Eksplorium *p*-ISSN 0854-1418
Volume 46 No. 2, June 2025: 1042–1051 *e*-ISSN 2503-426X

A Discriminative AI-Based Sequential Classification Model for Fake News Detection using BERT

Deepali Goyal Dev
Research Scholar, GGSIPU, New Delhi, India
*E-mail: ddepali@gmail.com

Article Received: 10 May 2025, Revised: 15 June 2025, Accepted: 24 June 2025

ABSTRACT

In today's world, there is a widespread flow of data and information due to immense usage of web applications. The usage of information available on web is difficult to be categorised and identified as true or Fake for a normal user. This creates a sort of doubt and confusion in the mind of normal users as how to verify and authenticate the news in the social media. This is essential in today's modern-day scenario as regulatory bodies do wants and enforce that the true news or information should be further circulated or propagated rather than fake news which at times is not known to common user and he may be mis leaded by fake news. Public confidence in widely disseminated news has been undermined by the ease with which disinformation may now reach large audiences due to immense growth of social media and online platforms. To overcome such situation, we need a well-defined system which ensure the identification of fake and true news with highest accuracy. The authors proposed a text-based sequential classification model for detecting fake news using Bidirectional Encoder Representations from Transformers (BERT). The model proposed is able to clearly classify the news as fake or true with enhanced accuracy and precision. The model will be able to accurately identify the Fake news and will benefit the society by helping to clearly distinguish between fake and true news. The model is built on usage of transformer model which has ability to learn from complex relationship in label data and enhances the performance of predicted model with faster training time due to parallel processing.

Keywords-: BERT, Fake News, True News, Transformers, Accuracy

1. Introduction

The development of technology has significantly altered the way that people use media to receive, disseminate, and engage with information [1]. Social media platforms have become central to news dissemination [2], enabling direct interaction between content and users while granting widespread access to information [3]. However, this increased accessibility has led to a significant drawback of rise and spread of false information and unverified stories, commonly referred to as "fake news" [4]. In addition to the production of false information, social media enable the occurrence of information overload in which the system environment is filled with inferior, unnecessary or deliberately misleading information [5]. This kind of data overload deforms users' perception of accurate sources of information and contributes to confusion, distrust, and contaminated information environment [6]. The prevalence of such contaminated materials only complicates the tasks of personal and systemic proactivity toward the sanctity of cyberspace discussion [7]. The phenomenon of false information has become most acute in social media as this platform is based on user-generated content and is predicated on algorithms that drive interactions that garner 'likes,' shares, and comments [8]. These algorithms intended to capture user attention further exacerbate sensational and misleading news, allowing the proliferation of false information [9]. Prior methods attempt to detect fake news using conventional architectures of the Machine Learning (ML) framework, mostly with an emphasis on feature engineering [10]. These approaches employ using handcrafted features that include language, metadata, reposting rates, etc., for classification models including Naïve Bayes (NB), SVM, and decision trees. These are also explainable methods, and fast in terms of computation and amenable to dealing with small samples as well [11]. They are especially useful in early assessment where the amount of information is relatively small. A major disadvantage is that the features used for model selection are usually domain dependent which restricts the usability of the method. Furthermore, they are poor in analysing higher order structures in multiscale data and hence have limited detection capacity for subtle misinformation. For example, the SVM model might run quite satisfactorily by analysing the headlines of fake news through n-gram detection of linguistic absurdities but lag when facing context-sensitive or multimodal inputs. Consequently, NLP techniques have enhanced fake news detection with an improved ability to analyse written text with regards to its semantic content [12]. These methods use NLP techniques for analysing syntax, semantics, as well as sentiment in news articles. The linguistic features of fake news can be identified by methods like sentiment analysis, TF-IDF, and Bag of Words [13]. Advanced techniques involve latent topic modelling (e.g., LDA) and dependency parsing to understand hidden structures in text. NLP approaches provide for efficient identification of the stylistic and semantic features of the text; distinguishing between truth and lies. Such relators are particularly useful when working with language pairs when fine differences are crucial. Still, common problems exist for NLP techniques based on their inability to work with contextually rich or multimodal fake news for example memes and videos. Furthermore, these methods have problems with deliberate undermining of intent and use of ambiguity and sarcasm. Advanced developments like the context-aware embeddings (Word2Vec, GloVe) have refined the NLP systems knowledge about word interactions [14]. However, the creation of models that can achieve human-like understanding of intent and meaning still pose a significant problem. DL approaches have significantly transformed fake news detection because it makes it possible for the model to extract abstract and high-level features from an extensive amount of data [15]. While in conventional methods feature extraction happens manually or required a lot of engineering, DL basically performs this task automatically. CNNs detect fake news based on images, while CNNs detect sequential textual data using RNN, including LSTM. Pseudo Transformed models including BERT and GPT have enhanced the evaluation of context and semantics in the fake news text [16]. Large end-to-end DL models perform well in identification of multimodal fake news, taking into account textual, image, and user interaction inputs. They perform with the best accuracy on large sample data and react to modifications in the dissemination of misleading information. One of the issues with DL is its huge demand for computational power and often significant annotated datasets. Furthermore, as most of these models use a black-box approach to decision-making, their interpretability continues to be a significant concern. For example, BERT-based models have gaining high accuracy in capturing systematic text features and correlations. Author proposes a BERT Sequential model for detection of fake news using the ISOT Dataset [28]. The dataset comprises real and fake news articles, enabling binary classification. Our approach involves data preprocessing, tokenization, model fine-tuning, and evaluation, achieving near-perfect classification performance. The results underscore the potential of BERT in high-stakes NLP applications.

2. Related Work

The social networks have revolutionized information dissemination, consumption, and use in a significant manner. It is platforms through which a user can socialize in a new unprecedented form but at a price of opening avenues for dissemination of fake news and disinformation that are information shared with a view of manipulating fellow users for social motives. Fake news detection has evolved over the past years. Early approaches relied on lexical and syntactic features, such as TF-IDF, n-grams, and part-of-speech tags, used with classifiers like SVM, logistic regression, and decision trees [17]. These methods achieved moderate success but struggled with semantic complexity and context-dependent patterns. Improvements were brought about by deep learning models that captured local dependencies and sequential patterns, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks [18]. However, their unidirectional processing limits their ability to understand full contextual relationships, particularly in lengthy news articles. The introduction of BERT [19] marked a paradigm shift in NLP. BERT's bidirectional architecture and performing tasks like Masked Language Modeling (MLM) and Next Sentence Prediction (NSP) [20] enable it to capture rich contextual representations, making it highly effective for text classification. Studies like [21] have applied BERT to fake news detection, reporting accuracies above 95% on datasets like ISOT [28]. Other transformer models, such as RoBERT, DistilBERT, and XLNet, have also been explored, but BERT remains a robust baseline due to its balance of performance, accessibility, and extensive community support. Recent work has investigated multimodal approaches, incorporating images and metadata [22], but text-based models remain dominant due to their simplicity and effectiveness. Our work builds on these advancements, providing a detailed implementation of BERT for fake news detection, with a focus on reproducibility, scalability, and practical deployment.

_p_-ISSN 0854-1418
_e_-ISSN 2503-426X

3. Methodology

3.1 Dataset

The ISOT Fake News Dataset [28] contains 21,417 real and 23,481 fake news articles collected from various sources. Real articles are sourced from reputable outlets, while fake articles are derived from unreliable websites. Each article includes a title, text, subject, and date. For this study, we combine the title and text fields and assign binary labels (1 for real, 0 for fake). We focus on the title and text fields, assigning binary labels (1 for real, 0 for fake) to create a unified dataset of 44,898 samples. The dataset's near-balanced distribution (48% real, 52% fake) makes it suitable for binary classification without requiring oversampling or under sampling. Figure 1 shows distribution of real and fake news. Figure 2 shows distribution of text length by labels and figure 3 shows top 20 most frequent word used in the dataset.

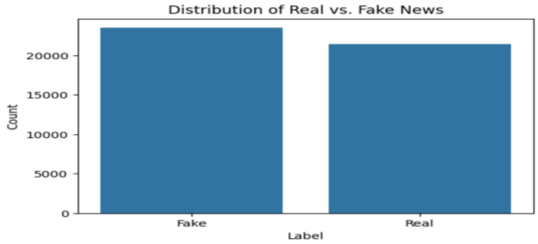

Figure 1: Distribution of Real vs. Fake News

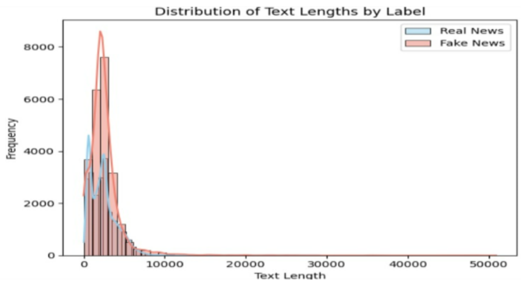

Figure 2: Distribution of Text Lengths by Label

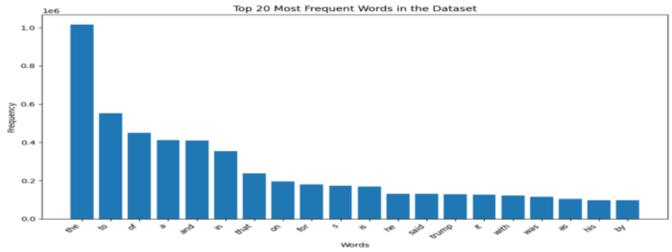

Figure 3: Top 20 Most Frequent Words in Dataset

3.2 Data Preprocessing

Preprocessing ensures data quality and compatibility with BERT. The steps include:

3.2.1 Data Consolidation: Load True.csv and Fake.csv into pandas data frames, assign labels (1 for real, 0 for fake), concatenate the data frames, and shuffle with a random seed (42) for reproducibility.

3.2.2 Text Cleaning: Combine title and text fields, handling missing values by filling with empty strings. Apply a cleaning function to remove URLs (regex r"http\S+"), special characters (retaining alphanumeric characters and spaces), and convert text to lowercase to standardize input.

3.2.3 Dataset Splitting: Perform stratified splitting to create training (80%, 35,918 samples), validation (10%, 4,489 samples), and test (10%, 4,489 samples) sets, ensuring balanced class distribution to prevent bias.

3.3Tokenization and Dataset Preparation

Authors used the BertTokenizerFast from the Hugging Face Transformers library, based on bert-base-uncased, with a maximum sequence length of 256 tokens to balance context retention and computational efficiency. Texts are tokenized with truncation for longer texts and padding for shorter texts, adding special tokens ([CLS], [SEP]) to mark sequence boundaries. The tokenized data is converted into TensorFlow datasets with a batch size of 16, shuffled with a buffer size of 2048 to prevent overfitting, and optimized with prefetching to reduce I/O bottlenecks.

3.4 Model Architecture

The bert-base-uncased model, a transformer-based architecture created by Devlin et al. [23], is utilized by the suggested fake news detection system. In contrast to unidirectional models like LSTMs or word2vec, BERT, or Bidirectional Encoder Representations from Transformers, is made to process text bidirectionally in order to capture rich contextual representations. In false news identification, where stylistic and semantic distinctions (such as sensationalist language and factual errors) are crucial signs, BERT's bidirectional methodology allows it to comprehend the context of each word based on both its previous and succeeding terms. Each of the 12 transformer layers (L=12) in the bert-base-uncased model has 12 attention heads (H=12) and a hidden size of 768 (D=768), for a total of about 110 million parameters. Every transformer layer uses a self-attention mechanism with many heads [24], defined as:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) \cdot V \tag{1}$$

where each attention head's dimension is dk = 64, and the query, key, and value matrices are Q, K, and V, respectively, obtained from input embeddings [24]. Multi-head attention captures a variety of associations between tokens by concatenating findings from 12 heads. A feed-forward neural network (FFN) with a ReLU activation and two linear transformations is included in each layer [24]:

$$FFN(x) = max(0, xW_1 + b_1)W_2 + b_2 \tag{2}$$

where b1 and b2 are biases and W1 and W2 are weight matrices [24]. Layer normalization and residual connections stabilize training. BERT is pre-trained on Masked Language Modeling (MLM), predicting 15% of masked tokens, and Next Sentence Prediction (NSP), learning sentence relationships, using large corpora (e.g., Wikipedia, BookCorpus). This pretraining equips BERT with generalizable language understanding, which we fine-tune for binary classification. We use the TFBert- ForSequenceClassification class from the Hugging Face Transformers library, configured for binary classification (num_labels=2). The model takes tokenized input sequences (maximum length 256 tokens) and outputs logits for two classes: real (label 1) and fake (label 0). The [CLS] token's final hidden state (768-dimensional vector) serves as the aggregate representation, passed through a classification head comprising:

- A dropout layer (p=0.1) to prevent overfitting.
- A linear layer mapping the 768-dimensional vector to a 2-dimensional output.
- A softmax activation to produce class probabilities [23]:

$$P(class|x) = softmax(W \cdot h_{[CLS]} + b) \tag{3}$$

where $h_{[CLS]}$ is the [CLS] token's hidden state, W is the weight matrix, and b is the bias [23].

The choice of bert-base-uncased balances computational efficiency and performance. The uncased model normalizes text to lowercase and removes accents, reducing the vocabulary size to 30,522 tokens, suitable for the ISOT dataset where case sensitivity is less critical. The 12-layer architecture captures nuanced linguistic patterns (e.g., exaggerated phrasing in fake news) without overfitting on the 44,898-sample dataset. Compared to bert-large (24 layers, 340M parameters), bert-base-uncased is more computationally manageable on standard GPUs (e.g., NVIDIA Tesla V100). Fine-tuning adjusts the pre-trained weights using sparse categorical cross entropy loss and the AdamW optimizer (learning rate 2e-5), optimizing for fake news detection. Table 1 shows hyperparameter values. Figure 4 shows proposed architecture.

Table 1: Hyperparameters values and experiments setting [21]

Hyperparameter	Value
Batch size	16
Learning rate	2×10^{-5}
Weight decay	0.0
Epoch count	3
Warmup ratio	0.10 (10 % of total training steps)
Hidden Activation	GELU (in each Transformer layer)
Classification Activation	Softmax (over final dense logits)
Loss function	Sparse Categorical Crossentropy (from logits)

Hyperparameter	Value
Optimizer	AdamW (with linear warmup & decay schedule)
Dropout value	0.10 (default BERT dropout on the pooled output)

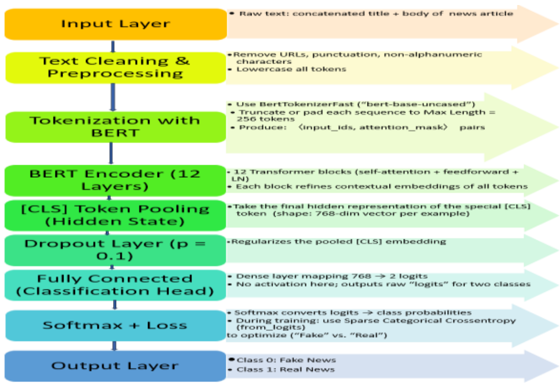

Figure 4: Proposed Methodology

3.5 Algorithm

The proposed algorithm is outlined below, with a corresponding flowchart in Figure 5.

Algorithm 1: Fake News Detection Using BERT

1: *Input: True.csv, Fake.csv*
2: *Output: Trained BERT model, Test accuracy, Classification report*
3: *Load True.csv and Fake.csv into df_true, df_fake*
4: *Assign labels: df_true["label"] = 1, df_fake["label"] = 0*
5: *Concatenate and shuffle: df = pd.concat([df_true, df_fake]).sample(frac=1, random_state=42)*
6: *Clean text: Remove URLs, special characters, convert to lowercase*
7: *Split dataset: 80% train, 10% validation, 10% test (stratified)*
8: *Load BERT tokenizer (bert-base-uncased)*
9: *Tokenize texts with max length 256, create TensorFlow datasets (batch_size=16)*
10: *Load BERT model (bert-base-uncased, num_labels=2)*
11: *Initialize AdamW optimizer (lr=2e-5, steps=len(train)//16*3, warmup_steps=steps//10)*
12: *Compile model (sparse categorical crossentropy, accuracy)*
13: *Train model for 3 epochs*
14: *Evaluate on test set, compute predictions*
15: *Calculate accuracy and classification report*
16: *Return: Model, accuracy, report*

Figure 5: BERT Fine Tuning Pipeline Flow Diagram

4. Results & Discussion

4.1 Evaluation Results

The following parameters have been used to evaluate the performance of BERT Sequential model [24] [25].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F_1 = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \quad (7)$$

In the above equations, A positive class's total correct predictions are denoted by TP, a negative class's total correct predictions by TN, a positive class's total wrong predictions by FP, and a negative class's total incorrect predictions by FN.

4.2 Discussion

We as authors presents the ideal four metrics (accuracy, F1-score, recall, and precision) on the ISOT datasets being as textual data to assess our BERT-based sequence-classification model. These findings are reported for the ISOT in Table 2. In existing literature, it is evident that significant amount of domain-specific pretraining is used while in our proposed work, Bert-base-uncased, the first step in our pipeline, has already been trained on billions of words getting more pre trained data which is ideally suited for better sequential classification. In order to obtain high accuracy, authors fine-tune for only three epochs as opposed to 150 epochs (usually evident in literature) as BERT's contextual embeddings naturally catch rich linguistic patterns. This results in significant cut in overall computation time and lowers the risk of overfitting. The batch size of the data in such a computational requirement is of significant importance. As compared to existing literature where usage of 256-batch size is common resulting in requirement of large amount of GPU RAM and frequently requires raising the learning rate or decreasing dropout in order to prevent overfitting. However, authors used a batch size of 16 to achieves a better balance and high accuracy. Smaller batches improve the model's ability to generalize, which is crucial when data sources are spread across several domains. Our proposed model has dropout of 0.1 as it employs BERT's internal dropout of 0.1 as compared to a custom dropout of 0.2(as per available literature), which may impede learning or necessitate more cautious LR tweaking. For classification problems, BERT's default of 0.1 has already received extensive validation and is adequate for regularization without impairing convergence.

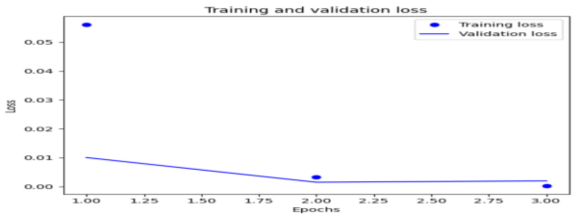

Figure 6(a): Training and Validation Loss

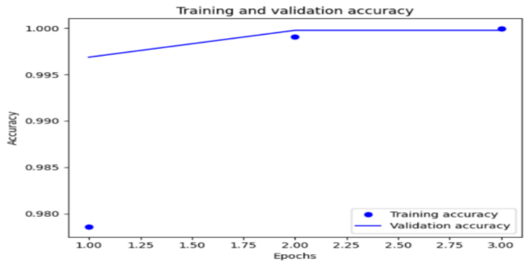

Figure 6(b): Training and Validation Accuracy

Figure 6 (a) and (b) shows, As the three epochs progress, the training and validation losses both gradually drop until they are almost zero by the third epoch. In a similar vein, training and validation accuracy increase from around 98% to about 99.9%, suggesting robust convergence and little overfitting.

Table 2: Result of Proposed Model on ISOT dataset

Class	Precision	Recall	F1-Score	Support
Fake (0)	0.9996	0.9996	0.9996	2347
Real (1)	0.9995	0.9995	0.9995	2141
Accuracy			0.9996	4490
Macro Avg	0.9996	0.9996	0.9996	4490
Weighted Avg	0.9996	0.9996	0.9996	4490

Table 2 On the ISOT test set, the proposed BERT-based model achieves nearly perfect discrimination, with precision, recall, and F1-scores of 0.9995–0.9996 for both the Fake and actual classes over 2,348 fake and 2,142 actual data. Overall accuracy is 0.9996, and both weighted and macro-averages support this finding, showing very consistent performance free from noticeable bias or class imbalance. Figure 7 shows confusion matrix of BERT Model.

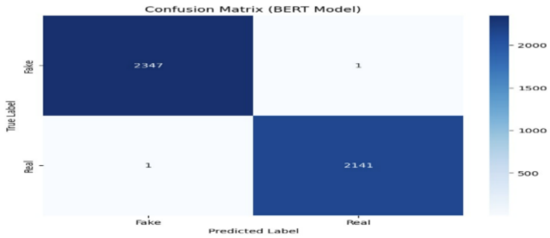

Figure 7: Confusion Matrix of BERT Model

Table 3 shows comparison of BERT Sequential model with BiLSTM (Partition), CNN-BiLSTM (Partition), CNN-PCNNBiLSTM-ATT (Partition), BiLSTM (Partition + Overlap), CNN-BiLSTM (Partition + Overlap), CNN-PCNNBiLSTM-ATT (Partition + Overlap), BiLSTM (No Partition), CNN-BiLSTM (No Partition) and CNN-PCNNBiLSTM-ATT (No Partition) [21]. Out of 2,348 fraudulent cases, 2,347 were correctly identified as fake, and only one was mislabelled as real, according to the confusion matrix. Similarly, out of 2,142 actual examples, 2,141 were correctly identified, with only one misclassification. The nearly flawless diagonal dominance demonstrates that the BERT model identifies phony news from real news on the ISOT test set with almost no errors.

Table 3 : Comparison of proposed BERT model with reference [21]

Model	Accuracy	Recall	Precision	F1-Score
BiLSTM (Partition)	94.14	93.02	93.1	94.03
CNN-BiLSTM (Partition)	95.92	94.85	93.93	95.91
CNN-PCNNBiLSTM-ATT (Partition)	98.32	98.17	98.23	98.33
BiLSTM (Partition + Overlap)	95.9	94.57	94.51	95.88
CNN-BiLSTM (Partition + Overlap)	96.43	95.32	95.2	96.41
CNN-PCNNBiLSTM-ATT (Partition + Overlap)	99.47	99.35	99.11	99.26
BiLSTM (No Partition)	90.38	90.01	90.07	90.3
CNN-BiLSTM (No Partition)	92.85	91.48	91.3	92.83
CNN-PCNNBiLSTM-ATT (No Partition)	95.88	94.97	94.99	95.86
BERT + Sequential	**99.96**	**99.95**	**99.96**	**99.96**

Figure 8, 9, 10 and 11 shows the graphical representation of comparison of accuracy, Recall, Precision and F1-Score with reference [21]. The table makes it very evident that using overlapping chunks and more complex structures continuously improve performance. The simple BiLSTM achieves 94.14% accuracy under the "Partition" scheme, which increases to 95.92% with the addition of a CNN front-end and to 98.32% with the full CNN-PCNNBiLSTM-ATT model. By adding overlaps (also known as "Partition + Overlap"), the network is able to acquire boundary information that non-overlapping chunks might miss. For example, CNN-PCNNBiLSTM-ATT with overlap achieves 99.47% accuracy. On the other hand, the model must truncate or process extremely long texts all at once when no partitioning is used ("No Partition"), which results in lower scores (95.88% accuracy for CNN-PCNNBiLSTM-ATT). Lastly, our suggested BERT + Sequential method achieves 99.96% accuracy (99.95% recall, 99.96% precision, 99.96% F1), significantly outperforming all prior combinations. This indicates that BERT has nearly perfect discrimination on this dataset by using global self-attention over each token in each (potentially chunked) input.

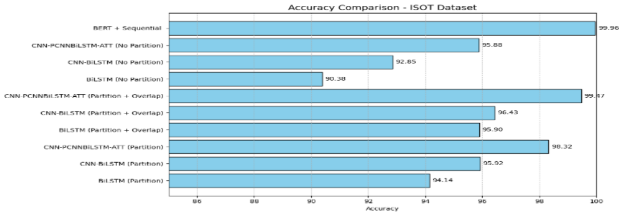

Figure 8: Comparison of Accuracy of BERT Model

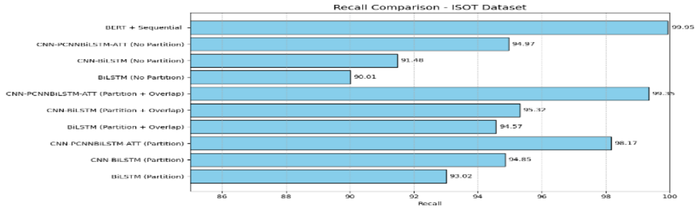

Figure 9: Comparison of Recall of BERT Model

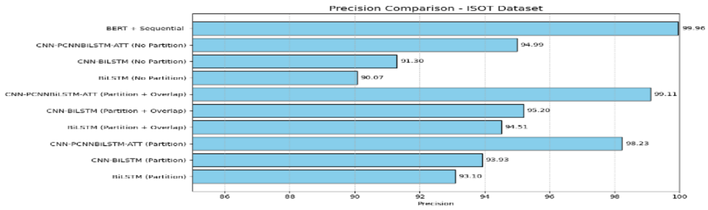

Figure 10: Comparison of Precision of BERT Model

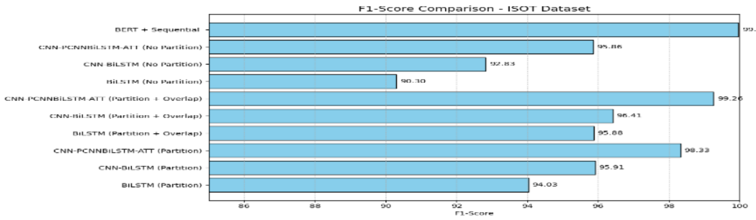

Figure 11: Comparison of F1-Score of BERT Model

The suggested model attains the best performance on the ISOT dataset, somewhat surpassing previous research ([27] 2024 and [21] 2025) with an accuracy and F1-score of 99.96%. In contrast to the accuracy of 99.95% and 99.27% reported by [27] and [21], our method shows a slight but significant gain in precision and recall, suggesting that the improvements (e.g., BERT's global context and adjusted hyperparameters) produce nearly flawless classification consistency. Table 4 shows comparison with baseline studies.

Table 4: Comparison with Baseline studies

Reference	Year	Data Set	Accuracy	Recall	Precision	F1-Score
[27]	2024	ISOT	99.95	99.95	99.95	99.95
[21]	2025	ISOT	99.27	99.35	99.11	99.26
Proposed	-	ISOT	99.96	99.95	99.96	99.96

Conclusion & Future Work

CONCLUSION

The work proposed by authors applied a deep learning method for identifying false news items using BERT (Bidirectional Encoder Representations from Transformers). The dataset of ISOT Fake News was pre-processed and optimized using a BERT-based model which resulted in test accuracy of 99.96% with precision, recall, and F1-scores above 99.96% for both classes that is true and fake news. The model performed exceptionally well. This study by authors offers a scalable methodology for detecting fake news and demonstrates the effectiveness of transformer-based models in thwarting disinformation.

FUTURE WORK

The work shown by the authors can be further extended by creating a "two-fold" training set that specifically compares and contrasts easy and challenging cases Instead of just combining all actual and all false articles. The social media data is always a streaming data which requires on real time analysis on such data to find as whether it is fake or true news which require high level of dynamic news feed to be analysed requiring an efficient model trained and validated on live feeds.

References:

[1] X. Zhang and A. A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," *Information Processing & Management*, 57, 102025, 2020.

[2] Y. Papanastasiou, "Fake news propagation and detection: A sequential model," *ManagSement Science*, 66, 1826–1846, 2020.

[3] U. Mertoğlu and B. Genç, "Automated fake news detection in the age of digital libraries," *Information Technology and Libraries*, 39, 1–19, 2020.

[4] M. K. Elhadad, K. F. Li, and F. Gebali, "Detecting misleading information on COVID-19," *IEEE Access*, 8, 165201–165215, 2020.

[5] M. Spradling, J. Straub, and J. Strong, "Protection from 'fake news': The need for descriptive factual labeling for online content," *Future Internet*, 13, 142, 2021.

[6] Z. Bastick, "Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation," *Computers in Human Behavior*, 116, 106633, 2021.

[7] M. Choraś, M. Pawlicki, R. Kozik, K. Demestichas, P. Kosmides, and M. Gupta, "Socialtruth project approach to online disinformation (fake news) detection and mitigation," *Proceedings of the 14th International Conference on Availability, Reliability and Security*, 68, 1–10, 2019.

[8] W. Shahid, B. Jamshidi, S. Hakak, H. Isah, W. Z. Khan, M. K. Khan, and K. K. Choo, "Detecting and mitigating the dissemination of fake news: Challenges and future research opportunities," *IEEE Transactions on Computational Social Systems*, 11, 4649–4662, 2022.

[9] S. Tufchi, A. Yadav, and T. Ahmed, "A comprehensive survey of multimodal fake news detection techniques: Advances, challenges, and opportunities," *International Journal of Multimedia Information Retrieval*, 12, 28, 2023.

[10] J. Alghamdi, S. Luo, and Y. Lin, "A comprehensive survey on machine learning approaches for fake news detection," *Multimedia Tools and Applications*, 83, 51009–51067, 2024.

[11] A. Altamimi, "Novel approach for predicting fake news stance detection using large word embedding blending and customized CNN model," *PLoS ONE*, 19, e0314174, 2024.

[12] M. Nadeem, P. Abbas, W. Zhang, S. Rafique, and S. Iqbal, "Enhancing Fake News Detection with a Hybrid NLP-Machine Learning Framework," *IECE Transactions on Intelligent Systematics*, 1, 203–214, 2024.

[13] Z. Khanam, B. N. Alwasel, H. Sirafi, and M. Rashid, "Fake news detection using machine learning approaches," *IOP Conference Series: Materials Science and Engineering*, 1099, 012040, 2021.

[14] V. I. Ilie, C. O. Truică, E. S. Apostol, and A. Paschke, "Context-aware misinformation detection: A benchmark of deep learning architectures using word embeddings," *IEEE Access*, 9, 162122–162146, 2021.

[15] C. K. Hiramath and G. C. Deshpande, "Fake news detection using deep learning techniques," *2019 1st International Conference on Advances in Information Technology (ICAIT)*, 411–415, 2019.

[16] P. Bharadwaj and Z. Shao, "Fake news detection with semantic features and text mining," *International Journal on Natural Language Computing*, 8, 17–22, 2019.

[17] H. Ahmed, I. Traore, and S. Saad, "Detecting Opinion Spams and Fake News Using Text Classification," *Security and Privacy*, 1, e9, 2017.

[18] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newsletter*, 19, 22–36, 2017.

[19] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake News Detection in Social Media with a BERT-Based Deep Learning Approach," *Multimedia Tools and Applications*, 80, 11765–11788, 2021.

[20] M. Shubhangi, A.Swati, B. Akanksha, S. Dilip, "Ethical and Societal Implication of Sentiment Analysis using NLP in Educational Feedback System," Journal of Information Systems Engineering and Management, 10 (15s), 742-749, 2025.

[21] I. Q. Abduljaleel and I. H. Ali, "Detecting Fake News Using BERT Word Embedding, Attention Mechanism, Partition and Overlapping Text Techniques," *TEM Journal*, 14, 1152-1165, 2025.

[22] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, and J. Gao, "EANN: Event adversarial neural networks for multi-modal fake news detection," *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 849–857, 2018.

[23] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1, 4171–4186, 2019.

[24] X. Song, Q. Niu, J. Liu, B. Peng, S. Zhang, M. Liu, … Z. Yuan, "Transformer: A Survey and Application", 2024. https://doi.org/10.31219/osf.io/5p2hu

[25] Elaf Ali Abbood and T. A. Al-Assadi, "A New Convolution Neural Layer Based on Weights Constraints," pp. 7–13, Nov. 2022, doi: https://doi.org/10.1109/icdsic56987.2022.10076083.

[26] L. abed Al. Hamza, H. A. Lafta, and S. Z. Al Rashid, "Exploring Predictive Models Utilizing Machine Learning and Deep Learning Techniques for Diabetes Mellitus: A Comprehensive Literature Review," *Lecture Notes in Networks and Systems*, pp. 615–631, Oct. 2024, doi: https://doi.org/10.1007/978-981-97-2671-4_46.

[27] Iman Qays Abduljaleel and I. H. Ali, "Deep Learning and Fusion Mechanism-based Multimodal Fake News Detection Methodologies: A Review," *Engineering Technology & Applied Science Research*, vol. 14, no. 4, pp. 15665–15675, Aug. 2024, doi: https://doi.org/10.48084/etasr.7907.

[28] DataSet : https://www.kaggle.com/datasets/aadiola/isot-fake-news-dataset