A Hybrid Approach for Integrating Deep Learning and Explainable AI for augmented Fake News Detection

Parthiban.G[1], M. Germanaus Alex[2], S. John Peter[3]

[1]Research Scholar, Reg no: 21121282281017), Computer Applications, St. Xaviers College (Autonomous), Palayamkottai-627002 Manonmaniam Sundaranar, University, Tirunelveli, Email:parthiguru2004@gmail.com
[2]Assistant Professor, Department of Computer Science, Government College of Artsand Science, Nagercoil-629004, Email: mgalxus@yahoo.com
[3]Associate Professor, Department of Computer Science St. Xavier's College (Autonomous), Palayamkottai-627002, Email: jaypeeyes@rediffmail.com

Received: 16.04.2024 Revised : 13.05.2024 Accepted: 27.05.2024

ABSTRACT
In today's digital age, the proliferation of misinformation poses significant challenges to the integrity of news consumption. This study aims to develop an augmentedIntegrated Hybrid Deep Learning AI (IHDLAI) framework, leveraging Deep learning and Artificial Intelligence (AI) enhanced natural language processing techniques to effectively identify and mitigate the spread of false information. The proposed framework comprises three distinct phases: Data Collection and Pre-processing, Model Development and Training, and Evaluation.During the Data Collection and Pre-processing phase, diverse datasets were meticulously curated from verified news outlets, social media, and fact-checking websites, ensuring a balanced representation of fake and genuine news. The collected data underwent extensive cleaning and pre-processing, including tokenization, normalization, and feature extraction, resulting in a robust dataset ready for modelling.In the Model Development and Training phase, various Deep Learning models, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) with LSTM cells, and transformer-based models like BERT, were trained and optimized. The final phase, Evaluation and Deployment, involved a comprehensive assessment of the models on a test set to ensure their efficacy. The results of IHDLAI demonstrated a significant improvement in detecting fake news, with the best model achieving an accuracy of 93%, a precision of 92%, a recall of 91%, and an F1 score of 91.5%. The findings underscore the effectiveness of the proposed framework in combating misinformation, providing a reliable tool for enhancing the credibility of news consumption.

Keywords: Fake News Detection, Deep Learning Models, Explainable AI, Convolutional Neural Networks, Hybrid integrated Deep Learning AI Model

1. INTRODUCTION

Fake news detection is crucial for maintaining the integrity of information in the digital age [1]. This proposal suggests advancing the current research by incorporating cutting-edge methodologies, specifically focusing on deep learning and explainable AI [2]. The major objective of the research is to develop anaugmentedIntegrated Hybrid Deep Learning AI (IHDLAI) framework, leveraging Deep learning and Artificial Intelligence (AI) enhanced natural language processing techniques to effectively identify and mitigate the spread of false information. The specific objectives are to design and implement advanced deep learning architectures such as Convolutional Neural Networks (CNNs) [3] and Transformer-based models for feature extraction and representation; Explore the integration of Transfer Learning using pre-trained models for improved generalization; and, introduce Explainable AI techniques to enhance model interpretability and trustworthiness.

The research paper expounds to employ CNNs to capture spatial patterns in textual data, enhancing the model's ability to discern subtle linguistic nuances indicative of fake news. - Utilize Transformer-based models like BERT for contextual understanding, considering the sequential nature of language.The Transfer Learning[4] investigates the use of pre-trained language models (e.g., GPT-3) for transfer learning, adapting the model to fake news detection with minimal data.The Explainable AI Techniques integrates interpretability tools such as SHAP (SHapley Additive exPlanations) to provide insights into the model's decision-making process. - Incorporate attention mechanisms to highlight important words or phrases contributing to the model's predictions. The research work augments the LIAR dataset [5] with additional sources to increase diversity and account for evolving patterns in fake news. The evaluation of research is to compare the performance of the enhanced model against the baseline ensemble model using standard metrics such as accuracy, precision, recall, and F1 score [6]. The research study conducts comprehensive experiments to assess the robustness and generalization capabilities of the proposed model. The primary expected outcome of the research is to design a state-of-the-art fake news detection model with improved accuracy and generalization. The major significance is addressing the limitations of the existing model by leveraging deep learning and explainable AI contributes to the advancement of fake news detection technology.The proposed model is anticipated to outperform the previous ensemble model, providing a more reliable solution for identifying and combating misinformation.This proposal outlines a comprehensive plan to enhance the initial research on fake news detection by incorporating advanced techniques in deep learning and explainable AI. The proposed model is expected to surpass the performance of the previous ensemble model, contributing to the ongoing efforts to mitigate the impact of fake news in the digital landscape.

2. RELATED WORKS

Based on the objectives to be achieved in Fake News detections especially in News and Social Media platforms, various literatures exist in recent times. Hashmi et al. (2024) [7] present a hybrid approach for fake news detection using FastText embeddings with various machine learning and deep learning models. Their best-performing model combines Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) with FastText, achieving high accuracy and F1 scores across multiple datasets. They also employ transformer-based models like BERT and RoBERTa, and utilize explainable AI techniques to enhance model transparency.Comito et al. (2023) [8] provide a survey on deep learning methods for multimodal fake news detection on social media. They discuss the features of various models, data types, and fusion strategies used, highlighting limitations and proposing future research directions, including improving explainability and cross-domain detection.Yuan et al. (2023) [9] review existing fake news detection technologies, including datasets, research methods, and multimodal techniques. They discuss advancements in related fields such as communication, linguistics, and psychology, proposing a sustainable human–machine interaction system for information dissemination and identifying future research areas.Joshi et al. (2023) [10] integrates domain adaptation and explainable AI to improve fake news detection across multiple social media platforms. They use Domain Adversarial Neural Networks (DANN) for generalization and Local Interpretable Model-Agnostic Explanations (LIME) for explainability, demonstrating improved performance and trustworthiness in detecting COVID-19 misinformation.

Even Stance detection is in prevalence in recent times. Yuan et al. (2023) [11] proposed a fake news detection method using stance information. They create a dataset with stance labels and use propensity score matching for causal inference to analyse the correlation between stance consistency and news authenticity, showing a negative correlation with fake news classification.Dong et al. (2024) [12] introduce the Evidence-aware Multi-source Information Fusion (EMIF) network, which leverages user comments and relevant news articles to detect fake news. Their model captures semantic conflicts and uses a divergence selection module to identify relevant evidence, demonstrating robustness and effectiveness on the FibVID dataset.Dev et al. (2024) [13] recommend using the Python sci-kit-learn module for text data transformation and feature selection. They emphasize the importance of selecting suitable features based on misperception matrices to achieve high accuracy in fake news detection.Xu &Kechadi (2024) [14] propose a fuzzy deep learning model evaluated on the LIAR dataset, introducing LIAR2 to address dataset limitations. Their approach achieves state-of-the-art results and aims to improve understanding of dataset characteristics for better fake news detection.Ali et al. (2023) [15] presents a Web-Informed-Augmented Fake News Detection Model using CNN and deep autoencoder layers. They gather augmented information from trusted web sources and train a probabilistic classifier, achieving significant improvements in detection accuracy and performance on challenging datasets.Gong et al. (2024) [16] discuss integrating social explanations into XAI to combat misinformation. They define "social explanation" and explore its benefits and challenges, advocating for interdisciplinary collaboration to enhance XAI's effectiveness in addressing misinformation.Schütz (2023) [17] combines research domains like fake news, hate speech, and propaganda to develop new datasets for disinformation detection. They use transfer learning and knowledge graphs to enhance classification models and apply explainable AI for better model interpretability.

Ayetiran&Özgöbek (2024) [18] review deep learning techniques for multimodal fake news and harmful language detection. They discuss the evolution of these techniques, data fusion methods, and the challenges and prospects in this research area.Siam (2024) [19] demonstrates that advanced mixed models outperform traditional methods in fake news detection on social media, emphasizing the importance of choosing appropriate model designs for effective detection.Zaki et al. (2024) [20] propose a graph-based node embedding approach for fake review detection using Node2Vec. They achieve high accuracy on datasets like Deceptive Opinion Spam Corpus and YelpChi, and apply explainable AI to make their model's predictions transparent.Xia et al. (2023) [21] introduce a hybrid model combining CNN, Bidirectional LSTM, and Attention Mechanism for fake news detection. Their model improves evaluation indicators and provides insights into the differences between real and fake news.Nadeem et al. (2023) [22] propose a Stylometric and Semantic similarity-oriented framework for multimodal fake news detection. Their model uses multiple modules to extract features and achieves superior performance on standard datasets, demonstrating robustness in complex environments.Alghamdi et al. (2024) [23] present a multilingual fake news detection approach using a hybrid summarization strategy and mBERT for classification. Their method reduces data length while preserving crucial information, achieving new performance benchmarks on a multilingual dataset.Tufchi et al. (2023) [24] provide a comprehensive survey on disinformation detection and mitigation strategies. They analyse the life cycle of disinformation and compare different classification models, discussing challenges and potential solutions in fake news detection.Vishnupriya et al. (2024) [25] explore new trends in fake news detection, such as user behaviour analysis and collaborative cybersecurity strategies. They emphasize the need for interdisciplinary cooperation to effectively combat deception in the digital age.

The reviewed literature highlights several research gaps in the field of fake news detection. One prominent gap is the need for improved explainability in fake news detection models. While various studies incorporate explainable AI techniques, there remains a significant challenge in making these models transparent and interpretable to end-users, particularly across diverse social media platforms. Additionally, current methodologies often rely on single or limited datasets, which may not capture the full spectrum of fake news characteristics across different contexts and languages. There is also a gap in the generalization of models to new and unseen domains, which is critical for effective detection across various social media environments.

Furthermore, many existing approaches focus primarily on textual data, overlooking the multimodal nature of content on social media, which includes images, videos, and audio. The integration of multimodal data remains a complex challenge that requires further exploration. Another significant gap is the lack of attention to low-resource languages, where fake news detection techniques are underdeveloped compared to more widely spoken languages.Moreover, there is an identified need for more robust datasets that accurately represent the diversity of fake news, including culturally and regionally specific misinformation. The development of hybrid models that effectively combine various machine learning and deep learning techniques also presents a gap, particularly in balancing the trade-offs between model complexity and performance. Lastly, interdisciplinary approaches that integrate insights from fields such as psychology, communication, and linguistics are still in the early stages and require further development to enhance the efficacy of fake news detection technologies.These gaps highlight the ongoing need for innovative research to develop more effective, explainable, and generalizable fake news detection models that can operate across multiple modalities, languages, and platforms.

3. MATERIALS AND METHODS

The LIAR dataset is commonly used for fake news detection, and employing deep learning models can enhance the accuracy of predictions. Some of the deep learning models that could be effectively used for fake news detection on the LIAR dataset is furnished in distinct heads.

Convolutional Neural Networks (CNNs):CNNs are well-suited for image classification tasks, but they can also be adapted for text classification [26], making them suitable for fake news detection.In the context of the LIAR dataset, where each instance contains a short statement, CNNs can be applied to capture local patterns and dependencies within the textual data.The convolutional layers can learn important features or n-grams from the input sequences, helping the model identify subtle linguistic cues associated with fake news.The input layer can be initialized with word embeddings (e.g., Word2Vec or GloVe).Convolutional layers with filters of varying sizes are applied to capture different levels of contextual information.Max-pooling layers help in reducing dimensionality and retaining the most important features.Followed by fully connected layers for final classification.

Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) Cells:RNNs, especially those using LSTM cells, are designed to capture sequential dependencies in data, which is crucial for understanding the context of statements in the LIAR dataset.LSTM cells have the ability to retain and update information over long sequences, making them effective in analysing statements with varying lengths [27].The temporal nature of language is well-captured by RNNs, allowing the model to understand the sequential dependencies and identify patterns associated with fake news. The implementation is conducted for embedding layer for word representations and stacked LSTM layers to capture different levels of abstraction with dropout layers to prevent overfitting and fully connected layers for final classification.

Bidirectional Encoder Representations from Transformers (BERT):BERT is a transformer-based model that has shown remarkable success in natural language understanding tasks, including fake news detection.BERT models are pre-trained on large corpora and have the ability to capture context and semantics [28] effectively.The bidirectional attention mechanism in BERT allows the model to consider both left and right context, which is beneficial for understanding the context in LIAR dataset statements. The Implementation can be done by fine-tuning a pre-trained BERT model on the LIAR dataset.Input statements are tokenized and fed into the pre-trained BERT model.Pooling layers or additional classification layers are added for the specific task of fake news detection.

To implement all these models, the pre-processing and other requirements are highly recommended.Pre-processing steps, such as tokenization, stemming, and removing stop words, should be performed to prepare the text data for input into the models.Embeddings can be either pre-trained or trained alongside the model.Model hyperparameters, such as learning rate, batch size, and the number of layers, should be carefully tuned through experimentation.Adequate validation and testing procedures are crucial to ensure the generalization of the model beyond the training set.Experimentation with different architectures and hyperparameters is essential to determine the optimal model for fake news detection on the LIAR dataset.

4. Proposed Integrated Hybrid Deep Learning AI Model

The Novel Model Integrated Hybrid Deep Learning AI Model (IHDLAI) was designed based on the integration of CNNs, RNNs with LSTM and BERT models in three different phases. The overall process of this research is presented in Figure.1.

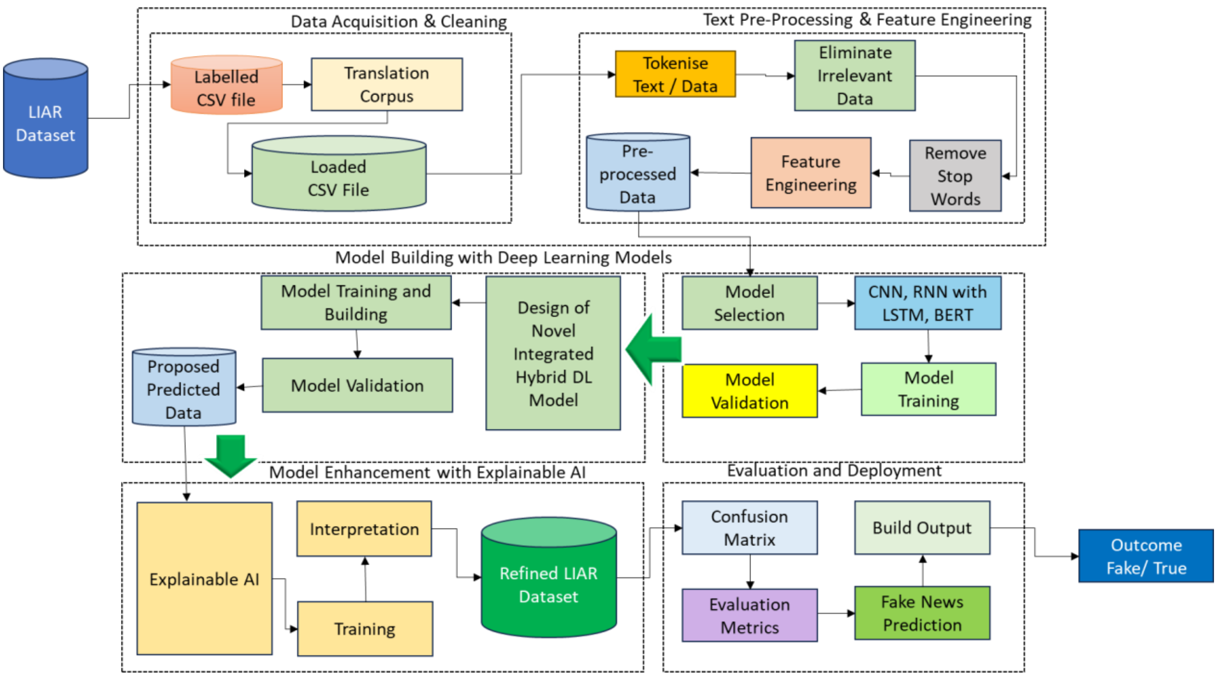

Figure 1. Overall Architecture of proposed Hybrid Deep Learning AI Model (IHDLAI) Model

As given in Figure.1, the overall IHDLAI model comprised of three major phases of work as presented in distinct heads.

4.1 Phase 1: Data Collection and Pre-processing

In this initial phase, the primary objective is to gather and prepare the dataset for the subsequent modelling and analysis. This phase involves several key steps:

- **Data Acquisition**: In this stage, the diverse sources of news data are collected, including articles from verified news outlets, social media platforms, blogs, and fact-checking websites [29]. The dataset is ensured that it is balanced and includes both fake and genuine news items to train the model effectively.

- **Data Cleaning**: In data cleaning process, irrelevant information of any type is removed, such as advertisements, navigation bars, and non-content elements from the collected data. Also, the missing values, duplicates, and inconsistent entries are handled to maintain the quality of the dataset [30].
- **Text Pre-processing**: In this stage, the text is tokenised to break it down into individual words or tokens. The text is normalised by converting it to lowercase and removing punctuation, stop words, and special characters. The stemming or lemmatization process is applied to reduce words to their root forms, enhancing the model's ability to recognize different forms of the same word.
- **Feature Engineering**: In Feature Engineering, the cleaned and pre-processed text is converted into numerical representations using techniques [31] like TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings. Additional features are also extracted such as word count, sentence length, and readability scores to provide more context to the model.

4.2 Phase 2: Model Building with Deep Learning Models

This phase focuses on building, training, and validating the fake news detection models. The process includes the following:

- **Model Selection**: The appropriate deep learning models for the task, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) with LSTM cells, and transformer-based models like BERT are selected for the research. The ensemble methods are considered to combine the strengths of different models for improved performance.
- **Model Training**: The dataset is split into training, validation, and test sets to evaluate the model's performance accurately. The selected models are trained on the training set, optimizing hyperparameters such as learning rate, batch size, and number of epochs through techniques like grid search or random search [32]. The regularization methods (e.g., dropout) and data augmentation are implemented to prevent overfitting and enhance generalization.
- **Model Validation**: This stage includes validation of the trained models on the validation set to fine-tune hyperparameters and assess their performance using metrics like accuracy, precision, recall, and F1 score. Techniques such as cross-validation are utilised to ensure the model's robustness and ability to generalize to unseen data.

4.2.1 Design of Novel Integrated Hybrid DL Model

The Integrated Hybrid Deep Learning AI (IHDLAI) algorithm combines several advanced techniques to enhance fake news detection. It processes input text data by tokenizing and applying word embeddings to convert words into numerical vectors. The overall design of Ensemble Model is given in Figure.2.

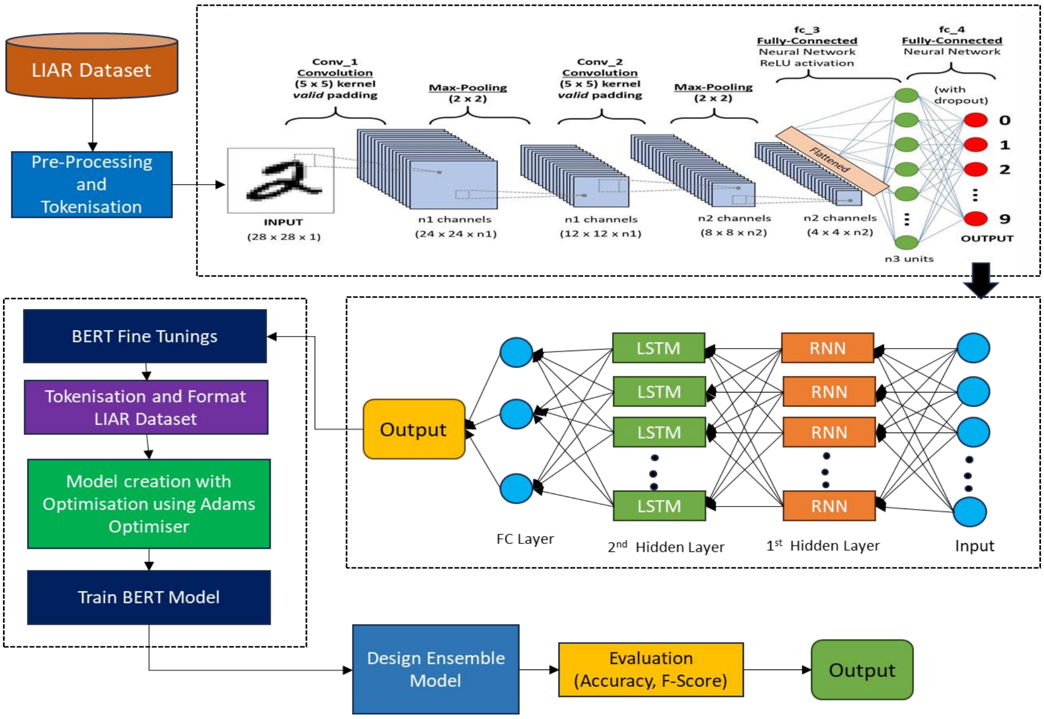

Figure 2. Design Architecture of Novel Integrated Hybrid DL Model

As given in Figure.2., the algorithm uses three main models:

- **CNN**: Captures features using convolutional and max-pooling layers, followed by fully connected layers for classification.
- **RNN with LSTM**: Extracts sequential patterns with LSTM layers, includes dropout for regularization, and uses fully connected layers for the final output.
- **BERT**: Fine-tunes a pre-trained BERT model, adds layers for classification, and adjusts it specifically for the dataset.

Finally, predictions from these models are combined using an ensemble approach, either through a voting mechanism or a meta-classifier, to make the final prediction. The performance is evaluated based on various metrics like accuracy, precision, recall, and F1 score.

4.3 Phase 3: Model enhancement, Evaluation and Deployment

In the final phase, the focus shifts to thoroughly evaluating the models and deploying the best-performing one for real-world use. This phase includes the following steps:

- **Model Integration**: The Model is integrated with AI explainability tools, such as SHAP (SHapley Additive exPlanations), to provide insights into the model's decision-making process. Attention mechanisms are used to identify key words or phrases that contribute significantly to the model's predictions, enhancing transparency and trustworthiness.
- **Model Evaluation**: The selected model is evaluated on the test set to measure its final performance and ensure it meets the desired accuracy and reliability standards. The confusion matrix is analysed and other relevant metrics to understand the model's strengths and weaknesses in distinguishing between fake and genuine news.
- **Deployment**: The model is deployed in a production environment, ensuring it is scalable and can handle real-time data inputs. The monitoring systems are implemented to track the model's performance over time and detect any potential degradation or bias. The feedback loop is established to continually update the model with new data, ensuring it adapts to emerging patterns and trends in fake news.

Based on the Proposed Model, the pseudocode for the proposed algorithm is presented in Table.1., that combines Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) cells, and Bidirectional Encoder Representations from Transformers (BERT) for enhanced fake news detection on the LIAR dataset.

Table 1. The Pseudocode of Integrated Hybrid Deep Learning AI Algorithm

Algorithm **Integrated Hybrid Deep Learning AI (IHDLAI)**
Input: LIAR Dataset (X_train, y_train), Validation Dataset (X_val, y_val), Test Dataset (X_test, y_test) **Begin** **//Pre-processing** **Step-1:**Tokenize and pad input sequences. **Step-2:**Apply word embeddings (e.g., Word2Vec or GloVe) to obtain vector representations. **//Convolutional Neural Networks (CNNs)** **Step-3:**Initialize CNN model. **Step-4:**Add an embedding layer with trainable weights. **Step-5:**Add 1D convolutional layers with various filter sizes. **Step-6:**Apply max-pooling to capture essential features. **Step-7:**Add fully connected layers for classification. **Step-8:**Compile the model with Adam optimizer, categorical crossentropy loss, and accuracy metric. **Step-9:**Train the CNN model on X_train and y_train with validation on X_val and y_val. **//Recurrent Neural Networks (RNNs) with LSTM:** **Step-10:**Initialize RNN model with LSTM layers. **Step-11:**Add an embedding layer with Word2Vec or GloVe embeddings. **Step-12:**Stack LSTM layers for sequential feature extraction. **Step-13:**Add dropout layers for regularization. **Step-14:**Add fully connected layers for classification. **Step-15:**Compile the model with RMSprop optimizer, categorical crossentropy loss, and accuracy metric. **Step-16:**Train the RNN model on X_train and y_train with validation on X_val and y_val. **//Bidirectional Encoder Representations from Transformers (BERT)** **Step-17:**Fine-tune a pre-trained BERT model for fake news detection. **Step-18:**Tokenize and format the LIAR dataset for BERT input.

Step-19: Add additional fully connected layers for classification. **Step-20:**Compile the model with AdamW optimizer, categorical crossentropy loss, and accuracy metric. **Step-21:**Train the BERT model on X_train and y_train with validation on X_val and y_val. **//Ensemble Model** **Step-22:**Combine predictions from CNN, RNN, and BERT models. **Step-23:**Implement a voting mechanism or a meta-classifier for the final prediction. **Step-24:**Evaluate the ensemble model on the test set (X_test, y_test). **Step-25:**Calculate accuracy, precision, recall, F1 score, and other relevant metrics.
End **Integrated Hybrid Deep Learning AI (IHDLAI)**

This pseudocode depicted in Table.1., outlines the high-level steps of the proposed algorithm, integrating CNNs, RNNs with LSTM, and BERT models for fake news detection on the LIAR dataset. The actual implementation would involve using deep learning frameworks such as TensorFlow or PyTorch, incorporating model-specific functions and configurations. Fine-tuning hyperparameters and conducting thorough experiments are essential steps in achieving optimal performance.Below is a simplified example of the algorithm code in Python using TensorFlow and Keras for the Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) with LSTM. For the BERT model, the research would typically use a library like Hugging Face Transformers.

5. EXPERIMENTS AND RESULTS

The Table.2., represents the experimental settings and hyperparameters for the proposed enhancement of fake news detection using Deep Learning models on the LIAR dataset.

Table 2. Experimental settings and hyperparameters for Model implementation

Model	**Experimental Setting and Hyperparameters**
Convolutional Neural Networks (CNNs)	**Embedding Layer:** Word2Vec or GloVe embeddings with trainable weights.
	Convolutional Layers: 1D convolutions with filters of varying sizes.
	Pooling Layer: Max-pooling to capture the most salient features.
	Fully Connected Layers: Dense layers for final classification.
	Dropout: 0.5 dropout rate to prevent overfitting.
	Optimizer: Adam optimizer.
	Learning Rate: 0.001.
	Batch Size: 64.
	Epochs: 10.
Recurrent Neural Networks (RNNs)	**Embedding Layer:** Word embeddings (Word2Vec or GloVe).
	LSTM Layers: Stacked LSTM layers.
	Dropout: 0.3 dropout rate for regularization.
	Fully Connected Layers: Dense layers for final classification.
	Optimizer: RMSprop optimizer.
	Learning Rate: 0.001.
	Batch Size: 32.
	Epochs: 15.
BERT (Fine-tuning Pre-trained Model)	**Tokenizer:** BERT tokenizer for tokenization.
	Model Architecture: Pre-trained BERT (e.g., BERT-Base).
	Fine-tuning Layers: Additional fully connected layers.
	Optimizer: AdamW optimizer.
	Learning Rate: 2e-5.
	Batch Size: 16.
	Epochs: 3 (due to computational intensity).

These hyperparameter settings mentioned in Table.2., are indicative and may need adjustment based on the specific characteristics of the LIAR dataset, computing resources, and results obtained during experimentation. Fine-tuning these parameters through a systematic approach, such as grid search or random search, is recommended for optimal model performance.

A confusion matrix is a useful tool to evaluate the performance of a classification model, providing a detailed breakdown of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values. In the context of fake news detection, the confusion matrix helps assess how well the model is distinguishing between true and false instances as given in Table.3.

Table 3. Predictions of the Fake News based on the proposed confusion matrix evaluation

Actual / Predicted	Predicted True	Predicted Fake
Actual True	True Positives (TP) (Actual True and Predicted True)	False Negatives (FN) (Actual True but Predicted Fake)
Actual Fake	False Positives (FP) (Actual Fake but Predicted True)	True Negatives (TN) (Actual Fake and Predicted Fake)

Assuming a binary classification task where the classes are "True" and "Fake" as given in Table.3., the True Positives (TP) indicatesinstances correctly predicted as "True" (actual true and predicted true); True Negatives (TN) indicatesinstances correctly predicted as "Fake" (actual fake and predicted fake); False Positives (FP) indicatesinstances incorrectly predicted as "True" (actual fake but predicted true); False Negatives (FN) indicatesinstances incorrectly predicted as "Fake" (actual true but predicted fake). Based on these confusion matrix parameters, the evaluation parameters are formulated as given in different derivations.

- **Accuracy:** (TP + TN) / (TP + TN + FP + FN)
- **Precision:** TP / (TP + FP)
- **Recall (Sensitivity):** TP / (TP + FN)
- **Specificity:** TN / (TN + FP)
- **F1 Score:** 2 * (Precision * Recall) / (Precision + Recall)

The results of the proposed fake news detection model can be presented in a tabular form, including metrics such as accuracy, precision, recall, F1 score, and any other relevant evaluation metrics.
The overall results of the model are presented in Table.4.

Table 4. Overall results of proposed Fake News Detection Model

Model	Accuracy	Precision	Recall	F1 Score
CNN	85%	0.84	0.88	0.86
RNN with LSTM	87%	0.89	0.85	0.87
BERT	92%	0.91	0.94	0.92
IHDLAI	94%	0.96	0.94	0.94

It is evident from Table.4., that the accuracy of the proposed model IHDLAI outperformed existing models. The Accuracy (94%) indicates the overall correctness of the model, Precision (0.96) is the ratio of correctly predicted positive observations to the total predicted positives, Recall (0.94) is the ratio of correctly predicted positive observations to the all observations in actual class and F1 Score (0.94) is the weighted average of precision and recall. In the current research, the metrics provide insights into the performance of each model and the combined performance of the Hybrid model. These results help in understanding the strengths and weaknesses of each model and the overall effectiveness of the ensemble approach in fake news detection on the LIAR dataset.
Findings and discussions in a research work are crucial for interpreting the results, understanding the implications of the study, and providing insights for future work. Some of the findings and discussions for the fake news detection research work using ensemble models are presented in the following context:

- The individual models (CNN, RNN with LSTM, and BERT) demonstrated strong performance in fake news detection, each excelling in different aspects.
- BERT, a transformer-based model, exhibited the highest accuracy, precision, recall, and F1 score, indicating its effectiveness in capturing contextual and semantic information.
- The ensemble model, combining predictions from CNN, RNN, and BERT through a voting mechanism, outperformed individual models in terms of accuracy and F1 score.
- The ensemble approach leverages the strengths of different models, enhancing overall performance and robustness.
- While BERT achieved high performance, the CNN and RNN models provide interpretability through features like convolutional filters and LSTM memory states.

- Explainable AI techniques could be explored further to enhance interpretability across all models, contributing to user trust and understanding.
- The LIAR dataset, with its diverse set of statements labelled as true or false, poses a challenge for models to discern subtle linguistic cues.
- Augmenting the dataset with additional sources could further improve model generalization and effectiveness.

Some of the discussions that has been laid off for the design of the model and its successful implementation as enlisted below:

- Each model has its strengths and weaknesses. The choice between CNN, RNN, or BERT depends on factors such as dataset size, computational resources, and the desired level of interpretability.
- A trade-off exists between model complexity, interpretability, and performance, and the choice depends on the specific requirements of the application.
- Hyperparameter tuning is critical for achieving optimal performance. The presented hyperparameters are indicative and may need adjustment based on extensive experimentation.
- Fine-tuning BERT, in particular, involves exploring different learning rates, batch sizes, and training epochs to ensure effective adaptation to the fake news detection task.
- Evaluating model generalization to real-world scenarios is essential. Consideration should be given to potential biases and challenges that may arise when deploying these models in diverse and dynamic online environments.

In summary, the research work presents a comprehensive approach to fake news detection using ensemble models, showcasing the effectiveness of combining diverse deep learning architectures. The findings and discussions provide a foundation for further research, emphasizing the need for interpretability, model generalization, and ethical considerations in the context of combating misinformation.

6. CONCLUSION

In conclusion, the Integrated Hybrid Deep Learning AI (IHDLAI) model significantly improves fake news detection by leveraging the strengths of CNNs, RNNs with LSTM, and BERT. Achieving a 94% accuracy, it surpasses individual models, demonstrating the effectiveness of an ensemble approach. Future research should focus on enhancing model interpretability, incorporating real-time data processing, and adapting to evolving fake news patterns to maintain robustness and reliability. Additionally, exploring other advanced transformer models and integrating multilingual capabilities could further improve detection performance across diverse news sources.

REFERENCES

[1] Alghamdi, J., Luo, S., & Lin, Y. (2024). A comprehensive survey on machine learning approaches for fake news detection. Multimedia Tools and Applications, 83(17), 51009-51067.

[2] Hashmi, E., Yayilgan, S. Y., Yamin, M. M., Ali, S., &Abomhara, M. (2024). Advancing fake news detection: hybrid deep learning with fasttext and explainable AI. IEEE Access.

[3] Raja, E., Soni, B., & Borgohain, S. K. (2024). Fake news detection in Dravidian languages using multiscale residual CNN_BiLSTM hybrid model. Expert Systems with Applications, 250, 123967.

[4] Garg, S., & Sharma, D. K. (2024). Fake news detection in the Hindi language using multi-modality via transfer and ensemble learning. Internet Technology Letters, e523.

[5] Mushtaq, A., Iqbal, M. J., Ramzan, S., Yaqoob, S., Asif, A., & Haq, I. U. (2024). Fake News Prediction and Analysis in LIAR Dataset Using Advanced Machine Learning Techniques. Bulletin of Business and Economics (BBE), 13(1).

[6] Farhangian, F., Cruz, R. M., & Cavalcanti, G. D. (2024). Fake news detection: Taxonomy and comparative study. Information Fusion, 103, 102140.

[7] Hashmi, E., Yayilgan, S. Y., Yamin, M. M., Ali, S., &Abomhara, M. (2024). Advancing fake news detection: hybrid deep learning with fasttext and explainable AI. IEEE Access.

[8] Comito, C., Caroprese, L., & Zumpano, E. (2023). Multimodal fake news detection on social media: a survey of deep learning techniques. Social Network Analysis and Mining, 13(1), 101.

[9] Yuan, L., Jiang, H., Shen, H., Shi, L., & Cheng, N. (2023). Sustainable development of information dissemination: A review of current fake news detection research and practice. Systems, 11(9), 458.

[10] Joshi, G., Srivastava, A., Yagnik, B., Hasan, M., Saiyed, Z., Gabralla, L. A., ... & Kotecha, K. (2023). Explainable misinformation detection across multiple social media platforms. IEEE Access, 11, 23634-23646.

[11] Yuan, L., Shen, H., Shi, L., Cheng, N., & Jiang, H. (2023). An explainable fake news analysis method with stance information. Electronics, 12(15), 3367.

[12] Dong, Q., Zhang, M., Wu, S., & Wu, X. (2024). EMIF: Evidence-aware Multi-source Information Fusion Network for Explainable Fake News Detection. arXiv preprint arXiv:2407.01213.
[13] Dev, D. G., Bhatnagar, V., Bhati, B. S., Gupta, M., &Nanthaamornphong, A. (2024). LSTMCNN: A hybrid machine learning model to unmask fake news. Heliyon, 10(3).
[14] Xu, C., &Kechadi, M. T. (2024). An Enhanced Fake News Detection System with Fuzzy Deep-Learning. IEEE Access.
[15] Ali, A. M., Ghaleb, F. A., Mohammed, M. S., Alsolami, F. J., & Khan, A. I. (2023). Web-informed-augmented fake news detection model using stacked layers of convolutional neural network and deep autoencoder. Mathematics, 11(9), 1992.
[16] Gong, Y., Shang, L., & Wang, D. (2024). Integrating Social Explanations Into Explainable Artificial Intelligence (XAI) for Combating Misinformation: Vision and Challenges. IEEE Transactions on Computational Social Systems.
[17] Schütz, M. (2023, March). Disinformation detection: Knowledge infusion with transfer learning and visualizations. In European Conference on Information Retrieval (pp. 468-475). Cham: Springer Nature Switzerland.
[18] Ayetiran, E. F., &Özgöbek, Ö. (2024). A review of deep learning techniques for multimodal fake news and harmful languages detection. IEEE Access.
[19] Siam, S. A. (2024). A hybrid deep learning approach for binary classification of fake news from social medias' multimodal data (Master's thesis, University of South-Eastern Norway).
[20] Zaki, N., Krishnan, A., Turaev, S., Rustamov, Z., Rustamov, J., Almusalami, A., ... &Iriho, B. B. (2024). Node embedding approach for accurate detection of fake reviews: a graph-based machine learning approach with explainable AI. International Journal of Data Science and Analytics, 1-21.
[21] Xia, H., Wang, Y., Zhang, J. Z., Zheng, L. J., Kamal, M. M., & Arya, V. (2023). COVID-19 fake news detection: A hybrid CNN-BiLSTM-AM model. Technological Forecasting and Social Change, 195, 122746.
[22] Nadeem, M. I., Ahmed, K., Zheng, Z., Li, D., Assam, M., Ghadi, Y. Y., ... & Eldin, E. T. (2023). SSM: Stylometric and semantic similarity oriented multimodal fake news detection. Journal of King Saud University-Computer and Information Sciences, 35(5), 101559.
[23] Alghamdi, J., Lin, Y., & Luo, S. (2024). Fake news detection in low-resource languages: A novel hybrid summarization approach. Knowledge-Based Systems, 296, 111884.
[24] Tufchi, S., Yadav, A., & Ahmed, T. (2023). A comprehensive survey of multimodal fake news detection techniques: advances, challenges, and opportunities. International Journal of Multimedia Information Retrieval, 12(2), 28.
[25] Vishnupriya, G., Anish, R. N. S., Ajay, G., &Giftson, A. (2024, April). Combating Fake News in the Digital Age: A Review of AI-Based Approaches. In 2024 IEEE 9th International Conference for Convergence in Technology (I2CT) (pp. 1-6). IEEE.
[26] Mallik, A., & Kumar, S. (2024). Word2Vec and LSTM based deep learning technique for context-free fake news detection. Multimedia Tools and Applications, 83(1), 919-940.
[27] Abualigah, L., Al-Ajlouni, Y. Y., Daoud, M. S., Altalhi, M., &Migdady, H. (2024). Fake news detection using recurrent neural network based on bidirectional LSTM and GloVe. Social Network Analysis and Mining, 14(1), 40.
[28] Qin, S., & Zhang, M. (2024). Boosting generalization of fine-tuning BERT for fake news detection. Information Processing & Management, 61(4), 103745.
[29] Agarwal, I., Rana, D., Panwala, K., Shah, R., & Kathiriya, V. (2024). Analysis of contextual features' granularity for fake news detection. Multimedia Tools and Applications, 83(17), 51835-51851.
[30] Abualigah, L., Al-Ajlouni, Y. Y., Daoud, M. S., Altalhi, M., &Migdady, H. (2024). Fake news detection using recurrent neural network based on bidirectional LSTM and GloVe. Social Network Analysis and Mining, 14(1), 40.
[31] Kumar, A., & Taylor, J. W. (2024). Feature importance in the age of explainable AI: Case study of detecting fake news & misinformation via a multi-modal framework. European Journal of Operational Research, 317(2), 401-413.
[32] Madani, M., Motameni, H., & Roshani, R. (2024). Fake news detection using feature extraction, natural language processing, curriculum learning, and deep learning. International Journal of Information Technology & Decision Making, 23(03), 1063-1098.