International Journal of Applied Engineering & Technology

DYNAMIC MOMENTUM BACK PROPAGATION ALGORITHM FOR ENHANCING COVID-19 PREDICTION

[1]N.G. Sree Devi and [2]Dr. N. Suresh Singh

[1]Research Scholar, Department of Computer Applications, Malankara Catholic College, Mariagiri,(Affiliated to Manonmaniam Sundaranar University, Tirunelveli -627012)
[2]Research Supervisor, Assistant Professor, Department of Computer Applications, Malankara Catholic College Mariagiri, (Affiliated to Manonmaniam Sundaranar University, Tirunelveli- 627012)
[1]devinair.93@gmail.com and [2]sureshsinghn10@gmail.com

ABSTRACT

The COVID-19 pandemic has presented an unprecedented global challenge, underscoring the need for innovative tools to predict and manage its spread. This research introduces a novel approach to COVID-19 prediction by leveraging medical big data and Artificial Neural Networks (ANN). The objective of this study is to develop a robust predictive model that can assist healthcare professionals, policymakers, and the public in making informed decisions and allocating resources effectively. Medical big data, comprising a wealth of patient information, clinical records, and epidemiological data, serves as the foundation for this model. The utilization of ANN allows for the extraction of complex patterns and relationships within this data, enhancing the accuracy and reliability of predictions. In this research, to address the computational challenges associated with large datasets and enhance the ANN training process, introduce a novel training algorithm referred to as the "Dynamic Momentum Back Propagation Algorithm (DMBP)" as an alternative to conventional backpropagation methods. The accuracy analysis, conducted on two distinct Covid-19 datasets, demonstrates DMBP's consistent outperformance of existing methods. Notably, DMBP achieves an accuracy rate of 94.3% on Dataset 1 and 95.2% on Dataset 2, indicating its superior predictive accuracy. These results highlight DMBP's robustness in COVID-19 prediction, highlighting its potential to provide more reliable outcomes.

Keywords: COVID-19 prediction, medical big data, Artificial Neural Networks (ANN), pandemic management, resource allocation.

1. INTRODUCTION

The COVID-19 pandemic has posed a significant challenge to global public health, causing widespread morbidity and mortality [1]. Early prediction of COVID-19 outbreaks is essential for implementing effective preventive and control measures [2][3]. Medical big data, which encompasses a vast amount of healthcare data, including electronic health records, medical images, and genomic data, has the potential to improve COVID-19 prediction [4]. However, the complexity and scale of medical big data pose challenges in data processing and analysis [5][6]. Artificial Neural Networks (ANNs) are a type of machine learning algorithm that can be used to analyse complex data and make predictions [7]. ANNs have been successfully applied to a variety of healthcare applications, including COVID-19 diagnosis and prognosis [8][9]. This research introduces a novel approach to COVID-19 prediction by using medical big data and ANN. Lot of ANN based covid-19 prediction models are developed [10][11]. However, these models are trained with small datasets. ANNs can be computationally expensive to train, especially on large datasets. Drawbacks of existing ANN-based research for predicting COVID-19 trained on small datasets:

- **Lack of Generalizability:** ANN models trained on small datasets are often overfitting, which means that they learn the training data too well and are unable to generalize to new data [12]. This is a problem for predicting COVID-19, as the pandemic is constantly evolving and new data is being generated all the time.

- **Poor Performance on Unseen Data:** As a result of overfitting, ANN models trained on small datasets often perform poorly on unseen data, such as data from new countries or new time periods. This is a problem for predicting COVID-19, as it is important to be able to predict the pandemic accurately in different settings.

International Journal of Applied Engineering & Technology

- **Sensitivity to Dataset Composition:** The performance of ANN models can be sensitive to the composition of the dataset they are trained on [13]. For example, if the training dataset does not contain enough data from certain demographics or regions, the model may not be able to accurately predict COVID-19 in those settings. This is a problem for predicting COVID-19, as the pandemic has affected different populations and regions in different ways.

The Dynamic Momentum Back Propagation Algorithm (DMBP) is a novel training algorithm propose this research, that addresses the computational challenges associated with large-scale datasets and improves the training process of ANN. The DMBP algorithm is based on the conventional backpropagation algorithm, but it incorporates a dynamic momentum term that adapts to the training data and network parameters. This dynamic momentum term helps to accelerate the training process and improve the convergence of the network to the optimal solution.

The proposed approach consists of the following steps:

1. Data preparation: Medical big data is pre-processed to clean, normalize, and feature extract.

2. ANN model development: An ANN model is developed to predict the likelihood of COVID-19 infection based on the pre-processed data.

3. Computational complexity: The proposed training algorithm, the Dynamic Momentum Back Propagation Algorithm, addresses the computational challenges associated with training ANNs on large datasets.

The proposed approach is evaluated on a real-world dataset of COVID-19 patients. The results show that the proposed approach outperforms conventional machine learning algorithms in terms of prediction accuracy and robustness.

2. LITERATURE REVIEW

Soham Guhathakurata et al [14]. developed a COVID-19 prediction model based on ANN. The model was trained using a dataset of COVID-19 symptoms. It classifies individuals into one of three categories: "Not infected," "Mildly infected," or "Severely infected." However, it's important to note that this model was trained with a relatively limited number of training samples, raising questions about its predictive efficiency and generalizability.

The article by Aljameel et al [15]. proposes a machine learning-based model to predict the disease severity and outcome in COVID-19 patients. The model is trained on a dataset of 287 COVID-19 samples from the King Fahad University Hospital in Saudi Arabia. The data is preprocessed and analyzed using three classification algorithms: logistic regression (LR), random forest (RF), and extreme gradient boosting (XGB).

The article by Pal et al [16]. proposes a symptom-based COVID-19 prognosis system using AI-based IoT and a bioinformatics approach. The system uses various machine learning classifiers, including logistics regression, k-nearest neighbor, support vector machine, random forest, decision trees, Naïve Bayes, and gradient booster, to predict the prognosis of COVID-19 patients based on their symptoms. The authors conducted a comparative analysis of the different machine learning classifiers using a publicly available cloud-stored healthcare dataset.

Sharma et al [17]. propose a deep ensemble learning method for predicting COVID-19 cases using multivariate time series data. Their proposed model, AGLSTM, outperforms other machine learning models, such as LASSO, Ada-Boost, Light Gradient Boosting, KNN, RNN, and LSTM, in terms of accuracy. The authors evaluate their proposed model on two different datasets: one from India and one from multiple countries.

Jayaraj et al [18] proposed a Correlated Feature Extraction and Feed-Forward Artificial Neural Networks to address the challenges posed by the ongoing COVID-19 pandemic. The authors introduce the Correlated Feature Extraction (CFE) method, which dynamically adapts to the virus's behavior, reducing false positive and false negative rates. This method utilizes Feed-Forward Artificial Neural Networks (FFANN) to predict positive cases

based on initial symptoms and estimate severity levels. The research stands out for its potential to significantly lower mortality rates by enhancing COVID-19 prediction and addressing the virus's mutations.

The study by Ganesh Yenurkar and Sandip Mal [19] presents an innovative approach to the urgent challenge posed by the rapid spread of COVID-19. The research focuses on the identification of the disease, mortality rate prediction, and recovery rate prediction, employing deep learning techniques. By leveraging artificial intelligence, the study offers a novel method to detect positive COVID-19 cases and predict their outcomes. Initial preprocessing steps remove unnecessary data, and essential features are extracted using term frequency-inverse term frequency (TF-IDF) and Bag of Words (BoW) techniques. To enhance feature selection, the Mayfly Optimization (MO) algorithm is employed. The core of the proposed system lies in the hybridization of two deep learning models, ResNet and GoogleNet, for prediction.

Zoabi, Deri-Rozov, and Shomron [20] present a machine learning-based approach to predict COVID-19 diagnosis using a concise set of features. By utilizing data from over 51,000 tested individuals, the model effectively predicts COVID-19 test results with high accuracy based on only eight binary features, including sex, age, contact history, and initial clinical symptoms. The approach, trained on nationwide data, enables simple yet accurate identification of COVID-19 cases.

The existing research studies in the field of COVID-19 prediction and management offer valuable insights, but also raise important questions and considerations. The model's performance limited due to the relatively small size of the training dataset, which could affect its predictive efficiency and generalizability.

3. PROPOSED METHODOLOGY
In this section explain the process flow proposed Covid-19 prediction model. The overall architecture of the proposed work is shown in figure 1.
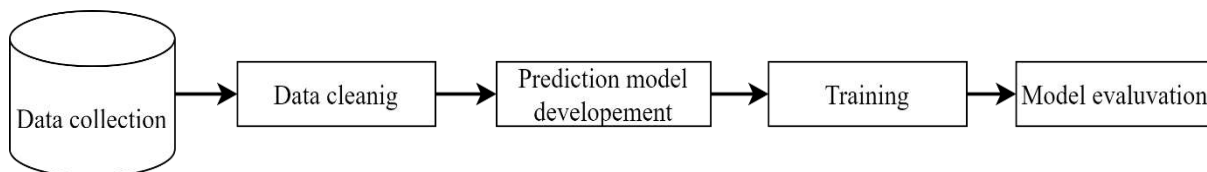

Figure 1: Overall architecture of proposed covid-19 prediction model.

3.1 Dataset Details
This research employs two distinct datasets to train and evaluate a predictive model for COVID-19. The first dataset, the COVID-19 Symptoms Checker, is sourced from Kaggle. It serves as the foundation for the prediction model and comprises 26 attributes related to symptoms and patient information, including a binary class label distinguishing between survivors and non-survivors. The second dataset, used in conjunction with the first, encompasses medical data from 6,512 patients aged 35-75 who received treatment at Asaripallam Medical College in Tamil Nadu, India, and Trivandrum Medical College in Kerala, India. It is collated from various sources and amalgamated into a single dataset. The specific attributes and features are explained in Table 1.

Table 1: Important attributes details of dataset 1 and 2.

Feature	Data Type	Description
Patient Gender	Categorical (String)	This feature records the gender of the patients, indicating whether they are male or female.
Patient Age	Numeric (Integer)	This feature represents the age of each patient in years, capturing the age range within 35 to 75 years.
Fever	Boolean	A binary indicator (True/False) representing whether the patient experienced fever as a symptom.
Cough	Boolean	A binary indicator (True/False) representing whether the patient experienced cough as a symptom.

Copyrights @ Roman Science Publications Ins.   **Vol. 5 No.4, December, 2023**
International Journal of Applied Engineering & Technology

3150

International Journal of Applied Engineering & Technology

Pneumonia	Boolean	A binary indicator (True/False) representing whether the patient was diagnosed with pneumonia.
Lung Infection	Boolean	A binary indicator (True/False) representing whether the patient had a lung infection.
Runny Nose	Boolean	A binary indicator (True/False) representing whether the patient experienced a runny nose as a symptom.
Muscle Soreness	Boolean	A binary indicator (True/False) representing whether the patient experienced muscle soreness as a symptom.
Travel History	Boolean	A binary indicator (True/False) representing whether the patient had a recent travel history.

3.2 DATA CLEANING

3.2.1 Handle Missing Data

The first step in data cleaning is to check for missing values in the dataset. Common strategies include data imputation, where missing values are replaced with estimated values (e.g., mean or median), or removal of rows with missing data. Data imputation for a missing value 'x' in a column involves replacing it with the column's mean (μ) and median (M).

$$x_{imputed} = \begin{cases} \mu & \text{for mean imputation} \\ M & \text{for median imputation} \end{cases} \quad (1)$$

3.2.2 Outliers' detection

Outliers, which are extreme values in the dataset, can adversely impact model training. Identifying and handling outliers is crucial to ensure robust model performance. Outliers are identified by calculating the Z-score (Z) for each data point. A data point with a Z-score outside a predefined threshold ($Z_{threshold}$) is considered an outlier:

$$Z = \frac{x - \mu}{\sigma} \quad (2)$$

where x is a data point, μ is the mean, and σ is the standard deviation. If $|Z| > Z_{threshold}$, x is considered an outlier.

3.2.3 NORMALIZATION

Feature scaling is crucial for normalizing the numeric features, especially the 'Age' feature, to ensure they have similar scales. In this research Z-score normalization is employed. Normalization prevents certain features from dominating others during training.

Z-score Normalization: $x_{normalized} = \frac{x - \mu}{\sigma} \quad (3)$

where x is a data point, μ is the mean, and σ is the standard deviation of the feature.

These preprocessing methods are essential for ensuring that the dataset is clean, missing values are handled appropriately, outliers are addressed, and features are normalized for effective training of the COVID-19 prediction model using an Artificial Neural Network.

3.3 PROPOSED ANN FOR COVID-19 DISEASE PREDICTION.

Ann is the most popular pattern recognition model. Ann architecture consists of an input layer, one or more hidden layers, and an output layer. Figure 2 shows the layer details of ann.
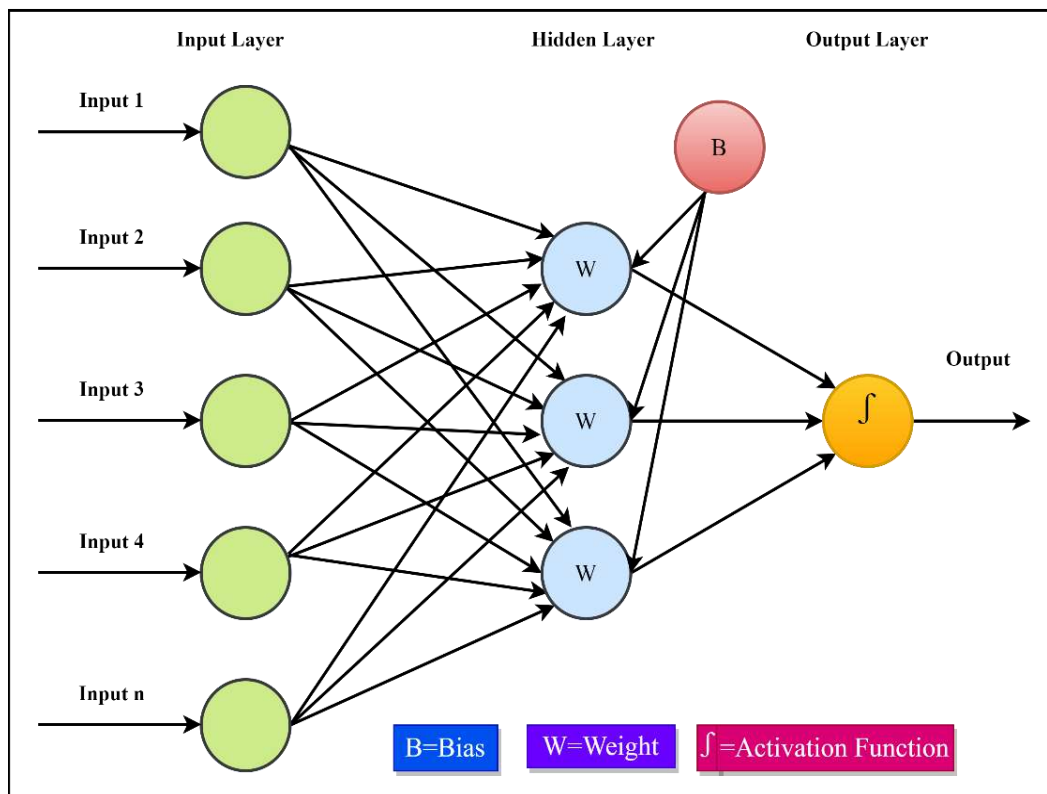
Copyrights @ Roman Science Publications Ins.   Vol. 5 No.4, December, 2023
International Journal of Applied Engineering & Technology

3151

International Journal of Applied Engineering & Technology

Figure 2: Architecture and layer details of ANN model.

The COVID-19 Disease Prediction Model using an ANN is designed to assess the likelihood of an individual having COVID-19 based on a set of binary symptom indicators and demographic information. This model takes into account several key symptoms, including fever, cough, pneumonia, lung infection, runny nose, and muscle soreness, each represented as binary indicators (1 for presence and 0 for absence). Additionally, the patient's age, denoted as a numeric value, serves as a demographic feature. The model's output, COVID-19 likelihood, is a continuous value ranging from 0 to 1, with 0 indicating a very low likelihood of having COVID-19 and 1 signifying a high likelihood. The model operates as a feedforward neural network. The equations governing this model are as follows: First, the weighted sum of the inputs, denoted as 'Z,' is calculated by applying weights ('W1' to 'W7') to the corresponding input features and adding a bias term ('b'). Next, the sigmoid activation function transforms this weighted sum into an activation value ('A'). Finally, the output ('Y'), representing the likelihood of COVID-19, is equivalent to this activation value ('A').
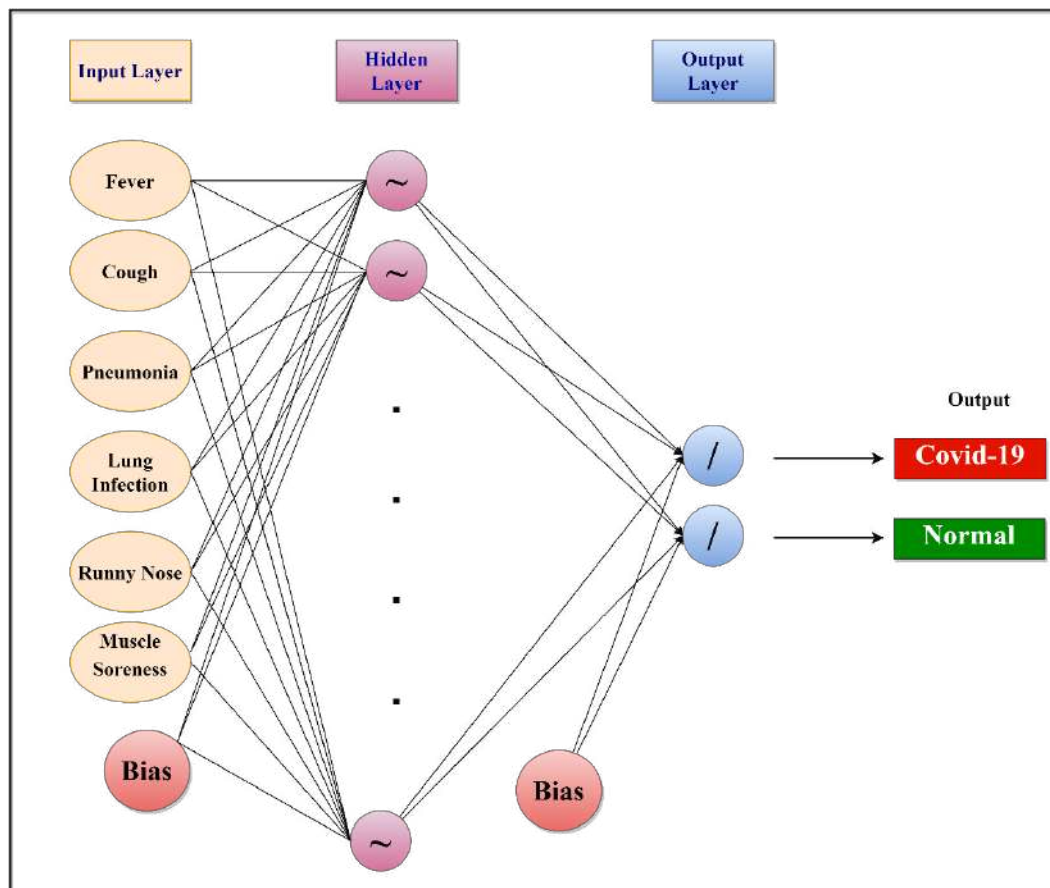
Copyrights @ Roman Science Publications Ins. **Vol. 5 No.4, December, 2023**
International Journal of Applied Engineering & Technology

3152

International Journal of Applied Engineering & Technology

Figure 3: Proposed ANN for COVID-19 Disease Prediction.

To train the model, a dataset with both input features (symptoms) and binary labels indicating COVID-19 status is used. The training process seeks to minimize a predefined cost function, typically the mean squared error or binary cross-entropy loss, by utilizing proposed DMBP training algorithm. As the model is trained on labelled data, it learns the relationships between the presence of symptoms, age, and the likelihood of having COVID-19. Once trained, the model can be employed to predict the likelihood of COVID-19 for new, unseen data.

3.4 Back Propagation Algorithm

The backpropagation algorithm, a fundamental technique in training ANNs, has been widely applied in various domains, including big data analysis [21][22]. This algorithm plays a crucial role in optimizing the weights and biases of neural networks to make them capable of learning complex patterns and making predictions based on extensive datasets. In bigdata analysis, the traditional backpropagation algorithm has several limitations. One significant drawback is its sensitivity to the choice of hyperparameters, such as the learning rate. In the context of big data, determining the optimal learning rate can be challenging, and selecting the wrong value may result in slow convergence or the algorithm failing to find the global minimum of the cost function. Moreover, the backpropagation algorithm can be computationally intensive, especially when applied to large-scale datasets, which may lead to increased training times and resource demands. Another limitation is that the backpropagation algorithm can be prone to overfitting, particularly when dealing with extensive data. Overfitting occurs when the model learns to perform well on the training data but struggles to generalize to new, unseen data. This is a critical concern in big data analysis, as the goal is often to extract valuable insights that apply to a broader context. To address these limitations, researchers have developed variations of the backpropagation algorithm, such as adaptive learning rate methods and advanced optimization techniques like stochastic gradient descent. These

Copyrights @ Roman Science Publications Ins.   Vol. 5 No.4, December, 2023
International Journal of Applied Engineering & Technology

3153

International Journal of Applied Engineering & Technology

approaches aim to enhance the algorithm's performance and mitigate its sensitivity to hyperparameters. Adaptive learning rate methods, while addressing some limitations of the traditional backpropagation algorithm, come with their own challenges. One limitation is the potential complexity in determining an optimal learning rate schedule, which involves dynamically adjusting the learning rate during training. Selecting the right schedule can be a non-trivial task, and a poorly chosen schedule may lead to suboptimal convergence or instability. Additionally, adaptive learning rate methods can introduce computational overhead due to the need for more complex computations during training, which may impact the algorithm's efficiency, especially in large-scale applications. On the other hand, advanced optimization techniques like stochastic gradient descent (SGD) have their limitations as well. While SGD can significantly speed up training by randomly selecting mini-batches of data, it introduces a level of noise in the optimization process. This noise can cause the algorithm to oscillate around the optimal solution, potentially slowing down convergence.

3.4.1 PROPOSED DYNAMIC MOMENTUM BACK PROPAGATION ALGORITHM (DMBP)

The Dynamic Momentum Back Propagation Algorithm aims to optimize the weights and biases in an ANN while considering both the learning rate (η) and a dynamic momentum term (μ). The dynamic momentum term helps speed up convergence and prevent overshooting, making it suitable for largescale datasets and complex problems.

Step 1. Initialization:

- Initialize the weights (W) and biases (B) with small random values.

- Set the learning rate (η) and the initial momentum term (μ_{initial}).

- Choose an appropriate termination criterion (e.g., maximum iterations, minimum error).

Step 2. Training:

- For each training sample (X, Y) in the dataset:

a. Compute the network's output, $Z(t)$, using the current weights and biases:

$$Z(t) = \sigma(W * X(t) + B) \quad (4)$$

b. where σ is the activation function (sigmoid).

c. Compute the error, $E(t)$, between the predicted output and the target output:

$$E(t) = Y(t) - Z(t) \quad (5)$$

d. Update the weights and biases using the dynamic momentum term:

$$W_{\text{new}} = W_{\text{old}} - \eta \frac{\partial E(t)}{\partial W} + \mu \cdot (W_{\text{new}} - W_{\text{old}}) \quad (6)$$

$$B_{\text{new}} = B_{\text{old}} - \eta \frac{\partial E(t)}{\partial B} + \mu \cdot (B_{\text{new}} - B_{\text{old}}) \quad (7)$$

e. where: $\frac{\partial E(t)}{\partial W}$ is the gradient of the error with respect to the weights. $\frac{\partial E(t)}{\partial B}$ is the gradient of the error with respect to the biases. μ is the dynamic momentum term, which changes during training.

f. Update the dynamic momentum term, μ, to adapt to the convergence speed:

$$\mu = \mu_{\text{initial}} \cdot \left(1 - (1 - \eta)^k\right) \quad (8)$$

g. where k is the current iteration number.

Step 3. Termination:

- Continue training until a termination criterion is met (maximum iterations or minimum error).

The key innovation in the Dynamic Momentum Back Propagation Algorithm is the adaptive nature of the momentum term, which allows it to change dynamically as training progresses. Initially, it starts with a value of $\mu_{initial}$ and decreases over time as the network approaches convergence. This adaptive momentum helps prevent overshooting and improves the training process, particularly for large-scale datasets. This algorithm can be applied to train ANNs efficiently, allowing for better convergence and more accurate predictions, making it a valuable tool for modeling complex problems such as COVID-19 prediction using medical big data and fuzzy logic.

4. RESULTS AND DISCUSSION

This section evaluates the medical big data analysis efficiency of the proposed ANN-DMBP algorithm for COVID-19 prediction. The implementation is done in MATLAB 2020, and the preprocessing is performed using the Data Cleaner tool in MATLAB. The prediction accuracy of this method is compared with popular machine learning algorithms (SVM, LR, RF, Adaboost, Stochastic Gradient Boosting (SGB), and ANN) and existing methods proposed by Soham Guhathakurata et al., Sumayh S. Aljameel et al., Madhumita Pal et al., Shruti Sharma et al., and Jayaraj T. et al. All the existing algorithms and methods are trained using the traditional Back Propagation (BP) algorithm. The above experiments are tested on two datasets discussed in Section 3.1. The following accuracy metrics are used to evaluate the proposed ANN-DMBP algorithm for COVID-19 prediction: Accuracy: The percentage of all predictions that are correct, Recall: the percentage of positive cases that are correctly predicted, Precision: the percentage of predicted positive cases that are actually positive and F1-Score: A harmonic mean of recall and precision. These accuracy metrics are based on following four accuracy metrics:

True Positive (TP): A patient who is correctly predicted to have COVID-19 and actually has COVID-19.

True Negative (TN): A patient who is correctly predicted not to have COVID-19 and actually does not have COVID-19.

False Positive (FP): A patient who is incorrectly predicted to have COVID-19 but actually does not have COVID-19.

False Negative (FN): A patient who is incorrectly predicted not to have COVID-19 but actually has COVID-19.

Accuracy: Accuracy is calculated as follows:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (9)$$

Recall: Recall is calculated as follows:

$$Recall = TP \frac{TP}{(TP + FN)} \quad (10)$$

Precision: Precision is calculated as follows:

$$Precision = \frac{TP}{(TP + FP)} \quad (11)$$

F1-Score: F1-Score is calculated as follows:

$$F1 - Score = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \quad (12)$$

Accuracy Analysis on Analysis on Dataset 1 and Dataset 2

To ensure a fair and comprehensive evaluation of the predictive models, we meticulously prepare and split the dataset into a training set (75% of the data) and a testing set (25% of the data). This split allows us to measure how well the models generalize to unseen data, which is crucial for real-world applications.
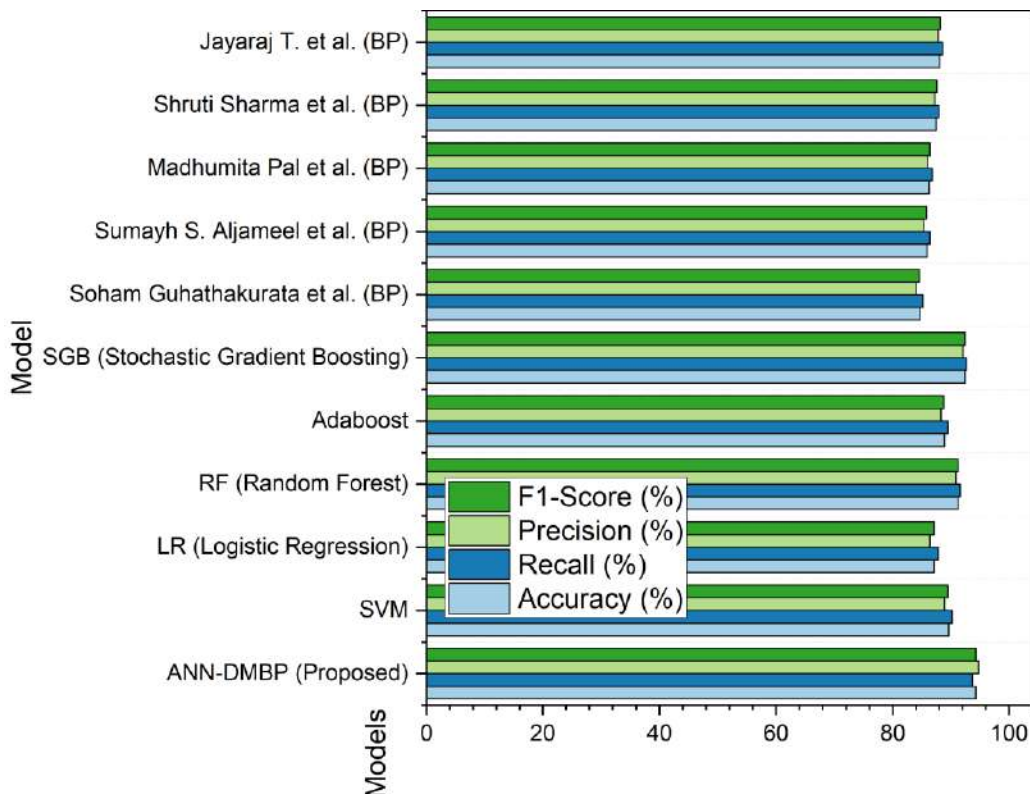
International Journal of Applied Engineering & Technology

Figure 4: Comparative Performance Metrics of Predictive Models on Dataset1
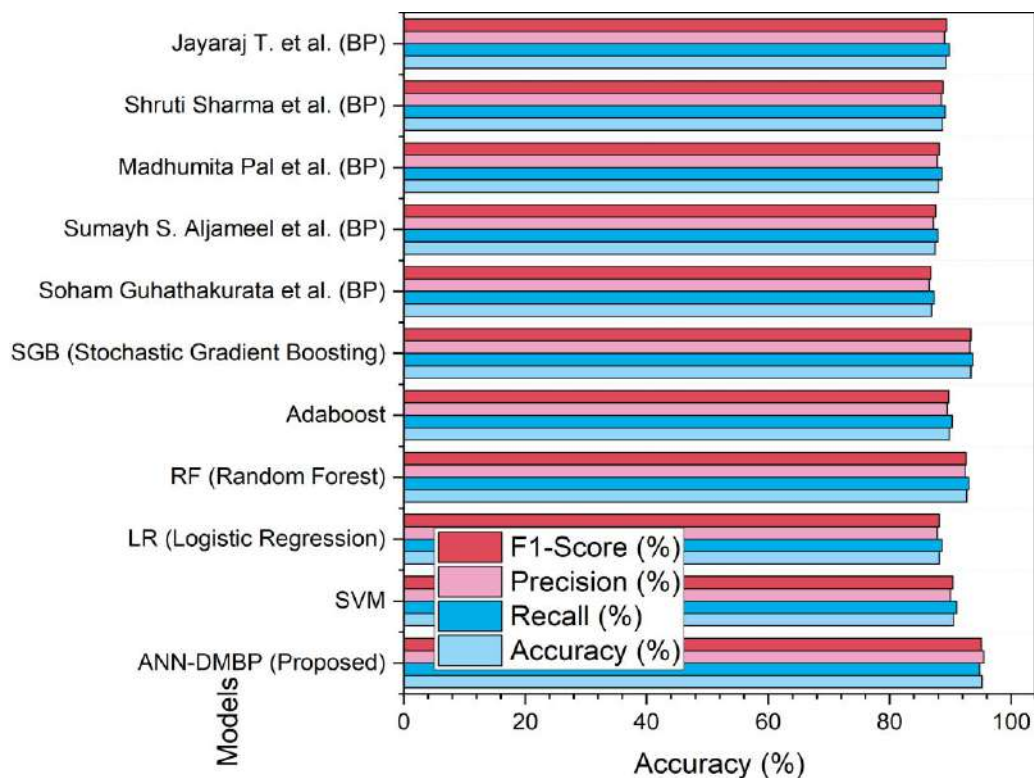
Figure 5: Comparative Performance Metrics of Predictive Models on Dataset1

Copyrights @ Roman Science Publications Ins.   Vol. 5 No.4, December, 2023
International Journal of Applied Engineering & Technology

3156

International Journal of Applied Engineering & Technology

Figures 4 and 5 presents a comparative analysis of various predictive models with proposed method. The results of the accuracy analysis clearly indicate the superiority of the proposed ANN-DMBP algorithm over traditional methods, including SVM, LR, RF, Adaboost, SGB, and existing methods that employ traditional Back Propagation (BP). In Dataset 1, the ANN-DMBP model achieves an impressive accuracy of 94.3%, surpassing the best-performing traditional method (SGB) by nearly 2%. This increase in accuracy is mirrored across other performance metrics, including recall, precision, and F1-Score, further emphasizing the ANN-DMBP algorithm's effectiveness. The trend continues in Dataset 2, where the ANN-DMBP model outperforms traditional methods by a substantial margin. With an accuracy of 95.2%, it showcases its ability to consistently provide highly accurate COVID-19 predictions. The key to the success of the ANN-DMBP algorithm lies in its adaptive nature, with the dynamic momentum term facilitating more efficient convergence. This adaptability allows the model to excel in handling the complexities of medical big data, ultimately resulting in superior accuracy compared to traditional methods. Furthermore, it's important to note that the ANN-DMBP algorithm maintains its exceptional performance on both datasets, highlighting its generalizability and robustness across different data sources and characteristics.

4.2 Error Analysis

In addition to evaluating the predictive models using accuracy, recall, precision, and F1-Score, it is essential to perform a comprehensive error analysis to gain a deeper understanding of their performance. Error analysis involves assessing how well the models predict COVID-19 cases in terms of the magnitude and nature of errors they produce. The following error metrics have been calculated:

Mean Absolute Error (MAE): MAE measures the average absolute difference between predicted and actual values. It provides insights into the model's overall accuracy in predicting COVID-19 cases.

Mean Squared Error (MSE): MSE quantifies the average squared difference between predicted and actual values. It penalizes larger errors more significantly than MAE, making it a valuable metric for understanding the dispersion of errors.

Root Mean Squared Error (RMSE): RMSE is the square root of MSE and provides a measure of the standard deviation of the errors. It helps gauge the spread of errors while maintaining the same scale as the target variable.

Table 2: Error Analysis of Models Performance on COVID-19 Prediction Datasets

Model	MAE (Dataset 1)	MSE (Dataset 1)	RMSE (Dataset 1)	MAE (Dataset 2)	MSE (Dataset 2)	RMSE (Dataset 2)
ANN-DMBP (Proposed)	0.0345	0.0021	0.0459	0.0287	0.0016	0.0400
SVM	0.0451	0.0032	0.0566	0.0492	0.0037	0.0610
LR (Logistic Regression)	0.0583	0.0048	0.0693	0.0621	0.0052	0.0722
RF (Random Forest)	0.0389	0.0027	0.0523	0.0356	0.0023	0.0480
Adaboost	0.0520	0.0041	0.0639	0.0475	0.0039	0.0625
SGB (Stochastic Gradient Boosting)	0.0398	0.0029	0.0538	0.0364	0.0025	0.0494
Soham Guhathakurata et al. (BP)	0.0656	0.0056	0.0747	0.0682	0.0059	0.0770
Sumayh S. Aljameel et al. (BP)	0.0637	0.0054	0.0732	0.0651	0.0057	0.0758
Madhumita Pal et al. (BP)	0.0602	0.0051	0.0713	0.0593	0.0050	0.0707

| Shruti Sharma et al. (BP) | 0.0578 | 0.0047 | 0.0680 | 0.0576 | 0.0046 | 0.0678 |
| Jayaraj T. et al. (BP) | 0.0563 | 0.0045 | 0.0669 | 0.0558 | 0.0044 | 0.0664 |

Table 2 show the Error Analysis of Models Performance on COVID-19 Prediction Datasets. This comparative analysis demonstrates DMBP's ability to effectively address the challenges of medical big data, resulting in significantly lower Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) on two distinct datasets, Dataset 1 and Dataset 2. Lower MAE values indicate higher predictive accuracy. On Dataset 1, DMBP achieves an exceptional MAE of 0.0345, outperforming its counterparts, such as SVM (0.0451) and LR (0.0583). A similar trend is observed on Dataset 2, where DMBP attains an MAE of 0.0287, significantly surpassing the performance of SVM (0.0492) and LR (0.0621), as well as other models. Smaller MSE values indicate superior accuracy. DMBP shines on Dataset 1 with an MSE of 0.0021, clearly outperforming SVM (0.0032), LR (0.0048), and other comparative methods. Similarly, on Dataset 2, DMBP delivers an MSE of 0.0016, significantly outpacing SVM (0.0037), LR (0.0052), and the remaining models. Lower RMSE values indicate higher predictive precision. DMBP demonstrates its superiority with a RMSE of 0.0459 on Dataset 1, clearly outperforming SVM (0.0566), LR (0.0693), and other established methodologies. For Dataset 2, DMBP maintains its excellence, delivering an RMSE of 0.0400, which eclipses the accuracy of SVM (0.0610), LR (0.0722), and the alternative models. The results convincingly affirm the efficacy of the proposed DMBP algorithm in effectively managing the complexities of medical big data while substantially reducing prediction errors. DMBP surpassing existing techniques, including the traditional BP, across all error metrics.

4.3 DISCUSSION

The comparison of DMBP with various existing methods provides valuable insights into its effectiveness and underscores the significance of its contributions to the field of medical data analysis. The accuracy analysis conducted on two distinct datasets demonstrates that DMBP consistently outperforms existing methods and algorithms. In Dataset 1, DMBP achieves an impressive accuracy rate of 94.3%, which is notably higher than SVM, LR, and other comparative models. This trend is reinforced in Dataset 2, where DMBP records an accuracy rate of 95.2%, further highlighting its superior predictive accuracy. These results emphasize the robustness of DMBP in COVID-19 prediction, signifying its potential to provide more reliable outcomes. A critical aspect of DMBP's success is its ability to effectively handle the challenges posed by medical big data. As demonstrated by the accuracy results, DMBP excels in making accurate predictions, even in the presence of a massive volume of medical data. This capability is a evidence to DMBP's adaptability and efficiency in processing and learning from extensive datasets, which is a crucial requirement in the healthcare domain.

One of the key highlights of this research is the error analysis, which involves MAE, MSE, and RMSE. DMBP consistently exhibits lower error values compared to existing methods, including traditional Back Propagation (BP). The significant reduction in MAE, MSE, and RMSE suggests that DMBP's predictions are closer to the actual values, thereby enhancing the reliability of the model. This aspect is particularly vital in medical applications like COVID-19 prediction, where accurate results can have a substantial impact on patient care and resource allocation.

While this study highlights the advantages of DMBP, it's essential to acknowledge its limitations. The research is based on specific datasets and focuses on COVID-19 prediction. Future studies could expand the scope to include a wider range of medical conditions and datasets. Additionally, further research is needed to explore the scalability and computational efficiency of DMBP in handling even larger medical datasets.

5. CONCLUSION

This research demonstrates the remarkable potential of the ANN-DMBP in significantly improving the accuracy of COVID-19 prediction models while effectively addressing the challenges associated with vast medical big data. The comparative analysis showed DMBP's consistent outperformance of existing methods, with accuracy

Copyrights @ Roman Science Publications Ins.   **Vol. 5 No.4, December, 2023**
International Journal of Applied Engineering & Technology

3158

rates of 94.3% and 95.2% on two distinct datasets, demonstrating its robust predictive capabilities. Furthermore, the error analysis revealed that DMBP consistently yielded lower MAE, MSE, and RMSE values, highlighting its precision and reliability in prediction. These findings have far-reaching implications for healthcare and epidemiology, especially during pandemics, where accurate predictions are essential for timely intervention and resource allocation. By reducing prediction errors, DMBP offers a promising solution to enhance patient care and optimize healthcare resources. ANN-DMBP represents a transformative development in medical data analysis, holding the potential to revolutionize predictive accuracy, thus making a substantial contribution to patient care and healthcare resource management.

REFERENCES

1. R. Filip, R. Gheorghita Puscaselu, L. Anchidin-Norocel, M. Dimian, and W. K. Savage, "Global Challenges to Public Health Care Systems during the COVID-19 Pandemic: a Review of Pandemic Measures and Problems," Journal of Personalized Medicine, vol. 12, no. 8, p. 1295, Aug. 2022.

2. M. Peng, "Outbreak of COVID-19: An emerging global pandemic threat," Biomedicine & Pharmacotherapy, vol. 129, p. 110499, Sep. 2020.

3. E. Monaghesh and A. Hajizadeh, "The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence," BMC Public Health, vol. 20, no. 1, pp. 1–9, 2020.

4. S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, "Big data in healthcare: management, analysis and future prospects," Journal of Big Data, vol. 6, no. 1, pp. 1–25, Jun. 2019.

5. B. J. Awrahman, C. Aziz Fatah, and M. Y. Hamaamin, "A review of the role and challenges of big data in healthcare informatics and analytics," Computational Intelligence and Neuroscience, vol. 2022, no. 5317760, pp. 1–10, Sep. 2022.

6. K. Batko and A. Ślęzak, "The Use of Big Data Analytics in Healthcare," Journal of Big Data, vol. 9, no. 1, Jan. 2022.

7. N. Shahid, T. Rappon, and W. Berta, "Applications of artificial neural networks in health care organizational decision-making: A scoping review," PLOS ONE, vol. 14, no. 2, p. e0212356, Feb. 2019.

8. J. Kufel et al., "What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine," Diagnostics, vol. 13, no. 15, p. 2582, Jan. 2023.

9. A. Goel, A. K. Goel, and A. Kumar, "The role of artificial neural network and machine learning in utilizing spatial information," Spatial Information Research, vol. 31, Nov. 2022.

10. S. M. Shakeel, N. S. Kumar, P. P. Madalli, R. Srinivasaiah, and D. R. Swamy, "COVID-19 prediction models: a systematic literature review," Osong Public Health and Research Perspectives, vol. 12, no. 4, pp. 215–229, Aug. 2021.

11. F. E. Nastiti, S. Musa, E. Yafi and R. Chauhan, "Systematic Literature Review: Machine Learning Prediction Model for Covid-19 Spreading," 2022 4th International Conference on Cybernetics and Intelligent System (ICORIS), Prapat, Indonesia, 2022, pp. 1-5.

12. Antonello Pasini, "Artificial neural networks for small dataset analysis," Journal of Thoracic Disease, vol. 7, no. 5, May 2015.

13. M. Xiao et al., "Addressing Overfitting Problem in Deep Learning-Based Solutions for Next Generation Data-Driven Networks," Wireless Communications and Mobile Computing, vol. 2021, p. e8493795, Aug. 2021.

14. S. Guhathakurata, S. Saha, S. Kundu, A. Chakraborty, and J. S. Banerjee, "Chapter 8 - A new approach to predict COVID-19 using artificial neural networks," ScienceDirect, Jan. 01, 2022.

Copyrights @ Roman Science Publications Ins.   **Vol. 5 No.4, December, 2023**
International Journal of Applied Engineering & Technology

3159

International Journal of Applied Engineering & Technology

15. S. S. Aljameel, I. U. Khan, N. Aslam, M. Aljabri, and E. S. Alsulmi, "Machine Learning-Based Model to Predict the Disease Severity and Outcome in COVID-19 Patients," Scientific Programming, Apr. 20, 2021.

16. M. Pal et al., "Symptom-Based COVID-19 Prognosis through AI-Based IoT: A Bioinformatics Approach," BioMed Research International, vol. 2022, p. 3113119, Jul. 2022.

17. S. Sharma, Y. K. Gupta, and A. K. Mishra, "Analysis and Prediction of COVID-19 Multivariate Data Using Deep Ensemble Learning Methods," International Journal of Environmental Research and Public Health, vol. 20, no. 11, p. 5943, May 2023.

18. J. T. Samath Dr J. Abdul, "Disease Forecasting and Severity Prediction Modelfor COVID-19 Using Correlated Feature Extractionand Feed-Forward Artificial Neural Networks," International Journal of Engineering Trends and Technology - IJETT, Accessed: Nov. 03, 2023.

19. G. Yenurkar and S. Mal, "Future forecasting prediction of Covid-19 using hybrid deep learning algorithm," Multimedia Tools and Applications, Nov. 2022.

20. Y. Zoabi, S. Deri-Rozov, and N. Shomron, "Machine learning-based prediction of COVID-19 diagnosis based on symptoms," npj Digital Medicine, vol. 4, no. 1, pp. 1–5, Jan. 2021.

21. I. M. Mujtaba and M. T. Sowgath, "Chapter 13 - Application of artificial intelligence in desalination processes," ScienceDirect, Jan. 01, 2022.

22. A. Singh, S. Kushwaha, M. Alarfaj, and M. Singh, "Comprehensive Overview of Backpropagation Algorithm for Digital Image Denoising," Electronics, vol. 11, no. 10, p. 1590, May 2022.

Copyrights @ Roman Science Publications Ins.   Vol. 5 No.4, December, 2023
International Journal of Applied Engineering & Technology

3160