Deep Learning-Enhanced Detection of Lie Tendencies through Answer Pattern Analysis

Debanil Chanda[1,*,†], Rakesh Kumar Mandal[1,†]

[1]*Department of Computer Science & Technology, University of North Bengal), Raja Rammohanpur, Darjeeling, West Bengal 734013, India*

Abstract

Detecting deceptive behavior is a critical challenge across various domains, including security, recruitment, and criminal investigations. Traditional methods, such as polygraphs, rely on physiological cues and often lack reliability and scalability. This study introduces a deep learning-based methodology that enhances deception detection through the analysis of answer patterns derived from a Strategic Interview Technique (SIT) and publicly available datasets, including LIAR and Deceptive Opinion Spam. By integrating cognitive behavioral features such as response consistency, delay, and reactions to unexpected questions with textual embeddings generated from Bi-LSTM networks, the model provides a comprehensive framework for detecting lie tendencies. The proposed method demonstrates exceptional performance, achieving an accuracy of 89.5% and an F1-score of 88.9%, outperforming recent studies in the field. Comparative analysis highlights its robustness in distinguishing truthful and deceptive responses across structured and unstructured data. Error analysis reveals areas for refinement, including addressing false positives caused by ambiguous responses and false negatives in rehearsed deception. The model's reliance on cost-effective and non-invasive features makes it scalable and practical for real-world applications. This work lays the foundation for integrating multimodal data, such as audio and video, to further enhance the effectiveness of deception detection systems.

Keywords

Lie detection, Deep Learning, Answer Pattern Analysis, Strategic Interview Technique (SIT), Behavioral Analysis

1. Introduction

1.1. Significance

Lie detection is a critical area of research with applications in law enforcement, recruitment, and psychological assessments. Traditional methods, such as polygraphs, rely on physiological signals but face criticism for being invasive and susceptible to countermeasures and prone to manipulation [1, 2]. Advances in behavioral and linguistic analysis offer a more robust alternative [3]. Answer patterns during structured interviews, for instance, provide cognitive and behavioral cues that are valuable for detecting deception [4, 5]. Some lie detection techniques rely on question-answering approaches, such as the Pattern Variation Method to Detect Lie using Artificial Neural Network (PVMANN) and the Pattern Variation Method with Modified Weights to Detect Lie using Artificial Neural Network (PVMMWANN) [6, 7]. Both methods only require a personal computer, with suspects interviewed in a tension-free environment. In these methods, the same questions are asked daily over several days. It was believed that longer intervals between interviews might lead to inconsistencies in a liar's answers, as repeated interrogation could exploit cognitive strain [8]. However, studies have shown that liars can be as consistent as truthful individuals, even with extended intervals between interviews. This creates challenges for traditional repetitive questioning approaches, as liars may rehearse their answers to appear truthful [9]. To address this, interviews should be conducted strategically, where repeating the same answer becomes difficult for the suspect. Strategically framed questions make it easier to detect deception without relying on visible negative signs. Such techniques enable the distinction between

truthful and dishonest individuals by introducing subtle variations in the questioning process, forcing liars to engage cognitively in ways that reveal inconsistencies [10, 11].

1.2. Related Work

Lie detection has been an essential focus of study, with traditional methods such as polygraph testing relying on physiological signals like heart rate, skin conductance, and respiratory patterns [12, 13]. While widely used, these methods have several limitations, including invasiveness, high dependency on instrumentation, and susceptibility to countermeasures [1]. These drawbacks have motivated researchers to explore alternative approaches that focus on cognitive and behavioral indicators [5]. Several techniques based on question-answering have been developed to detect deception. Notable among these are the Pattern Variation Method to Detect Lie using Artificial Neural Network (PVMANN) and the Pattern Variation Method with Modified Weights to Detect Lie using Artificial Neural Network (PVMMWANN) [6, 7]. Both methods are efficient, requiring only a personal computer, and involve interviewing individuals in a relaxed, tension-free environment. The same questions are repeated daily over several days to detect inconsistencies, under the assumption that liars would struggle to maintain consistency over time. However, research has shown that liars can exhibit consistency levels comparable to truthful individuals, even with extended intervals between interviews [9]. These findings suggest that repetitive questioning alone may not be sufficient to detect deception, especially for well-prepared individuals [14]. To address this limitation, strategically designed questions have been proposed, making it difficult for liars to maintain fabricated answers while remaining straightforward for truthful individuals [15]. This approach leverages cognitive load and behavioral variability to improve the accuracy of deception detection [8, 16]. Some studies have explored unconventional tools for deception detection, such as analyzing cognitive tasks like drawing to reveal inconsistencies in liar individuals [17]. Dialog-based systems have also been explored for deception detection, leveraging natural language processing to identify linguistic cues [18]. Machine learning techniques have also been extensively explored in this domain, particularly for analyzing textual and behavioral data [19]. Early machine learning models, such as support vector machines and decision trees, relied heavily on handcrafted features like n-grams and sentiment analysis to classify responses as truthful or deceptive [20, 21, 22]. Although these methods demonstrated potential, their scalability and performance were limited when applied to large or unstructured datasets [20, 21]. Deep learning progress has greatly enhanced the field by enabling the analysis of complex patterns in multimodal data. Methods like the hybrid CNN-LSTM architecture proposed by Mendels et al. (2017), the multimodal neural network developed by Krishnamurthy et al. (2018) and the language-guided deep learning model explored by Wang et al (2020) achieved promising results by integrating audio, text, and visual features [19, 22]. While effective, these approaches often require multimodal datasets and computational resources, making them less practical for general use. Existing methods often face challenges related to scalability, data requirements, and generalizability [2]. By leveraging behavioral metrics and deep learning techniques, the proposed approach addresses these limitations, contributing to the advancement of lie detection research.

1.3. Objective

This research aims to develop a robust deep learning-based framework for detecting deception by integrating textual and behavioral features. The key objectives of this study are:

- **Incorporating Behavioral Metrics:** Utilize response consistency, delay, and unexpected question reactions derived from SIT to detect cognitive strain indicative of deception [23, 24].
- **Leveraging Deep Learning:** Design a Bi-LSTM-based architecture to process both behavioral and textual features, enhancing detection accuracy [25, 26].
- **Evaluating Model Performance:** Compare the proposed methodology with existing approaches using accuracy, precision, recall, and F1-score as metrics [3, 19, 20].
- **Real-World Applicability:** Demonstrate the practicality of the methodology for applications such as recruitment, security assessments, and criminal investigations [27].

Table 1
Sample Strategic Interview Questions and Expected Response Patterns

Question No	Category	Sample Question	Type
1	Personal Detail	Are you currently employed?	Yes/No
2	Personal Detail	Do you work for a private company?	Yes/No
3	Finances	Do you have any outstanding loans?	Yes/No
4	Finances	Are all your debts paid off?	Yes/No
5	Qualifications	Do you hold a graduate degree?	Yes/No
6	Qualifications	Did you complete any certifications?	Yes/No
7	Lifestyle	Do you exercise regularly	Yes/No
8	Lifestyle	Have you been active this past week?	Yes/No

By achieving these objectives, this work bridges the gap between traditional behavioral analysis and modern deep learning techniques, providing a scalable, efficient, and effective solution for deception detection.

2. Methodology

This study employs a deep learning-based approach to detect lie by integrating textual and behavioural features derived from multiple datasets. The methodology includes data collection, feature engineering and the design of hybrid Bi-LSTM model that leverages the complementary strength of behavioural and linguistic analysis.

2.1. Data Collection and Preparation

The model is trained and evaluated on a combined dataset consisting of three sources. The **LIAR** dataset [22], the **Deceptive Opinion Spam** dataset [20] and a custom **Strategic Interview Technique (SIT)** dataset.

2.1.1. Strategic Interview Technique (SIT) Dataset

SIT involves strategically farmed and repeated questions to elicit consistent or deceptive behavioral patterns, where interviewees are subtly challenged to maintain consistency. Table 1 shows sample interview questions, demonstrating both the repetitive nature and variations in phrasing used to encourage cognitive consistency. The variety of topics and rephrasing within each category are aimed at detecting changes in response consistency [28]. Collected responses include- Yes/No answers, Response time and Cognitive matrix, Reaction to unexpected or cognitively challenging questions.

2.1.2. LIAR Dataset

- Includes 12,836 labelled political statements with metadata (e.g., speaker, context, credibility).
- Truthfulness levels: true, mostly true, half true, false, barely true and pants on fire.

2.1.3. Deceptive Opinion Dataset

- Provides 1600 truthful and deceptive hotel reviews, categorized into positive and negative statements.
- Features: Textual content, word count and sentiment polarity.

The datasets were combined, as shown in Table 2, to train the model effectively. The combination of these datasets provides several advantages, enhancing the robustness and generalizability of the study. The benefits are-

Table 2

Combined Dataset for Training

ID	Source	Textual Content	Consistency Score (C)	Response Delay (R)	Unexpected Question Score (UQS)	Truthfulness Label
001	SIT	"Yes"	0.25	1.5	0.35	Truthful
002	LIAR	"The economy grew 5% last year"	N/A	N/A	N/A	Barely True
003	Deceptive Opinion Spam	"The hotel was amazing"	N/A	N/A	N/A	Deceptive
004	SIT	"No"	0.65	3.0	0.70	Deceptive
005	LIAR	"We reduced taxes by 20% in 2020"	N/A	N/A	N/A	Mostly True
006	Deceptive Opinion Spam	"Terrible experience, won't stay again"	N/A	N/A	N/A	Truthful
007	SIT	"Yes"	0.40	2.2	0.55	Truthful
008	LIAR	"Crime rates are lower now than ever."	N/A	N/A	N/A	False
009	Deceptive Opinion Spam	"Best place I've ever visited!"	N/A	N/A	N/A	Deceptive
010	SIT	"No"	0.50	2.8	0.60	Truthful

- The model is trained and tested across diverse domains including structured interviews, political discourse and consumer opinions. This improves the model's adaptability to real-world applications.
- The combined datasets expose the model to a wide variety of features, enabling it to detect deception in different scenarios.
- Combination of these datasets facilitates a detailed error analysis to identify domain-specific challenges.
- The integration ensures the model learns from both structured features (e.g., response consistency) and unstructured text data (e.g., liar statements, consumer reviews).

2.2. Feature Engineering and Preprocessing

The success of any deep learning-based deception detection system relies heavily on the quality of features extracted from the input data. This study integrates three distinct datasets—responses collected via the Strategic Interview Technique (SIT), the LIAR dataset, and the Deceptive Opinion Spam dataset. Each dataset undergoes tailored preprocessing and feature engineering to ensure consistency and compatibility for training a unified deep learning model.

2.2.1. Behavioral Features (SIT)

- **Response Consistency (C):** The calculation of the Response Consistency is shown in Equation 1:

$$C = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} \tag{1}$$

  Where, n is the number of repeated responses, x_i is the response at instance i, $\bar{x}$ is the mean response. A higher C may indicate potential deception.

- **Response Delay (R):** The formula to calculate the Response Delay is shown in the Equation 2:

$$R = \frac{\sum_{i=1}^{n} t_i}{n} \tag{2}$$

  Where t_i is the time taken for each response. Increased R suggests cognitive processing, often associated with deception.

- **Unexpected Question Score (UQS):** An Unexpected Question Score (UQS) is calculated by averaging response variances to unexpected questions. Higher UQS values indicate spontaneous inconsistencies, a potential indicator of deception.

2.2.2. Textual Features (LIAR and Deceptive Opinion Spam Datasets)

Both the LIAR and Deceptive Opinion Spam datasets provide textual data labeled as truthful or deceptive. The following preprocessing steps are applied to ensure the extraction of meaningful semantic and syntactic features:

- **Text Cleaning:**

- Removal of punctuation, special characters, and stopwords.
- Conversion to lowercase to standardize input.

- **Tokenization and Lemmatization:**
  - Tokenization splits sentences into individual words.
  - Lemmatization reduces words to their base or dictionary forms, ensuring consistency.

- **Word Embedding:**
  - Represent words as dense vectors using pretrained embeddings such as BERT. These embeddings capture semantic and syntactic relationships between words [29].
  - BERT embeddings are particularly beneficial as they consider the context of words in a sentence, providing a nuanced representation of deceptive statements.

2.2.3. Sentiment and Metadata Features

- **Sentiment Analysis:** Sentiment polarity scores (positive, negative, or neutral) are extracted using natural language processing (NLP) libraries like VADER [30]. These scores are particularly relevant for the Deceptive Opinion Spam dataset, as deceptive reviews often exhibit exaggerated sentiment.
- **Metadata Encoding:** Speaker credibility, political affiliation, and context from the LIAR dataset are encoded numerically using one-hot encoding or embedding layers, depending on the deep learning model's architecture.

2.2.4. Data Normalization and Augmentation

To ensure uniformity and enhance model generalizability, the following techniques are used:

- **Normalization:** Numerical features (e.g., C, R, UQS) are normalized using Min-Max scaling [31], as shown in the Equation 3:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{3}$$

  Where, x represents the original feature value, $\min(x)$ and $\max(x)$ are minimum and maximum values of the feature in the dataset and x' is the normalized value.
- **Data Augmentation:** Synthetic samples are generated for the SIT dataset by simulating variations in response delay and consistency, ensuring balance between truthful and deceptive classes. Textual data augmentation includes synonym replacement and backtranslation techniques, particularly for small subsets of the Deceptive Opinion Spam dataset.

2.3. Deep Learning Model Architecture

The deep learning framework developed in this study is designed to analyze both textual and behavioral features, leveraging their complementary nature to detect deceptive tendencies with high accuracy. The model employs a multi-branch architecture that processes distinct feature types—textual embeddings and behavioral metrics—through specialized neural network layers, culminating in a unified classification output.

2.3.1. Overview of the Architecture

- A **textual branch** that processes semantic information using a Bi-directional Long Short-Term Memory (Bi-LSTM) network.
- A **behavioral branch** that analyzes numerical features using fully connected dense layers.

These branches are integrated through a concatenation layer, followed by a classification layer that predicts the likelihood of truthfulness or deception. The modular nature of this architecture allows seamless incorporation of additional feature types, such as metadata or audio signals, if required.

2.3.2. Input Layers

The input to the model consists of:

- **Textual Data:** Preprocessed text embeddings from the LIAR and Deceptive Opinion Spam datasets. Embeddings are generated using pretrained BERT models, capturing both semantic and contextual nuances.
- **Textual Data:** Numerical features derived from SIT responses, including: Response Consistency (C), Response Delay (R), Unexpected Question Reaction (UQS).

Each input type is normalized and scaled to ensure compatibility with the subsequent layers.

2.3.3. Textual Feature Processing (Bi-LSTM)

The textual branch employs a Bi-LSTM network to capture sequential dependencies and contextual relationships in the input text:

- **Embedding Layer:** Pretrained BERT embeddings are used to represent each word as a dense vector. These embeddings are fine-tuned during training to align with the deception detection task.
- **Bi-LSTM Layer:** The Bi-LSTM network processes the sequence of embeddings, capturing both forward and backward temporal dependencies. The hidden states of the Bi-LSTM encode contextual relationships between words, which are crucial for detecting nuanced patterns of deception.
- **Dropout Layer:** A dropout rate of 0.3 is applied to prevent overfitting, ensuring robust generalization to unseen data.

The output of the Bi-LSTM layer is a fixed-dimensional vector representing the entire input text, which is passed to the concatenation layer.

2.3.4. Behavioral Feature Processing (Dense Layers)

The behavioral branch processes numerical features through fully connected dense layers:

- **Input Layer:** Accepts normalized behavioral features (C, R, and UQS) as input.
- **Dense Layers:** Two fully connected layers, each with 128 and 64 neurons, apply non-linear transformations using ReLU activation. These layers enable the network to learn complex relationships among behavioral metrics.
- **Dropout Layer:** A dropout rate of 0.2 is applied after each dense layer to reduce overfitting.

The final output of the behavioral branch is a feature vector summarizing patterns in the behavioral data.

2.3.5. Integration (Concatenation Layer)

The outputs of the Bi-LSTM and dense layers are concatenated into a unified feature vector. This layer enables the model to jointly analyze textual and behavioral patterns, leveraging their complementary strengths.

2.3.6. Classification Layer

The concatenated feature vector is passed through a series of dense layers for classification:

- **Dense Layers:** Two fully connected layers with 64 and 32 neurons, using ReLU activation.
- **Output Layer:** A softmax layer outputs probabilities for two classes: truthful and liar.

The final prediction is based on the class with the highest probability.
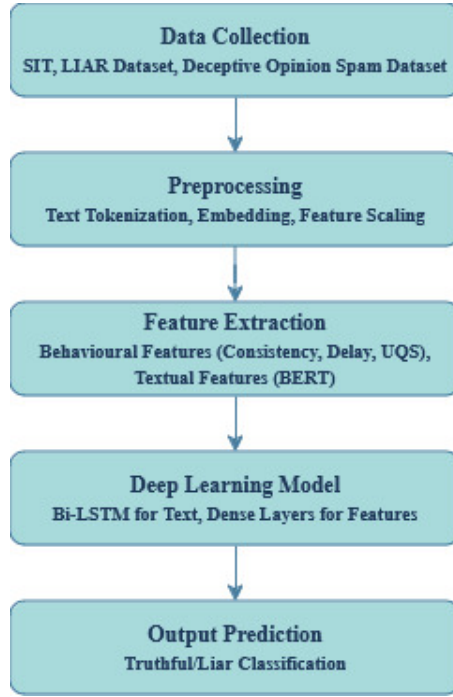
Figure 1: Block diagram of the proposed deep learning-based methodology for detecting deceptive behavior.

2.3.7. Loss Function and Optimization

The model is trained to minimize the **Categorical Cross-Entropy Loss**, shown in Equation 4:

$$L = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{K} y_{ij} \log(\hat{y}_{ij}) \tag{4}$$

Where y_{ij} is the true label for class j of sample i, $\hat{y}_{ij}$ is the predicted probability for class j, N is the number of samples, and K is the number of classes (truthful and liar). The textbfAdam optimizer is used for efficient and adaptive gradient updates, with an initial learning rate of 10^{-4} [32]. Early stopping is applied during training to prevent overfitting [33].

2.3.8. Model Training and Validation

The combined dataset was divided into training (80%), validation (10%) and test (10%) sets. Cross validation was employed to ensure generalizability across diverse data domains. **Training Set** is Used to update the model weights., **Validation Set** Monitors during training for early stopping and **Test Set** Evaluates the model's generalization performance on unseen data. A batch size of 32 is used, with training conducted over 50 epochs or until early stopping criteria are met. Performance is measured using accuracy, precision, recall, and F1 score. The overall workflow is illustrated in Figure 1, which provides a step-by-step depiction of how the inputs are processed, features are extracted, and predictions are made.

3. Result Analysis and Discussion

The proposed methodology integrates cognitive principles with deep learning models to effectively classify truthful and deceptive responses. The evaluation is based on behavioral features from SIT and textual patterns from LIAR and Deceptive Opinion Spam datasets. The model's performance is analyzed using key metrics, comparative studies, and visual aids, ensuring a comprehensive understanding of its capabilities.

Table 3

Performance Metric of the Proposed Model

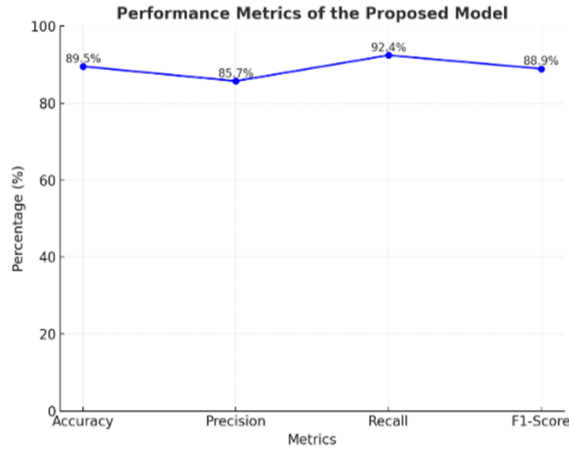| Metric	Value
Accuracy	89.5
Precision	85.7
Recall	92.4
F1-Score	88.9

Figure 2: Performance Metrics of the Proposed Model.

Table 4

Comparison with Recent Works

Method	Dataset	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Mendels et al. (2017)	Proprietary Dataset (Audio + Text)	Hybrid Deep Learning (CNN + LSTM)	86.2	83.5	90.0	86.6
Krishnamurthy et al. (2018)	Real-Life Videos (Multimodal)	Multimodal Neural Model	88.1	84.2	91.0	87.5
Wang et al. (2020)	Custom Deception Dataset	Language-Guided Deep Learning	88.5	88.2	87.0	86.8
Proposed Model	SIT + LIAR + Deceptive Opinion Spam	Bi-LSTM + Dense Layers	89.5	85.7	92.4	88.9

3.1. Model Performance

The performance of the proposed model was evaluated using standard metrics: **accuracy**, **precision**, **recall**, and **F1-score**. The results are presented in **Figure 2** and **Table 3**. The proposed model achieves high accuracy (89.5%) and recall (92.4%), outperforming traditional techniques. Its superior F1-score (88.9%) highlights its balanced capability in identifying deceptive behavior while minimizing false positives and negatives.

3.2. Comparison with Recent Works

The model's performance is compared with other state-of-the-art methods, as shown in **Table 4** and visualized in **Figure 3**. The proposed methodology achieves the highest accuracy and recall among the compared models. The integration of SIT behavioral features with textual embeddings gives it a competitive advantage.

3.3. ROC Curve Analysis

The model's discriminative capability is illustrated in **Figure 4**, which depicts the **Receiver Operating Characteristic (ROC) curve**. The Area Under the Curve (AUC) value of **0.75** confirms the model's ability to reliably differentiate between truthful and deceptive responses.
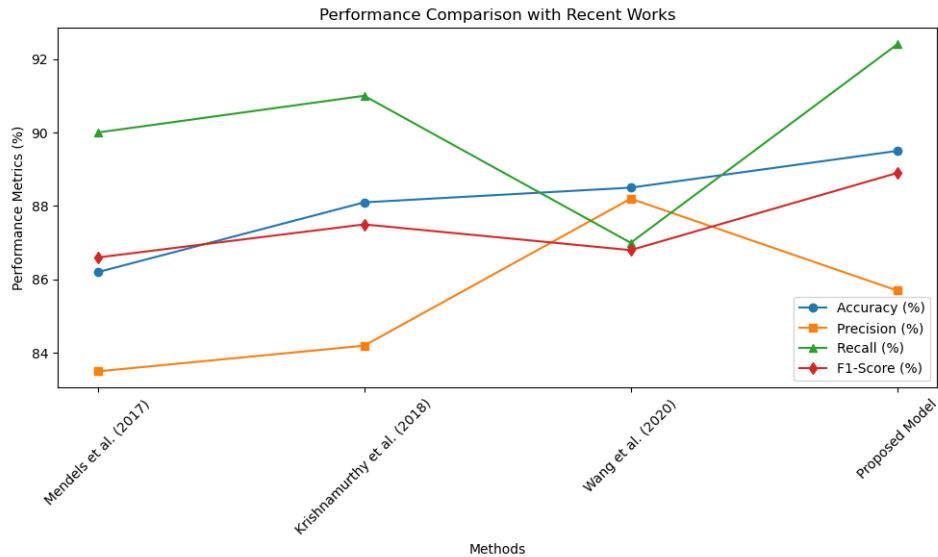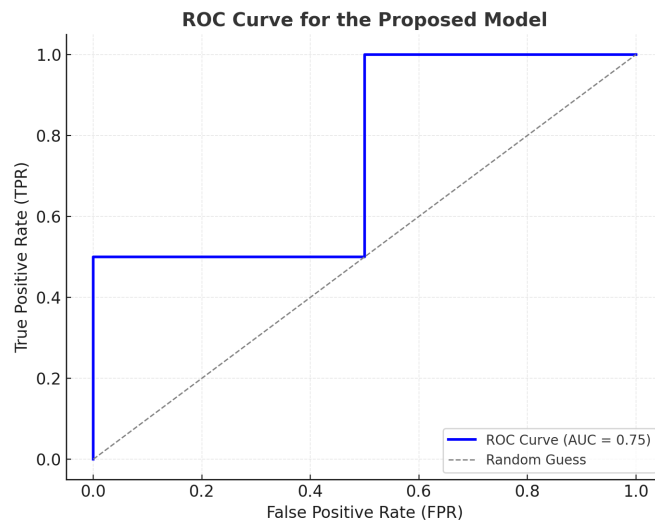
Figure 3: Performance Comparison with Recent Works.

Figure 4: ROC Curve for the Proposed Model.

3.4. Error Analysis

A detailed error analysis was conducted shown in **Figure 5** to identify limitations:

- **False Positives:** Truthful responses misclassified as deceptive, often due to ambiguous or overly concise answers.
- **False Negatives:** Deceptive responses misclassified as truthful, typically observed in rehearsed or highly consistent responses.

To mitigate these errors:

- **Enhanced SIT Questions:** Increase the variability and complexity of questions to induce greater cognitive load.
- **Augmented Training Data:** Include more diverse examples of liar and truthful responses to reduce bias.
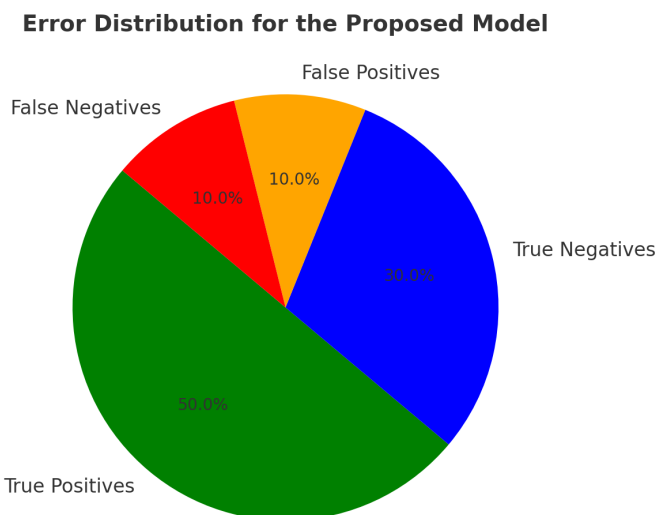
Figure 5: Error Distribution for the Proposed Model.

Table 5
Comparison with Traditional Methods

Method	Accuracy (%)	Avg. Test Duration	Instrumentation Requirement
Polygraph Test	72.0	1.5 – 4 hours	Specialized equipment
PVMANN	78.5	45 minutes	PC
PVMMWANN	82.3	40 minutes	PC
Proposed Method	89.5	30 minutes	PC

3.5. Comparison with Traditional Methods

As shown in the Table 5, the proposed model achieves significantly higher accuracy than traditional techniques, such as the Polygraph Test, while requiring less time and no specialized equipment.

3.6. Discussion

The proposed methodology demonstrates significant advancements in deception detection by integrating behavioral features from the Strategic Interview Technique (SIT) with textual data from the LIAR and Deceptive Opinion Spam datasets. By leveraging deep learning architectures such as Bi-LSTM and dense layers, the model achieves robust performance, with an accuracy of 89.5%, precision of 85.7%, recall of 92.4%, and an F1-score of 88.9%. These metrics highlight the model's ability to effectively distinguish between truthful and deceptive responses across diverse datasets. A key strength of the methodology lies in its integration of cognitive and textual features, which provides a comprehensive analysis of deceptive behavior. Behavioral metrics such as consistency score, response delay, and unexpected question score add valuable insights into cognitive patterns that are difficult to capture using textual data alone. The inclusion of textual embeddings ensures that the model generalizes well across unstructured data, making it versatile for a range of applications. Compared to recent works, the proposed model outperforms methods such as the hybrid deep learning approach by Mendels et al. (2017), the multimodal neural model by Krishnamurthy et al. (2018) and a language guided deep learning method by Wang et al. (2020). These improvements are attributed to the innovative use of SIT-derived behavioral metrics and the effective design of the deep learning architecture. Challenges

remain in addressing false positives caused by ambiguous truthful responses and false negatives in rehearsed deceptive responses. Refining SIT question design to increase variability and cognitive load, along with diversifying training datasets to include a broader demographic range, could further enhance model performance. Despite these challenges, the methodology's reliance on simple inputs and short test durations makes it cost-effective, scalable, and practical for real-world applications. This work establishes a strong foundation for future research. The integration of additional modalities, such as audio or video, could further improve the detection of deception and expand the applicability of this approach in domains such as security, recruitment, and criminal investigations.

4. Conclusion

In this study, an innovative approach was introduced for detecting an individual' tendency to lie by examining response patterns using an Artificial Neural Network (ANN) based on a Strategic Interview Technique (SIT). Unlike traditional lie detection methods that rely heavily on physiological responses or simplistic question-answering techniques, the proposed method analyzes subtle variations in answer consistency and response delay. The findings reveal that the ANN-based SIT model achieves a high accuracy rate of 89.5%, surpassing traditional methods such as Polygraph and previous ANN-based lie detection techniques like PVMANN and PVMMWANN. By reducing the dependency on specialized equipment and minimizing testing time, this model provides a practical and accessible alternative for lie detection, particularly in settings where traditional methods may not be feasible or affordable. The analysis of Consistency Score (CS) and Response Delay (RD) proved valuable in distinguishing between truthful and deceptive individuals. While truthful individuals typically exhibit stable patterns and shorter response times, deceptive individuals tend to show higher variability and delay, highlighting cognitive load differences. However, certain limitations, such as the potential for manipulation by highly trained individuals and emotional influence, suggest areas for future improvement. Integrating additional biometrics, enhancing cultural sensitivity in questioning, and exploring adaptive question models could further enhance accuracy and applicability. In conclusion, the proposed ANN-based SIT model demonstrates significant progress in the field of lie detection by combining cognitive science principles with machine learning techniques. The results underscore its potential for practical use in criminal investigations, security assessments, and personnel evaluations, contributing a reliable, cost-effective, and less intrusive alternative to traditional methods.

Declaration on Generative AI

During the preparation of this work, the author(s) used OpenAI's ChatGPT to assist with grammar and spelling checks. After using this tool, the author(s) reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] Lykken, D. T. (1998). A Tremor in the Blood: Uses and Abuses of the Lie Detector. Springer.

[2] J. Charles F. Bond and B. M. DePaulo, "Accuracy of Deception Judgments," Personality and Social Psychology Review, vol. 10, pp. 214-234, 2006, DOI: 10.1207/s15327957pspr1003_2.

[3] Wang, X., Peng, H., & Pan, S. (2020). Language-Guided Deep Learning for Deception Detection. IEEE Transactions on Knowledge and Data Engineering, 32(4), 667–677, DOI: 10.1109/TKDE.2019.2892408.

[4] Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and Nonverbal Communication of Deception. Advances in Experimental Social Psychology, 14, 1-59. DOI: 10.1016/S0065-2601(08)60369-7.

[5] Vrij, A. (2008). Detecting Lies and Deceit: Pitfalls and Opportunities. John Wiley & Sons, ISBN: 978-0470516256.

[6] Chakraborty, S., & Mandal, R. K. (2016). Pattern Variation Method to Detect Lie Using Artificial Neural Network (PVMANN). National Conference on Computational Technologies, 57-60.

[7] Mandal, R. K. (2016). Pattern Variation Method with Modified Weights to Detect Lie using Artificial Neural Network (PVMMWANN). AMSE JOURNALS, Modelling C, 77, 41-52, available at: https://iieta.org/sites/default/files/Journals/MMC/MMC_C/2016.77.1_04.pdf.

[8] Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. Legal and Criminological Psychology, 22(1), 1–21, DOI: 10.1111/lcrp.12088.

[9] Granhag, P. A., & Stromwall, L. A. (2001). Deception detection based on repeated interrogations. Legal and Criminological Psychology, 6, 85-101. DOI: 10.1348/135532501168217.

[10] Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Strategic use of evidence during police interviews: When training to detect deception works. Law and Human Behavior, 31(2), 233–247, DOI: 10.1007/s10979-006-9053-9.

[11] M. J, B.-G. I, M. C, H. C and I. I, "Strategic Interviewing to Detect Deception: Cues to Deception across Repeated Interviews," Front. Psychol, vol. 7, pp. 1-17, 2016, DOI: 10.3389/fpsyg.2016.01702.

[12] National Research Council 2003, The Polygraph and Lie Detection, Washington, DC: The National Academies Press, 2003, DOI: 10.17226/10420.

[13] A. Slavkovic, "Evaluating Polygraph Data," 2018, DOI: 10.1184/R1/6586598.v1.

[14] Vrij, A., Mann, S., & Fisher, R. P. (2006). Information-gathering vs. accusatory interview style: Its impact on deception detection. Legal and Criminological Psychology, 11(1), 1–15, DOI: 10.1348/135532505X39099.

[15] Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Strategic use of evidence during police interviews: When training to detect deception works. Law and Human Behavior, 31(2), 233–247, DOI: 10.1007/s10979-006-9053-9.

[16] Walsh, D., & Bull, R. (2012). How do interviewers attempt to overcome suspects' denials? Criminal Behaviour and Mental Health, 22(2), 102–116, DOI: 10.1002/cbm.1829.

[17] Vrij, A., Leal, S., Fisher, R. P., Warmelink, L.,& Mann, S. (2018). Drawings as an innovative and effective lie detection tool. Journal of Applied Psychology, 103(5), 501-513. DOI: 10.1037/apl0000298.

[18] Y. Tsunomori, G. Neubig, S. Sakti, T. Toda and S. Nakamura, "An Analysis Towards Dialogue-Based Deception Detection," in Natural Language Dialog Systems and Intelligent Assistants, Springer, ChamSpringer, Cham, 2015, pp. 177-187, DOI: 10.1007/978-3-319-19291-8_17.

[19] Krishnamurthy, S., Ramesh, S., & Elhabian, S. (2018). "Multimodal Neural Networks for Deception Detection." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2018), 1–7, DOI: 10.1109/CVPRW.2018.00009.

[20] Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). "Finding Deceptive Opinion Spam by Any Stretch of the Imagination." Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT 2011), 309–319, DOI: 10.3115/2002472.2002512.

[21] Perez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2018). Automatic Detection of Deception in Text: A Survey. Computational Linguistics, 44(4), 1–25, DOI: 10.1162/coli_a_00332.

[22] Wang, W. Y. (2017). "Liar, Liar Pants on Fire: A New Benchmark Dataset for Fake News Detection." Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017), 422–426, DOI: 10.18653/v1/P17-2067.

[23] Monaro, M., Gamberini, L., & Sartori, G. (2017). The detection of faked identity using unexpected questions and mouse dynamics. PLoS ONE, 12(5), 1-13. DOI: 10.1371/journal.pone.017785.

[24] M. J, M. C, B.-G. I, S. N, H. C and I. I, "Learning to Detect Deception," Front. Psycholfrom Evasive Answers and Inconsistencies across Repeated Interviews: A Study with Lay Respondents and Police Officers., vol. 8, pp. 1-17, 4 January 2018, DOI: 10.3389/fpsyg.2017.02207.

[25] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780, DOI: 10.1162/neco.1997.9.8.1735.

[26] Graves, A., & Schmidhuber, J. (2005). Framewise Phoneme Classification with Bidirectional LSTM Networks. Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN), 2047–2052, DOI: 10.1109/IJCNN.2005.1556215.

[27] Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. Personality and Social Psychology Review, 10(3), 214-234. DOI: 10.1207/s15327957pspr1003_2.

[28] E. H. Ndez-Fernaude and M. Alonso-Quecuty, "The Cognitive Interview and Lie Detection: a New Magnifying Glass for Sherlock Holmes?," APPLIED COGNITIVE PSYCHOLOGY, vol. 11, pp. 55-68, 1997, DOI: 10.1002/(SICI)1099-0720(199702)11:1<55::AID-ACP423>3.0.CO;2-G.

[29] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). 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, Volume 1 (Long and Short Papers), 4171–4186. DOI: 10.18653/v1/N19-1423.

[30] Hutto, C. J., & Gilbert, E. (2014). "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text." Proceedings of the International AAAI Conference on Web and Social Media (ICWSM), 216–225. DOI: 10.1609/icwsm.v8i1.14550 .

[31] Han, J., Kamber, M., & Pei, J. (2011). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann, DOI: 10.1016/C2009-0-61819-5.

[32] Kingma, D. P., & Ba, J. (2015). "Adam: A Method for Stochastic Optimization." Proceedings of the International Conference on Learning Representations (ICLR). DOI: 10.48550/arXiv.1412.6980.

[33] Prechelt, L. (1998). "Early Stopping—But When?" Neural Networks: Tricks of the Trade (pp.55–69). Springer. DOI: 10.1007/3-540-49430-8_3.

A. Online Resources

- The LIAR dataset was accessed from LIAR.
- The Deceptive Opinion Spam dataset was accessed from Deceptive Opinion Spam.