Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7766199

# ENHANCEDTHE PERFORMANCE OF ADAPTIVE PREDICTIVE ANALYSIS USING OPTIMISED DEEP LEARNING ALGORITHMS FOR BIG DATA

**Sini Shibu[1], Dinesh Kumar Sahu[2]**
[1]Department of Computer Science & Engg., SRK University, Bhopal, India
sini.shibu09@gmail.com
[2]Department Computer Science &Engg., SRK University, Bhopal, India
dinesh_sahu20@yahoo.com

**Abstract**
Predictive analysis and the accuracy of patterns in large amounts of data play critical roles in many disciplines. The process of prediction finds important information in a large amount of data that would otherwise be obscured. In the current decade, all business organisations store digital data in the form of databases and cloud storage. The stored data analyses business performance and forecasts future business planning. The planning of business depends on the accurate and precise pattern analysis of big data. Recently, several prediction-based linear and non-linear machine learning and data mining algorithms have been proposed. These algorithms face the problem of multiple features in big data and the declining accuracy of predictive algorithms. This paper proposes a hybrid algorithm based on convolutional neural networks (CNN) and BAT optimization algorithms. The BAT optimization algorithm reduces the variance of multiple features in big data. The optimised features data process of classification uses the CNN model. The proposed algorithm is tested on two standard datasets: KDDCUP2003 and heart disease data. The parametric analysis of the proposed algorithm shows better performance than existing algorithms such as RNN and CNN.
**Keywords: - Predictive Analysis, Bigdata, Bat, CNN, RNN, Machine Learning, KDD**

## Introduction

The fact that digital data is growing at a very fast rate is a great thing for the future market and technology. The growth of data gives pattern analysis of future product trends more variety and room for change. Predictive analysis is a way to produce patterns of data and analyse the patterns of behaviour [1,2,3,4]. The growth of data is an outcome of big data and accelerates several properties of big data, such as velocity, variety, and volume. The conventional machine learning and data mining approach faces a problem of building accurate models for predictive mining. Machine learning algorithms compromise with the bottleneck problems of overfitting and overfitting [5,6,7]. The problem of overfitting and underfitting decreases the capacity and efficiency of models. Adaptive learning approaches are a new area of data analysis and vital tools for evaluating patterns in high-volume big data. Due to their efficient processing of several sectors and their rapid and continual methodological enhancement. The several authors reporting on the survey suggest that the most critical and challenging issue in modern data analysis tools is onset prediction of patterns [8,9,10,11]. The application of big data analysis is found in several fields, such as health care data, security data, and cloud data. The evaluation

of patterns in this big data plays a vital role in the development of a country and business. In the current decade, several algorithms for predictive learning are proposed, such as linear, logistic, classification, and regression trees; learning vector quantization; support vector machines; boosting; and deep neural networks. The objectives of algorithms and their outcomes have major differences. The major role of a machine-learning-based algorithm is the analysis of the non-linear relationships among several variables [12]. The approach of ML-based algorithms is accurate for big data, but the training of massive data is time-consuming and expensive to obtain efficient patterns. Despite several approaches to machine learning algorithms, the hybrid model is another approach to predictive pattern analysis. The formation of hybrid models uses data mining techniques and optimization approaches. Machine learning and adaptive learning perform better than other approaches to predictive analysis, but with certain limitations [13,14]. The extension of machine learning algorithms such as deep learning, convolutional neural networks (CNN), and recurrent neural networks improves the performance of big data analysis[15]. In this paper, we propose optimised CNN models for the predictive analysis of big data. The proposed algorithm is very efficient, reduces the training time of sample data, and increases the efficiency of data processing. For the analysis of algorithms, I employed network security data and heart disease data. In real time, these two big data sets play a vital role in business processing in any developed country. The remainder of the paper describes related work in Section II, proposed methodology in Section II, experimental analysis in Section IV, and conclusions in Section V.

## II. Related Work

The diversity and size of big data make it a challenging task for machine learning and neural network algorithms for prediction and analysis. Recently, several authors proposed an adaptive model using deep learning and optimization algorithms for the predictive analysis of big data. This section explores recent work in the area of predictive analysis for big data.The author [1] A key objective for improving genomic prediction methods is the possibility of more effectively integrating physiology-based selection strategies on various agro-ecological and agro-management systems. When comparing genotype SY performance in row spacing treatments, 69% had superior SY response in 38cm, while 31% had greater SY in 76 cm row spacing.The author [2] Future research will concentrate on the intricate design of the partial models, including a thorough description model for corporate objectives. This paper promotes a methodology that is capable of assisting portfolio management in making the best decision to design the product portfolio in accordance with a holistic framework.The author [3] this position paper is an effort to provide a framework for predictive and prescriptive governance, with the goal of supplying information to help policymakers make well-informed and fact-based decisions.The author [4] the primary goal of this effort is to create a new LSTM tool for forecasting flood dispersion over. By the use of test scenarios related to the Kushiro-Tsunami, the performance of the LSTM-ROM produced here has been.The author [5] these tactics include, among other things, developing a predictive model to pinpoint workers who are likely to leave. This section provides a summary of some of the more important subjects that will probably come up when implementing various HRA initiatives.The author [6] the suggested smart-wearable uses biometrics to analyse a person's emotional state and forecast the likelihood of a subsequent impulsive outburst. In 93% of the cases, the experimental subject was within

10 min and in 82% within 5 min of the anticipated time.The author [7] we show that a variety of decision problems can be solved using our suggested solution approaches, which are modelled after machine learning (ML) techniques such local regression (LOESS), classification and regression trees (CART) and random forests (RF). According to the coefficient of prescriptiveness, this methodology accounts for an 88% improvement.The author [8] In this essay, we contend that next event forecasting is insufficient for practitioners. In light of this, this work-in-progress study suggests a method for selecting the process-representative next best actions.The author [9] this article should be interpreted as a methodology recommendation for the implementation of an autonomous forecasting model in a test bench for the detection of a machine failure and to support the creation of algorithms for preventive and descriptive maintenance.The author [10] The function of PPA in healthcare is also examined from the views of various stakeholders. Our findings show that the stakeholders pursue divergent interests, necessitating governmental control to allow PPA to spread widely.The author [11] Twelve ANN models were used in a comparative analysis to determine which ANN method would be most effective for estimating the target function. The outcomes demonstrated that 6 and 10 neurons in the first and second hidden layers, respectively, were the optimal topology for the two-hidden layer ANN.The author [12] Power plant's preventative maintenance and diagnosis of its flame tube serves as an illustration of the methodology. The neural network was trained using the physics-based model in a variety of malfunctioning flame tube scenarios, and it obtained an accuracy of greater than 0.95.The author [13] The method that is being given is based on GGNNs, which are ideal for describing a business process' sequential flow. After 548 epochs and using the previously mentioned setup, we had an accuracy by the total number of classifications of 77.00%.The author [14] the description of the current issue and suggested approach for figuring out the best order numbers for products in multi-item inventory management. The effectiveness of the suggested strategy is examined by various analyses, and the following crucial findings are noteworthy here.The author [15] The model outperformed the benchmark techniques, which comprised Buy and Hold trading strategies, multiple linear regression, trading, and analyst predictions. The ANN methodology used to make the predictions achieved accuracy performance scores of up to 60%.The author [16] A significant e-commerce platform provided two real-world datasets that were used to evaluate the proposed framework. In comparison to the standard industry practice, the results show that the decision support framework is beneficial in providing delivery cost savings of up to 10.6%.The author [17] We suggest a holistic end-to-end prescriptive maintenance Framework (HeePMF) that combines equipment and operational data, analysis of maintenance needs, and feedback with predictive technology to produce insights that can be put into practice.The author [18] We put forth a cutting-edge technique to forecast a business process' future performance at the level of the process model. To be more specific, we build an annotated transition system and use it to produce a process representation matrix.The author [19] The control-flow conformity of the suggested actions is ensured by this technique's utilisation of business process simulation. The recommendation of more sophisticated next best actions can be facilitated by more powerful multi-tasking DNN systems.The author [20] Hence, if an unplanned outage happens, a measure is taken, and a few hours later the same or a related fault manifest, for instance. As a result, plant availability rises by about 2% to a total of

94%.The author [21] The suggested methodology can be used as a support software tool for recruiters and HR managers and can be applied directly by HR professionals without the requirement for more in-depth technical or machine learning understanding.The author [22] I am able to avoid generic functional and non-functional needs for prescriptive recommender systems because of this. The applicability of various AI and ML algorithms for the anticipated PRS criteria needs to be further assessed in future studies.The author [23] Shift winding operation data parameterization and combining with configuration data enabled behaviour-based management of robotic remanufacturing using vector mapping and deep neural learning. For random wires, the dense orthocyclic winding provides the highest packing density of 90.7%.The author [24] Prescriptive models were created and evaluated in an offset well in the North Sea using a strict data processing and modelling methodology. Yet, in comparison to the earlier casing departure tasks, the TOB values often dropped by only 2%.The author [25] The methodology includes data processing, cleaning, and ingestion into databases; automated machine learning (Auto ML) application to produce an accurate machine-learning model; and numerical optimization of decision parameters to minimise an economic objective. Savings of 5 to 32% on well completion costs were attainable while keeping production levels the same.The author [26] It contributes to IS-related developments as a prescriptive science by identifying meaning-sends-connections. We define a collection of integrative research techniques that take place at the explanation-prescription nexus, the intersection of explanatory, predictive, and prescriptive science.The author [27] the empirical study examines the accuracy of the suggested strategy on various benchmark outcome prediction problems in comparison to leading-edge rival approaches, demonstrating its viability.The author [28] The research approach used entails a survey review looking into various machine learning and deep learning algorithms and how they are used in AI to get an advantage.The author [29] The strong ability of Machine Learning (ML) algorithms to explore nonlinear correlations among data variables is one reason why they are being used more and more for predictive analytics.The author [30] this study suggests a novel approach to optimization in the WWTP area: long short-term memory (LSTM) artificial neural networks (ANN) combined with genetic algorithms (GA).The author [31] we outline our method's structure and use simulated data to compare its performance to more established techniques like point-estimate-based optimization, stochastic optimization, and recently discovered machine learning-based optimization techniques.The author [32] Also, the suggested model can automatically create the daily plan for the interpreter, saving the hospital a lot of time. This study will be expanded in subsequent research by taking into account the travel time between various hospital locations when the distance is considerable.The author [33] As real-world use cases made possible by a classification model for the detection of product movements on the store floor, we look at electronic article monitoring and automated checkouts. Ensemble techniques and different algorithmic strategies may assist develop a more trustworthy detection system by increasing prediction power even further.The author [34] says: „For identifying the type of fluid and determining saturation in geological formations, this method is very sensitive and successful. If electrodes are placed in the production/injection wells with close proximity to one another, resolution will improve.To more effectively handle the inherent stochasticity in predictions, the suggested approach assesses customer interruptions in terms of probability distributions. Contrary to the traditional

BMA technique, the suggested algorithm bases the base learner weights on a multinomial logistic function of the data.The author [36] to advance prediction analysis utilising prescriptive analytics and to offer real-time effective solutions to any stock market user. Even though the present systems predict stock utilising multiple methodologies and algorithms. Our proposed system's prediction methodology, however, has been successful in reducing the accuracy gap by a maximum of 0.93%.The author [37] recently has the focus switched from model-based information systems and process models, which were the centre of conventional BPM research, to data-driven techniques like process mining.The author [38] says, our technique uses a straightforward recurrent neural network with some creative tweaks that incorporate Causal Graph modelling. We outline the fundamentals of causality and causal graphs and develop a Learning Causal Digital Twin (LCDT) solution.The author [39] Max UpTM Fleet, created by LSA, is a vehicle fleet asset management system that will reduce unplanned breakdowns by 70% to 75% and give a 35% to 45% reduction in downtime.The author [40] For our dataset, the top-performing pipeline had an area under the receiver operating characteristic curve of 0.82 and an average precision of 84.2% in predicting fraudulent behaviour.

## III. Methodology

This section describes the proposed methodology of predictive analysis for big data. The proposed algorithm is an encapsulation of the bat optimization algorithm and the convolutional neural network (CNN). The employed Bat algorithm works on multiple variables of features and reduces the variance of features in big data. The CNN algorithm performs the classification task for big data. Here we explore the bat optimization algorithm and the CNN algorithm and finally describe the proposed algorithm.

### A. Bat optimization algorithm

Bat optimization algorithm is bio-inspired meta-heuristic function for global optimal solution[19]. The approach of bat algorithm is population-based for enthralling actions of bat group such as determining the location of food source and categories numerous types of insects in whole dark atmosphere. The researchers are motivated to analyses the bat algorithm because of its unique echolocation capability. The entire bat group uses sonar, also known as echolocation, to locate the location of the food source and to avoid obstacles. The bat group can locate a food source by sending low- and high-frequency sound pulses, which hit and bounce back to the bat. The processing of bat algorithm based on three rules as

1. All bats use echolocation to detect distance and know difference between food and progressive obstacles
2. Bats fly randomly at velocity Vi, at position xi, with constant frequency of freqi and different wavelength λ and loudness of Ao for hunting prey. Also they can automatically set emitted waves and sent pulse rates (r ϵ[0.1]) according to proximity to their hunts.
3. Given the loudness may vary in many different ways, consider that loudness varies form Ro(maximum value) to Rmin(minimum value).

By the rule the position $x_i^t$ with velocity $v_i^t$ for each artificial bat I in iteration t and frequency frqi is estimated as

$$freq_i = freq_{min} + (freq_{max} - freq_{min}).\beta \ldots\ldots\ldots\ldots\ldots (1)$$
$$v_i^t = v_i^{t-1} + (x_i^{t-1} - X^*).freq_i \ldots\ldots\ldots\ldots\ldots\ldots (2)$$
$$x_i^t = x_i^{t-1} + v_i^t \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Here $\beta$ ϵ[o,1] is a random vector with uniform distribution, $X^*$ is the best current position that is selected in each iteration and after comparison with the position of the artificial bats. Now freqi is selected between freqmin=0 and freqmax=100. In each iteration of local search, one solution is selected as the best solution(BS), and new position of each bat is updated with a random step as follows

$$X_{new} = x_{old} + \epsilon.A^t \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Here ϵ ϵ[-1,1] is a random number and $A^t$ =<$A_i^t$> is the average loudness of bats in iteration t. loudness $A_i$ and pulse rate r are updated as

$$A_i^{t+1} = \alpha.A_i^t r_i^{t+!} = r_i^o.exp(-y.t) \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Here $\alpha$ and Y are constants and for each 0<$\alpha$<*1 and r>0 when* t→∞, we have

$A_i^t --> 0\ r_i^t \rightarrow r_i^0\ t \rightarrow\infty$……(6)

## B. Convolution neural network

Convolutional neural network (CNN), is deep learning algorithm for image data classification[23]. The processing of CNN is linear and non-linear function of fully connected layer (FC), the CNN model consists of nonlinear function with an activation function. Every element of the input and pooling layer is affected by this nonlinear function, which also reduces the size of the final results. To analyse the image inputs, multiple perceptron's are used, and they are trained with bias values and learnable weights that can be used to separate pixel values in different parts of an image. The fact that CNN uses a local spatial domain for the input images, along with a small number of shareable parameters and fewer weights, is one of its main advantages. Due to its simpler computations and less memory usage than other models, this method generally performs better. The CNN is set of input layer, convolutional layer, pooling layer, fully connected layer and output layer. The varying capacity o layers robust the CNN classifier for the classification and detection of data. consider that the input features of CNN are map of layer x is Mx(M0=F). now the convolutional process can be expressed as

$$Mx = f\left(M_{x-1} \bigotimes Wx + bi\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

Here Wx is the convolutional kernel weight vector of the x layer, the symbol ⊗ represents convolutional approach, bi is the offset vector of x layer. F(x) is the activation function. By providing various window values, the convolutional layer extracts various feature information from the data matrix Mi1 and various feature information from the data using various convolution kernels. By sharing the same weight and offset throughout the convolution operation, the same convolution kernel adheres to the notion of "parameter sharing," significantly reducing the number of parameters used by the complete neural network. Following the convolutional layer, the pooling layer typically samples the feature map using various sampling algorithms. The pooling layer may be written as follows if Mx is the input and Mx+1 is the output of the pooling layer.

$$M_{x+1} subsampling(Mx) \ldots\ldots\ldots\ldots\ldots\ldots (8)$$

The window region's mean or maximum value is typically chosen by the sampling criterion. The pooling layer primarily minimizes the feature's size, which lessens the impact of redundant features on the model.

### C. Proposed Methodology

The processing of algorithm describes as the feature of data is f1,f2,……,fn, the B is set of population of artificial Bat and b is set of sample , N is total set of data, OFS is optimized set of features, P is pattern, W is wight of vector, Z is space of mapping of features.

1. Input: feature data $F = \{f_1, \dots\dots, f_n\}$
   Population of artificial bat $B\ (select\ A:\ N\ =\ B\ -\ b)$
2. Output: set of $optimized\ features(OFS) = \{OPS(F_1), \dots\dots\dots, OFS(p_n)\}$
3. $i\ \leftarrow 0;$ {step}
4. for all $f_t \in F$do
5. $OFS(f_t) \leftarrow$ iteration $(t_i)$
6. if the total number of features vector $=\ b$ then
7. $CNN^i \leftarrow \{f_{i\frac{b}{2}+m}\ B \in \{1, \dots\dots, \frac{b}{2}\}\}$
8. $\left(max^i, N^i\right) \leftarrow train - patterns\ (p^i)$
9. for all $P_j^i\ \in\ RLU^i$ do
10. estimate $P - stride\ (p_j^i), loudness(p_j^i), OFS(p_j^i)$
11. end for
12. update vector $CNN^i$
13. if $i >\ 0$ the
14. $(Z;\ W)$Weighted $pattern(P^i\ \cup\ P^{i-1}, N^i\ \cup\ N^{i-1})$
15. for all $Z_j \in Z$ do
    Estimate $P - variance\ (Z_j)$
16. end for
17. $P^i \leftarrow Z$
18. $N^i \leftarrow \{\sum_{wI \in W_1} w_I, \dots\dots\dots, \sum_{wI \in W_c,} W_I\}$
19. Update $P^{i-1}, Z$
20. Update $set\ of\ OFS$
21. end if
22. $i\ \leftarrow i + 1$ ;
23. end if
24. end for
25. prediction of patterns

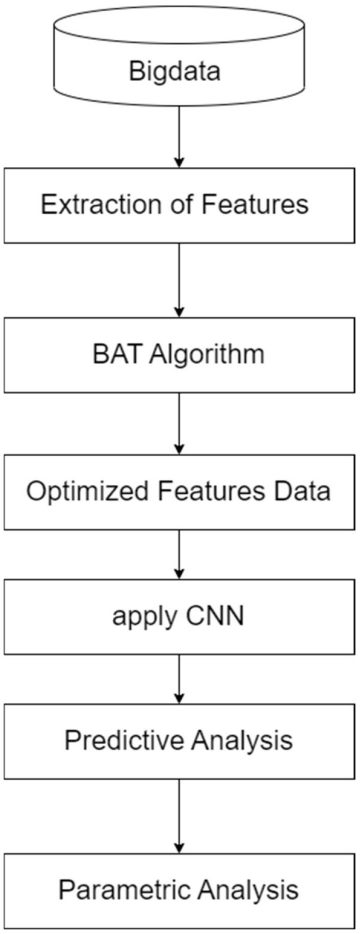


Figure 1 proposed model of predictive analysis of bigdata using BAT algorithm and CNN

## IV. Experimental analysis

To evaluate the performance of the proposed algorithm, MATLAB software is used for the simulation process. The efficiency of the proposed algorithm is evaluated using parametric measures like accuracy, specificity, and sensitivity. By the confusion matrix, estimate the value of prediction as true positive (TP), true negative (TN), false positive (FP), and false negative (FN)[24,25,26].

Sensitivity- Precision measures the proportion of predicted positives/negatives which are actually positive/negative.

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \text{.....(9)}$$

Specificity -It is the proportion of actual positives/negatives which are predicted positive/negative.

$$Specificity = \frac{TN}{TN + FP} \times 100 \text{.....(10)}$$

Accuracy-It is the proportion of the total number of predictions that were correct or it is the percentage of correctly classified instances.

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \text{.....(10)}$$

**Dataset**

To validate the proposed algorithm, used several datasets of heart disease and intrusion detection system. The source of dataset is UCI machine Learning Repository. All dataset free available for study purpose. The description of dataset mention below[30].

Cleveland

This database contains 76 attributes, but consider only of 14 of them. The total number of instances is 303.

Z-Alizadeh Sani[31]

This dataset contains 270 instance and 13 attributes. Each patient could be in two possible categories CAD or Normal. A patient is categorized as CAD, if his/her diameter narrowing is greater than or equal to 50%, and otherwise as Normal.

Statlog[33]

This dataset has been 270 instance and 13 attributes. The missing attribute of dataset is null.

KDDCUP2003[35]

KDDCUP2003 is an upgraded version of the KDD cup99 data set that suggests solutions to some of the previous version's issues. This data collection is a useful benchmark for researchers to compare different types of intrusion detection system (IDS) methodologies, create an intrusion detection system (Host based or Network based), and conduct experiments in the cyber security field.

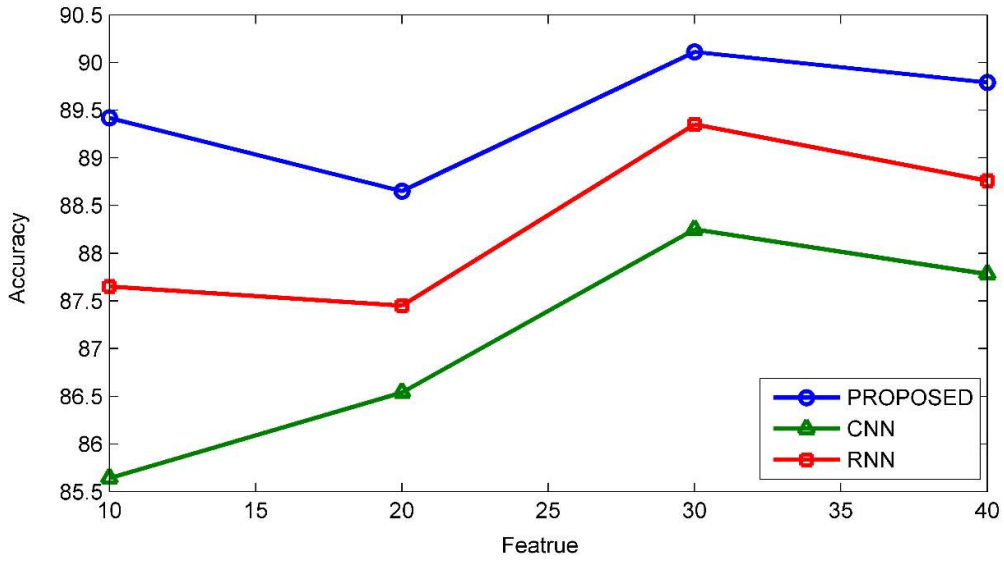


Figure:2 Comparative performance of result analysis of CNN, RNN, and Proposed algorithm for accuracy

Here we can see that the value of proposed is better than the value of RNN and CNN which is such that the lowest value of proposed is 88.56 and highest value is 90.11 the lowest value of RNN is 87.45 and the highest is 88.76 and the lowest value of CNN is 85.64 which is the best result proposed by most CNN and RNN.

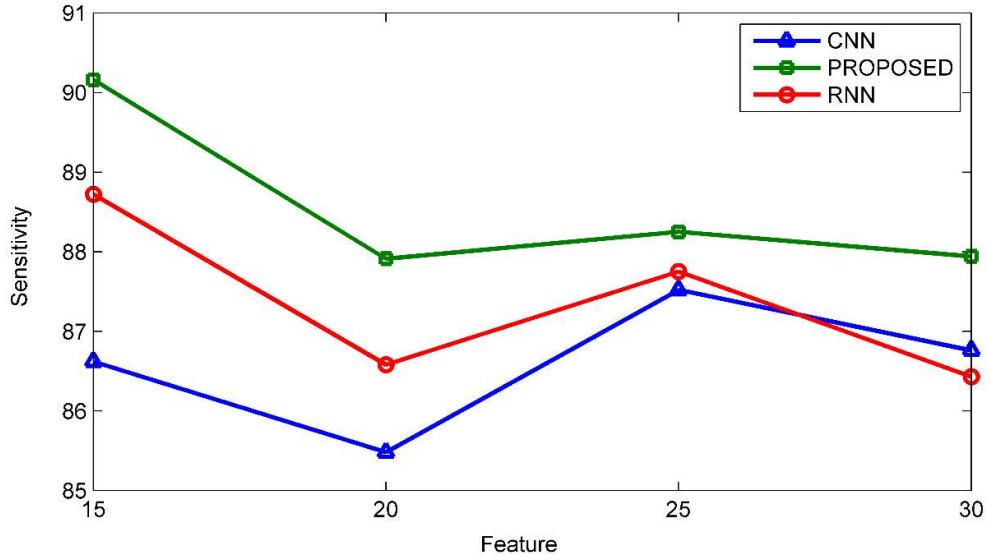


Figure:3 Comparative performance of result analysis of CNN, RNN, and Proposed algorithm for sensitivity

Here we can see that the value of proposed is better than the value of RNN and CNN which is such that the lowest value of proposed is 87.91 and highest value is 90.16 the lowest value of RNN is 86.43 and the highest is 88.72 and the lowest value of CNN is 85.48 which is the best result proposed by most CNN and RNN.

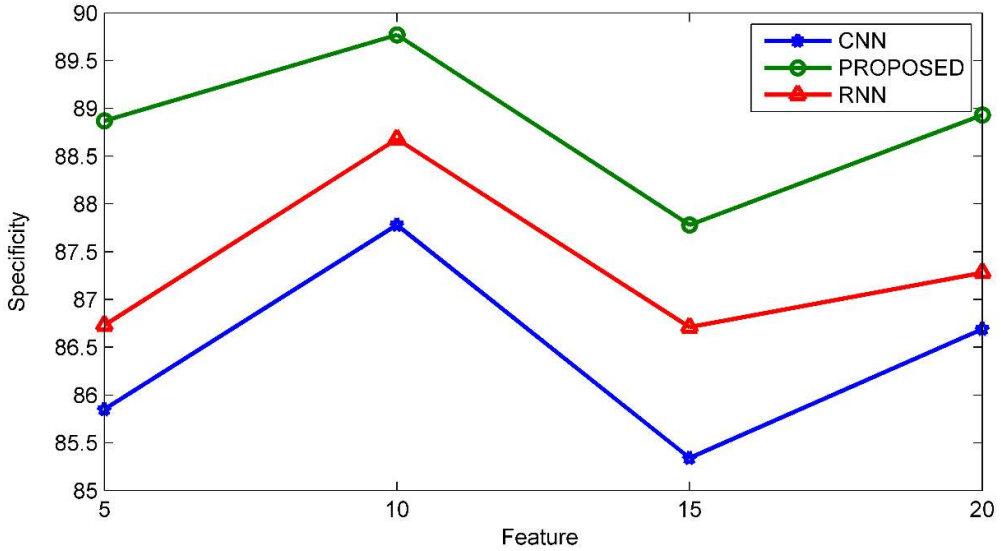


Figure:4 Comparative performance of result analysis of CNN, RNN, and Proposed algorithm for specificity

Here we can see that the value of proposed is better than the value of RNN and CNN which is such that the lowest value of proposed is 87.78 and highest value is 89.77 the lowest value of RNN is 86.71 and the highest is 88.68 and the lowest value of CNN is 85.34 which is the best result proposed by most RNN and CNN.

## V. Conclusion & Future Work

The efficiency of the proposed algorithm is proved by solving several issues related to the predictive analysis of big data. The proposed algorithm overcomes the limitation of local feature selection for the mapping of feature space. The proposed algorithm employed bat optimization for the optimization of multiple attributes of big data. Since these metaheuristic optimization algorithms are robustly practicable, BAT is used. After that, a convolutional neural network (CNN) is created, and the proposed algorithm is used to carry out the classification. This categorization distinguished between normal and attack data during the classification process. The dataset has been updated to include the attack range, which is categorised as denial-of-service attack, Prob, U2R, and R2L. heart disease datasets from Cleveland and other sources are used to validate the classification of other data. The findings show that the proposed algorithm outperforms existing methods for every parameter that was taken into account. The proposed technique yields accuracy of 93%, sensitivity of 91%, specificity of 90%, and improvements over RNN and CNN. The comparison findings are explained as follows: The suggested algorithm outperformed CNN and RNN in terms of accuracy by 3% and 2%, respectively. The achieved specificity is greater than 2% for RNN and 3% for CNN. The obtained sensitivity is greater than 2% for RNN and 3% for CNN.

## References

[1]. Parmley, Kyle A., Race H. Higgins, Baskar Ganapathysubramanian, Soumik Sarkar, and Asheesh K. Singh. "Machine learning approach for prescriptive plant breeding." Scientific reports 9, no. 1 (2019): 17132.

[2]. Jank, Merle-Hendrikje, Christian Dölle, and Günther Schuh. "Product portfolio design using prescriptive analytics." In Advances in Production Research: Proceedings of the 8th Congress of the German Academic Association for Production Technology (WGP), Aachen, November 19-20, 2018 8, pp. 584-593. Springer International Publishing, 2019.

[3]. Misra, Santosh K., and Dibu John Philip. "Predictive and prescriptive governance-improved delivery of government services."

[4]. Hu, Rong, Fangxin Fang, C. C. Pain, and I. M. Navon. "Rapid spatio-temporal flood prediction and uncertainty quantification using a deep learning method." Journal of Hydrology 575 (2019): 911-920.

[5]. Fischer, Lucas Santos. "A predictive and prescriptive framework for employee churn." NOVA University of Lisbon (2019).

[6]. Raychaudhuri, Soumya Jyoti, Soumya Manjunath, Chithra Priya Srinivasan, N. Swathi, S. Sushma, K. N. Nitin Bhushan, and C. Narendra Babu. "Prescriptive analytics for impulsive behaviour prevention using real-time biometrics." Progress in Artificial Intelligence 10 (2021): 99-112.

[7]. Bertsimas, Dimitris, and Nathan Kallus. "From predictive to prescriptive analytics." Management Science 66, no. 3 (2020): 1025-1044.

[8]. Weinzierl, Sven, Sandra Zilker, Matthias Stierle, Martin Matzner, and Gyunam Park. "From predictive to prescriptive process monitoring: Recommending the next best actions instead of calculating the next most likely events." In Wirtschaftsinformatik (Zentrale Tracks), pp. 364-368. 2020.

[9]. Martins, João Pedro Serrasqueiro, Filipe Martins Rodrigues, and Nuno Henriques. "Modeling system based on machine learning approaches for predictive maintenance applications." KnE Engineering 2020 (2020): 857-871.

[10]. Oesterreich, Thuy Duong, Christian Fitte, Alina Behne, and Frank Teuteberg. "Understanding the role of predictive and prescriptive analytics in healthcare: a multi-stakeholder approach." (2020).

[11]. Rahmanifard, Hamid, Hamzeh Alimohamadi, and Ian Gates. "Well Performance Prediction in Montney Formation Using Machine Learning Approaches." In SPE/AAPG/SEG Unconventional Resources Technology Conference. OnePetro, 2020.

[12]. Belov, Sergei, Sergei Nikolaev, and IghorUzhinsky. "Hybrid Data-Driven and Physics-Based Modeling for Gas Turbine Prescriptive Analytics." International Journal of Turbomachinery, Propulsion and Power 5, no. 4 (2020): 29.

[13]. Harl, Maximilian, Sven Weinzierl, Mathias Stierle, and Martin Matzner. "Explainable predictive business process monitoring using gated graph neural networks." Journal of Decision Systems 29, no. sup1 (2020): 312-327.

[14]. Punia, Sushil, Surya Prakash Singh, and Jitendra K. Madaan. "From predictive to prescriptive analytics: A data-driven multi-item newsvendor model." Decision Support Systems 136 (2020): 113340.

[15]. Weckman, Gary R., Ronald W. Dravenstott, William A. Young II, Ehsan Ardjmand, David F. Millie, and Andy P. Snow. "A prescriptive stock market investment strategy for the restaurant industry using an artificial neural network methodology." In Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications, pp. 217-237. IGI Global, 2020.

[16]. Kandula, Shanthan, Srikumar Krishnamoorthy, and Debjit Roy. A Predictive and Prescriptive Analytics Framework for Efficient E-Commerce Order Delivery. No. WP 2020-11-01. Indian Institute of Management Ahmedabad, Research and Publication Department, 2020.

[17]. Choubey, Suresh, Ryan G. Benton, and Tom Johnsten. "A holistic end-to-end prescriptive maintenance framework." Data-Enabled Discovery and Applications 4 (2020): 1-20.

[18]. Park, Gyunam, and Minseok Song. "Predicting performances in business processes using deep neural networks." Decision Support Systems 129 (2020): 113191.

[19]. Weinzierl, Sven, Sebastian Dunzer, Sandra Zilker, and Martin Matzner. "Prescriptive business process monitoring for recommending next best actions." In Business Process Management Forum: BPM Forum 2020, Seville, Spain, September 13–18, 2020, Proceedings 18, pp. 193-209. Springer International Publishing, 2020.

[20]. Biebl, Fabian, Robert Glawar, Anahid Jalali, Fazel Ansari, Bernhard Haslhofer, Peter de Boer, and Wilfried Sihn. "A conceptual model to enable prescriptive maintenance for etching equipment in semiconductor manufacturing." Procedia CIRP 88 (2020): 64-69.

[21]. Pessach, Dana, Gonen Singer, Dan Avrahami, Hila Chalutz Ben-Gal, Erez Shmueli, and Irad Ben-Gal. "Employees recruitment: A prescriptive analytics approach

via machine learning and mathematical programming." Decision Support Systems 134 (2020): 113290.

[22]. Eider, Markus, and Andreas Berl. "Requirements for Prescriptive Recommender Systems Extending the Lifetime of EV Batteries." In 2020 10th International Conference on Advanced Computer Information Technologies (ACIT), pp. 412-417. IEEE, 2020.

[23]. Matenga, Alice, EriyetiMurena, and Khumbulani Mpofu. "Prescriptive Modelling System Design for an Armature Multi-coil Rewinding Cobot Machine." Procedia CIRP 91 (2020): 284-289.

[24]. Popp, Timothy, Mehrdad G. Shirangi, Ole PetterNipen, and Anders Berggreen. "Prescriptive data analytics to optimize casing exits." In IADC/SPE International Drilling Conference and Exhibition. OnePetro, 2020.

[25]. Feder, Judy. "Prescriptive Analytics Aids Completion Optimization in Unconventionals." Journal of Petroleum Technology 72, no. 04 (2020): 52-53.

[26]. Seidel, Stefan, and Richard T. Watson. "Integrating explanatory/predictive and prescriptive science in information systems research." Communications of the Association for Information Systems 47, no. 1 (2020): 49.

[27]. Pasquadibisceglie, Vincenzo, Annalisa Appice, Giovanna Castellano, Donato Malerba, and Giuseppe Modugno. "ORANGE: outcome-oriented predictive process monitoring based on image encoding and CNNs." IEEE Access 8 (2020): 184073-184086.

[28]. Radanliev, Petar, David De Roure, Kevin Page, Max Van Kleek, Omar Santos, La'Treall Maddox, Pete Burnap, Eirini Anthi, and Carsten Maple. "Design of a dynamic and self-adapting system, supported with artificial intelligence, machine learning and real-time intelligence for predictive cyber risk analytics in extreme environments–cyber risk in the colonisation of Mars." Safety in Extreme Environments 2 (2020): 219-230.

[29]. Tinoco, Joaquim, Manuel Parente, António Gomes Correia, Paulo Cortez, and David Toll. "Predictive and prescriptive analytics in transportation geotechnics: Three case studies." Transportation Engineering 5 (2021): 100074.

[30]. Arismendy, Luis, Carlos Cárdenas, Diego Gómez, Aymer Maturana, Ricardo Mejía, and Christian G. Quintero M. "A Prescriptive Intelligent System for an Industrial Wastewater Treatment Process: Analyzing pH as a First Approach." Sustainability 13, no. 8 (2021): 4311.

[31]. Kolcu, Mehmet, and Alper E. Murat. "Integrated Optimization of Predictive and Prescriptive Tasks." arXiv preprint arXiv:2101.00354 (2021).

[32]. Ahmed, Abdulaziz, and Elizabeth Frohn. "A predictive and prescriptive analytical framework for scheduling language medical interpreters." Health Care Management Science 24 (2021): 531-550.

[33]. Hauser, Matthias, Christoph M. Flath, and Frédéric Thiesse. "Catch me if you scan: Data-driven prescriptive modeling for smart store environments." European Journal of Operational Research 294, no. 3 (2021): 860-873.

[34]. Dell'Aversana, Paolo. "Reservoir prescriptive management combining electric resistivity tomography and machine learning." AIMS Geosciences 7, no. 2 (2021): 138-161.

[35]. Kabir, Elnaz. "Predictive and Prescriptive Analytics for Managing the Impact of Hazards on Power Systems." PhD diss., 2021.

[36]. Meenakshi, N., A. Kumaresan, R. Nishanth, R. Kishore Kumar, and A. Jone. "Stock market predictor using prescriptive analytics." Materials Today: Proceedings (2021).

[37]. Kratsch, Wolfgang. "Data-driven Management of Interconnected Business Processes Contributions to Predictive and Prescriptive Mining."

[38]. Madhavan, P. G. "Evidence-based Prescriptive Analytics, CAUSAL Digital Twin and a Learning Estimation Algorithm." arXiv preprint arXiv:2104.05828 (2021).

[39]. Walsh, John D. Price-Colin S. "Condition-Based Maintenance (CBM) Prescriptive Analytics for Military Vehicles."

[40]. Farrugia, David, Christopher Zerafa, Tony Cini, Bruno Kuasney, and Karen Livori. "A real-time prescriptive solution for explainable cyber-fraud detection within the iGaming industry." Sn computer science 2, no. 3 (2021): 215.