Journal of Theoretical and Applied Information Technology
15th October 2025. Vol.103. No.19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMPROVED BINARY DIFFERENTIAL EVOLUTION BASED CLASSIFIER SELECTION IN STACKED ENSEMBLE FRAMEWORK FOR EFFECTIVE DECEPTION SYSTEM

VIJAYASREE BODDU[1] , PRAKASH KODALI[2]

[1]Research Scholar, Department of ECE, NIT Warangal, Telangana, India

[2]Assistant Professor, Department of ECE, NIT Warangal, Telangana, India

E-mail: [1]bv22ecr1r04@student.nitw.ac.in, [2]kprakash@nitw.ac.in

ABSTRACT

Deception is a state that refers to acting in a way that causes another person to believe something that is not true. Deception is a national security concern when investigating crimes. Accurate deception system is a critical challenge in criminal analysis necessitating the development of efficient predictive models. Despite advancements in ensemble-based machine learning models, selecting diverse classifiers to enhance model performance remains a significant hurdle. Past efforts to boost classification accuracy using ensemble learning encountered constraints. To enhance deception performance in this study a two level stacking framework is proposed. Selecting best classifier combination for this stacking framework is a hard problem. This problem is formulated as combinatorial optimization problem and attempted to solve using Binary Differential Evolution (BDE) algorithm. For effective solution and better results, in the proposed approach, base learners are encoded using binary encoding, while meta learners are encoded using one-hot encoding. Further, the proposed BDE uses dynamic mutation scaling factor and cross over rate. Finally, the continuous solution space is converted using sigmoid transfer function followed by thresholding. In this study nine diverse base learners and four meta learners are used to construct the stacking ensemble model. The proposed framework optimizes the combination of these classifiers using BDE, aiming to improve predictive performance. Our optimization process relies on a fitness function derived from the ensemble accuracy score and number of classifiers. A Concealed Information Test (CIT) is performed to collect the deception dataset. This dataset is used to evaluate the proposed model. The proposed model outperformed in terms of accuracy, sensitivity, specificity and F1-score when compared with State-Of-The-Art (SOTA) models in the literature. Next, when the proposed model was compared with state-of-the-art (SOTA) ensemble models, it not only achieved the best performance in terms of accuracy and sensitivity, but also reduced the ensemble model complexity drastically. Our findings demonstrate that improvements in performance across many metrics, highlighting the effectiveness of the BDE-based ensemble approach in designing a more accurate deception system.

Keywords: *Binary Differential Evolution, Deception Detection, Ensemble Learning, Optimization, Stacking*

1. INTRODUCTION

Accurate deception detection system performance is paramount in criminal analysis to enable timely investigations Polygraph tests are the most common methods to determine whether someone is guilty or innocent. These tests record involuntary nervous reactions such as pulse rate, breathing rate, electro dermal response, and so on to study and set them [1]. The innocent person's heart rate or respiration rate may increase if they are apprehensive about the questions answered because deception detection is a matter of ethics and morality, the results of polygraph examinations cannot be trusted. A more direct picture of brain reactions created while a person is lying is developed as a more preferable method to the polygraph test. The Brain-Computer Interface (BCI) is a tool that allows you to evaluate and determine how your brain works. A variety of invasive and non-invasive methods have been used to measure and record these processes. An electroencephalogram (EEG) is a popular, non-invasive acquisition technique that is easy to use and provides a cost-effective means to record neuronal activity in the brain as EEG signals. Depending on the individual's activity, EEG signals have distinct features. Different areas of the brain are engaged while a person is going about his or her daily activities, such as participating in sports or meditating. EEG electrodes are positioned in precise areas to record specific activity [2, 3]. Previous research efforts have explored the use of machine learning (ML) algorithms such as Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbour (KNN), and Decision Tree (DT) for deception detection [4-7]. While these algorithms show promise in specific scenarios, they often fail in accurate predictions due to their limited scope and individual classifier performance. To address these limitations, researchers have turned to ensemble learning approaches that leverage the strengths of multiple classifiers to enhance predictive accuracy and robustness. Ensemble-based models are built on the principle that diverse classifiers contribute unique perspectives and error patterns, leading to improved overall performance compared to individual classifiers. Stacked ensemble learning, also known as stacked generalization or stacking, is a powerful robust strategy in machine learning that integrates the predictions from various foundational learners to enhance the overall efficacy [8]. The process of stacking involves several key steps:

1. Pool Generation: In this stage heterogeneous pool of base learners are created. Then these learners are created by varying the hyper parameters of learners [9].
2. Classifier Selection: In this step best learners that contribute to overall ensemble is identified based on the performance metrics during training.
3. Classifier Aggregation: After the base learners are chosen, their predictions are aggregated through base leaners. Then the meta learner is learned from the predictions of the base learners to produce the ensemble prediction [10].

Stacked ensemble learning tries to exploit the strengths of individual classifiers while eliminating their weaknesses. By aggregating a variety of models and learning from their combined predictions, stacked ensembles tend to be more accurate and robust than a single model [11].

To enhance deception performance in this study a two level stacking framework is proposed. For effective solution and better results, in the proposed approach base learners are encoded using binary and meta learners are using one-hot encoding. Further, the proposed BDE uses dynamic mutation scaling factor and cross over rate. Finally, the continuous solution space is converted using sigmoid transfer function followed by thresholding

2. RELATED WORK

To capture replies from subjects a Guilty Knowledge Test (GKT) is conducted using EEG device. This approach is similar to polygraph examinations [12]. Also, several authors have conducted Concealed Information Tests (CIT) to understand the subject behaviour. This behaviour analysis helps us to categorize the subjects into innocent or guilty. The same approach is used in polygraphs our EEG based approach is non-invasive.

Bootstrapping is a technique used for oversampling purpose in the literature. The same technique is applied in EEG research area to increase the samples of stimulus. In [13] authors have used same bootstrapping technique for increase 3 types of oddball stimulus. This stimulus is used for psychological studies.

In EEG-based deception detection, the Event-Related Potential (ERP) stimuli that occur during innocent sessions are used. Authors have conducted CIT test to collect ERP stimuli data on various subjects [14]. Authors in [15-16], used CIT approach to detect deception. As EEG contains biological data contains lot of artifacts in [17] used blind source separation for its removal.

In [18] authors have proposed handcrafted Empirical Mode Decomposition (EMD) technique for feature extraction from P-300 EEG signal in lie detection application. Even though this approach performs better it suffers from real time applicability and scalability. However, others used EMD for fault detection using vibration signals and time series data analysis [19-20].

A subject specific analysis using spatial spiking neural networks is proposed for lie detection. The CIT test collected data filtered using Finite Impulse Response (FIR) filter and applied Common Spatial Pattern (CSP) to extract spatial components. The proposed approach attained a peak accuracy of 90.15%. However, authors didn't applied channel selection approach to enhance the performance [21].

From the literature study, we observe that no authors have used stacking ensemble approach for deception performance enhancement. Moreover, no one attempted to meta heuristics to solve combinatorial problem in stacked ensemble for effective deception classification.

3. PROPOSED METHOD

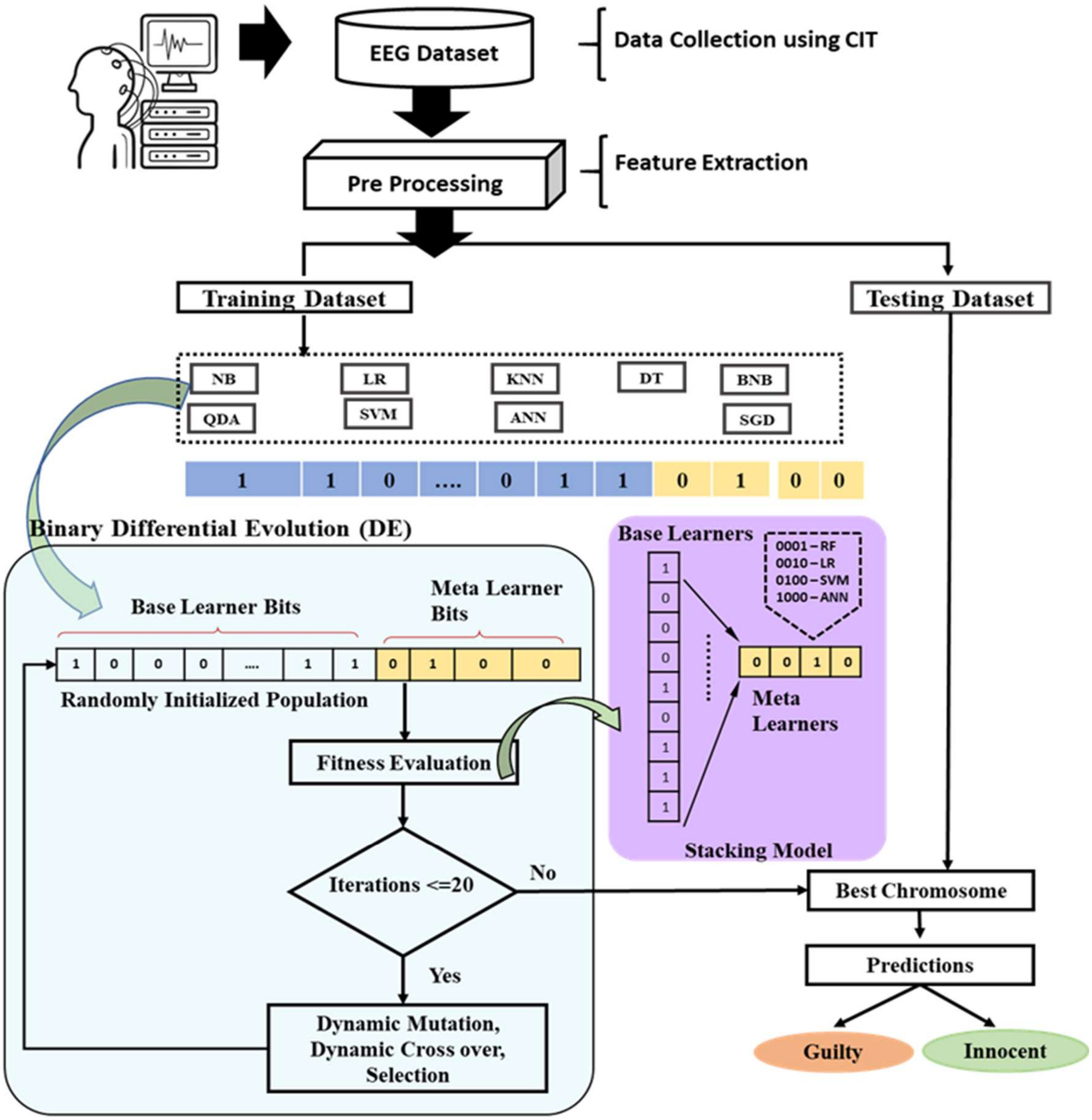

Figure 1: The Proposed Model

Deception detection using EEG provides a direct, non-invasive approach. However traditional methods lack performance. So, ensemble learning has shown promise but selecting diverse base and meta learners is a combinatorial challenge. Binary Differential Evolution (BDE) offers an efficient mechanism to explore the classifier combination space and optimize predictive accuracy. By integrating BDE with a two-level stacking framework, the study aims to maximize accuracy, reduce ensemble complexity.

The workflow of proposed techniques in this paper illustrated in Figure 1. The workflow includes the following steps:

- Pre-processing of the dataset using WT.
- Utilization of grid search for optimizing the hyperparameters of each classifier.
- Creation of nine base learners comprising Naïve Bayes, Logistic Regression, Support Vector Machine, K-Nearest Neighbours, ANN, Decision Tree, QDA, Bernoulli Naive Bayes, and SGD classifiers trained using 5-fold cross-validation. Additionally,

four models are designated for meta-learners, namely Logistic Regression, Random Forest, KNN, and ANN.

- Application of Binary version DE to determine the optimal ensemble configuration.
- Proposal of a novel fitness function tailored for Binary DE.

These steps collectively contribute to enhancing the accuracy and robustness of the model proposed in this study.

3.1 Ensemble Stacking Method

The stacking ensemble method constitutes a sophisticated technique that blends statistical principles with machine learning expertise. It involves assembling a collection of base learners alongside a meta-learner, as depicted in Figure 1, to harness the strengths of each algorithm for superior predictive accuracy. Base learners contribute diversity to the ensemble, while the meta-learner combine their outputs effectively. The stacking process begins by training various individual models using K-fold cross-validation on the training dataset at the foundational level. The predictions from these base learners along with additional computed features such as standard deviation and confidence score (mean) of these outputs are used as input data for the meta-learner. This meta-learner is trained to integrate the diverse insights from the base learners, resulting in a more robust and accurate overall model. Choosing the base learners is a critical aspect when building a stacking ensemble, considering their accuracy and heterogeneity to achieve optimal performance. Strong learners are typically chosen as base learners to ensure accuracy while using a variety of algorithms to enhance model diversity. This approach delves into the underlying relationships within the data from various perspectives, thereby enhancing prediction capabilities. On the other hand, the meta-learner is a stable model adopt at effectively synthesizing inputs from different base learners.

3.2 Ensemble Stacking Method

The DE algorithm [22] is a strong optimization technique developed by taking cues from evolutionary theory and natural selection. It works by keeping track of a population of possible solutions in a multidimensional search space, represented as vectors. The idea of differential mutation and crossover, which creates novel solutions by merging differences among randomly selected individuals within the population, is essential for better solution. These methods make it possible to search the search space more effectively while aiming for improvements in the value of the objective function.

In our BDE algorithm these probabilities are adjusted dynamically over the generations using following formulas:

$$Mp = 0.4 + 0.5 \times (\mathrm{g/Ngen})$$
$$Cp = 0.9 - 0.3 \times (\mathrm{g/Ngen})$$

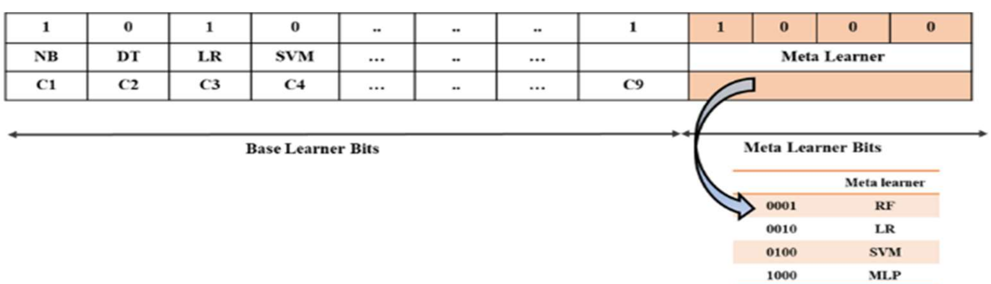

Figure 2 Chromosome Encoding schema

The Binary DE algorithm selects, recombines (using cross over), and mutates candidate solutions through iterative refinement based on a predetermined fitness function. This process produces consecutive generations of possible solutions. The algorithm is guided towards optimal or nearly optimal solutions through this iterative process, demonstrating its efficacy in handling complicated optimization problems.

3.3. Learners

In order to improve the performance of the stacking framework, base and meta learners are chosen carefully.

Base Learners: Naive Bayes, Logistic Regression (LR), Support Vector Classifier (SVC), K-Nearest

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Neighbors (KNN), Artificial Neural Network (ANN), Decision Tree, Quadratic Discriminant Analysis (QDA), Bernoulli Naive Bayes, SGD Classifier

Meta-Learners: Logistic Regression (LR), Random Forest (RF), K-Nearest Neighbors (KNN), Artificial Neural Network (ANN)

3.4. DE Stacking Ensemble Learning

Stacking has two fundamental steps: Identifying the base and meta-models and integrating them. It's very much essential to balance accuracy and diversity while selecting the learners. The model combination has a significant effect on predictive results. In our research, we utilized a differential evolution algorithm to identify the most suitable model blend for stacking, resulting in the creation of the DE stacking algorithm. Following data pre-processing and partitioning into training and testing sets, we trained our nine base models on the training data using k-fold cross-validation with k = 5. Subsequently, we constructed another dataset containing the prediction probabilities from the base learners along with additional features such as standard deviation and confidence score (mean) computed using base learner outputs. Next, we applied the BDE algorithm to identify the best combination of base and meta learners represented as an 13-bit chromosome. Upon finding the optimal chromosome, we utilized it to make predictions for our testing data, leveraging the identified combination for enhanced prediction accuracy and performance.

3.4.1. Chromosome representation

In this study, we tackle the challenge of combining models through two steps: assembling base learners and choosing meta-learners. In stacking ensemble learning, the combination of base learners can be visualized as a fusion of nine novel attributes, represented by an 13-bit binary code where 1 signifies selection and 0 indicates exclusion. The selection of the meta-learner entails choosing the superior model among four alternatives, encoded using a 4-bit one hot encoding format. For instance, 0001 corresponds to Random Forest, 0010 to logistic regression, 0100 to SVC, and 1000 to ANN. An 13-bit binary code forms a chromosome, solving the base and meta-learner mix for stacking ensemble learning. The same is shown in Figure 2.

3.4.2. DE parameters and operations

The DE algorithm begins with an initial population of 50 chromosomes, each representing a potentially useful solution. Such diversity allows for exploration across multiple base and meta-learner combinations. DE relies on important operators selection, mutation and crossover. After some number of iterations, the population converges on the best possible solutions. The final solution combining the base learners and the meta learner is the best individual in the population based on fitness. Our mutation strategy was DE/RAND/1 with dynamic mutation scaling factor. Additionally, the recombination operation of crossover operates with a probability of ranging from 0.9 to 0.6 produces new individuals that will helps for effective solution space exploration. Finally, the continuous space is converted into binary by using sigmoid function followed by thresholding.

3.5. Fitness Function

The binary vector is passed to fitness function for evaluation**.** The fitness function is pivotal in steering the BDE algorithm towards optimal solutions. For this study, we adopt the accuracy score and number of classifiers as our fitness function.

$$\begin{aligned} &Objective\ function \\ &= w1 * Stacking\ Accuracy - w2 \\ &* \frac{Number\ of\ Classifiers\ Selected}{Total\ Classifiers\ in\ Ensemble} \end{aligned} \quad (1)$$

For better results, ensemble accuracy has to be maximized, and ensemble complexity has to be minimized. Overall, the objective function has to be maximized.

4. RESULTS

4.1. Experimental Environment Setup

The proposed model is implemented using Anaconda distribution, Jupiter notebook, and Python programming language. All the experiments were carried out on Intel(R) i7 (TM) 9700 CPU @ 3.00GHz processor with 32GB RAM and 64bit Windows 10 operating system.

4.2. Acquisition of EEG Data and Concealed Information Test (CIT)

The EEG data is collected with an acquisition device called a brain vision recorder. Brain vision analyzer 2.1 is used to evaluate EEG signals obtained from the brain vision recorder. As the EEG signals are recorded in a real-world setting, they contain a lot of artifacts. Hence, the signal is passed via a band pass filter, which removes the unwanted frequency band without decreasing the signal's quality. For our research purpose, a band-pass filter with a frequency range from 0.3 Hz to 30 Hz is used. This is the most commonly used frequency range for performing a mental task

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

experiment [23]. In this research, a Concealed Information Test (CIT) is performed to investigate human behavior of whether he/she is lying or not. This posed a binary classification problem, and data is divided into two categories, namely "guilty" and "innocent". To conduct an effective evaluation, the 5-Fold Cross-Validation (5-FCV) approach is applied to EEG data.

The lie detection experiment begins with the subject seeing the stimuli images for 31 seconds on a display screen. The subject must recognize the visuals and respond with a "yes" or "no" response. Each image will be shown for 1.1 seconds, followed by a 2-second blank image. So, the subject is shown 10 images in total, seven of which are irrelevant, two of which are target images, and one of which is the probing image. The subject is shown in these images at random. This process is shown in Figure 3.

The experiment takes place over two sessions (guilty and innocent), each consisting of 30 trials for a single person. A single trail structure is shown in Figure 4. So, for ten people, we get 600 trials or samples, which are sampled at 250Hz frequency.

In two sessions, a CIT experiment was conducted with ten participants. The EEG data is recorded with 16 electrodes such as FC1, Cz, C4, CP5, FC2, C3, CP1, CP2, CP6, P3, Oz, Fz, Pz, P4, O1, and O2 sites using the 10-20 international electrode placement system as like in [28].

4.2.1. Guilty Session

During the guilty session, subjects are told to react "yes" for the target image and "no" for the probing image and irrelevant image. The investigation examines the suspect, who disputes the charges, and demonstrates that he is indeed the genuine perpetrator. Because the probe is known to the subject, but he/she is willfully denying it, the probe is generated in the guilty brain for both target and probe. A P300 [25] response is elicited for the guilty person as soon as the probe and target are introduced.

4.2.2. Innocent Session

Subjects are asked to react "yes" for the target image and "no" for the probing image and irrelevant image during the innocent session. Because both stimuli are unfamiliar to the innocent, the response evoked in the innocent brain for irrelevant and probing is the same. In the innocent subject's brain, only target stimuli will evoke P300. The P300 (P3) wave is a component of the event-related potential (ERP) that is elicited during the decision-making process.

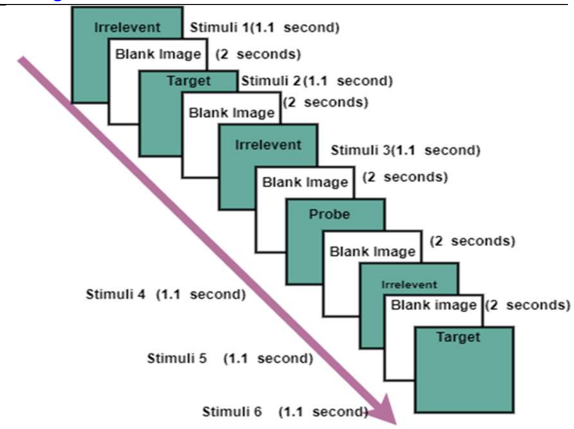

Figure 3: Experiment Process

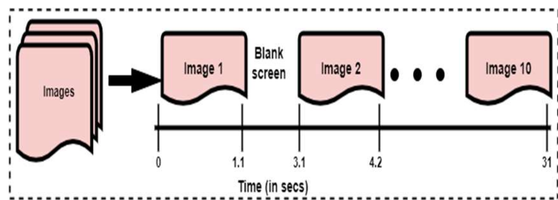

Figure 4:. A Trail Structure

4.3. Dataset Pre-Processing

Before constructing the proposed ensemble model, the EEG dataset is pre-processed. The following steps were performed during the pre-processing phase: EEG data must be preprocessed and examined after data acquisition. Pre-processing entails a number of procedures intended to increase the data's signal-to-noise ratio and make it easier to spot any experimental effects that may be present.

4.3.1. MNE

The MNE library was employed for the preprocessing step. For examining, displaying, and interpreting human neurophysiological data, including EEG, MNE is an open-source Python tool. The most popular EEG systems are included in the MNE collection of data reading and conversion utilities, which may be used to import and process data from various hardware systems. An electrode-capturing brainwave activity is called an EEG channel. Data is converted by MNE into a raw object, which contains the labels for each data channel and other metadata. It converts raw continuous EEG Data to segmented Epochs. Then, the model is applied to the EEG dataset, and the results are collected and analyzed.

4.4. Performance Measures

Performance is measured in terms of accuracy, specificity, G-measure, precision, F1-score, and sensitivity. All performance measures are shown in Equations 2, 3, 4, 5, and 6.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$Accuracy\ (acc) = \frac{TP + TN}{TP + TN + FP + FN}$$	(2)
$$Specificity\ (spec) = \frac{TN}{TP + FP}$$	(3)
$$Precision\ (prec) = \frac{TP}{TP + FP}$$	(4)
$$F1 - Score\ (F1) = 2 * \frac{Precision * Recall}{Precision + Recall}$$	(5)
$$Sensitivity(Recall) = \frac{TP}{TP + FN}$$	(6)

Table 1: Confusion Matrix

		Predicted	
		Guilty	Innocent
Actual	Guilty	TP	TN
	Innocent	FP	FN

• TP (True Positive) represents the subjects that are classified as lying who actually lied
• TN (True Negative) represents the subjects that are classified as innocent who are actually innocent,
• FP (False Positive) represents the subjects that are classified as a lie but they are innocent
• FN (False Negative) represents the subjects that are classified as innocent, but there are guilty

4.5. Results Analysis

In this process, a wavelet transform (WT) is used, which decomposes data into four levels of approximation and detail coefficients using Python-based PyWavelets [26]. PyWavelets is an open-source wavelet transform library for Python. Each subject has 16 channels, and data is used to train 9 base learners. For better performance, the hyperparameters of these classifiers are fine-tuned using grid search. These hyperparameter values are tabulated in Table 2. Such fine-tuned classifiers are evaluated in terms of accuracy, sensitivity, specificity, and F1 score on the test dataset. These results are shown in Table 3. The effectiveness of the proposed DE-based stacking model relies on several key parameters like Crossover rate, Mutation rate, number of generations, μ, λ, etc. These parameters critically influence the algorithm's performance, balancing the exploration of the solution space with the exploitation of promising areas. Importantly, optimal parameter values are problem-specific. When properly tuned, these parameters enable the DE algorithm to lead to improved overall performance, faster convergence, and a higher likelihood of discovering high-quality solutions. Hence, these values are fine-tuned using grid search, and the finalized values are listed in Table 4.

Table 2: Fine-tuned parameters of Base Learners

Base Learner	Hyper Parameters	Base Learner	Hyper Parameters
LR	C:100 Penalty: L1	SVM	C: 10 Gamma: Auto Kernel: RBF
KNN	# Neighbours: 3 P: 1 Weights: Uniform	DT	MaxDepth: 10 Min sample in leaf: 2 min sample to split: 5
QDA	Reg_Param: 0	NB Bernoulli	Alpha: 0.01
SGD	Alpha: 0.0001 MaxIterations: 1000 Penality: L1	MLP	Optimizer: Adam Hidden Layer Sizes: 64, 32 Activation Function: ReLu Loss: Binary Cross Entropy

Table 3: Performance of Base Learners before Applying the Proposed Method

Learner	Acc (%)	Sen(%)	Spec (%)	F1-score (%)
NB	60.63	89.000	32.00	69.00
LR	69.40	64.43	69.74	69.30
SVM	87.50	100	75.27	88.77
KNN	94.22	100	88.56	94.47
MLP	72.57	70.18	74.90	71.67
DT	95.71	100	92.61	96.36
NB(Bernoulli)	61.94	63.01	59.40	61.62
SGD	73.51	80.37	66.78	75.00
QDA	92.72	87.92	97.41	97.42

Table 4: Optimized Parameters in BDE

Parameter	Value
Mutation rate ($\boldsymbol{M_P}$)	0.02
Cross over rate (cp)	0.9
Number of generations (Ngen)	15
Population size	50
Maximum no of iterations	80
Number of individuals for next generation (μ)	15
number of children to produce at each generation (λ)	50

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 5: Performance of Proposed Model over 10 runs

Runs	Ensemble complexity	Test results				
		Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC
1	7	96.66	100	92.88	97.50	0.980
2	5	96.52	94.6	96.59	95.23	0.9640
3	7	96.66	100	92.85	97.50	0.9914
4	9	96.66	100	92.85	97.50	0.9971
5	5	95.00	100	89.28	95.96	0.9628
6	5	96.66	100	92.85	97.50	0.9740
7	6	96.66	100	92.85	97.50	0.9914
8	5	95.00	100	89.28	95.96	0.9628
9	6	95.00	93.46	97.60	89.33	0.9569
10	6	98.33	100	96.96	98.46	0.9530
Min	5	95.00	93.46	89.28	89.33	0.9530
Median	6.1	96.36	98.80	93.39	96.24	0.9733
Max	9	98.33	100	97.60	98.46	0.9914
Average	6.1	96.36	98.80	93.39	96.24	0.9733

Table 6: Performance of Proposed Model.

Ensemble Complexity	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC
6.1	**96.36**	**98.80**	**93.39**	**96.24**	**0.9733**

Table 7: Comparison between Proposed Model and SOTA Models

S. No.	Classifier	FE	Channels	Acc (%)	Sen (%)	Spe (%)	Ref
1	**SVM**	EMD	16	81.71	78.33	83.47	2018 [7]
2	**KNN**	Hjorth	16	81.90	78.90	85.10	2018 [27]
3	**LDA**	WPT	16	91.67	90.32	93.10	2019 [26]
4	**Ensemble**	WT	16	84.70	82.50	83.90	2019 [28]
5	**K means + FNN**	WT	16	83.10	95.00	-	2020 [4]
6	**ELM**	STFT	16	88.33	86.66	90	2020 [60]
7	**SLNN**	CSP	16	90.15	92.72	88.06	2024 [21]
8	**PROPOSED**	**WT**	**16**	**96.36**	**98.80**	**93.39**	

Our analysis revealed that the average number of required classifiers is minimal, indicating efficient ensemble construction. For the final performance evaluation, we used the mean values obtained from the ten runs. This approach provides a comprehensive assessment of the model's stability and effectiveness across multiple initializations, offering a more reliable representation of its overall performance. Finally, the performance of the proposed model in terms of ensemble complexity, accuracy, sensitivity, specificity, and F1-score over 10 runs is given in Table 6.

4.6. Comparative Analysis

There are various classifiers that have already been used, such as KNN, SVM, LDA, Ensemble, K-means, and MLFNNs. Also, various feature extraction techniques like EMD, Hjorth parameters, Wavelet Transform, Fourier, and Short Fourier Transforms have been used already. A few studies proposed channel selections using optimization algorithms. Finally, our proposed with SOTA models in the domain. These results are shown in Table 7.

From the table, it is observed that our proposed model outperforms the SOTA models in terms of accuracy, sensitivity, specificity, and F1 score. These results are highlighted. Further, to prove the robustness of the proposed model, it is compared with SOTA ensemble models in the literature. These results are shown in Table 8. From the results, it is proved that our model achieved the best performance in terms of all the metrics.

Learner	**Acc (%)**	**Sen (%)**	**Spec (%)**	**F1 (%)**	**Complexity**
Ada Boost	95.38	96.00	94.83	97.42	50
XG Boost	94.76	95.33	95.57	97.78	200
Light GB	94.38	96.34	94.83	97.42	200
GBC	93.32	96.45	93.67	94.33	200
RF	95.57	94.33	95.20	97.60	50
Extra Tree Classifier	95.32	93.56	95.67	96.33	50

4.6. Discussion

Even though, our proposed stacking model using BDE for optimizing base and meta learners reduced model complexity when compared with naïve stacking. However, it's still be computationally intensive. Applying the proposed framework for large-scale screening (e.g., border security) may require additional validation. Further, EEG signals vary significantly between subjects, so the model may require retraining or fine-tuning for new subjects for real-time deployment.

5. CONCLUSION

In conclusion, the novel BDE stacking algorithm presented in this study harnesses the power to optimize stacked ensemble framework for effective deception prediction. It integrates 9 diverse base learners and 4 meta learners that are trained and evaluated on a dataset collected using CIT. The BDE algorithm, guided by the ensemble accuracy score and ensemble complexity as the fitness function, that efficiently searches for optimal integration of base and meta learners. The proposed model is tested using CIT dataset and compared with SOTA models. The results proved that proposed model outperforms in terms of accuracy, f1-score, sensitivity, and specificity when compared with SOTA models. Further, the proposed model is compared with ensemble methods and it is outperformed in terms of all metrics. Hence, our results show the potential of the BDE stacking algorithm in enhancing deception prediction accuracy. By leveraging evolutionary algorithm and ensemble learning techniques, we have achieved notable advancements over existing methods, highlighting the significance of a well-tuned model combination configuration.

REFERENCES:

[1]. National Research Council, Division of Behavioral, Social Sciences, Committee on National Statistics, Board on Behavioral, Sensory Sciences, & Committee to Review the Scientific Evidence on the Polygraph. (2003). The polygraph and lie detection. National Academies Press.

[2]. "Ramadan, R. A., & Vasilakos, A. V. (2017). Brain computer interface: control signals review. *Neurocomputing*, *223*, 26-44.

[3]. Lykken, D. T. (1959). The GSR in the detection of guilt. *journal of Applied Psychology*, *43*(6), 385.

[4]. Bablani, A., Edla, D. R., Kuppili, V., & Ramesh, D. (2020). A multi stage EEG data classification using k-means and feed forward neural network. *Clinical Epidemiology and Global Health*, *8*(3), 718-724.

[5]. Bablani, A., Edla, D. R., Tripathi, D., Dodia, S., & Chintala, S. (2019). A synergistic

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

concealed information test with novel approach for EEG channel selection and SVM parameter optimization. *IEEE Transactions on Information Forensics and Security*, *14*(11), 3057-3068.

[6]. Dodia, S., Edla, D. R., Bablani, A., & Cheruku, R. (2020). Lie detection using extreme learning machine: A concealed information test based on short-time Fourier transform and binary bat optimization using a novel fitness function. *Computational Intelligence*, *36*(2), 637-658.

[7]. Bablani, A., Edla, D. R., & Tripathi, D. (2018). Subject based deceit identification using empirical mode decomposition. *Procedia computer science*, *132*, 32-39.

[8]. Wolpert, D. H. (1992). Stacked generalization. *Neural networks*, *5*(2), 241-259.

[9]. Dietterich, T. G. (2000, June). Ensemble methods in machine learning. In *International workshop on multiple classifier systems* (pp. 1-15). Berlin, Heidelberg: Springer Berlin Heidelberg.

[10]. Van der Laan, M. J., Polley, E. C., & Hubbard, A. E. (2007). Super learner. *Statistical applications in genetics and molecular biology*, *6*(1).

[11]. Ting, K. M., & Witten, I. H. (1999). Issues in stacked generalization. *Journal of artificial intelligence research*, *10*, 271-289."

[12]. Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology*, *28*(5), 531-547.

[13]. Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology*, *41*(2), 205-219.

[14]. Rosenfeld, J. P., Labkovsky, E., Winograd, M., Lui, M. A., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): A new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology*, *45*(6), 906-919.

[15]. Kubo, K., & Nittono, H. (2009). The role of intention to conceal in the P300-based concealed information test. *Applied Psychophysiology and Biofeedback*, *34*, 227-235.

[16]. Meixner, J. B., & Rosenfeld, J. P. (2011). A mock terrorism application of the P300-based concealed information test. *Psychophysiology*, *48*(2), 149-154.

[17]. Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., Mckeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*(2), 163-178.

[18]. Arasteh, A., Moradi, M. H., & Janghorbani, A. (2016). A novel method based on empirical mode decomposition for P300-based detection of deception. *IEEE Transactions on Information Forensics and Security*, *11*(11), 2584-2593..

[19]. Wang, F., & Fang, L. (2019). A fault diagnosis method for automaton based on morphological component analysis and ensemble empirical mode decomposition.

[20]. Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, *454*(1971), 903-995.

[21]. Edla, D. R., Bablani, A., Bhattacharyya, S., Dharavath, R., Cheruku, R., & Boddu, V. (2024). Spatial spiking neural network for classification of EEG signals for concealed information test. *Multimedia Tools and Applications*, *83*(33), 79259-79280.

[22]. Singh, N., & Singh, P. (2020). Stacking-based multi-objective evolutionary ensemble framework for prediction of diabetes mellitus. *Biocybernetics and Biomedical Engineering*, *40*(1), 1-22..

[23]. Farahani, E. D., & Moradi, M. H. (2017). Multimodal detection of concealed information using genetic-SVM classifier with strict validation structure. *Informatics in Medicine Unlocked*, *9*, 58-67..

[24]. AMINIAN, M. A., & Kafaee, R. M. (2019). Investigating The Effects of Modem Electromagnetic Waves (2. 4 GHz) on Electroencephalogram.

[25]. Frantzidis, C. A., Bratsas, C., Papadelis, C. L., Konstantinidis, E., Pappas, C., & Bamidis, P. D. (2010). Toward emotion aware computing: an integrated approach using multichannel neurophysiological recordings and affective visual stimuli. *IEEE transactions on Information Technology in Biomedicine*, *14*(3), 589-597.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[26]. Dodia, S., Edla, D. R., Bablani, A., Ramesh, D., & Kuppili, V. (2019). An efficient EEG based deceit identification test using wavelet packet transform and linear discriminant analysis. *Journal of neuroscience methods*, *314*, 31-40.

[27]. Bablani, A., Edla, D. R., & Dodia, S. (2018). Classification of EEG data using k-nearest neighbor approach for concealed information test. *Procedia computer science*, *143*, 242-249.

[28]. Bablani, A., Edla, D. R., Tripathi, D., & Kuppili, V. (2019). An efficient concealed information test: EEG feature extraction and ensemble classification for lie identification. *Machine Vision and Applications*, *30*, 813-832.