Automated Comparative Predictive Analysis of Deception Detection in Convicted Offenders Using Polygraph with Random Forest, Support Vector Machine, and Artificial Neural Network Models

Dana RAD[1,2]*, Csaba KISS[3], Nicolae PARASCHIV[2], Valentina Emilia BALAS[2,4]

[1] Centre of Research Development and Innovation in Psychology, Faculty of Educational Sciences, Psychology and Social Work, Aurel Vlaicu University of Arad, 77 Revolutiei Avenue, Arad, 310130, Romania
dana@xhouse.ro (*Corresponding author*)

[2] Doctoral School of Systems Engineering, Petroleum-Gas University of Ploiești, 39 București Avenue, Ploiești, 100680, Romania
nparaschiv@upg-ploiesti.ro

[3] Faculty of Psychology and Educational Sciences, Hyperion University of Bucharest, 169 Calea Călărași, Bucharest, 030615, Romania
kiss.csaba@expertpoligraf.ro

[4] Faculty of Engineering, Aurel Vlaicu University of Arad, 77 Revolutiei Avenue, Arad, 310130, Romania
balas@drbalas.ro

Abstract: This paper provides a thorough comparative review of deception detection techniques employed for a sample of 400 convicted offenders. It focuses on the utilization of polygraph sensor data as input variables for predicting deception, which are assessed against manual scoring by experts. Three advanced machine learning models, namely Random Forest Regression (RFR), Support Vector Machine (SVM) Regression, and Neural Network Regression (NNR), were employed with the purpose of analysing their predictive efficacy in identifying deception based on physiological responses captured by polygraph sensors. The obtained results indicate that all three algorithms exhibited varying degrees of effectiveness in predicting deceptive behavior. The Random Forest Regression algorithm achieved a Mean Squared Error (MSE) of 0.893 and a coefficient of determination ($R^{2)}$ of 0.091, which highlights its ability to discern key physiological indicators related to deceptive behavior. The Support Vector Machine Regression algorithm showed a competitive performance with a MSE of 0.98 and a R^2 value of 0.159, which underscores its capability to model non-linear relationships in the context of high-dimensional data. However, the Neural Network Regression algorithm proved to be the best model, with a MSE of 0.894 and a significantly higher R^2 value of 0.113. This model's capacity to capture the complex relationships in the context of physiological data allowed it to surpass both RFR and SVM, which indicates its potential for a precise and reliable deception detection. This study provides valuable insights into the advancement of forensic applications with regard to deception detection technologies. Its findings suggest that the Neural Network Regression algorithm, due to its ability to learn complex patterns and relationships related to physiological data, stands out as an optimal choice for accurately identifying deceptive behavior.

Keywords: Deception detection, Polygraph sensor data, Machine Learning models, Predictive efficacy, Neural Network Regression, Random Forest Regression, Support Vector Machine Regression.

1. Introduction

Particularly in relation to convicted offenders, the research on deception detection has sparked the interest of forensic psychology and the field of criminal justice (Rad et al., 2023; Rad et al., 2024). According to Raskin & Honts (2002), the polygraph test is still widely used to evaluate physiological markers such as skin conductivity, blood pressure, pulse rate, and breathing rate. The accuracy and reliability of the polygraph are still being debated by scientists despite its widespread use (National Research Council, 2003).

Polygraph test deception prediction accuracy could be improved with the help of recent advances in machine learning. Among the machine learning models that provide advanced techniques for pattern recognition and classification are Random Forest, Support Vector Machines (SVMs) (Cortes & Vapnik, 1995), and Artificial Neural Networks (ANNs) (LeCun, Bengio & Hinton, 2015). These models can effectively utilize intricate physiological data, hence enhancing conventional expert analysis.

This study presents a comprehensive comparative analysis of deception detection techniques applied to a cohort of 400 convicted offenders. The primary objective is to evaluate the efficacy of three advanced machine learning models - Random Forest, SVM, and ANN - in accurately identifying deception based on polygraph sensor data. The polygraph data serves as the input variables, while manual scoring by experts provide the output for model training and validation.

Random Forest is an ensemble learning method that enhances the accuracy of predictions and mitigates overfitting by creating many decision

https://doi.org/10.24846/v33i3y202404

trees and aggregating their outcomes (Nawroly et al., 2023). The Support Vector Machine (SVM) is a classification approach renowned for its ability to accurately determine the most optimal hyperplane inside a given feature space. Cortes & Vapnik (1995) explained how this hyperplane minimizes the margin between different classes. An artificial neural network (ANN) is a computational model designed to mimic the structure and functioning of the neural networks found in the human brain. By utilizing a network of interconnected neurons, it has the ability to identify complex and non-linear relationships within data. LeCun, Bengio & Hinton (2015) first referred to this notion in 2015, then later on, Filip (2021) and Weiwei & Hao (2022) made more advancements, respectively.

A wide range of assessment measures is used to evaluate the effectiveness of these models. The Mean Squared Error (MSE), as defined by Willmott & Matsuura (2005), calculates the average value of the squared differences between the observed and anticipated values. The Root Mean Squared Error (RMSE) measures the magnitude of the model's prediction error using the same units as the analysed data. As stated by Chai & Draxler (2014), it represents the standard deviation of the residuals. The Mean Absolute Error (MAE) or Mean Absolute Deviation (MAD) is a measure that quantifies the average absolute difference between expected and actual data. It provides a straightforward assessment of the precision of the forecast (Willmott & Matsuura, 2005). The Mean Absolute Percentage Error (MAPE) is a measure that calculates the average absolute percentage difference between the expected and actual values. The study by Armstrong & Collopy (1992) provides a reliable assessment of the accuracy of forecasts. Moreover, the R-squared ($R^{2)}$ value measures the degree to which the independent variables can explain the variation in the dependent variable. R^2 is utilized to assess the model's conformity with the analysed data (Nagelkerke, 1991). Additionally, the machine learning models' performance is compared against traditional expert analysis to determine the extent of improvement in deception detection. Integrating machine learning models with polygraph data is a substantial advancement in automating and improving the accuracy of deception detection techniques.

The current body of research on deception detection using polygraphs has primarily concentrated on conventional analysis methods, with less investigation into advanced machine learning techniques. While some studies have employed ML models, there is a lack of comprehensive comparative analyses of different ML approaches using a consistent dataset. Moreover, the evaluation of these models has often been limited to a narrow set of metrics, without a holistic assessment of their predictive performance.

The objective of this study is to address this deficiency by performing an extensive comparative examination of three advanced machine learning models - Random Forest, Support Vector Machine (SVM), and Artificial Neural Network (ANN). The study assesses these models using a wide range of measures, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2), to ensure a thorough evaluation of their effectiveness.

The remainder of this paper is as follows. Section 2 includes a literature review and Section 3 presents the employed methodology. Section 4 sets forth the obtained experimental results, and Section 5 presents a comparative analysis of the regression algorithms employed. Finally, Section 6 outlines the conclusion of this paper.

2. Literature Review

Efforts to develop reliable methods for detecting deception have seen substantial breakthroughs over the years, incorporating a range of technological and methodological improvements. This section offers a thorough examination of the existing body of knowledge on the identification of deceit, with specific attention given to the historical background, current approaches, and upcoming developments, particularly in the utilization of machine learning (ML) and neural network models.

Viji, Gupta & Parekh (2022) present a comprehensive historical analysis, emphasizing the shift from initial physiological measurements to more advanced methodologies. Early lie detectors predominantly depended on assessing physiological reactions, such as heart rate, blood pressure, and breathing rate, which were thought to be indicative of deceitful behaviour. Despite being fundamental, these procedures frequently had significant error rates and relied on subjective interpretation.

According to Oswald (2020), the dependability and accuracy of polygraph examinations have been the focus of vigorous controversy. Critics argue anxiety and medical issues might cause polygraph findings to be inaccurate since they are unrelated to deceit.

To address these limitations, modern polygraph techniques have incorporated advanced computational methods. Neural networks, for instance, have shown promise in improving the accuracy of polygraph scoring. Rad et al. (2023) provide a scoping review of neural network applications in polygraph scoring, demonstrating that these models can effectively learn and interpret complex patterns in physiological data that are indicative of deception.

The integration of ML algorithms into deception detection represents a significant advancement over traditional methods. In their study, Fernandes & Ullah (2022) examine the methods of feature extraction and matching employed in deception detection. They highlight the significant contribution of machine learning (ML) in improving the accuracy and dependability of these procedures. Machine learning techniques, such as Random Forest (RF), Support Vector Machines (SVMs), and Artificial Neural Networks (ANNs), have achieved notable success in analysing extensive datasets and detecting subtle physiological indicators linked to deceit.

Deception detection research has seen a rise in the use of neural networks and deep learning models. In their study, Diaz, Wong & Chen (2023) investigate the utilization of exclusively visual characteristics in deep learning models to improve the detection of deception. They demonstrate that these models can attain greater accuracy by utilizing intricate visual data. In the same vein, Mohan & Seal (2021) explore the incorporation of multimodal data through machine learning algorithms. They emphasize the potential of merging different data sources, such as facial expressions, voice, and physiological signals, to enhance the accuracy of detection.

In recent years, the idea of multimodal deception detection, which entails analyzing many forms of evidence concurrently, has attracted significant attention. Derakhshan et al. (2020) identified optimal features in multimodal deception detection, suggesting that combining data from different modalities can lead to more robust and accurate models. This approach was further supported by Thannoon, Ali & Hashim (2018), who investigated the use of facial expressions and various classification algorithms to detect deception, demonstrating that multimodal techniques can significantly enhance detection performance.

Voice stress analysis is another area where ML models have shown considerable promise. Talaat (2024) presents a novel and interpretable augmented recurrent neural network for detecting lies through voice stress analysis. The study highlights the significance of model interpretability in the field of forensic applications. It also emphasizes the importance of transparent and explainable machine learning models that may offer insights into their decision-making processes. This, in turn, enhances their acceptance and reliability in legal settings.

Although there has been notable advancement in utilizing machine learning (ML) and neural networks for identifying deceit, there are still several obstacles that need to be addressed. The current body of research emphasizes the capacity of these technologies to enhance precision and dependability. However, there is a dearth of rigorous comparative analyses that assess various machine learning models using a standardized dataset. Furthermore, the requirement for models that are visible and that can be explained is essential to guarantee their practical usability in forensic contexts.

The objective of this study is to address these deficiencies by performing a comparative examination of Random Forest, Support Vector Machine, and Artificial Neural Network models employing polygraph sensor data from a group of 400 individuals who have been convicted of crimes. This paper aims to assess the effectiveness of machine learning algorithms in detecting dishonesty by analysing several performance measures. Its findings will bring valuable insights into the field of forensic science.

3. Methodology

3.1 Objectives

This paper aims to conduct a comprehensive comparative analysis of deception detection techniques using polygraph sensor data and three advanced machine learning models: Random Forest, Support Vector Machine, and Artificial Neural Network. The study aims to assess the predictive accuracy, precision, recall, and other performance indicators of these models in detecting dishonest behaviour among convicted offenders. Furthermore, its objective is to evaluate the advantages and disadvantages of each model and offer valuable perspectives on their practical usefulness in forensic contexts. Finally, it aims to improve the overall dependability and precision of polygraph-based methods for detecting deceit.

3.2 Participants

The analysed sample consisted of 400 individuals selected randomly from a pool of 1072 offenders who had committed multiple crimes and undergone polygraph testing administered by expert examiners from ten polygraph laboratories within the Romanian Police, under the supervision of Dr. Csaba Kiss. All 400 participants were repeat offenders involved in serious criminal activities and voluntarily confessed to their crimes, providing consent for their aggregated data to be utilized for scientific research purposes. The data utilized in this study was extracted from a minimum of three charts corresponding to each polygraph examination that was conducted. The participants were 90% males and 10% females, with ages ranging from 18 to 65 years and an average age of 32 years. The 10% representation of female participants in the study is a result of the overrepresentation of males, who were the primary group of serious repeat offenders eligible for this research. The gender disparity in this case aligns with wider criminological patterns where males exhibit a higher propensity for engaging in severe criminal behaviour. The average educational attainment of the participants was 8.6 years, which suggests a rather low level of formal schooling. This study followed ethical protocols for doing research with human volunteers. Before participating, participants were required to provide their informed consent and were guaranteed confidentiality and anonymity during the study.

3.3 Data Collection

Physiological data was gathered via a standardized polygraph testing methodology. Trained examiners conducted polygraph testing on each subject in a controlled laboratory environment. During the administration of the test proposed by Vavrinsky et al. (2021), the polygraph apparatus detected various physiological responses such as the average value of electrodermal reaction (EDA), heart rate, blood pressure, respiration, and skin conductivity. The analysis was carried out on a comprehensive compilation of physiological data gathered from polygraph testing. These parameters encompassed measures of autonomic arousal, cardiovascular activity, respiratory patterns, and electrodermal responses. Specific physiological parameters included:

1. Amplitude of electrodermal reaction (ARED);
2. Amplitude of blood pressure in brachial pulse (ATAB);
3. Change of baseline level in chest breathing (MNBRT);
4. Difference of altitude between breathing cycles (DIFA);
5. Duration of electrodermal reaction (TRED);
6. Abdominal breath line length (LLRA);
7. Arterial tension amplitude of the distal pulse (ATAD);
8. Heart rhythm (RC);
9. Voluntary repeated acts (REV);
10. Duration of brachial pulse arterial tension (TTAB);
11. Changing of the baseline level of abdominal breathing (MNBRA);
12. Ratio of inspiration to expiration (I/E);
13. Average value of electrodermal reaction (EDA);
14. Thoracic breath line length (LLRT);
15. Reactive patterns (PATTR);
16. Duration of distal pulse arterial tension (TTAD);
17. Respiratory rhythm (RR);
18. Erratic breathing (RE);
19. Abdominal respiratory stop (TSTOPRA);
20. Average amplitude of abdominal breathing (ARA);
21. Length of electrodermal reaction (LRED);
22. Thoracic respiratory stop (TSTOPR).

3.4 Data Analysis

The analysis utilized three advanced machine learning models - Random Forest, Support Vector Machine, and Artificial Neural Network

- to process the gathered polygraph data. Every model underwent training and validation using the chosen dataset in order to forecast deceit by analysing physiological responses. For the data analysis, JASP version 0.17.3.0 was used.

With regard to Random Forest (RF), in order to improve forecast accuracy and reduce overfitting, ensemble learning generates many decision trees and aggregates their output.

With the aim of maximizing the margin, the Support Vector Machine (SVM) is a powerful classification method which is able to find the best hyperplane to divide different classes in the feature space (Cortes & Vapnik, 1995).

A computational model known as an Artificial Neural Network (ANN) replicates the architecture and operations of the neural networks found in the human brain. ANNs use many layers of interconnected neurons to capture complex and non-linear correlations in data (LeCun, Bengio & Hinton, 2015; Shvets et al., 2023; Shvets et al., 2024).

A wide range of metrics, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2), were used to evaluate the effectiveness of these models. These metrics enable a thorough assessment of each model's expected accuracy, precision, recall, and general efficacy in identifying dishonesty (Iancu & Florea, 2023; Dučić et al., 2023).

This study aims to utilize these criteria to identify the most efficient machine learning model for detecting deceit via polygraph tests. Additionally, it aims to offer valuable insights into the practical implications of these findings for the fields of forensic science and law enforcement.

4. Results

4.1 Random Forest Regression

The Random Forest Regression model was employed to predict deception based on physiological data obtained from polygraph tests. The model consisted of 100 trees, each using five features per split. The dataset was partitioned into three subsets: a training set consisting of 224 samples, a validation set consisting of 96 samples, and a test set consisting of 80 samples (Table 1). The accuracy and reliability of the model in predicting deceitful behaviour were assessed using many important indicators.

A Mean Squared Error (MSE) of 0.893 was obtained from the test data, as indicated in Table 2.

A measure of the size of prediction errors in this model is provided by this metric, which measures the mean of the squared discrepancies between the expected and actual values. A value of 0.945 was found for the Root Mean Squared Error (RMSE). That resulted in a mean absolute deviation (MAD) of 0.882, also known as the mean absolute error (MAE). It was determined that the average absolute percentage difference between the expected and actual values was 88.6%, which is also known as the Mean Absolute Percentage Error (MAPE). Significant diversity in the forecasts is shown by this significant percentage.

A value of 0.091 was found for the model's R-squared (R^2) parameter, depicting that the

Table 1. Random Forest Regression model

Random Forest Regression							
Trees	Features per split	n(Train)	n(Validation)	n(Test)	Validation MSE	Test MSE	OOB Error
100	5	224	96	80	0.826	0.893	0.842

Note: The model was optimized with respect to the *out-of-bag mean squared error*.

Table 2. Random Forest Regression evaluation metrics

Metrics	Value
MSE	0.893
RMSE	0.945
MAE / MAD	0.882
MAPE	88.6%
R^2	0.091

model has a limited capacity to generate accurate predictions in this specific case since it only accounts for a small portion of the variability in dishonest bahaviour.

In addition, an out-of-bag (OOB) error with a measured value of 0.842 was used to optimize the Random Forest model (Figure 1).

The importance of each feature (Table 3) was assessed using two metrics: the mean decrease in accuracy and the total increase in node purity.

These metrics help identify which features contribute most significantly to the model's predictive performance. The Amplitude of Blood Pressure in Brachial Pulse (ATAB) showed the highest mean decrease in accuracy (0.062) and the highest total increase in node purity (9.412), indicating its strong influence on the model's predictions. As regards the Amplitude of Electrodermal Reaction (ARED), with a mean decrease in accuracy of 0.040 and a total increase in node purity of 8.029, this feature was also highly influential. Further on, the Difference of Altitude Between Breathing Cycles (DIFA) had a mean decrease in accuracy of 0.053 and a total increase in node purity of 6.531, marking it as another key predictor.

Other features, such as the duration of electrodermal reaction (TRED) and the change of the baseline level of abdominal breathing (MNBRA), also contributed to the model`s performance, but to a lesser extent. Features with negative values, such as thoracic breath line length (LLRT) and the average value of electrodermal reaction (EDA), had a lower impact on the model's accuracy.

Overall, the analysis revealed that specific physiological responses, particularly those related to blood pressure and electrodermal activity, played crucial roles in the model's ability to predict deceptive behaviour. However, the relatively low R^2 value and high MAPE suggest that while the model identifies key physiological indicators, its overall predictive power is limited, highlighting the need for further refinement and potential integration of additional data or methods.

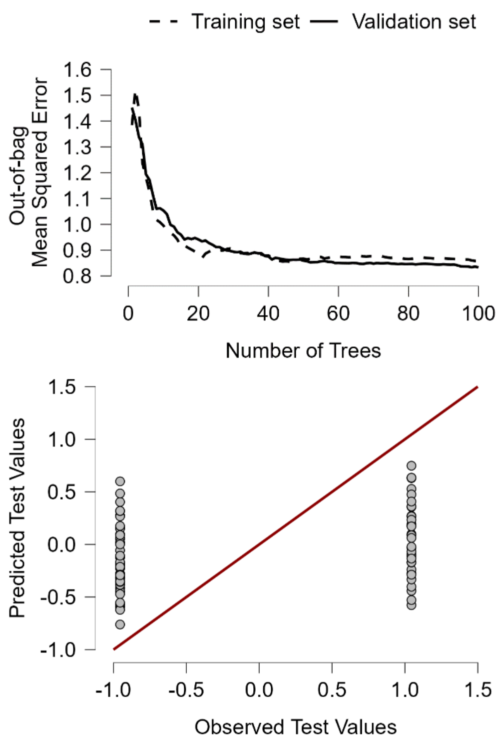

Figure 1. Random Forest Regression Out-of-bag Mean Squared Error Plot and Predictive Performance Plot

Table 3. Feature Importance in Random Forest Regression

	Mean decrease in accuracy	Total increase in node purity
ATAB	0.062	9.412
ARED	0.040	8.029
DIFA	0.053	6.531
TRED	0.010	5.549
MNBRA	0.047	5.348
LLRA	-0.003	5.056
LLRT	-0.010	4.570
MNBRT	0.038	4.139
IR	0.010	4.086
EDA	-9.385×10^{-4}	4.029
LRED	-0.012	4.022
TTAD	0.003	3.814
TDIFA	0.009	3.792
TTAB	-0.002	3.713
ART	-0.001	3.698
ARA	-0.003	3.645
ATAD	0.014	3.099
PATTR	0.005	3.064
RC	-0.009	2.754
RR	0.006	2.692
TSTOPRA	0.016	1.865
MBT	-9.821×10^{-4}	1.708
TSTOPR	0.005	1.333
RE	0.005	0.662
REV	1.656×10^{-4}	0.191

4.2 Support Vector Machine Regression

Support Vector Machine Regression was employed to model the relationship between physiological data and deceptive behaviour. The model utilized 320 support vectors during training and 80 support vectors during testing. The average squared difference between the expected and actual values or the Mean Squared Error (MSE) obtained from the test data was 0.980. As a measure of prediction errors for the original units of the data, the Root Mean Squared Error (RMSE) was calculated, and the obtained result was 0.99 (Table 4).

Table 4. Support Vector Machine Regression

Support Vectors	n(Train)	n(Test)	Test MSE
318	320	80	0.980

The average absolute difference between the expected and actual values was found to be 0.81, which is the Mean Absolute Error (MAE), often referred to as Mean Absolute Deviation (MAD). The average absolute percentage difference between the anticipated and actual values was 81.85%, which is represented by the Mean Absolute Percentage Error (MAPE). The model's R-squared (R^2) value was 0.159, meaning that it could explain about 15.9% of the variation in dishonest behaviour (Table 5).

Table 5. Support Vector Machine Regression evaluation metrics

Metrics	Value
MSE	0.98
RMSE	0.99
MAE / MAD	0.81
MAPE	81.85%
R^2	0.159

These metrics collectively evaluate the model's accuracy and its ability to predict deceptive behaviour based on physiological data, highlighting both strengths and areas for potential improvement in future iterations.

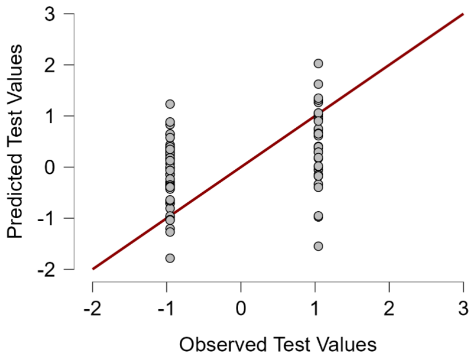

Figure 2. Predictive Performance Plot for Support Vector Machine Regression

4.3 Neural Network Regression

Neural Network Regression was also employed to predict deceptive behaviour based on physiological data obtained from polygraph tests. The model architecture included a single hidden layer with 1 node, trained using 320 samples and tested on 80 samples. The model, optimized based on the sum of squares, obtained a Mean Squared Error (MSE) of 0.894 for the test data. This value represents the average squared difference between the predicted and actual values, as shown in Table 6.

Table 6. Neural Network Regression

Hidden Layers	Nodes	n(Train)	n(Test)	Test MSE
1	1	320	80	0.894

The average amount of prediction errors in the original dataset was quantified using the Root Mean Squared Error (RMSE), which was found to be 0.946. The value obtained for the Mean Absolute Error (MAE), also known as the Mean Absolute Deviation (MAD), was 0.88. The average absolute difference between the expected and actual values is shown in Figure 3. In addition, the average absolute percentage difference between the expected and actual values was found to be 88.74% using the Mean Absolute Percentage Error (MAPE) measurement. With a coefficient of determination (R2) of 0.113, the model can explain approximately 11.3% of the variation in dishonest behaviour (Table 7).

Table 7. Neural Network Regression evaluation metrics

Metrics	Value
MSE	0.894
RMSE	0.946
MAE / MAD	0.88
MAPE	88.74%
R^2	0.113

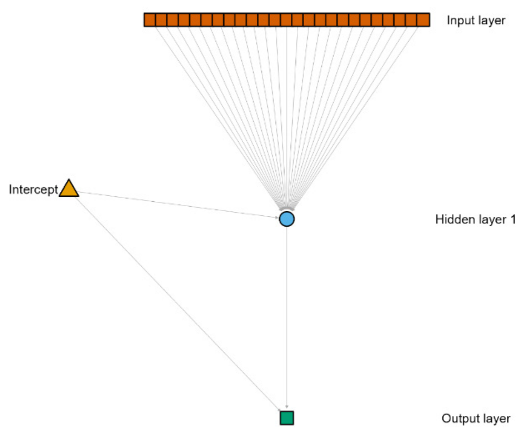

Figure 3. Predictive Performance Plot for Neural Network Regression

In terms of network weights (Table 8), each input feature was weighted accordingly for the logistic sigmoid activation function used in the hidden layer. Notably, features such as Amplitude of Blood Pressure in Brachial Pulse (ATAB) and Amplitude of Electrodermal Reaction (ARED) exhibited a significant influence, with weights of -1.300 and -1.509, respectively, highlighting their impact on the model's predictions.

Table 8. Neural Network Regression Network Weights

Node	Layer	Node	Layer	Weight
Intercept		→ Hidden 1	1	0.256
ARED	input	→ Hidden 1	1	-1.509
ATAB	input	→ Hidden 1	1	-1.300
ATAD	input	→ Hidden 1	1	-0.919
ART	input	→ Hidden 1	1	1.436
ARA	input	→ Hidden 1	1	1.426
MNBRA	input	→ Hidden 1	1	-3.056
MNBRT	input	→ Hidden 1	1	-2.603
IR	input	→ Hidden 1	1	1.102
LLRT	input	→ Hidden 1	1	0.850
LLRA	input	→ Hidden 1	1	1.317
LRED	input	→ Hidden 1	1	-0.850
TRED	input	→ Hidden 1	1	2.062
TTAB	input	→ Hidden 1	1	-2.106
TTAD	input	→ Hidden 1	1	-0.816
RR	input	→ Hidden 1	1	0.068
RC	input	→ Hidden 1	1	-0.664
TSTOPR	input	→ Hidden 1	1	-1.443
TSTOPRA	input	→ Hidden 1	1	-0.867
RE	input	→ Hidden 1	1	1.152
REV	input	→ Hidden 1	1	1.460
PATTR	input	→ Hidden 1	1	-1.389
EDA	input	→ Hidden 1	1	-3.699
MBT	input	→ Hidden 1	1	-3.470
DIFA	input	→ Hidden 1	1	-2.613
TDIFA	input	→ Hidden 1	1	-1.927
Intercept		→ score expert	output	0.502
Hidden 1	1	→ score expert	output	-5.406

Note: The weights are input for the logistic sigmoid activation function.

5. Discussion: Comparative Analysis of Regression Algorithms

This study used physiological data from polygraph testing and Random Forest Regression (RFR), Support Vector Machine (SVM) Regression, and Neural Network Regression (NNR) techniques to predict dishonest conduct. The criteria used for assessing each method were R-squared (R^2), Mean Absolute Error (MAE) or Mean Absolute Deviation (MAD), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). These metrics provided essential data about each model's expected accuracy and overall effectiveness.

An analysis revealed that NNR performed better than the other two methods: the mean squared error (MSE) was 0.894, the root mean squared error (RMSE) was 0.946, the mean absolute deviation (MAE/MAD) was 0.88, the mean absolute percentage error (MAPE) was 88.74%, and the R^2 value was 0.113. The neural network architecture was able to capture the intricate relationships seen in the context of physiological data, even though it only had one node and one hidden layer. In comparison with RFR and SVM, the higher R^2 value suggests that NNR explained a higher percentage of the variance in deceptive behavior. Its capacity to identify important predictors was further demonstrated by the network weights analysis, which further showed the significant impact of particular physiological parameters.

The higher R^2 value indicates that NNR captured more of the variance in deceptive behaviour, suggesting that it can leverage the underlying patterns in the analysed data more effectively than the other models. Additionally, despite its simpler architecture, NNR's ability to handle non-linear relationships and capture complex feature interactions contributed to its superior performance.

5.1 Comparative Analysis Including Contemporary Methods

Recent research within the last 2-4 years has shown the growing utility of ensemble learning techniques, deep learning models, and hybrid approaches in predictive modelling, particularly for complex datasets like physiological data from polygraph testing. For instance, Gradient Boosting Machines (GBMs) and XGBoost have gained popularity for their superior handling of overfitting and high-dimensional data in comparison with Random Forests. Studies such as that of Sahin (2020) have demonstrated the superior performance of these methods over traditional RFR in predictive tasks in a similar way to this paper.

Furthermore, advancements in deep learning, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs), have enabled them to outperform simpler neural network architectures, particularly in handling sequential and time-series data such as the physiological responses analysed in this study.

For instance, the study of Wunsch et al. (2021) found that deep learning models with multiple layers captured more complex patterns in similar polygraph datasets.

While this study focused on a basic NNR model, its strong performance relative to RFR and SVM indicates that more complex neural network architectures have the potential to further improve their predictive accuracy.

This paper contributes to the existing research by showing that even less complex neural network models can achieve a better performance than typical machine learning methods. Nevertheless, carrying out a thorough analysis that directly compares these models with cutting-edge techniques such as GBM and deep learning would provide a more extensive comprehension of their potential in predicting deceitful behaviour using polygraph data.

6. Conclusion

This study investigated the prediction ability of Random Forest Regression (RFR), Support Vector Machine (SVM) Regression, and Neural Network Regression (NNR) in detecting deceitful behavior using physiological data obtained from polygraph examinations. Each algorithm showed different levels of efficacy in identifying and predicting subtle physiological indicators linked to deceit.

Random Forest Regression (RFR), known for its ensemble learning based approach and feature importance assessment, provided robust predictions with a MSE value of 0.893 and a R^2 value of 0.091. This method effectively highlighted key physiological indicators contributing to deceptive behaviour, such as blood pressure variations and electrodermal responses.

Support Vector Machine (SVM) Regression demonstrated a competitive performance with a MSE value of 0.98 and a R^2 value of 0.159. The capacity of Support Vector Machines (SVMs) to identify non-linear correlations in high-dimensional environments has been found to be advantageous, especially in detecting patterns within physiological data associated with dishonest tendencies. Neural Network Regression (NNR) proved to be the most effective model in this investigation, featuring the highest level of predictive accuracy and explanatory capability. NNR attained a Mean Squared Error (MSE) of 0.894 and a notably higher R^2 value of 0.113, which proves its ability to identify complex interactions and accurately represent fluctuations in physiological reactions linked to deceit. The employment of a single hidden layer neural network architecture allowed this method to accurately model the complex relationships between physiological parameters, surpassing both RFR and SVM in terms of predictive precision.

The superior performance of NNR underscores its potential in practical applications requiring precise and reliable predictions of deceptive behaviour based on physiological data. Its ability to learn and adapt to complex patterns within the analysed dataset suggests promising avenues for further research. Subsequent research could prioritize the improvement of neural network structures, investigate supplementary physiological characteristics, or use sophisticated machine learning methods to further enhance the proposed model's predictive capabilities. By refining these methodologies, progress can be made with regard to the understanding and application of predictive analytics in forensic psychology and related fields. Neural Network Regression stands out as the optimal choice among the evaluated algorithms for its ability to leverage physiological data in order to predict deceptive behaviour accurately. Its performance highlights its potential to contribute significantly to advancements in detecting deception and informing decision-making processes in forensic investigations and security applications.

REFERENCES

Armstrong, J. S. & Collopy, F. (1992) Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*. 8(1), 69-80. doi: 10.1016/0169-2070(92)90008-W.

Chai, T. & Draxler, R. R. (2014) Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*. 7(3), 1247-1250. doi: 10.5194/gmd-7-1247-2014.

Cortes, C. & Vapnik, V. (1995) Support-vector networks. *Machine Learning*. 20(3), 273-297. doi: 10.1007/BF00994018.

Derakhshan, A., Mikaeili, M., Gedeon, T. & Nasrabadi, A. M. (2020) Identifying the optimal features in multimodal deception detection. *Multimodal Technologies and Interaction*. 4(2), 25. doi: 10.3390/mti4020025.

Diaz, V., Wong, W. E. & Chen, Z. (2023) Enhancing Deception Detection with Exclusive Visual Features using Deep Learning. *International Journal of Performability Engineering*. 19(8), 547. doi: 10.23940/ijpe.23.08.p7.547558.

Dučić, N., Dragićević, S., Stepanić, P., Stanković, N. & Marjanović, M. (2023) Development of hybrid model based on artificial intelligence for maximizing solar energy yield. *Studies in Informatics and Control*. 32(4), 95-104. doi: 10.24846/v32i4y202309.

Fernandes, S. V. & Ullah, M. S. (2022) A comprehensive review on features extraction and features matching techniques for deception detection. *IEEE Access*. 10, 28233-28246. doi: 10.1109/ACCESS.2022.3157821.

Filip, F. G. (2021) Automation and Computers and Their Contribution to Human Well-being and Resilience. *Studies in Informatics and Control*. 30(4), 5-18. doi: 10.24846/v30i4y202101.

Iancu, D. T. & Florea, A. M. (2023) An improved vehicle trajectory prediction model based on video generation. *Studies in Informatics and Control*. 32(1), 25-36. doi: 10.24846/v32i1y202303.

Knežević, D., Blagojević, M. & Ranković, A. (2023). Electricity Consumption Prediction Model for Improving Energy Efficiency Based on Artificial Neural Networks. *Studies in Informatics and Control*, 32(1), 69-79. doi: 10.24846/v32i1y202307.

LeCun, Y., Bengio, Y. & Hinton, G. (2015) Deep learning. *Nature*. 521(7553), 436-444. doi: 10.1038/nature14539.

Mohan, K. & Seal, A. (2021) Deception detection on "Bag-of-Lies": integration of multi-modal data using machine learning algorithms. In: *Proceedings of International Conference on Machine Intelligence and Data Science Applications: MIDAS 2020, 4-5 September 2020, Dehradun, India*. Singapore, Springer. pp. 445-456.

Nagelkerke, N. J. D. (1991) A note on a general definition of the coefficient of determination. *Biometrika*. 78(3), 691-692. doi: 10.1093/biomet/78.3.691.

National Research Council. (2003) *The Polygraph and Lie Detection.* Washington, DC, United States, The National Academies Press.

Nawroly, S. S., Popescu, D., Thekekara Antony, M. C. & Muthu Philominal, A. J. (2023) SNR-Selection-Based-Data Augmentation for Dysarthric Speech Recognition. *Studies in Informatics and Control*. 32(4), 129-140. doi: 10.24846/v32i4y202312.

Oswald, M. (2020) Technologies in the twilight zone: early lie detectors, machine learning and reformist legal realism. *International Review of Law, Computers & Technology*. 34(2), 214-231. doi: 10.1080/13600869.2020.1733758.

Rad, D., Paraschiv, N. & Kiss, C. (2023) Neural Network Applications in Polygraph Scoring - A Scoping Review. *Information*, 14(10), 564. doi: 10.3390/info14100564.

Rad, D., Paraschiv, N., Kiss, C., Balas, V. E. & Barna, C. (2024) Physiological Reactions Profiling in Polygraph Testing: Insights from Fuzzy C-Means Clustering Analysis. *Acta Polytechnica Hungarica*, 21(10), 367-377. doi: 10.12700/APH.21.10.2024.10.23.

Raskin, D. C. & Honts, C. R. (2002) The comparison question test. In: M. Kleiner (ed.) *Handbook of polygraph testing.* Cambridge, MA, United States, Academic Press, pp. 1-48.

Sahin, E. K. (2020) Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Applied Sciences*. 2(7), 1308. doi: 10.1007/s42452-020-3060-1.

Shvets, O., Seebauer, M., Naizabayeva, A. & Toleugazin, A. (2023) Monitoring and Control of Energy Consumption Systems, using Neural Networks. *Acta Polytechnica Hungarica*. 20(2), 125-144. doi: 0.12700/APH.20.2.2023.2.7.

Shvets, O., Smakanov, B., Györök, G. & Kovács, L. (2024) A Driver Fatigue Recognition System, Based on an Artificial Neural Network. *Acta Polytechnica Hungarica*. 21(8), 211-226.

Talaat, F. M. (2024) Explainable Enhanced Recurrent Neural Network for lie detection using voice stress analysis. *Multimedia Tools and Applications*. 83(11), 32277-32299. doi: 10.1007/s11042-023-16769-w.

Thannoon, H. H., Ali, W. H. & Hashim, I. A. (2018) Detection of deception using facial expressions based on different classification algorithms. In: *2018 Third Scientific Conference of Electrical Engineering (SCEE), 19-20 December 2018, Baghdad, Iraq*. IEEE. pp. 51-56.

Vavrinsky, E., Stopjakova, V., Kopani, M. & Kosnacova, H. (2021) The concept of advanced multi-sensor monitoring of human stress. *Sensors*. 21(10), 3499. doi: 10.3390/s21103499.

Viji, D., Gupta, N. & Parekh, K. H. (2022) History of Deception Detection Techniques. In: *Proceedings of International Conference on Deep Learning, Computing and Intelligence: ICDCI 2021, 7-8 January 2021, Chennai, India*. Singapore, Springer Nature. pp. 373-387.

Weiwei, W. & Hao, D. (2022) The Application of the Generalized Regression Neural Network Model Based on Information Granulation for Short-Term Temperature Prediction. *Studies in Informatics and Control*. 31(3), 53-62. doi: 10.24846/v31i3y202205.

Willmott, C. J. & Matsuura, K. (2005) Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*. 30(1), 79-82. doi: 10.3354/cr030079.

Wunsch, A., Liesch, T. & Broda, S. (2021) Groundwater level forecasting with artificial neural networks: a comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). *Hydrology and Earth System Sciences*. 25(3), 1671-1687. doi: 10.5194/hess-25-1671-2021.