment analysis

Among the most critical techniques employed is sentiment analysis, which infers the feeling or emotion the piece of text conveys regarding the intent behind the information [30, 31]. For instance, most fake news stories are overcooked or couched in emotionally loaded language to worm their way into readers’ sensibilities. In low-resource languages, sentiment analysis is a significant challenge due to the lack of annotated datasets and the complexity of their linguistic structure. The intricacy of handling nuanced language features such as puns and wordplay adds another layer of difficulty, as these elements can dramatically alter the sentiment conveyed.

Recent progress in cross-lingual transfer in sentiment analysis uses models trained in high-resource languages to classify sentiment in low-resource languages directly. This might be based on a method like zero-shot learning, wherein a model trained in one language can make inferences regarding another untrained language. However, the results are not as promising when low-resource languages are considered, as accuracy is generally lower than in high-resource languages, which usually indicate the state of the art and thus warrant actual further advances in research and development [32, 33].

4.1.3. Named Entity Recognition (NER)

It identifies and classifies the entities mentioned in the text, such as people, organizations, locations, and so on [34, 35]. Fake news detection would leverage NER to identify from the text key players and locations, which it would then cross-check against fact-based databases for authenticity. Nevertheless, NER [36, 37] for low-resource languages is quite difficult because no annotated corpora exist and language-specific entities are pretty complicated.

Recent approaches have relied on transfer learning from high-resource languages to improve the performance of NER in low-resource conditions. Similarly, multilingual models of NER [38] trained on texts from multiple languages have proved quite effective for identifying entities from low-resource languages. They, however, usually lose their way on language-specific nuances and cultural references that creep in, hence leading to inaccuracies [39].

4.2. State-of-the-Art Approaches in Low-Resource Languages

4.2.1. Transfer learning

Recently, transfer learning has been very successful in NLP. In most cases, especially for low-resource languages, this has been the case. Essentially, transfer learning applies knowledge from high-resource-language or domain-trained models to help in the performance on tasks in the low-resource setting [40]. This typically means pre-training a model on a large corpus from a high-resource language like English and fine-tuning it on a smaller dataset from the target low-resource language.

Transfer learning has shown reasonable promise in the domain of fake news detection. Pretraining on huge amounts of data enables models to learn linguistic patterns, syntactic structures, and semantic

nuances common across languages [41, 42]. These models have been demonstrated to detect fake news with much better performance than from-scratch models, after fine-tuning with targeted data in a low-resource language. Besides, transfer learning removes the requirement of large annotated datasets, which are very hard to come by in low-resource languages; hence, the solution is pretty workable for multilingual fake news detection.

b) Multilingual embeddings

Another important component dealing with low-resource languages challenges is multilingual embeddings. Word representations learn semantic relationships across languages. Multilingual embedding spaces map words from different languages into a common vector space and let models make use of cross-lingual information, enabling transfer from high- to low-resource languages [43].

In fake news detection, multilingual embeddings enable the integration of data from several languages to be utilized with greater generalization across linguistic boundaries. This can become very helpful in multilingual societies where fake news is propagated in many languages at once. In this way, a model could handle fake news in low-resource languages using the rich information available from high-resource languages and improve the system performance and strength for fake news detection in general [44].

c) Cross-linguistic adaptation

Cross-linguistic adaptation is the method for adapting or tuning existing NLP models [45, 46] for good performance over a different portion of language usually those for low-resource languages. Based on linguistic similarity, either in base roots, cases of grammar, or cognates, among others, then such models trained over one language will have great performance over another. When detecting fake news, the linguistic subtleties of misinformation may vary from one language to the next but will still share some core features.

The first proposes that training should be done on a pivot language near the target low-resourced language. The model can thereafter be fine-tuned [47] to the target language with minimal additional effort to ensure that the approach is cost-effective in fake news detection. Also, other techniques, such as zero-shot learning and a few-shot, which allow generalization from very few examples, can be applied to enhance cross-linguistic adaptation further. Thereby, the model would be improved with respect to fake news detection in low-resource languages.

The flowchart diagram of approaches in Low-Resource Languages is shown in Fig. 3.

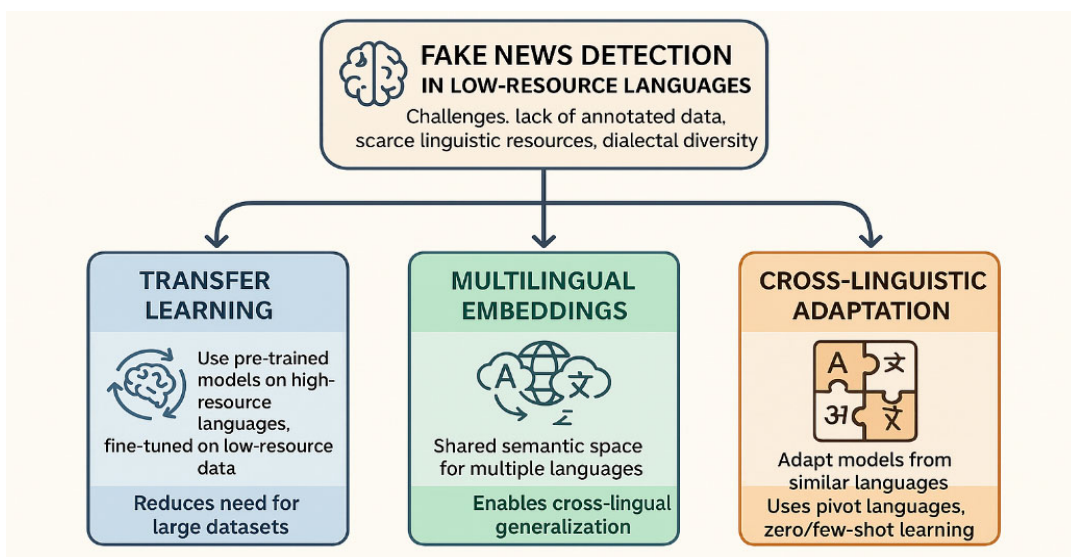


Fig. 3. Advanced NLP approaches for low-resource languages in fake news detection.

4.3. Critical Analysis of Existing Models

4.3.1. Strengths

4.3.1.1. Cross-linguistic generalization

One of the major strengths of current NLP models for fake news detection in low-resource languages rests in their ability to generalize across languages. Transfer learning and multilingual embeddings equip such models to borrow knowledge from high-resource languages and apply it in a low-resource setting, hence improving performance. Cross-linguistic transfer has contributed decisively to the enrichment of semantic knowledge and has thus provided better accuracy of detection in target languages, even when limited training data is available [48, 49]. AI-driven techniques will further enhance the ability of capturing complex linguistic patterns, thus helping cross-linguistic generalization.

4.3.1.2. Scalability and flexibility

Most contemporary fake news detection models are both scalable and flexible. As such, they are easily adaptable across many languages and social media platforms. Modular architectures and resorting to pre-trained language models like BERT [50, 51] and XLM-R allow fine-tuning in language-specific and domain-specific scenarios, hence applicability in very diversified linguistic settings without the need for extensive retraining from scratch. The integration of AI enables these models to learn and adapt continuously, making them scalable to new languages and platforms with minimal intervention.

4.3.1.3. Improved textual feature extraction

The integration of advanced NLP techniques such as sentiment analysis, named entity recognition (NER), and contextualized embeddings has significantly enhanced the capability of models to extract relevant textual features from multilingual data. These features are crucial for distinguishing between fake and real news, especially in languages where traditional lexical resources are scarce or non-existent [52], [53]. AI enhances this by making the process deeper and more subtle, which allows for capturing subtleties that would be lost through traditional methods. XAI, in turn, enhances that with insight into what features were most influential to the model's decision-making process, offering a level of transparency and thus trustworthiness.

4.3.1.4. Enhanced model robustness

A great number of strategies to improve the robustness against adversarial attacks and noise in social media data are already included in existing models. Those models are more and more powerful in maintaining accuracy across different languages and dialects, due to focusing on feature stability and resilience against linguistic variations [54, 55]. AI plays a critical role in enhancing this robustness, as it allows models to learn from diverse and noisy data environments. XAI contributes by making the decision-making process more interpretable, helping identify and mitigate potential vulnerabilities in the model, thereby improving overall robustness and reliability.

4.3.2. Limitations

4.3.2.1. Data scarcity and quality

Even after all the developments, high-quality annotated data remains one of the most important limitations to NLP models in low-resource languages [56, 57]. In this respect, a lack of comprehensive datasets hinders not only model training but also generalizability and reliability with regard to fake news detection across different linguistic contexts. This is much more pressing when dealing with underrepresented languages, wherein not even the basic linguistic resources may be available.

4.3.2.2. Limited cultural context understanding

Current models often miss an understanding of the cultural and contextual nuances underlying such fake news content. This becomes even more critical in multilingual and low-resource settings where cultural references, idiomatic expressions [58, 59], and local knowledge are crucial in correctly detecting fake news.

Their inability to understand such subtleties properly might hence result in misclassification and loss in effectiveness. Moreover, cross-culture challenges arise when models trained on one cultural context are applied to another, often failing to capture the unique socio-cultural dynamics that influence how fake news is spread and perceived.

4.3.2.3. Inadequate adaptation to low-resource languages

While transfer learning and multilingual embeddings have enhanced model performance, this is not without limitations. For instance, adaptation may still be out of reach for truly low-resource languages, in which knowledge transfer from high-resource languages still proves too inadequate to capture the richness of the language [60]. This often results in suboptimal performance and thereby puts a pointer to the fact that more tailored approaches are thus needed. Additionally, uneven economic development across regions exacerbates these issues, as countries with limited access to technological resources are unable to deploy and benefit from advanced NLP models, leading to a digital divide in fake news detection capabilities.

4.3.2.4. High computational requirements

However, one drawback of using state-of-the-art models with sophisticated architectures is their high computational costs, making them less deployable in resource-constrained environments. This particularly falls under relevant conditions in the low-resource language setting, where low access to state-of-the-art computational infrastructure clearly hinders the practical application of those models [61]. Political turmoil in certain regions can further complicate the deployment of these models, as instability may limit the availability of resources and the ability to maintain consistent data collection efforts, further hampering the effectiveness of fake news detection initiatives.

4.3.3. Opportunities for improvement

4.3.3.1. Development of multilingual datasets

The most promising line of improvement lies in the creation of large, high-quality multilingual datasets for fake news detection. Because of where models are at the moment both performing and being evaluated badly the extension of the availability of annotated data to a greater number of languages will help researchers better train and evaluate their models [62].

4.3.3.2. Incorporation of cultural and contextual knowledge

Increasing the models with cultural and contextual understanding can improve their accuracy in fake news detection [63]. For instance, it can be done by incorporating some forms of external knowledge sources, such as cultural ontologies or archives of local news, within the training process for a model. On the other hand, the creation of models able to learn dynamically from context-specific cues at inference time will further increase their adaptability.

4.3.3.3. Optimization for low-resource environments

In view of computational constraints, it is future research that shall orient NLP models to become more suitable for low-resource environments. This can include more efficient algorithms, reducing model sizes without losing too much in performance, or new methods for training that require less of the same resource. These efforts will make advanced techniques in NLP available to a wider audience and more applicable in low-resource scenarios [64].

The summary of a literature review is shown in Table 2.

Table 2. Summary of Literature Review

| Author, Year | Target Variable | Input | Architecture | Pre-Processing | Dataset | Outcome | Output Result |
|---------------------------------|---|--------------------|---|--|--------------------------------|----------------------------------|--|
| Alnabhan <i>et al.</i> , 2023 | Fake News Detection | Social Media Text | CNN, Bi(LSTM), BERT, Hybrid Models | Data Augmentation, Class Balancing | 178 articles | Fake News Detection | High Accuracy with Class Imbalance Issues Addressed |
| Alghamdi <i>et al.</i> , 2024 | Fake News Detection | Social Media Text | Hybrid Extractive-Abstractive Summarization | Domain Transfer, Data Reduction | Various Domains | Cross-Domain Transfer | Improved Model Versatility |
| Fagundes <i>et al.</i> , 2024 | Fake News Detection | Social Media Text | Deep Syntactic Representations | Feature Combination | 20 papers | Syntactic Information Usage | Enhanced Detection with Deep Syntactic Features |
| Krasadakis <i>et al.</i> , 2024 | Fake News Detection | Legal Documents | NER, EL, RelEx, Coref | Custom Dataset Challenges | Legal Texts | Legal Citation Management | Improved Multilingual NLP Applications |
| Meesad <i>et al.</i> , 2021 | Fake News Detection in Thai | Thai News Articles | Naïve Bayes, Logistic Regression, LSTM | Word Segmentation | 41,448 samples | Fake News Detection | LSTM Model Most Accurate, Addressed Thai Language Challenges |
| Bahmanyar <i>et al.</i> , 2024 | Fake News Detection in Multiple Languages | Multilingual Text | Transformer-based Models | Data Filtering, Zero-Shot Learning | 16 Languages | Multilingual Fake News Detection | 97.38% Accuracy, Cross-Lingual Transfer |
| Schmidt <i>et al.</i> , 2024 | Fake News Detection | Statement Text | LLaMA Models (Various Sizes) | Stylistic Feature Tuning | Snopes and FCTR1000 Datasets | Fake News Detection | High Accuracy with Contextual Improvement |
| Abboud <i>et al.</i> , 2022 | Fake News Detection in Low-Resource Languages | Multilingual Text | Cross-Lingual Transfer Learning | Machine-Translated Data | Various Low-Resource Languages | Low-Resource Fake News Detection | Improved Accuracy with Cross-Lingual Features |
| Masis <i>et al.</i> , 2024 | Fake News Detection | News Titles | Logistic Regression | Title-Based Feature Extraction | 19,000 entries | Fake News Detection | Effective with Title-Only Input |
| Ziyaden <i>et al.</i> , 2024 | Fake News Detection | Social Media Text | Text Augmentation Techniques | Back-Translation, Neural Network-Based | Various Languages | Enhanced Model Performance | Improved Accuracy with Augmentation Strategies |

5. Challenges

5.1. Scarcity of Annotated Data

Multilingual fake news detection using multilingual approaches is a big pain in the neck because of data scarcity with manual annotation [65]. High-quality labeled datasets are the backbone for training and validating NLP models; however, in big parts of the world, they remain limited or nonexistent for languages with low resources. The process of coming up with manual annotations is time-consuming and costly, and most importantly, for effective machine learning, it is difficult to achieve huge datasets [66]. Such an absence of annotated data leads to requirements of either transfer learning or unsupervised methods, causing generally low attainment of desired accuracies in low-resource settings [67]. This also prevents robust models from being developed.

5.2. Linguistic Diversity

Another huge challenge to multilingual fake news detection is linguistic diversity. Most of the low-resource languages indicate high variation at the dialect, orthographic, and syntax levels, making the development of NLP models really challenging [68]. Another way of looking at this is that diversity implies a model trained on one dialect or variant of a language may not work well on another. Thirdly, many low-resource languages are mainly oral, and hence there is little in terms of written forms; this therefore presents a challenge to the development of text-based models in NLP. With the languages varying both intra- and interlingually to handle, high flexibility is required from the models when considering general phenomena covered by a model against the limited resources available [69].

5.3. Model Adaptability

This has been a critical challenge in fake news detection: the adaptability of NLP models to low-resource languages. The effectiveness of most state-of-the-art models developed and fine-tuned for high-resource languages like English is quite limited when applied to low-resource languages [70]. This is because it is hard to generalize across the models due to the linguistic differences and scarcity of training data in the target language [71]. Moreover, models trained on high-resource languages often exhibit degraded performance when transferred to low-resource languages. This can be improved by techniques that make the models more effective learners of limited data, as well as by architectures that accommodate features unique to low-resource languages.

Table 3 highlights the major challenges in multilingual fake news detection for low-resource languages, including data scarcity, linguistic diversity, and limited model adaptability, all of which negatively impact system performance. Limited annotated data and high dialectal variation reduce model accuracy and generalization, while difficulties in transferring models from high-resource languages emphasize the need for adaptable architectures and multilingual learning strategies.

Table 3. Key Challenges in Multilingual Fake News Detection in Low-Resource Languages

| Challenge | Description | Implications | Potential Approaches |
|----------------------------|---|---|--|
| Scarcity of Annotated Data | Limited availability of manually labeled datasets in low-resource languages due to high costs and time required for annotation. | Hinders model training and validation; leads to poor performance and low accuracy. | Use of transfer learning, data augmentation, and unsupervised learning methods. |
| Linguistic Diversity | High variation in dialects, orthography, syntax, and presence of primarily oral languages. | Models trained on one dialect may not generalize to others; limited written data hinders the development of these models. | Develop dialect-specific resources; leverage multilingual embeddings and speech-to-text technologies. |
| Model Adaptability | Difficulty in adapting NLP models trained on high-resource languages to low-resource contexts. | Reduced performance and generalization; failure to capture unique linguistic features. | Design lightweight, adaptable architectures; train models on multilingual datasets; integrate meta-learning. |

6. Future Directions

Inevitably, the future of multilingual fake news detection lies in low-resource languages. As the digital world deepens, the urge for sophisticated and accurate tools to fight misinformation will heighten across diverse linguistic contexts [72]. Certainly, one of the promising avenues of future research lies in developing more advanced models of NLPs that can be fine-tuned seamlessly to the nuances of low-resource languages. This will require similar efforts in the creation of new datasets and in augmenting models with techniques such as transfer learning and multilingual embeddings.

Another important line of research is to make techniques in transfer learning more effective for enabling models fit for use in high-resource languages to be adapted for use in low-resource languages. This could be in the form of more efficient algorithms that transfer knowledge at minimal loss of accuracy or the

development of intermediate languages that bridge the gap between high- and low-resource languages. Otherwise, unsupervised learning methods would reduce dependency on large annotated datasets often unavailable for low-resource languages [73].

Other key directions would be to include techniques for cross-linguistic adaptation that would enable models to learn from and generalize in one go across multiple languages. This may involve the development of totally new architectures that would capture language similarities and differences better or, conversely, meta-learning strategies whereby models learn how to adapt rapidly to new languages with minimum retraining [74]. This way, enhancing the models' ability to generalize across languages, researchers will then establish much more robust systems capable of detecting fake news in most linguistic and cultural contexts.

Future work must also take into socio-cultural considerations regarding the detection of fake news in low-resource languages. This is generally done by creating more contextually sensitive models that consider, first of all, the regional cultural and political environments in which they must work. By doing so, it can make sure that these tools designed by researchers are effective and culturally sensitive [75]. This collaboration between academia, industry, and government would play an important role in the future of shaping multilingual fake news detection. Such collaborative efforts are considered to enable the sharing of resources, data, and expertise that are necessary to address the specific challenges of low-resource languages [76]. At the same time, international cooperation can be very instrumentally important for setting global standards and best practices on fake news detection to ensure that benefits from these technologies are realized on a global scale.

More important in the future will be the rise of user-centric approaches to fake news detection, where end-users take an active role in identification and reporting [77]. It may involve the development of new interfaces and tools that empower people to be part of the detection process or even the creation of community-driven platforms using collective intelligence naïve ways to stand against fake news on a regular basis [78]. By more directly including users in the process, researchers can create more effective and scalable solutions to the problem of misinformation.

In other words, the technical innovation on the method of detecting multilingual fake news, considering ethical aspects, the socio-culture awareness, and collaborative effort, charts the future directions of this field of low-resource languages. These will be extremely important efforts at the core as researchers continue to push boundaries on what is possible to make sure the digital information space remains reliable and trustworthy to all, no matter their linguistic background.

7. Conclusion

Using a specific emphasis on low-resource languages, this study offered a methodical analysis of NLP-based methods for multilingual bogus news detection on social media. The review analyzed major approaches, datasets, and modeling techniques, especially the application of transfer learning, multilingual embeddings, and cross-lingual models to reduce data paucity and language diversity, by integrating recent literature. The study indicated that while transfer learning has dramatically increased the portability of models from high-resource to low-resource languages, performance gaps remain owing to small annotated datasets, dialectal variety, and insufficient model adaptability.

A major conclusion of this study is the extreme scarcity of publicly available, high-quality multilingual datasets for fake news detection in low-resource contexts, therefore restricting model accuracy and generalizability. Furthermore, current study reveals an overemphasis on a limited number of languages and NLP tools, therefore under investigating several languages, dialects, and sociocultural contexts. Furthermore, few studies use a holistic approach combining several NLP activities including text classification, sentiment analysis, and named entity recognition, despite their ability to increase detection

robustness.

Future research ought to focus mostly on the development of open multilingual datasets as well as more adaptable and light model structures capable of handling linguistic diversity. Dependency on annotated data is decreased by more cross-linguistic adaptation, meta-learning, unsupervised learning, and user-centered detection methods. Adding socio-cultural context and encouraging cross-sector cooperation between academia, business, and legislators will also help to ensure ethical, scalable, and internationally effective fake news detection systems.

Overcoming these technical, grammatical, and contextual barriers will be entirely necessary for the advancement of multilingual false news detection and the preservation of digital information ecosystems throughout several linguistic groups.

Conflict of Interest

All authors declare that they have no conflicts of interest.

Author Contributions

Sibgha Munir developed the research, created the review approach, carried out the literature search and screening, critically examined the chosen papers, and prepared the original manuscript. Data organization, supported synthesis of results, helped to improve the technique and discussion parts, and evaluated the paper for technical correctness and clarity all contributed by Haris Munir. Both writers helped in modifying the manuscript; accepted the final version; and consent to be responsible for all parts of the work.

Acknowledgment

The authors want to recognize the academic assistance offered by the Department of Artificial Intelligence at the University of Engineering and Technology (UET) in Pakistan as well as the Department of Computer Science at COMSATS University Islamabad, Sahiwal Campus. First author appreciates faculty members and colleagues for their helpful comments during the creation of this review. This project received no external money.

References

- [1] Dhiman, B. (2023). The rise and impact of misinformation and fake news on digital youth: A critical review. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4438362>
- [2] De, A., Bandyopadhyay, D., Gain, B., & Ekbal, A. (2022). A transformer-based approach to multilingual fake news detection in low-resource languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 21(1), 1. <https://doi.org/10.1145/3472619>
- [3] Jand, H. (2018). Natural language processing with speech recognition. *International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)*.
- [4] Bhardwaj, A., Khanna, P., Kumar, S., & Pragma. (2020). Generative model for NLP applications based on component extraction. In *Procedia Computer Science*. <https://doi.org/10.1016/j.procs.2020.03.391>
- [5] Tsai, C. M. (2023). Stylometric fake news detection based on natural language processing using named entity recognition: In-domain and cross-domain analysis. *Electronics*, 12(17), 3676. <https://doi.org/10.3390/electronics12173676>
- [6] Mohawesh, R., Liu, X., Arini, H. M., Wu, Y., & Yin, H. (2023). Semantic graph based topic modelling framework for multilingual fake news detection. *AI Open*, 4. <https://doi.org/10.1016/j.aiopen.2023.08.004>
- [7] Zhang, Y., et al. (2023). Stance-level sarcasm detection with BERT and stance-centered graph attention

- networks. *ACM Transactions on Internet Technology*, 23(2), 20. <https://doi.org/10.1145/3533430>
- [8] Sharifani, K., Amini, M., Akbari, Y., & Godarzi, J. A. (2022). Operating machine learning across natural language processing techniques for improvement of fabricated news model. *International Journal of Science and Information System Research*, 12(9).
- [9] Alnabhan, M. Q., Branco, P., & Alnabhan, M. (2023). Fake news detection using deep learning: A systematic literature review. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2023>
- [10] Alghamdi, J., Lin, Y., & Luo, S. (2024). Fake news detection in low-resource languages: A novel hybrid summarization approach. *Knowledge-Based Systems*, 296, 111884. <https://doi.org/10.1016/j.knosys.2024.111884>
- [11] Fagundes, M. J. G., Roman, N. T., & Digiampietri, L. A. (2024). The use of syntactic information in fake news detection: A systematic review. *SBC Reviews on Computer Science*, 4(1), 1–10. <https://doi.org/10.5753/reviews.2024.2718>
- [12] Krasadakis, P., Sakkopoulos, E., & Verykios, V. S. (2024). A survey on challenges and advances in natural language processing with a focus on legal informatics and low-resource languages. *Electronics*, 13(3), 648. <https://doi.org/10.3390/electronics13030648>
- [13] Meesad, P. (2021). Thai fake news detection based on information retrieval, natural language processing and machine learning. *SN Computer Science*, 2(6), 434. <https://doi.org/10.1007/s42979-021-00775-6>
- [14] Agras, K., & Atay, B. (2024). A novel transformer-based deep learning pipeline for multilingual fake news detection. *International Journal of Applied Mathematics Electronics and Computers*, 6(4), 12. <https://doi.org/10.36838/v6i4.12>
- [15] Cekineli, R. F., Karagoz, P., & Coltekin, C. (2024). Cross-Lingual Learning vs. Low-Resource Fine-Tuning: A Case Study with Fact-Checking in Turkish. arXiv preprint, arXiv:2403.00411
- [16] Han, S. (2022). Cross-Lingual Transfer Learning for Fake News Detector in a Low-Resource Language. arXiv preprint, arXiv:2208.12482
- [17] Zovikoğlu, M., & Çetin, U. (2024). Detecting misinformation on social networks with natural language processing. *DergiPark*. Retrieved from <https://dergipark.org.tr/>
- [18] Ziyaden, A., Yelenov, A., Hajiyev, F., Rustamov, S., & Pak, A. (2024). Text data augmentation and pre-trained language model for enhancing text classification of low-resource languages. *PeerJ Computer Science*, 10, e1974. <https://doi.org/10.7717/peerj-cs.1974>
- [19] PMookdarsanit, P., & Mookdarsanit, L. (2021). The COVID-19 fake news detection in Thai social texts. *Bulletin of Electrical Engineering and Informatics*, 10(2), 981–988. <https://doi.org/10.11591/eei.v10i2.2745>
- [20] Farhangian, F., Cruz, R. M. O., & Cavalcanti, G. D. C. (2024). Fake news detection: Taxonomy and comparative study. *Information Fusion*, 103, 102140. <https://doi.org/10.1016/j.inffus.2023.102140>
- [21] Hu, B., Mao, Z., & Zhang, Y. (2024). An overview of fake news detection: From a new perspective. *Fundamental Research*. <https://doi.org/10.1016/j.fmre.2024.01.017>
- [22] Nasir, J. A., & Din, Z. U. (2021). Syntactic structured framework for resolving reflexive anaphora in Urdu discourse using multilingual NLP. *KSI Transactions on Internet and Information Systems*, 15(4), 1460–1479. <https://doi.org/10.3837/tiis.2021.04.012>
- [23] Ranathunga, S., Lee, E. S. A., Skenduli, M. P., Shekhar, R., Alam, M., & Kaur, R. (2023). Neural machine translation for low-resource languages: A survey. *ACM Computing Surveys*, 55(11), 229. <https://doi.org/10.1145/3567592>
- [24] Reddy, R. B., Tejaswi, T., Naveen, P., Gires, J., & Vamsi, B. (2023). Optimising the detection of fake news in multilingual exposition using machine learning techniques. *Proceedings of the 7th International*

- Conference on Intelligent Computing and Control Systems (ICICCS 2023)*. IEEE. <https://doi.org/10.1109/ICICCS56967.2023.10142901>
- [25] Wijayanti, R., Khodra, M. L., Surendro, K., & Widyantoro, D. H. (2023). Learning bilingual word embedding for automatic text summarization in low resource language. *Journal of King Saud University—Computer and Information Sciences*, 35(4), 101–112. <https://doi.org/10.1016/j.jksuci.2023.03.015>
- [26] Mahmud, T., Ptaszynski, M., Eronen, J., & Masui, F. (2023). Cyberbullying detection for low-resource languages and dialects: Review of the state of the art. *Information Processing & Management*, 60(5), 103454. <https://doi.org/10.1016/j.ipm.2023.103454>
- [27] Na, S., Na, P., Na, T., & Nasim, Z. (2022). Verify: Breakthrough accuracy in the Urdu fake news detection using text classification. *Proceedings of the 36th Pacific Asia Conference on Language, Information and Computation*.
- [28] Azhar, S. A. F., Hidayat, F., Azfarezat, M. H., Nabiilah, G. Z., & Rojali. (2023). Efficiency of fake news detection with text classification using natural language processing. *Journal of Theoretical and Applied Information Technology*, 101(22).
- [29] Yang, Y., et al. (2020). Multilingual universal sentence encoder for semantic retrieval. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*. <https://doi.org/10.18653/v1/2020.acl-demos.12>
- [30] Cuenca, W., González-Fernández, C., Fernández-Isabel, A., Martín de Diego, I., & Martín, A. G. (2022). Combining conceptual graphs and sentiment analysis for fake news detection. In *Studies in Computational Intelligence* (Vol. 955). Springer. https://doi.org/10.1007/978-3-030-88817-6_15
- [31] Mohamed, B., Haytam, H., & Abdelhadi, F. (2022). Applying fuzzy logic and neural network in sentiment analysis for fake news detection: Case of Covid-19. In *Studies in Computational Intelligence* (Vol. 1001). Springer. https://doi.org/10.1007/978-3-030-90087-8_19
- [32] Taherdoost, H., & Madanchian, M. (2023). Artificial intelligence and sentiment analysis: A review in competitive research. *Computers*, 12(2), 37. <https://doi.org/10.3390/computers12020037>
- [33] Ligthart, A., Catal, C., & Tekinerdogan, B. (2021). Systematic reviews in sentiment analysis: A tertiary study. *Artificial Intelligence Review*, 54(7), 4997–5053. <https://doi.org/10.1007/s10462-021-09973-3>
- [34] Li, J., Sun, A., Han, J., & Li, C. (2022). A survey on deep learning for named entity recognition. *IEEE Transactions on Knowledge and Data Engineering*, 34(1), 50–70. <https://doi.org/10.1109/TKDE.2020.2981314>
- [35] Ehrmann, M., Hamdi, A., Pontes, E. L., Romanello, M., & Doucet, A. (2023). Named entity recognition and classification in historical documents: A survey. *ACM Computing Surveys*, 56(2), 44. <https://doi.org/10.1145/3604931>
- [36] Budi, I., & Suryono, R. R. (2023). Application of named entity recognition method for Indonesian datasets: A review. *Bulletin of Electrical Engineering and Informatics*, 12(2), 1089–1098. <https://doi.org/10.11591/eei.v12i2.4529>
- [37] Sun, Z., & Li, X. (2023). Named entity recognition model based on feature fusion. *Information*, 14(2), 133. <https://doi.org/10.3390/info14020133>
- [38] Wang, H., Zhou, L., Duan, J., & He, L. (2023). Cross-lingual named entity recognition based on attention and adversarial training. *Applied Sciences*, 13(4), 2548. <https://doi.org/10.3390/app13042548>
- [39] Hernández, M. S. (2023). Beliefs and attitudes of canarians towards the Chilean linguistic variety. *Lenguas Modernas*, 62, 183–209. <https://doi.org/10.13039/501100011033>
- [40] Dhyani, B. (2021). Transfer learning in natural language processing: A survey. *Mathematical Statistician and Engineering Applications*, 70(1). <https://doi.org/10.17762/msea.v70i1.2312>
- [41] Palani, B., & Elango, S. (2023). CTrL-FND: Content-based transfer learning approach for fake news

- detection on social media. *International Journal of System Assurance Engineering and Management*, 14(3), 1013–1025. <https://doi.org/10.1007/s13198-023-01891-7>
- [42] Ghayoomi, M., & Mousavian, M. (2022). Deep transfer learning for COVID-19 fake news detection in Persian. *Expert Systems*, 39(8), e13008. <https://doi.org/10.1111/exsy.13008>
- [43] Schwarz, S., Theophilo, A., & Rocha, A. (2020). Emet: Embeddings from multilingual-encoder transformer for fake news detection. *Proceedings of ICASSP 2020—IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. <https://doi.org/10.1109/ICASSP40776.2020.9054673>
- [44] Kasim, S. (2023). One true pairing: Evaluating effective language pairings for fake news detection employing zero-shot cross-lingual transfer. In *Communications in Computer and Information Science* (Vol. 1780). Springer. https://doi.org/10.1007/978-3-031-27609-5_2
- [45] Han, X., Zhao, W., Ding, N., Liu, Z., & Sun, M. (2022). PTR: Prompt tuning with rules for text classification. *AI Open*, 3, 27–37. <https://doi.org/10.1016/j.aiopen.2022.11.003>
- [46] Yang, J., Hu, X., Xiao, G., & Shen, Y. (2024). A survey of knowledge enhanced pre-trained language models. *ACM Transactions on Asian and Low-Resource Language Information Processing*.
- [47] Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Long Papers)*. <https://doi.org/10.18653/v1/P18-1031>
- [48] Gururaja, S., Dutt, R., Liao, T., & Rosé, C. (2023). Linguistic representations for fewer-shot relation extraction across domains. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. <https://doi.org/10.18653/v1/2023.acl-long.414>
- [49] Wang, Z., & Hershcovich, D. (2023). On evaluating multilingual compositional generalization with translated datasets. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. <https://doi.org/10.18653/v1/2023.acl-long.93>
- [50] Farokhian, M., Rafe, V., & Veisi, H. (2024). Fake news detection using dual BERT deep neural networks. *Multimedia Tools and Applications*, 83(15), 44245–44268. <https://doi.org/10.1007/s11042-023-17115-w>
- [51] Ighamdi, J., Lin, Y., & Luo, S. (2023). Towards COVID-19 fake news detection using transformer-based models. *Knowledge-Based Systems*, 274, 110642. <https://doi.org/10.1016/j.knosys.2023.110642>
- [52] G Kuling, G., Curpen, B., & Martel, A. L. (2022). BI-RADS BERT and using section segmentation to understand radiology reports. *Journal of Imaging*, 8(5), 131. <https://doi.org/10.3390/jimaging8050131>
- [53] Wang, Y., et al. (2018). A comparison of word embeddings for the biomedical natural language processing. *Journal of Biomedical Informatics*, 87, 12–20. <https://doi.org/10.1016/j.jbi.2018.09.008>
- [54] Shi, X., Hu, M., Deng, J., Ren, F., Shi, P., & Yang, J. (2023). Integration of multi-branch GCNs enhancing aspect sentiment triplet extraction. *Applied Sciences*, 13(7), 4345. <https://doi.org/10.3390/app13074345>
- [55] Hao, Y., Xue, T., & Liu, G. (2023). Research on the robustness of neural machine translation systems in word order perturbation. *Chinese Journal of Network and Information Security*, 9(5), 120–130. <https://doi.org/10.11959/j.issn.2096-109x.2023078>
- [56] Mohammed, I., & Prasad, R. (2023). Building lexicon-based sentiment analysis model for low-resource languages. *MethodsX*, 11, 102460. <https://doi.org/10.1016/j.mex.2023.102460>
- [57] Haq, I., Qiu, W., Guo, J., & Tang, P. (2023). NLPashto: NLP toolkit for low-resource Pashto language. *International Journal of Advanced Computer Science and Applications*, 14(6). <https://doi.org/10.14569/IJACSA.2023.01406142>

- [58] Zeng, Z., & Bhat, S. (2021). Idiomatic expression identification using semantic compatibility. *Transactions of the Association for Computational Linguistics*, 9, 154–170. https://doi.org/10.1162/tacl_a_00442
- [59] Singh, S., & Singh, P. P. (2023). Transitive and intransitive verb analysis for idiomatic expression understanding: An NLP-based framework. *International Journal of Advanced Research in Science, Communication and Technology*, 123–130. <https://doi.org/10.48175/ijarsct-11682>
- [60] Raffel, C., *et al.* (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21, 1–67.
- [61] Careem, R., Johar, G., & Khatibi, A. (2024). Deep neural networks optimization for resource-constrained environments: Techniques and models. *Indonesian Journal of Electrical Engineering and Computer Science*, 33(3), 1843–1854. <https://doi.org/10.11591/ijeecs.v33.i3.pp1843-1854>
- [62] Kreutzer, J., *et al.* (2022). Quality at a glance: An audit of web-crawled multilingual datasets. *Transactions of the Association for Computational Linguistics*, 10, 50–72. https://doi.org/10.1162/tacl_a_00447
- [63] Fuchs, C. (2020). Cultural and contextual affordances in language MOOCs: Student perspectives. *International Journal of Online Pedagogy and Course Design*, 10(4). <https://doi.org/10.4018/IJOPCD.2020040104>
- [64] Wang, L., Feng, M., Zhou, B., Xiang, B., & Mahadevan, S. (2015). Efficient hyper-parameter optimization for NLP applications. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. <https://doi.org/10.18653/v1/D15-1253>
- [65] Dementieva, D., & Panchenko, A. (2020). Fake news detection using multilingual evidence. *Proceedings of 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE. <https://doi.org/10.1109/DSAA49011.2020.00111>
- [66] Tan, Z., *et al.* (2024). *Large Language Models for Data Annotation and Synthesis: A Survey*. arXiv preprint, arXiv:2402.13446.
- [67] Alzubaidi, L., *et al.* (2023). A survey on deep learning tools dealing with data scarcity: Definitions, challenges, solutions, tips, and applications. *Journal of Big Data*, 10(1), 46. <https://doi.org/10.1186/s40537-023-00727-2>
- [68] Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. <https://doi.org/10.18653/v1/2020.acl-main.560>
- [69] Edwards, J. (2012). *Multilingualism: Understanding Linguistic Diversity*. Continuum. <https://doi.org/10.5860/choice.50-1302>
- [70] Abdedaiem, A., Dahou, A. H., & Cheragui, M. A. (2023). Fake news detection in low resource languages using SetFit framework. *Inteligencia Artificial*, 26(72), 178–201. <https://doi.org/10.4114/intartif.vol26iss72pp178-201>
- [71] Raja, E., Soni, B., & Borgohain, S. K. (2023). Fake news detection in Dravidian languages using transfer learning with adaptive finetuning. *Engineering Applications of Artificial Intelligence*, 126, 106877. <https://doi.org/10.1016/j.engappai.2023.106877>
- [72] Gereme, F., Zhu, W., Ayall, T., & Alemu, D. (2021). Combating fake news in ‘low-resource’ languages: Amharic fake news detection accompanied by resource crafting. *Information*, 12(1), 20. <https://doi.org/10.3390/info12010020>
- [73] Zoph, B., Yuret, D., May, J., & Knight, K. (2016). Transfer learning for low-resource neural machine translation. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. <https://doi.org/10.18653/v1/D16-1163>

- [74] Chen, Y., et al. (2022). Cross-modal ambiguity learning for multimodal fake news detection. *Proceedings of the ACM Web Conference 2022*. <https://doi.org/10.1145/3485447.3511968>
- [75] Shu, K., Wang, S., & Liu, H. (2019). Beyond news contents: The role of social context for fake news detection. *Proceedings of the 12th ACM International Conference on Web Search and Data Mining*. <https://doi.org/10.1145/3289600.3290994>
- [76] Dementieva, D., Kuimov, M., & Panchenko, A. (2023). Multiverse: Multilingual evidence for fake news detection. *Journal of Imaging*, 9(4), 77. <https://doi.org/10.3390/jimaging9040077>
- [77] Park, M., & Chai, S. (2023). Constructing a user-centered fake news detection model by using classification algorithms in machine learning techniques. *IEEE Access*, 11, 78945–78958. <https://doi.org/10.1109/ACCESS.2023.3294613>
- [78] Dou, Y., Shu, K., Xia, C., Yu, P. S., & Sun, L. (2021). User preference-aware fake news detection. *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. <https://doi.org/10.1145/3404835.3462990>

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).