Improving Cross-Lingual Fake News Detection in Indonesia with a Hybrid Model by Enhancing the Embedding Process

Jihan Nabilah Hakim, Yuliant Sibaroni
School of Computing, Telkom University, Bandung, Indonesia

***Abstract*—In the digital age, the spread of false information across languages in digital form threatens the authenticity and credibility of information. This study aims to develop an efficient hybrid deep learning model for detecting cross-lingual fake news, particularly in resource-constrained environments, by enhancing the embedding process. It proposes a lightweight model that combines MUSE embeddings with CNN, LSTM, and LSTM-CNN architectures to evaluate performance across various language pairs with Indonesian as the source language. Experiments show that linguistic similarity significantly influences classification performance. CNN achieves an F1-score of 82% for the Indonesian–Malay pair, a similar language pair. While LSTM achieves 97% for the Indonesian–German language pair (a structurally different language pair). These findings highlight the effectiveness of hybrid architectures and multilingual embeddings in improving cross-lingual fake news detection, especially when English is not the source language. The proposed method provides a reliable yet computationally efficient solution for multilingual misinformation detection in resource-constrained environments.**

Keywords—Cross-lingual; fake news detection; hybrid learning; MUSE embeddings; digital misinformation

I. Introduction

Research on cross-lingual fake news detection attempts to overcome the obstacles raised by language diversity in global news exchange. The main challenge is to make the model able to understand and analyze news text in multiple languages without changing the original meaning. In addition, most fake news datasets currently support only one language, making them less effective in multilingual environments [1]. In addition, the lack of labeled data for low-resource languages can hinder the development of fake news detection methods.

To overcome this problem, cross-lingual research utilizes methods such as cross-lingual embedding or transfer learning that enable cross-lingual comparison and analysis without having to translate the entire dataset. As research in [2] shows, cross-lingual knowledge transfer is becoming a focus of research as it can help overcome the limitations of training data in the target language. Furthermore, research in [3] further explains that cross-lingual embedding can help deal with cross-lingual challenges and transfer learning to align texts from different languages into the same vector space, thus enabling interlanguage comparison and analysis without translating the entire dataset. This approach emphasizes language diversity and the need for models that can adapt to various multilingual datasets. In line with these challenges, the widespread use of online social media presents both opportunities and difficulties in terms of news spread. Social media platforms disseminate news from diverse sources, including real and fake news, significantly shaping our perspectives on rapidly circulating information. However, this information is not always reliable, and it becomes increasingly difficult to assess the accuracy of news content. As discussed in recent studies, social media platforms target users' emotions through news dissemination often manipulating perceptions [4]. Therefore, distinguishing between real and fake news in this domain has become a crucial research problem. Given the unique characteristics of social media and the strategies employed by fake news authors, existing detection algorithms may not be effective, highlighting the need for more robust models that can explain not only the content but also the context of information [5].

Through a hybrid learning approach, this research integrates several machine learning models with techniques such as transfer joint learning and MUSE embeddings. These approaches can enable cross-lingual analysis without the entire dataset needing to be translated. The transfer joint learning approach allows the model to utilize knowledge from the source language (Indonesian) and apply it to other target languages, while MUSE embeddings allow alignment across different languages, ensuring better contextual understanding. It is expected that the use of hybrid learning can improve the adaptability of fake news detection models in Indonesian and other languages, providing more accurate and robust solutions, especially for languages with limited datasets [6], [7].

Additionally, using MUSE embeddings as the main component of this research enables words from different languages to be represented in the same space. This significantly improves the model's ability to understand different language variations and enhances its cross-lingual fake news detection capabilities. MUSE embeddings provides a richer understanding for the model to learn from Indonesian source data and transfer that knowledge to other target languages. This approach contributes not only to fake news detection, but also to the development of more effective methods for overcoming the challenges of multilingual research [8].

Despite the growing interest in cross-lingual fake news detection, many existing approaches rely heavily on translation-based pipelines or pre-trained transformer models that are computationally expensive and language-biased. Most studies still focus on high-resource languages such as English, neglecting the unique challenges faced by low-resource languages like Indonesian. Furthermore, few studies have systematically explored the effect of linguistic similarity on hybrid models using classical neural architectures across a wide set of language pairs.

To address these limitations, this study investigates the following research question: “Can the integration of MUSE embeddings with hybrid LSTM-CNN architectures improve the effectiveness of fake news detection across linguistically diverse language pairs, particularly in low-resource settings?” We aim to fill the gap by proposing a lightweight, translation-free model that leverages cross-lingual embedding and joint learning without requiring expensive resources. Our approach prioritizes adaptability across multiple languages while preserving performance in structurally distinct linguistic contexts.

Therefore, this research aims to improve cross-lingual fake news detection by implementing a hybrid model that combines multiple neural architectures (CNN, LSTM, and LSTM-CNN) along with embedding-based cross-lingual embedding. The main contributions of this research are as follows: (1) investigate the impact of linguistic similarity on model performance in a cross-lingual setting, (2) evaluate a hybrid learning approach for detecting fake news across structurally distinct languages, and (3) demonstrate that using non-English languages as sources in a cross-lingual scenario with limited resources can provide competitive results. These findings are expected to provide new insights into cross-lingual fake news detection and aid the development of more inclusive and adaptable detection systems [9].

The rest of the study is organized as follows. Section II discusses the related works, Section III describes the proposed methodology, Section IV presents the experimental results and discussion in Section V, and Section VI concludes with conclusions and future research directions.

II. Related Works

The spread of misleading information, often referred to as "fake news," has been a major area of research, particularly in understanding the effect of linguistic similarity on hybrid models [7], [10]. Several researchers are interested in developing methods for effective detection and classification of fake news. One of the popular approaches is NLP (Natural Language Processing), which involves analyzing text, such as tweets or posts [11], [12]. Building on this challenge, research has explored methods for detecting fake news. One such study used the FakeNewsNet dataset [13]. This data set, which contains 260,000 news articles related to Twitter, is categorized into two subsets: PolitiFact and GossipCop. The researchers evaluated various machine learning models on this dataset, including Support Vector Machines (SVM), Logistic Regression, Naive Bayes, and Convolutional Neural Networks (CNN). The results of the PolitiFact test for SVM testing show an accuracy of 58% and a GossipCop showed 47%. PolitiFact's test results using logistic regression are 64%, whereas GossipCop's results were 82%. Naive Bayes testing returned results of 70% for GossipCop and 61% for PolitiFact. Lastly, the results of CNN's test on PolitiFact was 62% and 70%, respectively, for GossipCop.

Fake news detection is often conducted using monolingual datasets, but it can also utilize multilingual or cross-lingual datasets. Researchers have explored cross-lingual approaches, as seen in study [14], which studied fake news detection in a Mandarin-English language dataset using CNN, LSTM, Transformer, and BERT, and compared it with the MST-FaDe model. The research results indicated that MST-FaDe achieved the highest accuracy (88%) with an 80% training data proportion. Similarly, [15] developed a BERT-based model for multilingual fake news detection, effectively addressing linguistic differences in Indonesian, Hindi, and Swahili, which proves beneficial for languages with limited resources. Meanwhile, the study in [16] implemented the Passive Aggressive Classifier, Bi-LSTM, and RoBERTa, with Bi-LSTM achieving an accuracy of 61% on the best dataset. Although these studies highlight the effectiveness of Transformer-based models, LSTMs and CNNs remain widely used due to their efficiency and ability to handle multilingual data with fewer resources. The study developed these findings by integrating LSTM and CNN with MUSE embeddings, which aligns cross-lingual word representations without requiring full dataset translation.

In addition to the use of universal languages such as English, Indonesian has also been explored in cross-lingual research. As was done in the study [7], utilizing multilingual word representations without requiring the translation of a full dataset, allowed the model trained on English and German to generalize well to Indonesian. Interestingly, Dutch performs better than English in some cases, most likely due to the morphological and phonetic similarities to Indonesian. A study comparing monolingual and multilingual BERT-based models found that BiLSTM-CRF combined with IndoBERT (which had been previously trained) achieved an F1-Score of 94.90, proving its effectiveness in addressing word ambiguity in Indonesian [17].

The findings suggest that this approach is effective for enhancing the detection of fake news in Indonesian, achieving a high precision rate of 90% through transfer learning in multilingual models. Several approaches have been explored for detecting fake news, including the use of deep learning models, which can be used independently or in a hybrid configuration. Although architectures such as BERT and RoBERTa have demonstrated advanced performance in NLP tasks [14], [18], requires massive computational resources and large-scale multilingual training data, these might not always be available in real-world scenarios.

Moreover, LSTM and CNN provide a more efficient alternative for detecting fake news, especially in resource-constrained LSTM excels in modeling long-range dependencies, making it effective for processing sequential data, while CNN is particularly well-suited for capturing spatial and local patterns in text [19]. Further, the study in [20] introduced the sMemNN model, which integrates a similarity-based matrix mechanism and achieves an accuracy of 88.57%, outperforming CNN+LSTM (48.54%) and LSTM+CNN (65.36%). The lower performance of CNN+LSTM could be due to the dataset balancing technique and the missing pooling layer in CNN.

Although many previous studies have examined fake news detection using neural models and multilingual data, several limitations remain [9], [14], [15]. First, most approaches focus on high-resource languages such as English, neglecting performance on low-resource languages such as Indonesian. Second, many studies rely on large transformer models that require intensive computation, limiting their application in resource-constrained environments. Finally, few studies explicitly analyze the influence of linguistic similarity between language pairs on model performance. This study addresses these issues by proposing a lightweight hybrid model based on CNN, LSTM, and MUSE embeddings, enabling efficient cross-lingual fake news detection without translation and with better adaptation to structurally diverse languages.

III. Methodology

This research focuses on cross-lingual fake news detection using a hybrid learning approach. The model must handle text in multiple languages and integrate information from various sources. It uses Baseline and Hybrid Learning (CNN, LSTM) to model text classification on fake news datasets, improving accuracy and performance. The flow of the modeling process in this research is illustrated in Fig. 1.

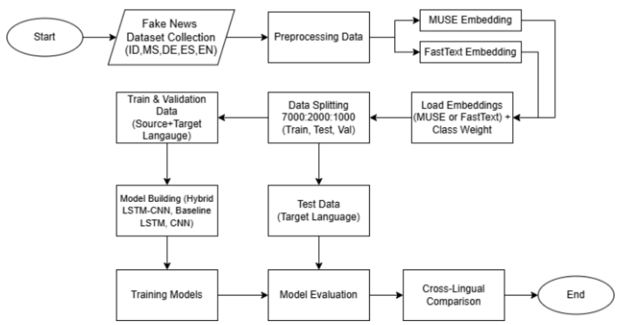

Fig. 1. Modeling process.

A. Data and Preprocessing

In the initial stage, we conducted data collection by selecting Indonesian language datasets [21], [22], [23] as the source dataset and several other languages as the target dataset. The languages selected as targets are Malay [24], German [25], Spanish [26], and English [27]. Each dataset consists of 10,000 data entries that have been labeled real or fake to support the classification process in this study. The selection of datasets from different languages aims to explore the effectiveness of the model in detecting fake news in a multilingual environment through a cross-lingual approach, where the model is trained in the source language and evaluated in the target language. This dataset selection aligns with the objective of evaluating fake news detection in low-resource multilingual contexts. The variation in linguistic similarity among the target languages also enables an analysis of how language structure influences cross-lingual model performance.

Data preprocessing was performed using techniques adapted to the characteristics of each language. For Indonesian and Malay datasets, we used the Sastrawi stopword technique [28] to eliminate common words that are not important to the classification. Meanwhile, for German, Spanish, and English, NLTK stopword are used in study [29] combined with the Snowball stemmer to maintain the base form of the word. In addition, there is an additional step for special character cleaning in German and Spanish, such as the characters ä, ö, ü, Ä, Ö, Ü, ß in German and á, é, í, ó, and ú in Spanish. This character-cleaning process is performed so that the model can focus on the important features without being affected by character variations.

B. Cross-Lingual Embedding

Cross-lingual embedding is a vector representation of words that can be used across different languages. The goal is to map each word in different languages into the same embedding vector space, so that words that have the same meaning in different languages will be close together in the embedding space. According to [30], cross-lingual embedding enables cross-lingual comparison of word meanings, which is very important for machine translation, bilingual lexicon formation, or cross-lingual information retrieval. In addition, this model allows model transfer between different languages. For example, between a high-resource and a low-resource language, providing the same representation space. The majority of cross-lingual embedding models are based on monolingual word embedding models and have been extended for bilingual settings. They also use monolingually trained models. Some of the monolingual embeddings used in cross-lingual research include CBOW (Continuous Bag-of-Word) [31] and GloVe (Global Vectors) [32].

There are two classes of method approaches for cross-lingual embedding: Mapping and Joint methods [33]. However, some of the cross-lingual embedding models developed are mostly for English Indonesian is often underrepresented in such models. According to study [34] for the Indonesian language often uses MUSE [35]. Regarding the embeddings used, FastText is obtained through unsupervised learning from monolingual text, while MUSE used in this research is based on bilingual dictionaries and falls under supervised learning as it utilizes parallel word pairs across languages. Although there are differences in how the embeddings are obtained, both serve only as initial word representations before entering the main supervised learning model (LSTM-CNN). Thus, the primary approach of this research remains within the realm of supervised learning. To represent words in various languages, this study utilizes MUSE (Multilingual Unsupervised and Supervised Embeddings) as the primary embedding method.

MUSE aligns monolingual FastText embeddings into a shared vector space using a high-quality bilingual dictionary, allowing semantically similar words from different languages to be positioned close together. This alignment supports cross-lingual learning without the need for translations, which is particularly useful when working with low-resource languages such as Indonesian. In this study, the supervised version of MUSE with an accurate bilingual dictionary is employed to align Indonesian with English, Malay, German, and Spanish. Additionally, FastText serves as a baseline embedding method for performance comparison. Unlike MUSE, FastText does not align embeddings across languages but instead captures subword information to handle rare and morphologically rich words. While FastText is effective in a monolingual context, MUSE provides a more robust cross-lingual representation, which is crucial for enabling model transfer for detecting fake news across languages.

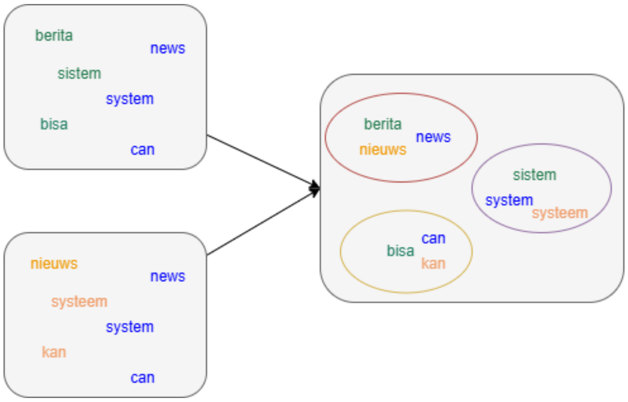

Fig. 2. Example of MUSE embeddings.

The MUSE embeddings combination in this study, illustrated in Fig. 2, integrates ID-EN (Indonesian-English) with MS-EN, DE-EN, ES-EN, and EN-EN to align Indonesian with Malay, German, Spanish, and English. This approach helps the model learn cross-lingual patterns by mapping word embeddings into a shared vector space. For instance, in the DE-EN pair, systeem (German) maps to system (English), while in ID-EN, system (Indonesian) maps to the same English word. Cross-lingual dictionaries ensure accurate alignment, such as news (Indonesian) with news (English) or news (German). The unified embedding space, formed after mapping, enhances multilingual NLP tasks, improving cross-lingual understanding and efficiency. However, challenges such as data imbalance, mapping errors, and computational demands must be addressed to optimize performance.

C. Model Architecture

CNN [36] consists of several important layers, namely the convolution layer, non-linearity layer, pooling layer, and fully-connected layer. The convolution and fully-connected layers have parameters, in contrast to the pooling and non-linearity layers which do not. CNNs perform excellently in machine learning problems, especially in applications involving natural language processing. Although there have been many studies applying hybrid CNN-LSTM, this study adopts the LSTM-CNN model instead. This is because the application of LSTM first captures the sequential dependencies in the text, after which CNN filters the important features identified by LSTM. Thus, this becomes suitable for text-based tasks because LSTM captures the context of the sentence before CNN performs feature extraction.

The choice of LSTM-CNN compared to CNN-LSTM is based on the characteristics of the text in fake news detection, where the order of words plays an important role in understanding the context before feature patterns can be identified. Therefore, the LSTM-CNN approach is more appropriate than CNN-LSTM, which is more commonly used in spatial data processing. Long Short-Term Memory (LSTM) [19] is a type of artificial neural network (ANN) that has been widely used for various text classification tasks. LSTMs are designed to overcome the missing gradient problem in traditional RNNs by introducing memory cells that maintain information over a long period of time. These memory cells are controlled by three gates: input gate, forget gate, and output gate. The input gate controls the flow of new information into the memory cell, the forget gate controls the flow of old information out of the memory cell, and the output gate controls the flow of information from the memory cell to the output.

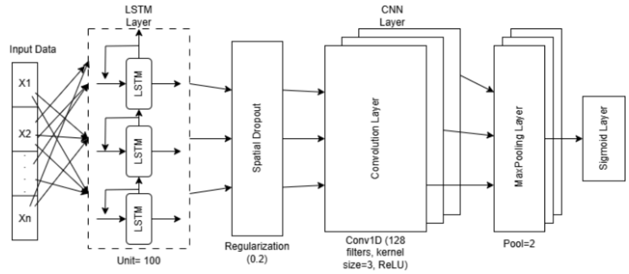

Fig. 3. LSTM-CNN architecture.

A hybrid model of LSTM with other models such as CNN has been suggested to improve the performance of text classification tasks. In the hybrid model shown in Fig. 3, CNN is used to extract static features from the input text, while LSTM is used to capture contextual features. The combination of these two models has been shown to improve the accuracy of text classification tasks. For example, a recent study proposed a hybrid model based on LSTM and CNN for text classification [37]. This study builds a CNN model on top of LSTM, where the text feature vectors generated from LSTM are further processed by CNN. The research found that the hybrid model outperformed the existing baseline methods.

D. Cross-Lingual Joint Learning

Building modeling that can overcome the challenges of multiple language diversity issues in cross-lingual research has been conducted in multiple studies that apply various methods that are considered capable of handling the context of language diversity. One of them is through state-of-the-art research [7], Therefore, our research aims to extend the applicability of the method to address the challenges of cross-lingual research without requiring language translation. The Joint Learning approach in Fig. 4 involves training a model across multiple tasks or languages to produce a representation that is mutually transferable between languages [38]. In the context of cross-lingual, it focuses on training the model with data from multiple languages (source and target) simultaneously so that the model can understand the linguistic patterns that are common between the languages. Cross-lingual joint learning techniques allow the model to be trained using training data from both languages at once, but the evaluation is done only in the target language [39], [40], [41]. For example, we labeled the training data in Indonesian and German as L_{id} and L_{de}. Where using L_{id} and L_{de} as training models and classifying the text into the target language L_{de}. This is because the combined representation can help in retaining the meaning and key information from the source language to the target.

$$L = \sum_{i=1}^{M} L_{Source} \cup \sum_{j=1}^{N} L_{Target} \tag{1}$$

Where L_{Source} and L_{Target} are the loss functions for the source and target languages respectively, while M and N are the amount of data in the source and target languages. Using this combined loss function, the model is trained to minimize errors in both the source and target languages, creating a more general representation that can be applied to the target language. This technique uses semantic alignments between languages through related embeddings.

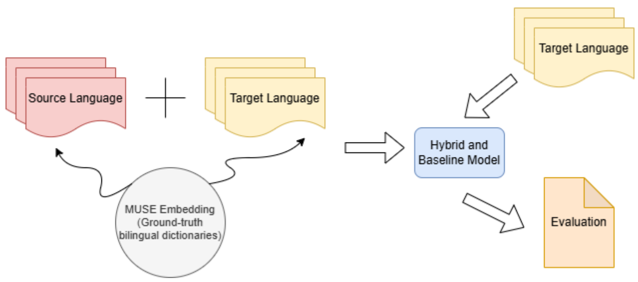

Fig. 4. Joint learning.

E. Experimental Setup

In this study, the LSTM-CNN hybrid learning architecture as well as the two base models, LSTM and CNN, are applied for performance comparison purposes. Hyperparameter settings were determined based on a combination of references from previous literature and initial experiments through Grid Search, with the aim of obtaining optimal results and fair evaluation between models. Some of the parameters adjusted include learning rate, batch size, and network structure, with the final configuration including an embedding dimension of 300, maximum input length of 100 tokens, LSTM units of 100, spatial dropout of 0.2, and CNN layer with 128 filters and kernel size of 3 (for applicable models). All models were trained using a combination of pooling methods such as MaxPooling for CNN and LSTM-CNN and GlobalMaxPooling for all models. One dense layer contains 100 units with sigmoid activation, learning rate 0.01, Adam optimization, class weights {0:1.0, 1:2.0}, number of epochs 10, and batch size 64. We also performed sensitivity analysis on learning rate (0.001, 0.005, 0.01) and batch size (32, 64, 128), and the results show that the combination of learning rate 0.01 and batch size 64 provides stable training and reduces overfitting. Evaluation was conducted using accuracy, precision, recall, and F1-score metrics to assess the model's performance in detecting fake news across languages. The accuracy metric reflects the total correct predictions, while precision assesses the accuracy of the model in identifying fake news without too many errors. Recall measures how many fake news were successfully detected out of all cases. F1-score, as the harmonic mean of precision and recall, was chosen as the main indicator because it is able to provide a balanced assessment, especially when dealing with unbalanced data.

IV. Experimental Results

A. Cross-Lingual Performance Analysis Across Language Pairs

This study evaluates the performance of LSTM-CNN, LSTM, and CNN with MUSE embeddings for cross-lingual fake news detection using a cross-lingual joint learning approach. This experiment aims to see the extent to which models trained in one language can adapt to detect fake news in another language in the context of cross-lingual learning. In this study, the model was trained using data in Indonesian as the source language, which was processed simultaneously with the target language data. The models were then tested on various target languages, including Malay, English, Spanish, and German. Indonesian was chosen as the source language to take advantage of its unique linguistic characteristics, while the target language was chosen to test the model's generalization ability on linguistically different languages. The study examines accuracy, precision, recall, and F1-score to determine model effectiveness. The results provide insights into the best approach for cross-lingual fake news detection.

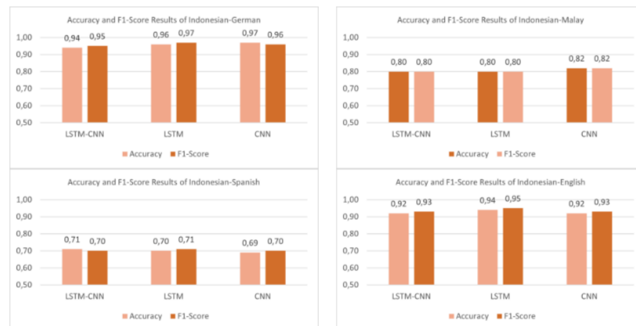

Fig. 5. Accuracy and F1-score performance across language pairs.

As shown in Fig. 5, the evaluation results indicate that model performance varies depending on the language pair. For the Indonesian-German pair, the CNN model recorded the best results with an accuracy of 0.97 and F1-score of 0.96, slightly better than LSTM which achieved an accuracy of 0.96 and F1-score of 0.97, while LSTM-CNN was slightly below with an accuracy of 0.94 and F1-score of 0.95. All three models performed very well. In contrast, for the Indonesian-Malay pair, all three models performed relatively similarly with accuracy and F1-score values ranging from 0.80 to 0.82, with CNN slightly ahead. For the Indonesian-Spanish pair, the results tended to be lower. The LSTM model recorded the highest F1-score of 0.71 with an accuracy of 0.70, slightly better than LSTM-CNN and CNN which were in a similar value range. This indicates a challenge in the transfer of joint learning to Spanish. Meanwhile, for the Indonesian-English pair, LSTM showed optimal performance with an accuracy of 0.94 and F1-score of 0.95, slightly ahead of CNN and LSTM-CNN which had slightly lower values. Overall, CNN is superior for more cognate language pairs, such as Indonesian-German and Indonesian-Malay, while LSTM is more effective for structurally different language pairs such as Indonesian-English and Indonesian-Spanish.

B. Cross-Lingual Model Comparison Across All Target Languages with MUSE Embedding

Table I presents the performance of LSTM-CNN, LSTM, and CNN in cross-lingual classification using MUSE embeddings with Indonesian as the source language. Accuracy, F1-score, precision, and recall were used as evaluation metrics. CNN achieved the highest accuracy of 0.97 on Indonesian-German, highlighting its strength in cross-lingual tasks. F1-score, a key measure combining precision and recall, also reached 0.97 in both LSTM and CNN for Indonesian-German, indicating balanced performance. Precision and recall were analyzed for class 0 (True) and class 1 (False), with Indonesian-German consistently achieving the highest values. While precision measures the model's correctness in identifying a class, recall assesses its ability to capture all relevant instances. Among all metrics, F1-score serves as the final benchmark of model performance. Although the Indonesian-German combination demonstrated the best performance when tested with the German language, its performance significantly declined when retested with the Indonesian language.

TABLE I. THE RESULTS OF CROSS-LINGUAL JOINT LEARNING IN CLASSIFICATION REPORT

Model	Source→Target	Accuracy	F1-Score
LSTM-CNN	ID-MS→MS	0.80	0.80
	ID-DE→DE	**0.94**	**0.95**
	ID-ES→ES	0.71	0.70
	ID-EN→EN	0.92	0.93
	ID-MS→ID	**0.94**	0.92
	ID-DE→ID	0.88	0.88
	ID-ES→ID	0.93	0.93
	ID-EN→ID	0.91	0.91
LSTM	ID-MS→MS	0.80	0.80
	ID-DE→DE	**0.96**	**0.97**
	ID-ES→ES	0.70	0.71
	ID-EN→EN	0.94	0.95
	ID-MS→ID	**0.93**	0.92
	ID-DE→ID	0.90	0.91
	ID-ES→ID	0.92	0.93
	ID-EN→ID	0.92	**0.94**
CNN	ID-MS→MS	0.82	0.82
	ID-DE→DE	**0.97**	**0.97**
	ID-ES→ES	0.69	0.70
	ID-EN→EN	0.92	0.93
	ID-MS→ID	0.90	0.90
	ID-DE→ID	0.89	0.90
	ID-ES→ID	**0.95**	**0.95**
	ID-EN→ID	0.93	0.94

[a] ID=Indonesian, MS=Malay, ES=Spanish, DE=German, EN=English

The Indonesian-Spanish pair showed the worst results after being tested with the target language, Spanish, due to significant linguistic differences, such as verb conjugations in Spanish that do not exist in Indonesian. The Indonesian-English pair showed high performance stability in both languages, ranging from 0.91 to 0.95. This is influenced by the large number of aligned words from English into Indonesian and the richer embedding quality of MUSE for English. The more accurate vector representation and abundant data allow the model to effectively capture cross-lingual semantic relationships without sacrificing accuracy in either language. The Indonesian-Malay pair, which is linguistically similar, showed different performance based on the evaluation language. When tested in Malay, F1-scores were lower in the range of 0.80 to 0.82 compared to when tested in Indonesian (0.92-0.94). This is due to the ambiguity of similar vocabulary with different meanings, word distribution, and data quality differences between the two languages. Although better than Indonesian-Spanish, the model remains biased towards Indonesian because it takes in more patterns from the source language.

C. *Language-Wise and Embedding Across Language Performance*

In the process of analyzing the differences in model performance across language pairs, it is observed that model performance is heavily influenced by linguistic similarity across language pairs the linguistic similarity between the source language (Indonesian) and the target language. For example, the better performance on Indonesian-German and Indonesian-English language pairs suggests that the model is more effective when applied to languages with more similar syntactic structures or grammatical patterns. This allows MUSE embeddings to perform optimally, as the resulting word vector representations can better capture these similarities. In contrast, the lower performance on the Indonesian-Spanish pair is likely due to greater grammatical and lexical differences, which reduces the effectiveness of MUSE embeddings' cross-lingual representation.

TABLE II. F1-SCORE RESULTS (MUSE EMBEDDING AND FASTTEXT)

Cross-Lingual	MUSE Embedding			FastText		
	LSTM-CNN	*LSTM*	*CNN*	*LSTM-CNN*	*LSTM*	*CNN*
Indonesian-Malay	0.80	0.80	0.82	0.68	0.50	0.52
Indonesian-German	**0.95**	**0.97**	**0.97**	**0.70**	**0.62**	**0.60**
Indonesian-Spanish	0.70	0.71	0.70	0.50	0.51	0.50
Indonesian-English	0.93	0.95	0.93	0.54	0.50	0.49

From a cross-lingual perspective, the use of MUSE embeddings assists in aligning word representations across languages; however, its effectiveness remains influenced by linguistic similarity, dataset quality, and training exposure. Languages that are linguistically close to Indonesian, such as Malay, demonstrate more stable performance compared to languages that are far apart, like Spanish, which experiences a decline in accuracy due to structural language gaps. Based on the results in Table II, MUSE consistently yields a higher F1-score compared to FastText, particularly for similar language pairs such as Indonesian-Malay (0.80–0.82). Even for more distant pairs like Indonesian-Spanish, MUSE's performance remains superior (0.70–0.71) compared to FastText, which only reaches 0.50–0.70. This indicates that MUSE excels in aligning interlinguistic representations that possess linguistic proximity. However, its effectiveness is still highly dependent on the quality of the embedding alignment and the availability of data. Therefore, the selection of embeddings and models should be customized by considering the characteristics of the target language, and for highly distinct language pairs, additional approaches may be required to achieve more accurate classification results.

D. *Comparison with State-of-the-Art Models*

Table III presents a comparative analysis of the results of this study against previous studies in cross-lingual fake news detection. Several previous studies have examined the performance of various models in cross-lingual fake news detection tasks. For example, on the Chinese-English language pair dataset, CNN and BERT each achieved an F1-score of 0.85, while the Transformer model scored 0.82. On the English-Other language pair dataset, BERT achieved an F1-score of 0.77, while RoBERTa outperformed with 0.87. The English-Indonesian pair dataset with Bi-LSTM reached 0.87. Some studies reported results using macro F1-scores, such as 0.35 (RoBERTa) and 0.17 (Bi-LSTM) on the English-English pair dataset, and 0.17 (RoBERTa) and 0.29 (Bi-LSTM) on the English-German dataset.

TABLE III. COMPARISON WITH STATE-OF-THE-ART MODELS USING F1-SCORE

Dataset	LSTM	CNN	LSTM-CNN	BERT	RoBERTa	Bi-LSTM
English-Other (CoNLL-2003) [7]	-	-	-	.77	**.87**	-
Chinese-English (Covid-19 FND) [14]	.85	.85	-	.85	-	-
English-Indonesian (TALLIP Fake News Dataset) [9]	-	-	-	-	-	**.87**
English-German (clef-2021) [16]	-	-	-	-	.17 (macro)	.29 (macro)
Indonesian-German (Our Study)	**.97**	**.97**	.95	-	-	-

This research focuses on cross-lingual fake news using LSTM, CNN, and LSTM-CNN models across multiple language pairs. The results show strong performance in various cross-lingual settings. Indonesian-German achieved 0.97 in both LSTM and CNN, and 0.95 in LSTM-CNN, showing highly competitive results compared to previous studies. It is important to note that although all the studies in Table III. Include in the category of cross-lingual fake news detection, the datasets and methodologies are different. Some previous studies used different multilingual corpora, including machine-translated datasets, whereas this study specifically focused on cross-lingual adaptation using MUSE embeddings without translation. Despite these differences, this comparison remains relevant as all approaches aim to address the same fundamental challenge, transferring knowledge between languages for fake news classification. The results show that our LSTM-CNN model achieves competitive performance with existing Transformer-based models while maintaining computational efficiency.

The experimental results highlight that model performance varies based on linguistic similarities and structural differences between languages. LSTM outperformed other models, particularly in Indonesian-German and Indonesian-English, indicating its strength in capturing long-range dependencies and cross-lingual representations. LSTM-CNN also performed well, demonstrating that hybrid architectures can maintain stable precision and recall. Meanwhile, CNN excelled in Indonesian-Malay suggesting that CNN is more effective for language pairs with high lexical similarity. Overall, these results highlight the fact that hybrid models with MUSE embeddings can improve cross-lingual news classification performance, especially for language pairs with structural similarities. Discussion

V. DISCUSSION

A. Influence of Linguistic Similarity on Model Performance

One notable finding was a decline in performance when models trained using German data were retested on Indonesian data. This is thought to be due to an imbalance in cross-lingual vector representation in MUSE embeddings, which are better at aligning languages with high-quality training data, such as German. Remapping to Indonesian may not be optimal, resulting in a reduced ability to capture Indonesian language patterns. On the other hand, the Indonesian-Spanish and Indonesian-Malay combinations did not show a significant decline in performance when retested on Indonesian. This suggests that models trained with German data are more difficult to adapt back to Indonesian due to greater structural and morphological differences. These findings emphasize that maintaining stability in cross-lingual representations remains a significant challenge, especially when the linguistic structures between the source and target languages differ substantially.

Cross-lingual representations such as MUSE are not optimal for language pairs that have considerable linguistic distance and limited parallel data. However, when tested with Indonesian, the performance of the model improved significantly, achieving the highest score among all language pairs. Interestingly, when the model trained with Spanish data was re-evaluated on Indonesian, its performance improved significantly. This suggests that the model adopts more robust linguistic patterns from the source language (Indonesian) but struggles to generalize them when transferred to a language with rich morphology like Spanish. This reinforces the importance of considering both linguistic distance and transfer direction in cross-lingual modeling.

B. Effectiveness of Embedding Alignment in Cross-Lingual Tasks

One of the main advantages of MUSE embeddings over traditional approaches is its ability to perform supervised bilingual alignment using high-quality dictionaries. Unlike unsupervised methods, this allows MUSE to project semantically similar words from different languages into a shared vector space without relying on large-scale parallel corpora. This makes MUSE particularly suitable for languages with limited resources, such as Indonesian, where parallel data is scarce. When compared to FastText, which operates solely in a monolingual space using subword representations, MUSE demonstrates significantly better performance in cross-lingual fake news classification. FastText is effective in handling morphologically rich languages but lacks cross-lingual semantic alignment, making it less suitable for multilingual transfer learning tasks.

The results show that MUSE consistently achieves higher F1-scores across all language pairs. Notably, the Indonesian-German (0.95–0.97) and Indonesian-English (0.93–0.95) pairs outperform the others, likely due to the availability of rich linguistic resources and effective bilingual mapping. Conversely, the Indonesian-Spanish pair recorded the lowest performance, highlighting the limitations of cross-lingual embeddings in handling linguistically distant language pairs with minimal training exposure.

These observations confirm that the success of cross-lingual models depends not only on model architecture but also on the quality of the underlying embedding alignment. These results emphasize the critical role of embedding strategies in enabling accurate multilingual NLP, especially when targeting languages with limited resources or diverse structures.

C. Advantages of the Proposed Hybrid Model

The results of this study show that the proposed hybrid model, which combines LSTM-CNN with MUSE embeddings, achieves competitive or even superior performance compared to previous studies in cross-lingual fake news detection. Specifically, experiments on the Indonesian-German and Indonesian-English language pairs yielded very high F1-scores (up to 0.97), demonstrating the model's strength in leveraging linguistic alignment and transfer learning. Unlike transformer-based architectures, this approach offers computational efficiency while maintaining robust performance on both similar and relatively distant language pairs.

However, the Indonesian-Spanish language pair poses significant challenges, likely due to substantial morphological and syntactic differences. This finding highlights the main limitation of embedding alignment strategies when applied to linguistically distant languages. To address this, future research could explore the integration of advanced multilingual embeddings, such as XLM-R or mBERT, and investigate fine-tuning techniques to enhance cross-lingual adaptability. Additionally, incorporating target language training data and evaluating models on informal or noisy text (e.g., social media content) could enhance their robustness in real-world applications.

VI. CONCLUSION

The evaluation results show that the best performance, based on F1-Score, is achieved by the LSTM model with MUSE embeddings, particularly for the Indonesian-German pair (0.97) and Indonesian-English pair (0.95). This confirms LSTM's effectiveness in capturing sequential dependencies, making it well-suited for cross-lingual tasks. The LSTM-CNN model produced comparable results, with its highest F1-Score of 0.95 on the Indonesian-German pair, but struggled in the Indonesian-Spanish pair, indicating that hybrid models may struggle to generalize effectively to structurally divergent languages. FastText embeddings performed significantly worse across all models and language pairs, with the Indonesian-Spanish pair only reaching an F1-Score of 0.50. This highlights the importance of multilingual embeddings like MUSE, which consistently delivered superior results. CNN performed competitively on Indonesian-German (0.97) but showed a decline in Indonesian-Spanish (0.70), indicating its limitations in capturing complex cross-linguistic relationships. These findings emphasize that model effectiveness depends on both linguistic characteristics and embedding selection. MUSE embeddings, combined with LSTM-based models, provide a lightweight yet effective alternative for cross-lingual fake news detection, making them suitable for real-world applications in low-resource language settings. Future research should explore the possibility of fine-tuning MUSE embeddings with additional target-language data to improve generalization across languages with highly divergent structures. Additionally, evaluating model robustness against adversarial attacks and noisy data, such as spelling variations, informal language, and manipulated text, would be valuable in assessing its real-world applicability. However, this study has several limitations. First, the scope of language pairs is limited to five, which may not fully represent global linguistic diversity. Second, the quality and balance of the dataset can affect model performance, especially for languages with limited resources. Additionally, the experiments were limited to structured news articles, excluding informal formats such as social media posts. These factors limit the generalizability of the findings. Future research could address these limitations by integrating more diverse datasets, languages, and text genres.

ACKNOWLEDGMENT

This research is part of the output of the Fundamental Research Grant Scheme, funded by Telkom University and the Ministry of Education, Science, and Technology, Indonesia with grant agreement number: 7925/LL4/PPG/2025, 121/LIT 07/PPM-LIT/2025.

The authors gratefully acknowledge the support and funding provided for the successful completion of this research.

REFERENCES

[1] D. Dementieva, M. Kuimov, and A. Panchenko, "Multiverse: Multilingual Evidence for Fake News Detection," J Imaging, vol. 9, no. 4, p. 77, Mar. 2023, doi: 10.3390/jimaging9040077.

[2] S. Han, "Cross-lingual Transfer Learning for Fake News Detector in a Low-Resource Language," Wiley Journal, Aug. 2022, [Online]. Available: http://arxiv.org/abs/2208.12482

[3] K. Goswami, S. Dutta, H. Assem, T. Fransen, and J. P. McCrae, "Cross-lingual Sentence Embedding using Multi-Task Learning," in Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Stroudsburg, PA, USA: Association for Computational Linguistics, 2021, pp. 9099–9113. doi: 10.18653/v1/2021.emnlp-main.716.

[4] S. K. Uppada, P. Patel, and S. B., "An image and text-based multimodal model for detecting fake news in OSN's," J Intell Inf Syst, Nov. 2022, doi: 10.1007/s10844-022-00764-y.

[5] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media," ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, Sep. 2017, doi: 10.1145/3137597.3137600.

[6] I. F. Putra and A. Purwarianti, "Improving Indonesian Text Classification Using Multilingual Language Model," 2020 7th International Conference on Advance Informatics: Concepts, Theory and Applications (ICAICTA), Sep. 2020, [Online]. Available: http://arxiv.org/abs/2009.05713

[7] S. O. Khairunnisa, Z. Chen, and M. Komachi, "Dataset Enhancement and Multilingual Transfer for Named Entity Recognition in the Indonesian Language," ACM Transactions on Asian and Low-Resource Language Information Processing, vol. 22, no. 6, pp. 1–21, Jun. 2023, doi: 10.1145/3592854.

[8] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep Learning Approach for Intelligent

Intrusion Detection System," IEEE Access, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[9] R. Mohawesh, S. Maqsood, and Q. Althebyan, "Multilingual deep learning framework for fake news detection using capsule neural network," J Intell Inf Syst, vol. 60, no. 3, pp. 655–671, Jun. 2023, doi: 10.1007/s10844-023-00788-y.

[10] D. A. Erku et al., "When fear and misinformation go viral: Pharmacists' role in deterring medication misinformation during the 'infodemic' surrounding COVID-19," Research in Social and Administrative Pharmacy, vol. 17, no. 1, pp. 1954–1963, Jan. 2021, doi: 10.1016/j.sapharm.2020.04.032.

[11] M. Tarczewska, A. Marciniak, and A. Giełczyk, "Fake or Real? The Novel Approach to Detecting Online Disinformation Based on Multi ML Classifiers," 2021, pp. 18–27. doi: 10.1007/978-3-030-77970-2_2.

[12] M. Choraś et al., "Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study," Appl Soft Comput, vol. 101, p. 107050, Mar. 2021, doi: 10.1016/j.asoc.2020.107050.

[13] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A Data Repository with News Content, Social Context and Spatialtemporal Information for Studying Fake News on Social Media," Sep. 2018, [Online]. Available: http://arxiv.org/abs/1809.01286

[14] Z. Guo, Q. Zhang, F. Ding, X. Zhu, and K. Yu, "A Novel Fake News Detection Model for Context of Mixed Languages Through Multiscale Transformer," IEEE Trans Comput Soc Syst, vol. 11, no. 4, pp. 5079–5089, Aug. 2024, doi: 10.1109/TCSS.2023.3298480.

[15] A. De, D. Bandyopadhyay, B. Gain, and A. Ekbal, "A Transformer-Based Approach to Multilingual Fake News Detection in Low-Resource Languages," ACM Transactions on Asian and Low-Resource Language Information Processing, vol. 21, no. 1, pp. 1–20, Jan. 2022, doi: 10.1145/3472619.

[16] M. Arif et al., "CIC at CheckThat! 2022: Multi-class and Cross-lingual Fake News Detection," CEUR Workshop Proc, vol. 3180, pp. 434–443, 2022.

[17] S. Wu and M. Dredze, "Beto, Bentz, Becas: The Surprising Cross-Lingual Effectiveness of BERT," in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Stroudsburg, PA, USA: Association for Computational Linguistics, 2019, pp. 833–844. doi: 10.18653/v1/D19-1077.

[18] C. Dulhanty, J. L. Deglint, I. Ben Daya, and A. Wong, "Taking a Stance on Fake News: Towards Automatic Disinformation Assessment via Deep Bidirectional Transformer Language Models for Stance Detection," Nov. 2019, [Online]. Available: http://arxiv.org/abs/1911.11951

[19] N. Mohd, H. Singhdev, and D. Upadhyay, "TEXT CLASSFICATION USING CNN AND CNN-LSTM," Webology, vol. 18, no. 4, 2021, doi: 10.29121/WEB/V18I4/149.

[20] M. Mohtarami, R. Baly, J. Glass, P. Nakov, L. Marquez, and A. Moschitti, "Automatic Stance Detection Using End-to-End Memory Networks," Apr. 2018, [Online]. Available: http://arxiv.org/abs/1804.07581

[21] M. G. Muharam, "Indonesia False News(Hoax) Dataset," Kaggle. Accessed: Apr. 24, 2024. [Online]. Available: https://www.kaggle.com/datasets/muhammadghazimuharam/indonesiafalsenews/data

[22] Pierobeat, "Fake News (Hoax) classification Using XGBoost," GitHub. Accessed: Apr. 24, 2024. [Online]. Available: https://github.com/pierobeat/Hoax-News-Classification/?tab=readme-ov-file

[23] I. Y. R. Pratiwi, R. A. Asmara, and F. Rahutomo, "Study of hoax news detection using naïve bayes classifier in Indonesian language," in 2017 11th International Conference on Information & Communication Technology and System (ICTS), IEEE, Oct. 2017, pp. 73–78. doi: 10.1109/ICTS.2017.8265649.

[24] H. Zolkepli, "Malay-Dataset.," GitHub. Accessed: Jun. 27, 2024. [Online]. Available: https://github.com/mesolitica/malaysian-dataset

[25] J. Risch, A. Stoll, L. Wilms, and M. Wiegand, "Overview of the GermEval 2021 Shared Task on the Identification of Toxic, Engaging, and Fact-Claiming Comments," in Proceedings of the GermEval 2021 Shared Task on the Identification of Toxic, Engaging, and Fact-Claiming Comments, Association for Computational Linguistics, 2021, pp. 1–12. Accessed: Jun. 27, 2024. [Online]. Available: https://aclanthology.org/2021.germeval-1.1

[26] J. O. Vizoso, "Spanish Political Fake News," Kaggle. Accessed: Aug. 15, 2024. [Online]. Available: https://www.kaggle.com/datasets/javieroterovizoso/spanish-political-fake-news/data

[27] L. Cui and D. Lee, "CoAID: COVID-19 Healthcare Misinformation Dataset," May 2020.

[28] A. Librian, M. Krishnan, A. Putra, R. Fadlillah, and P. Adi, "Sastrawi: Sistem Stemming Bahasa Indonesia," GitHub. Accessed: Nov. 05, 2024. [Online]. Available: https://github.com/sastrawi/sastrawi

[29] S. Bird, E. Klein, and E. Loper, Natural Language Processing with Python, 1st ed. O'Reilly Media, Inc., 2009.

[30] S. Ruder, I. Vulić, and A. Søgaard, "A Survey of Cross-lingual Word Embedding Models," Journal of Artificial Intelligence Research, vol. 65, pp. 569–631, Aug. 2019, doi: 10.1613/jair.1.11640.

[31] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," Jan. 2013, [Online]. Available: http://arxiv.org/abs/1301.3781

[32] J. Pennington, R. Socher, and C. Manning, "Glove: Global Vectors for Word Representation," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Stroudsburg, PA, USA: Association for Computational Linguistics, 2014, pp. 1532–1543. doi: 10.3115/v1/D14-1162.

[33] S. Sannigrahi and J. Read, "Isomorphic Cross-lingual Embeddings for Low-Resource Languages," Mar. 2022, [Online]. Available: http://arxiv.org/abs/2203.14632

[34] R. Wijayanti, M. L. Khodra, K. Surendro, and D. H. Widyantoro, "Learning bilingual word embedding for automatic text summarization in low resource language," Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 4, pp. 224–235, Apr. 2023, doi: 10.1016/j.jksuci.2023.03.015.

[35] G. Lample and A. Conneau, "Cross-lingual Language Model Pretraining," Jan. 2019, [Online]. Available: http://arxiv.org/abs/1901.07291

[36] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in 2017 International Conference on Engineering and Technology (ICET), IEEE, Aug. 2017, pp. 1–6. doi: 10.1109/ICEngTechnol.2017.8308186.

[37] J. Zhang, Y. Li, J. Tian, and T. Li, "LSTM-CNN Hybrid Model for Text Classification," in 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), IEEE, Oct. 2018, pp. 1675–1680. doi: 10.1109/IAEAC.2018.8577620.

[38] Z. Wang, J. Xie, R. Xu, Y. Yang, G. Neubig, and J. Carbonell, "Cross-lingual Alignment vs Joint Training: A Comparative Study and A Simple Unified Framework," Oct. 2019, [Online]. Available: http://arxiv.org/abs/1910.04708

[39] D. Van Thin, H. Quoc Ngo, D. Ngoc Hao, and N. Luu-Thuy Nguyen, "Exploring zero-shot and joint training cross-lingual strategies for aspect-based sentiment analysis based on contextualized multilingual language models," Journal of Information and Telecommunication, vol. 7, no. 2, pp. 121–143, Apr. 2023, doi: 10.1080/24751839.2023.2173843.

[40] Y. Cao et al., "Joint Representation Learning of Cross-lingual Words and Entities via Attentive Distant Supervision," in Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Stroudsburg, PA, USA: Association for Computational Linguistics, 2018, pp. 227–237. doi: 10.18653/v1/D18-1021.

[41] N. van der Heijden, S. Abnar, and E. Shutova, "A Comparison of Architectures and Pretraining Methods for Contextualized Multilingual Word Embeddings," Dec. 2019, [Online]. Available: http://arxiv.org/abs/1912.10169.