iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: XI Month of publication: November 2025

DOI: https://doi.org/10.22214/ijraset.2025.74798

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

Forecasting of Monkeypox New Cases in North and South America for 2023: A Comparative Analysis Using Machine Learning

Sharmin Akther[1], Amrin Binte Ahmed[2], Sanjida Tasnim[3]
[1, 2, 3]Department of Statistics and Data Science, Jahangirnagar University, Savar, Dhaka-1342

Abstract: The recent spread of monkeypox throughout the world highlights the critical need for trustworthy forecasting models to guide public health actions. This study evaluates several machine learning techniques for forecasting new monkeypox cases in North and South America in 2023. We applied five models Decision Tree (DT), Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), XGBoost, and Random Forest (RF) and evaluated them using four key error metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). For South America, the Decision Tree model produced the most accurate forecasts, with the lowest MAE (10.37), RMSE (13.28), MSE (176.25), and MAPE (28.62%). Meanwhile, XGBoost proved to be the best-performing model for North America, with the lowest MAE (17.06), RMSE (19.62), MSE (384.79), and MAPE (25.34%). Our results highlight the significance of region-specific modeling strategies since various machine learning methods performed better in multiple environments. This research lays the groundwork for creating more precise forecasting instruments that will aid in halting the monkeypox epidemic and improve response and readiness throughout North and South America.
Keywords: Monkeypox, Machine Learning, ANN, Decision Tree, Random Forest, KNN, MLP, XGBoost.

I. INTRODUCTION

Monkeypox virus (MPXV) was first detected in humans in 1970, and since then outbreaks have been managed mainly in endemic regions of Central and West Africa (1). Compared to SARS-CoV-2, MPXV is easier to contain due to its transmission through close contact and after the onset of symptoms. There isn't any convincing proof of asymptomatic or far-reaching airborne transmission (1). An atypically extensive human monkeypox outbreak was documented in August 1996, and cases persisted throughout 1997, reaching their highest points in August 1996, March 1997, and August 1997 (2). Since smallpox was eradicated, monkeypox (Mpox), a zoonotic illness, has the potential to grow into one of the most dangerous orthodox virus infections in humans (3). Similar to smallpox, Mpox is a zoonotic illness that is endemic to Central and West Africa, with a fatality rate of roughly 11% (4). However, since May 2022, non-endemic countries particularly those in North America and Europe have seen an exponential rise in the number of cases of monkeypox infections (5). For the first time, multiple cases and clusters of monkeypox have been reported simultaneously in both endemic and nonendemic countries across a broad range of geographic regions (6). Blood samples and clinical and epidemiologic data were gathered from cases defined by the standard definition and household contacts of cases after the U.S. monkeypox outbreak in 2003. The purpose of this was to assess the contribution of acquired immunity and preexisting immunity derived from the smallpox vaccine to disease susceptibility and clinical outcomes (7).

The viral disease monkeypox, which has serious consequences for public health, has resurfaced as a serious health issue. Its dispersion throughout several continents, especially North and South America, presents particular difficulties for healthcare systems. Disease surveillance and local monkeypox epidemic control must propose a monkeypox epidemic trend prediction model based on real-time monitoring data (8,9). Although it can also be sexually transmitted through intercourse with an infected person's bodily fluids, it is thought that the monkeypox virus primarily spreads among humans through respiratory secretions during extended face-to-face interactions(10). The respiratory system, the mucosa of the eyes, nose, and mouth, as well as cuts or wounds (even invisible ones), are all potential entry points for the virus into the body. Although the MPX virus was initially identified in monkeys, it is believed that squirrels and striped grass mice served as the virus's primary reservoir(11,12). The West African clade of the monkeypox virus is responsible for the 2022–2023 mpox outbreak in humans, which is occurring in South America. In May 2022, the first case of mpox was reported in Argentina, marking the beginning of the outbreak in South America. By August 14, 2022, eight countries and territories in South America had confirmed cases(13).

The human mpox outbreak that is being caused by the West African clade of the monkeypox virus includes the North American outbreak of mpox in 2022–2023. On May 18, 2022, the United States reported the first case of mpox, marking the beginning of the outbreak in North America. Twenty countries and territories in North America have confirmed cases as of August 23, 2022(14).

Recent studies on monkeypox have been published, but they have not provided precise measurements because the majority of the studies focused on a single country rather than comparing performance in different locations using machine-learning techniques(15). Machine learning (ML) techniques provide an additional approach to improve mpox prediction..Massive amounts of data, including contact tracking information, travel history, and clinical symptoms, can be analyzed by these algorithms(16). The analysis was conducted using conventional machine learning techniques like decision trees, random forests, XGBoost, KNN, and neural networks, including artificial neural networks and multi-layer perceptions (MLPs), which are made up of several layers of neurons. To assess the performance of these models, we have relied on statistical metrics like mean absolute error, mean squared error, and mean absolute percentage error(17).

The main contributions in this paper are fourfold: This includes artificial neural networks (ANNs), k-nearest neighbors (KNN), decision trees, random forests, multilayer perceptions (MLPs), and extreme gradient boosting (XGBoost) in the prediction of novel monkeypox cases in North and South America. Novel Predictions Model for Monitoring Mpox Trend Forecasting in 2023. It concentrates on improved accuracy in forecasting to assist public health authorities in deploying better targeted interventions. It then evaluates the effect of patterns in the region-wise outbreak on model performance and predictions.

II. MATERIALS AND METHODS

A. Data Source

The study employed daily data from the North American and South American Our World in Data databases that were made publicly available. The data covered the period from January 1, 2023, to December 31, 2023 (18). Because it was the most recent and comprehensive dataset available for both regions at the time of the study, this time range was selected to ensure reporting uniformity. New cases were chosen as the target variable for forecasting because they are most relevant for predicting short-term outbreak dynamics. The final dataset, consisting of daily new cases, was used for exploratory data analysis, temporal trend visualization, and as input to multiple machine learning models to determine the best-performing forecasting approach.

B. Study Design

To properly create and assess the machine learning models, the available dataset was split into training and testing subsets for each of the two subcontinents, North America and South America. In particular, 75% of the data was set aside for training, which enabled the models to discover hidden links and patterns in the data. The remaining 25% was reserved for test data, which the models did not see while they were being trained. An objective evaluation of each model's forecasting ability on fresh, untested data is ensured by this separation. We can gain a better understanding of the models' robustness and generalizability in predicting weekly new cases of monkeypox in various places by assessing them on this limited test set. For the data analysis, we use R software version 4.4.1.

1) Machine Learning Model

A machine learning model is an algorithm designed to identify trends or draw conclusions from a dataset that has never been seen before. Machine learning models, for instance, can parse and accurately identify the meaning behind words or sentences that have never been heard before in natural language processing (19). The term "learning" in the context of machine learning refers to a machine's capacity for data analysis and the output that results from that analysis (20). ANN, Decision Tree, Random Forest, KNN, and XGBoost are some popular machine learning algorithms for prediction.

2) Artificial Neural Network (ANN)

Inspired by a simplified representation of neurons in the brain, an artificial neural network consists of a network of interconnected nodes (21). A neural network, also known as an artificial neural network or neural network and abbreviated ANN or NN, is a machine learning model that draws inspiration from the architecture and operation of biological neural networks found in animal brains(22,23). Usually, empirical risk minimization is used to train neural networks. The foundation of this approach is the notion that the network's parameters should be optimized to reduce the empirical risk of the difference between the target values in a given dataset and the output that is predicted (24). The ANN functions similarly to the numerous connections between neurons in the brain, where each node (point) is connected to every other node in the form of a pathway for mutual interaction. All nodes in a neural structure can have weights assigned to them by the ANN using a single hidden layer (25).

3) Decision Tree

Although decision trees are a supervised learning technique, they are primarily used to solve classification problems. However, they can also be used to solve regression problems. This classifier is tree-structured, with internal nodes standing in for dataset features, branches for decision rules, and leaf nodes for each outcome (26). The objective is to build a model that, by utilizing basic decision rules deduced from the data features, predicts the value of a target variable. A piecewise constant approximation can be thought of as a tree(27). To achieve smaller trees and improved generalization, decision trees are routinely pruned either during or after training. During node splitting, this algorithm performs basic calculations that represent the purity of the node(28).

4) Random Forest

Random forests, also known as random decision forests, are an ensemble learning technique that generates a large number of decision trees during training for tasks like regression, classification, and other applications. The random forest's output for classification tasks is the class that the majority of the trees chose. The output of a regression task is the mean of the tree predictions(29,30). Regression problems with numerical target variables and classification problems with categorical target variables can both be solved with random forests (31). The method creates multiple decision trees for each subpart after randomly dividing the dataset into smaller segments. Then, each decision tree's anticipated output is combined to produce a forecast that is more accurate and dependable. The average of the values predicted by multiple decision trees is the output value for each input or subset in a random forest regression. In random forest regression, an n-tree bootstrap sample is produced from the actual input dataset(32).

5) Multilayer Perception Classifier (MLP)

Multi-layer Perceptron classifier, or MLPClassifier for short, is a name that implies a connection to a neural network. In contrast to other classification algorithms like Naive Bayes Classifier or Support Vector, MLPClassifier depends on an underlying Neural Network to carry out the classification task (33). For handling a range of classification tasks, such as text classification and image recognition, it is a versatile and efficient approach (34). Multiple language processing (MLP) and image and audio recognition, as well as time-series prediction, have all benefited from the effective use of MLPs (35). An essential component of MLP's ability to tackle challenging problems is its activation function (36).

6) k-Nearest Neighbor

The k-nearest neighbor algorithm, or KNN, is a machine learning algorithm that compares a single data point with a set of training data that it has learned to make predictions based on proximity(37). Being non-parametric, that is, not assuming anything about the data distribution, it is highly disposable in real-world situations (in contrast to other algorithms like GMM, which assume a Gaussian distribution of the given data) (38). The KNN algorithm is generally used as a classification algorithm, operating under the presumption that similar points can be found close to one another, though it can be applied to regression or classification problems instead (39).

7) Extreme Gradient Boosting (XGBoost)

A powerful machine-learning algorithm called XGBoost can assist you in better understanding your data and decision-making. An application of gradient-boosting decision trees is called XGBoost. Researchers and data scientists from all over the world have been using it to improve their machine-learning models(40). XGBoost is made to be quick, simple to use, and effective with big datasets. It can be used immediately after installation without requiring additional configuration because it doesn't need to optimize or tune parameters(41). Decision trees are generated sequentially in this algorithm. An important part of XGBoost is the weight. Each independent variable is given a weight before being fed into the decision tree to make predictions about the outcome(42). The algorithm, XGBoost, which stands for Extreme Gradient Boosting, is an ensemble machine learning technique that combines the predictions of several decision trees to create a reliable model with a higher prediction accuracy(43).

C. Forecast Accuracy Measures

By comparing expected values with actual observations and calculating the resulting errors, forecast accuracy was evaluated. The following metrics are frequently used: Mean Absolute Error (MAE), which shows the average size of errors; Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), which highlight greater errors; and Mean Absolute Percentage Error (MAPE), which shows errors as a percentage of real values. These metrics offer a thorough assessment of the model's performance(44).

III. RESULTS

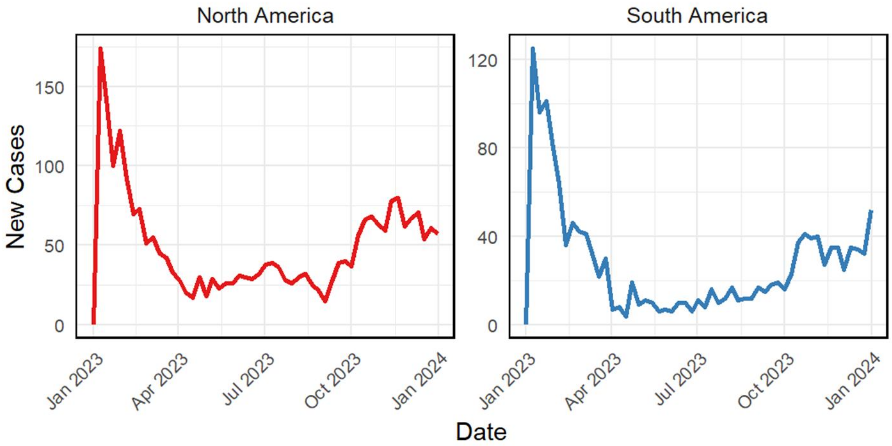


Figure 1: Daily New Monkeypox Cases in North America and South America, January 2023–December 2023

Figure 1 displays two-line plots, one on the left for North America and one on the right for South America, each showing new cases over time. Dates from January 2023 to the beginning of 2024 are represented on the x-axis, and the number of new cases is displayed on the y-axis. In North America (left), the number of cases starts to drop off sharply in early 2023 and then gradually increases and fluctuates until 2024. Towards the beginning of 2023, South America (right) likewise sees a sharp decline in new cases, which gradually increases by the end of the period. A legend designating the line colors teal for South America and red for North America is included with each plot.

From Figure 2, according to the monthly boxplot for North America, January 2023 saw the largest number of daily new cases of monkeypox, with a wide range and multiple extreme values. Narrower boxplots and lower medians show a sharp drop in February and March, followed by comparatively low and consistent numbers from April through September. But starting in October, the median and variability of daily cases increased considerably, and November saw another period of high counts until December saw a tiny decline. An early-year epidemic peak, a protracted period of control, and a late-year comeback are all suggested by this pattern.

For South America, the trend follows a similar pattern but with lower overall case numbers compared to North America. January saw the highest daily counts with considerable variation, followed by a steep drop through February and March. The mid-year period from April to September remained stable and low, with minimal spread in daily cases. A moderate rise began in October, peaking in November with both higher medians and increased variability, before declining slightly in December. This indicates that while South America experienced the same early-year peak and late-year increase, the magnitude of cases was consistently smaller than in North America.

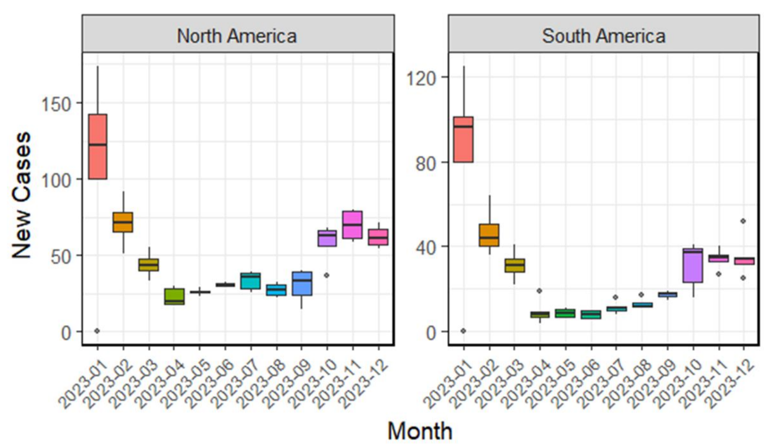


Figure 2: Monthly Distribution of Daily New Monkeypox Cases in North and South America During 2023, Highlighting Early-Year Peaks, Mid-Year Stability, and Late-Year Resurgence

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

Monkeypox cases peaked in January of 2023 in both North and South America, but sharply declined by March and then remained low until the middle of the year. After October, the number of cases increased once more, reaching a peak in November and then declining somewhat in December. North America continuously reported higher numbers than South America.

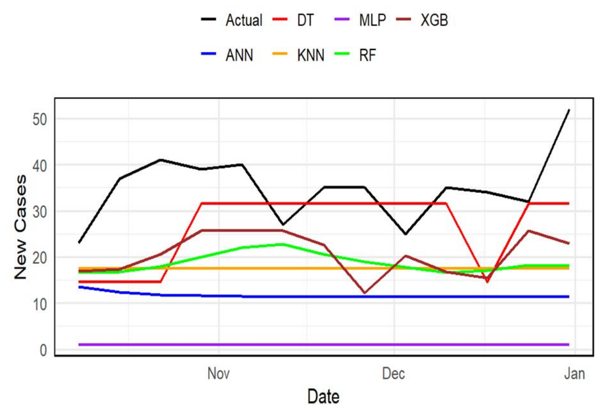


Figure 3: Line plot for South America

A line plot comparing actual and predicted new cases using different machine learning models over time is displayed in the image. The dates, which span the months of November through January, are represented by the x-axis. The number of new cases is shown on the y-axis, which goes from 0 to 50. The actual number of new cases is represented by the black line, which exhibits some fluctuations before sharply increasing toward the end of the period. The model predictions appear smoother, while the actual cases fluctuate more sharply. While some models, like XGB and DT, keep a flatter prediction throughout, other models, like ANN and KNN, track the trends more closely.

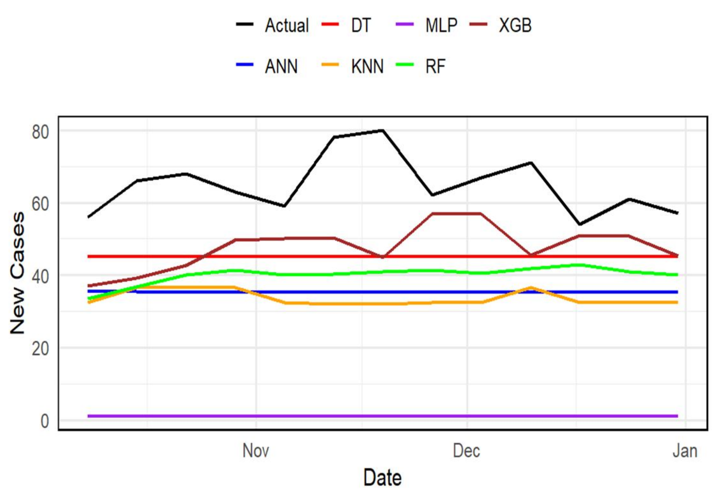


Figure 4: Line plot for North America

The graphic displays a line plot that contrasts the number of cases of monkeypox that actually occurred over time with the values predicted by different machine learning models. The date is represented by the x-axis, which spans the months of November through January. The number of new cases is shown on the y-axis, which runs from 0 to 80. The real number of new cases, which varies from 40 to 80, is shown by the black line. While most model predictions stay largely stable, the actual case trend exhibits more variability. While some models, like MLP and KNN, provide flatter predictions, others, like DT and XGB, capture more variation.

Four essential metrics were used to thoroughly assess the performance of the six machine learning models: Mean Squared Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE). The models included Decision Tree, Random Forest, KNN, MLP, and XGBoost. The results are confidently summarized in Table 1.

Table 1: Performance Metrics for Machine Learning Models on Monkeypox New Case Data Across South America.

Model	MSE	RMSE	MAE	MAPE
ANN	607.9858	24.65737	23.50599	65.54560
Decision Tree	176.2520	13.27599	10.36996	28.62056
Random Forest	358.8820	18.94418	17.56646	48.23856
KNN	359.1731	18.95186	17.50000	47.74494
MLP	1208.9231	34.76957	34.00000	97.01400
XGBoost	267.6725	16.36070	14.38917	38.88693

With the lowest MSE, RMSE, MAE, and MAPE, the Decision Tree is the model that performs the best across the board. This implies that it offers the most precise forecasts in this particular situation. With the highest MSE, RMSE, MAE, and MAPE values, MLP is the least effective model and shows low predictive accuracy. The extremely high MAPE (97.01%) indicates that prediction errors are likely to be large. KNN, XGBoost, and Random Forest (RF) all exhibit mediocre results; XGBoost performs marginally better overall. The results are confidently summarized in Table 2.

.

Table 2: Performance Metrics for Machine Learning Models on Monkeypox New Case Data Across North America.

Model	MSE	RMSE	MAE	MAPE
ANN	924.3275	30.40276	29.40116	44.64432
Decision Tree	446.4408	21.12915	19.66923	29.41782
Random Forest	732.5014	27.06476	26.01495	39.44557
KNN	1028.0577	32.06334	31.11538	47.39394
MLP	4126.0769	64.23455	63.76923	98.43498
XGBoost	384.7932	19.61615	17.05730	25.34135

With the lowest MSE, RMSE, MAE, and MAPE, XGBoost is the best-performing model. Being the most accurate model for predicting cases of monkeypox in North America, it has the lowest overall error metrics (XGB) and the smallest percentage error (25.34%), indicating that it provides the most accurate forecasts among the models. MLP is not appropriate for this forecasting task, as evidenced by its highest errors. While XGBoost is more accurate, Decision Tree (DT) and Random Forest (RF) perform about as well. KNN and ANN exhibit lower effectiveness and higher error rates.

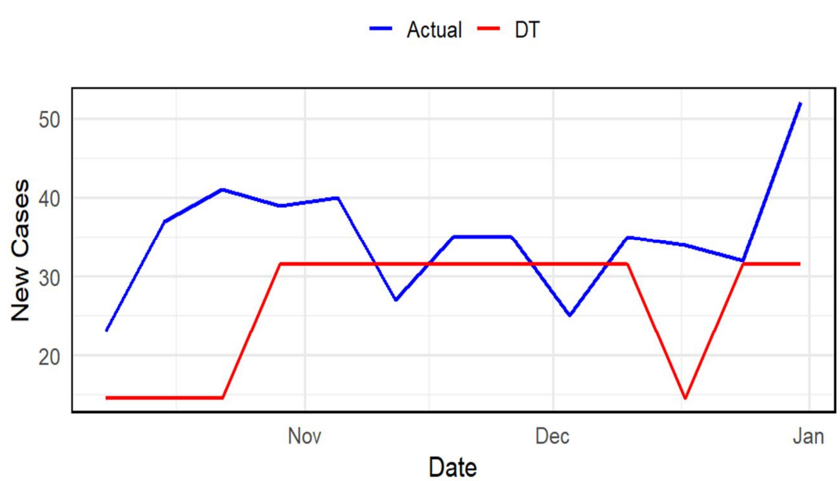


Figure 5: Forecast vs. Actual Daily New Monkeypox Cases in South America Using the Decision Tree (DT) Model, November 2023 to January 2024

The graphic displays a line plot that contrasts the number of cases of monkeypox that occurred over time with those predicted by a Decision Tree (DT) model. The date is represented by the x-axis, which spans the months of November through January. The number of new cases is shown on the y-axis, which goes from 0 to 50. The actual number of new cases is represented by the blue line, which varies between 30 and 50 during the period before sharply increasing toward the end. The Decision Tree (DT) model's predictions are represented by the red line, which exhibits a sharp decline in early January and is comparatively flat and less responsive to changes in the actual case numbers. Some of the more dynamic changes observed in the actual data seem to be absent from the Decision Tree model, which seems to only capture broad trends.

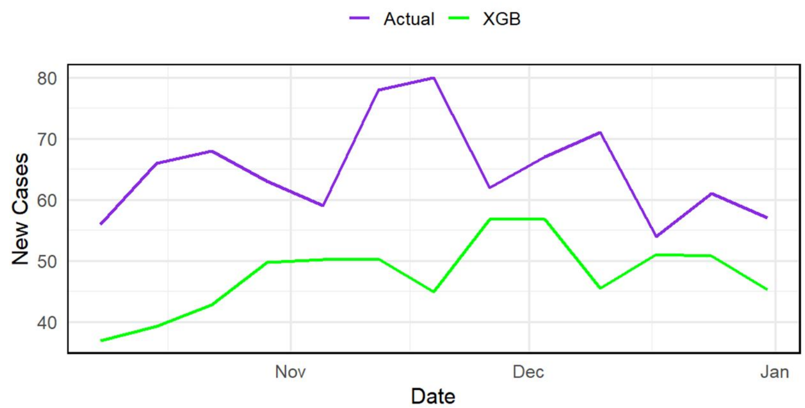


Figure 6: Forecast vs. Actual Daily New Monkeypox Cases in North America Using the Decision Tree (DT) Model, November 2023 to January 2024

The figure displays a line plot that contrasts the number of cases of monkeypox that occurred over time with those predicted by the XGBoost model. The date is represented by the x-axis, which spans the months of November through January. The number of new cases is shown on the y-axis, which runs from 40 to 80. The actual number of new cases is represented by the purple line; it varies from 50 to 80, with significant peaks in December. The XGBoost (XGB) model predicts values that are consistently lower than the actual cases and have fewer peaks than the actual cases; the green line illustrates this more gradually and smoothly. The higher peaks seen in the real case data are underestimated by the XGBoost model, even though it does capture some trends.

IV. DISCUSSION

This study's main goal was to create and evaluate machine learning models that could be used to predict the number of new cases of monkeypox in North and South America in 2023. Time series data of cases of monkeypox were subjected to five popular machine learning algorithms: Random Forest, Extreme Gradient Boosting (XGBoost), K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), and Decision Tree (DT). To determine the most accurate forecasting method for each region, the performance of each model was assessed using a variety of error metrics, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

The Decision Tree (DT) model yielded the lowest values for most performance metrics, making it the best performer for South America. This shows that maybe as a result of its ability to handle non-linear relationships within the data effectively, DT was able to capture the trend and pattern in the number of cases of monkeypox in this area. On the other hand, XGBoost proved to be the best model for North America, surpassing Decision Tree, Random Forest, and ANN. The superior performance of XGBoost can probably be attributed to its regularization techniques and its capacity to manage intricate interactions, which make it more appropriate for the type of data found in North American monkeypox cases. Although these models showed good forecasting abilities, their performance varied depending on the region, indicating that differences in complexity, trends, and noise may exist between the underlying data characteristics in North America and South America, which could have an impact on model performance.

Here, we go over the findings of a few more studies. Stacking Ensemble Learning (SEL) outperforms other machine learning techniques in predicting the rate of monkeypox transmission (15). When analyzing public sentiment regarding the monkeypox outbreak, the model developed using TextBlob annotation + Lemmatization + CountVectorizer + SVM produced the highest accuracy, approximately 0.9348 (45). For predicting daily cumulative confirmed cases of monkeypox, ARIMA (2, 2, 1) and exponential smoothing models were the most effective; however, Xgboost was the most effective for North America (5).

When it comes to predicting cases of monkeypox, machine learning techniques like the multilayer perceptron model (MLP) perform better than conventional time series models (46). Based on reviews, the most accurate method for identifying symptoms of monkeypox is XGBoost, with an accuracy of 1.0 (47). In terms of cumulative case counts, the NNETAR model performed better in Spain, the UK, and the USA, while DT was most effective in South America (48).

Potential Limitations of the Study: The exclusion of important contributing factors, such as demographic profiles, environmental factors, and vaccination coverage, may have biased the model fit. Secondly, some monkeypox event data were missing or incomplete, which may have introduced bias into the analysis.

As such, the models in this study may not generalize well to other regions that have different health care systems, population behaviors, or environmental conditions. Furthermore, model optimization may have been affected by sensitivity to hyperparameter tuning, particularly for Artificial Neural Networks (ANN) and XGBoost.

In the context of future research, other explanatory features such as mobility, climate conditions, quality of healthcare infrastructure, or vaccination coverage could be included in forecasting models in order to improve their predictive performance. Ensemble approaches that combine the best from different models might offer possible solutions to enhance prediction performance and regional robustness. However, the models also performed to various levels by region, although some individual models did relatively well on average, with North and South America amongst regions producing the best models. This expresses the call to create and iterate on region-specific models tailored to the unique epidemiological characteristics of each area.

Although no out-of-sample predictions were made, the in-sample forecasts offer valuable insights into short-term outbreak patterns. These results can help public health authorities implement timely interventions, use resources efficiently, and improve surveillance systems. By spotting potential increases in cases during the studied period, policymakers can make informed decisions to reduce the immediate impact of the outbreak.

Future research could examine time series-specific models like Prophet, Long Short-Term Memory (LSTM) networks, and hybrid ARIMA techniques to better manage intricate temporal dynamics. From a policy standpoint, more precise forecasting models that are tailored to a certain region could facilitate proactive public health initiatives, prompt outbreak response, and targeted resource allocation.

V. CONCLUSION

This study examined various machine learning models, including ANN, KNN, Decision Tree, Random Forest, and XGBoost, to forecast new cases of monkeypox in North and South America for 2023. We found that XGBoost delivered the best performance in North America, whereas the Decision Tree model yielded the most accurate predictions in South America.

The variation in model performance between North and South America highlights how regional factors can influence the choice of the best forecasting model. In South America, the Decision Tree model worked well due to its straightforward nature, while in North America, XGBoost’ s ability to capture more complex trends made it the better option.

These results show how important it is to adjust forecasting models based on the unique characteristics of each region. Although the results are promising, there are some limitations to consider. The small amount of case data, the lack of environmental and demographic factors, and the sensitivity to hyperparameters highlight the need for further improvements. By choosing the right models, health officials in North and South America can make more accurate predictions about monkeypox cases and act quickly. In the future, combining different models or trying out hybrid approaches could help improve the accuracy of forecasts across various regions.

VI. ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to Our World in Data for providing open access to the monkeypox incidence dataset used in this study. We also thank the Department of Statistics and Data Science for academic support and constructive feedback during the research process special thanks to all collaborators and contributors who provided valuable insights throughout the development of this work.

Funding

There is no funding for this work.

Competing Interest

The authors declare no competing interests.

Availability of Data

The dataset used in this study was obtained from publicly available sources provided by *Our World in Data* (https://ourworldindata.org/). All data used for analysis are open access and can be freely downloaded. The processed dataset used in this study is available from the corresponding author upon reasonable request.

Author’s Contribution

Sharmin Akther and Amrin Binte Ahmed contributed to the conceptualization, data curation, formal analysis, visualization, mapping, and preparation of the original draft. Sanjida Tasnim provided supervision, validation, and contributed to the review and editing of the manuscript.

REFERENCES

[1] Nuzzo JB, Borio LL, Gostin LO. The WHO Declaration of Monkeypox as a Global Public Health Emergency. JAMA. 2022 Aug 16;328(7):615–7.

[2] Heymann DL, Szczeniowski M, Esteves K. Re-emergence of monkeypox in Africa: a review of the past six years. Br Med Bull. 1998;54(3):693–702.

[3] Durski KN. Emergence of Monkeypox — West and Central Africa, 1970–2017. MMWR Morb Mortal Wkly Rep [Internet]. 2018 [cited 2024 Oct 14];67. Available from: https://www.cdc.gov/mmwr/volumes/67/wr/mm6710a5.htm

[4] Bankuru SV, Kossol S, Hou W, Mahmoudi P, Rychtář J, Taylor D. A game-theoretic model of Monkeypox to assess vaccination strategies. PeerJ. 2020 Jun 22;8:e9272.

[5] Wei W, Wang G, Tao X, Luo Q, Chen L, Bao X, et al. Time series prediction for the epidemic trends of monkeypox using the ARIMA, exponential smoothing, GM (1, 1) and LSTM deep learning methods. J Gen Virol [Internet]. 2023 Apr 6 [cited 2024 Oct 14];104(4). Available from: https://www.microbiologyresearch.org/content/journal/jgv/10.1099/jgv.0.001839

[6] Kmiec D, Kirchhoff F. Monkeypox: a new threat? Int J Mol Sci. 2022;23(14):7866.

[7] Karem KL, Reynolds M, Hughes C, Braden Z, Nigam P, Crotty S, et al. Monkeypox-induced immunity and failure of childhood smallpox vaccination to provide complete protection. Clin Vaccine Immunol CVI. 2007 Oct;14(10):1318–27.

[8] Zhang WD, Zu ZH, Xu Q, Xu ZJ, Liu JJ, Zheng T. Optimized strategy for the control and prevention of newly emerging influenza revealed by the spread dynamics model. PloS One. 2014;9(1):e84694.

[9] J Z, Ml L, Zf L, Mw S, Yn X, Fp J. Transmission patterns of COVID-19 in the mainland of China and the efficacy of different control strategies: a data- and model-driven study. Infect Dis Poverty [Internet]. 2020 Jul 6 [cited 2024 Oct 14];9(1). Available from: https://pubmed.ncbi.nlm.nih.gov/32631426/

[10] Jr W, Sn I. Monkeypox virus and insights into its immunomodulatory proteins. Immunol Rev [Internet]. 2008 Oct [cited 2024 Oct 14];225. Available from: https://pubmed.ncbi.nlm.nih.gov/18837778/

[11] Ngbolua K te N, Ngambika GK, Mbembo-wa-Mbembo B, Séraphin KP, Fabrice KK, Bongo GN, et al. First Report on Three Cases of Monkey pox in Nord Ubangi Province (Democratic Republic of the Congo). Br Int Exact Sci BIoEx J. 2020 Jan 9;2(1):120–5.

[12] Status of human monkeypox: clinical disease, epidemiology and research - ScienceDirect [Internet]. [cited 2024 Oct 14]. Available from: https://www.sciencedirect.com/science/article/pii/S0264410X1100524X

[13] 2022–2023 mpox outbreak in South America. In: Wikipedia [Internet]. 2024 [cited 2024 Oct 14]. Available from: https://en.wikipedia.org/w/index.php?title=2022%E2%80%932023_mpox_outbreak_in_South_America&oldid=1236886127

[14] 2022–2023 mpox outbreak in North America. In: Wikipedia [Internet]. 2024 [cited 2024 Oct 14]. Available from: https://en.wikipedia.org/w/index.php?title=2022%E2%80%932023_mpox_outbreak_in_North_America&oldid=1236886126

[15] Dada EG, Oyewola DO, Joseph SB, Emebo O, Oluwagbemi OO. Ensemble Machine Learning for Monkeypox Transmission Time Series Forecasting. Appl Sci. 2022 Jan;12(23):12128.

[16] Ncube B, Dzikiti M, Nyoni A, Ncube M, Ndlovu BM, Dube S. Effectiveness of Machine Learning algorithms in predicting Monkey Pox (Mpox): A Systematic Literature Review.

[17] Priyadarshini I, Mohanty P, Kumar R, Taniar D. Monkeypox Outbreak Analysis: An Extensive Study Using Machine Learning Models and Time Series Analysis. Computers. 2023 Feb 7;12(2):36.

[18] Mathieu E, Spooner F, Dattani S, Ritchie H, Roser M. Mpox. Our World Data [Internet]. 2024 Mar 22 [cited 2024 Oct 14]; Available from: https://ourworldindata.org/mpox

[19] Databricks [Internet]. 2022 [cited 2024 Oct 15]. What are Machine Learning Models? Available from: https://www.databricks.com/glossary/machine-learning-models

[20] Forbytes [Internet]. 2023 [cited 2024 Oct 15]. Top 6 Machine Learning Techniques for Predictive Modeling. Available from: https://forbytes.com/blog/main-machine-learning-techniques/

[21] Neural network (machine learning). In: Wikipedia [Internet]. 2024 [cited 2024 Oct 15]. Available from: https://en.wikipedia.org/w/index.php?title=Neural_network_(machine_learning)&oldid=1250750800

[22] Artificial Neural Network - an overview | ScienceDirect Topics [Internet]. [cited 2024 Oct 15]. Available from: https://www.sciencedirect.com/topics/neuroscience/artificial-neural-network

[23] Yang Z, Yang Z (2014). Comprehensive Biomedical Physics. Karolinska Institutet, Stockholm, Sweden: Elsevier. p. 1. ISBN 978-0-444-53633-4. Archived from the original on 28 July 2022. Retrieved 28 July 2022 - Google Search [Internet]. [cited 2024 Oct 15]. Available from: https://www.google.com/search?client=firefox-b-d&sca_esv=abee16815fc6fe5b&q=Yang+Z,+Yang+Z+(2014).+Comprehensive+Biomedical+Physics.+Karolinska+Institutet,+Stockholm,+Sweden:+Elsevier.+p.+1.+ISBN+978-0-444-53633-4.+Archived+from+the+original+on+28+July+2022.+Retrieved+28+July+2022&sa=X&ved=2ahUKEwix1qi6k5CJAxW6fGwGHQEYNoAQ7xYoAHoECAoQAQ&biw=1366&bih=643&dpr=1

[24] (01) The Nature of Statistical Learning Theory.pdf [Internet]. [cited 2024 Oct 15]. Available from: http://39.108.104.114/res/(01)%20The%20Nature%20of%20Statistical%20Learning%20Theory.pdf

[25] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Netw. 1989 Jan 1;2(5):359–66.

[26] www.javatpoint.com [Internet]. [cited 2024 Oct 15]. Decision Tree Algorithm in Machine Learning - Javatpoint. Available from: https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm

[27] scikit-learn [Internet]. [cited 2024 Oct 15]. 1.10. Decision Trees. Available from: https://scikit-learn/stable/modules/tree.html

[28] Ben-Haim Y, Tom-Tov E. A Streaming Parallel Decision Tree Algorithm.

[29] Abdulhafedh A. Comparison between Common Statistical Modeling Techniques Used in Research, Including: Discriminant Analysis vs Logistic Regression, Ridge Regression vs LASSO, and Decision Tree vs Random Forest. OALib. 2022;09(02):1–19.

[30] Tin Kam Ho. The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell. 1998 Aug;20(8):832–44.

[31] Random Forest Classification with Scikit-Learn [Internet]. [cited 2024 Oct 15]. Available from: https://www.datacamp.com/tutorial/random-forests-classifier-python

[32] Oyewola D, Hakimi D, Adeboye K, Shehu MD. Using Five Machine Learning for Breast Cancer Biopsy Predictions Based on Mammographic Diagnosis. Int J Eng Technol IJET. 2017 Apr 20;2(4):142–5.

[33] Nair A. AIM. 2019 [cited 2024 Oct 15]. MLP Classifier - A Beginner's Guide To SKLearn MLP Classifier. Available from: https://analyticsindiamag.com/ai-mysteries/a-beginners-guide-to-scikit-learns-mlpclassifier/

[34] Classification Using Sklearn Multi-layer Perceptron - GeeksforGeeks [Internet]. [cited 2024 Oct 15]. Available from: https://www.geeksforgeeks.org/classification-using-sklearn-multi-layer-perceptron/

[35] Deepchecks [Internet]. [cited 2024 Oct 15]. What is Multilayer Perceptron. Available from: https://www.deepchecks.com/glossary/multilayer-perceptron/

[36] Dutta S. Multi-Layer Perceptron and Backpropagation: A Deep Dive [Internet]. Medium. 2024 [cited 2024 Oct 15]. Available from: https://medium.com/@sanjay_dutta/multi-layer-perceptron-and-backpropagation-a-deep-dive-8438cc8bcae6

[37] What is k-Nearest Neighbor (kNN)? | A Comprehensive k-Nearest Neighbor Guide [Internet]. [cited 2024 Oct 15]. Available from: https://www.elastic.co/what-is/knn

[38] GeeksforGeeks [Internet]. 2017 [cited 2024 Oct 15]. K-Nearest Neighbor(KNN) Algorithm. Available from: https://www.geeksforgeeks.org/k-nearest-neighbours/

[39] What is the k-nearest neighbors algorithm? | IBM [Internet]. 2021 [cited 2024 Oct 15]. Available from: https://www.ibm.com/topics/knn

[40] xgboost algorithm - Google Search [Internet]. [cited 2024 Oct 16]. Available from: https://www.google.com/search?q=xgboost+algorithm&client=firefox-b-d&sca_esv=28f8fab5923385a6&ei=jK8OZ-OTB_rZseMP0vn0qQc&oq=XGBoost&gs_lp=Egxnd3Mtd2l6LXNlcnAiB1hHQm9vc3QqAggBMgsQABiABBiRAhiKBTILEAAYgAQYkQIYigUyChAAGIAEGEMYigUyChAAGIAEGEMYigUyChAAGIAEGEMYigUyBRAAGIAEMgUQABiABDIFEAAYgAQyBRAAGIAESJFRUIAVWOA3cAN4AZABAJgBuwGgAbwIqgEDMC43uAEByAEA-AEBmAIKoALJCqgCCsICChAAGLADGNYEGEfCAhQQABiABBiRAhi0AhiKBRjqAtgBAcICFhAAGAMYtAIY5QIY6gIYjAMYjwHYAQHCAgsQABiABBixAxiDAcICCxAuGIAEGLEDGIMBwgIOEAAYgAQYsQMYgwEYigXCAhEQLhiABBixAxjRAxiDARjHAcICEBAuGIAEGNEDGEMYxwEYigXCAggQLhiABBixA5gDDOIDBRIBMSBAiAYBkAYIugYECAEYB5IHBTMuNi4xoAfgMA&sclient=gws-wiz-serp

[41] Simplilearn.com [Internet]. 2022 [cited 2024 Oct 16]. What is XGBoost? An Introduction to XGBoost Algorithm in Machine Learning | Simplilearn. Available from: https://www.simplilearn.com/what-is-xgboost-algorithm-in-machine-learning-article

[42] GeeksforGeeks [Internet]. 2021 [cited 2024 Oct 16]. XGBoost. Available from: https://www.geeksforgeeks.org/xgboost/

[43] Torres LF. XGBoost: The King of Machine Learning Algorithms [Internet]. LatinXinAI. 2023 [cited 2024 Oct 16]. Available from: https://medium.com/latinxinai/xgboost-the-king-of-machine-learning-algorithms-6b5c0d4acd87

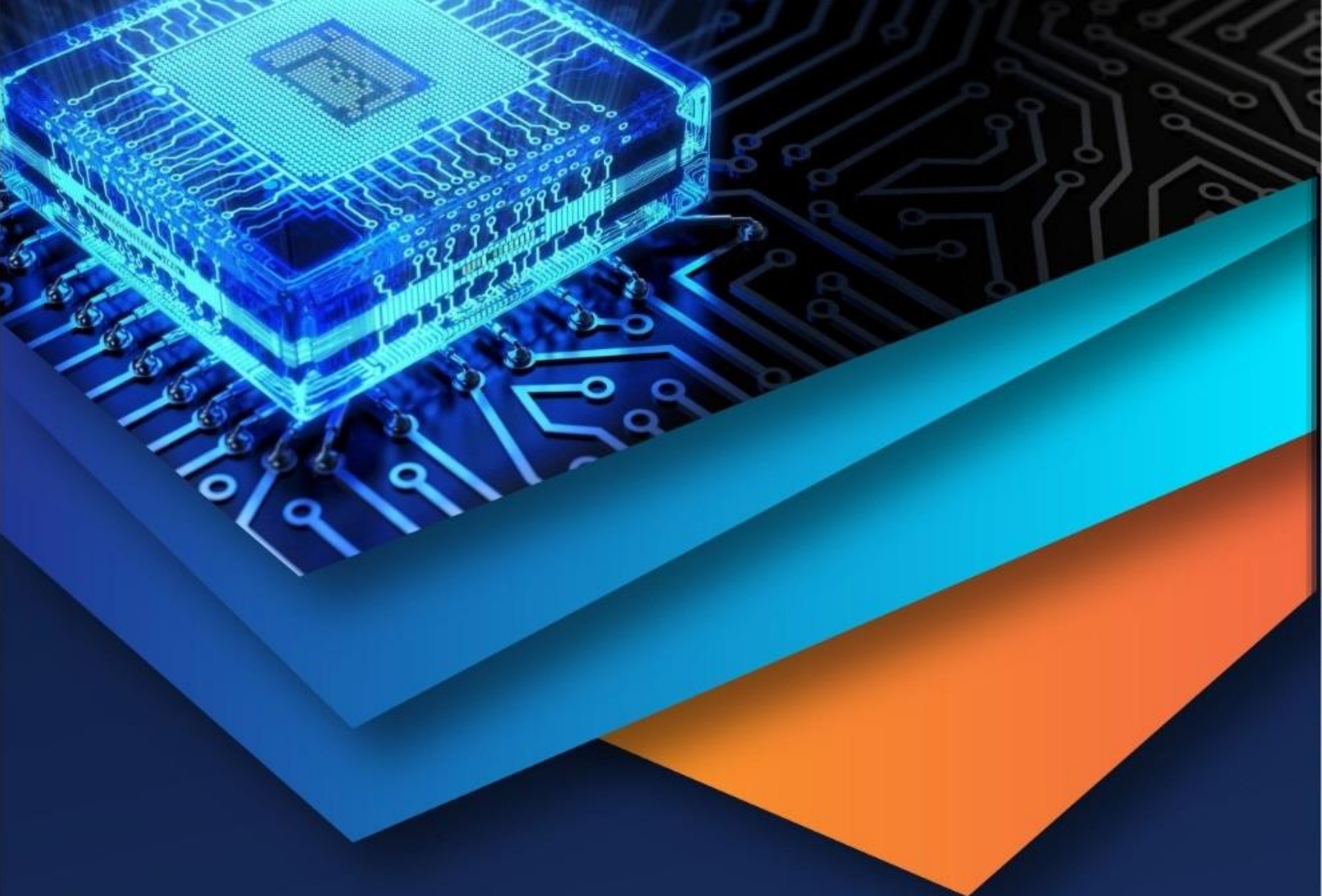
[44] Brożyna J, Mentel G, Szetela B, Strielkowski W. Multi-Seasonality in the TBATS Model Using Demand for Electric Energy as a Case Study. Econ Comput Econ Cybern Stud Res Acad Econ Stud. 2018 Mar 20;52:229–46.

[45] Bengesi S, Oladunni T, Olusegun R, Audu H. A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion From Twitter Tweets. IEEE Access. 2023;11:11811–26.

[46] Qureshi M, Khan S, Bantan RAR, Daniyal M, Elgarhy M, Marzo RR, et al. Modeling and Forecasting Monkeypox Cases Using Stochastic Models. J Clin Med. 2022 Jan;11(21):6555.

[47] Detection of Monkeypox Cases Based on Symptoms Using XGBoost and Shapley Additive Explanations Methods [Internet]. [cited 2024 Oct 17]. Available from: https://www.mdpi.com/2075-4418/13/14/2391

[48] Akinola SO, Wang QG, Olukanmi P, Marwala T. Early Prediction of Monkeypox Virus Outbreak Using Machine Learning. IETI Trans Data Anal Forecast ITDAF. 2023 Jul 6;1(2):14–29.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY