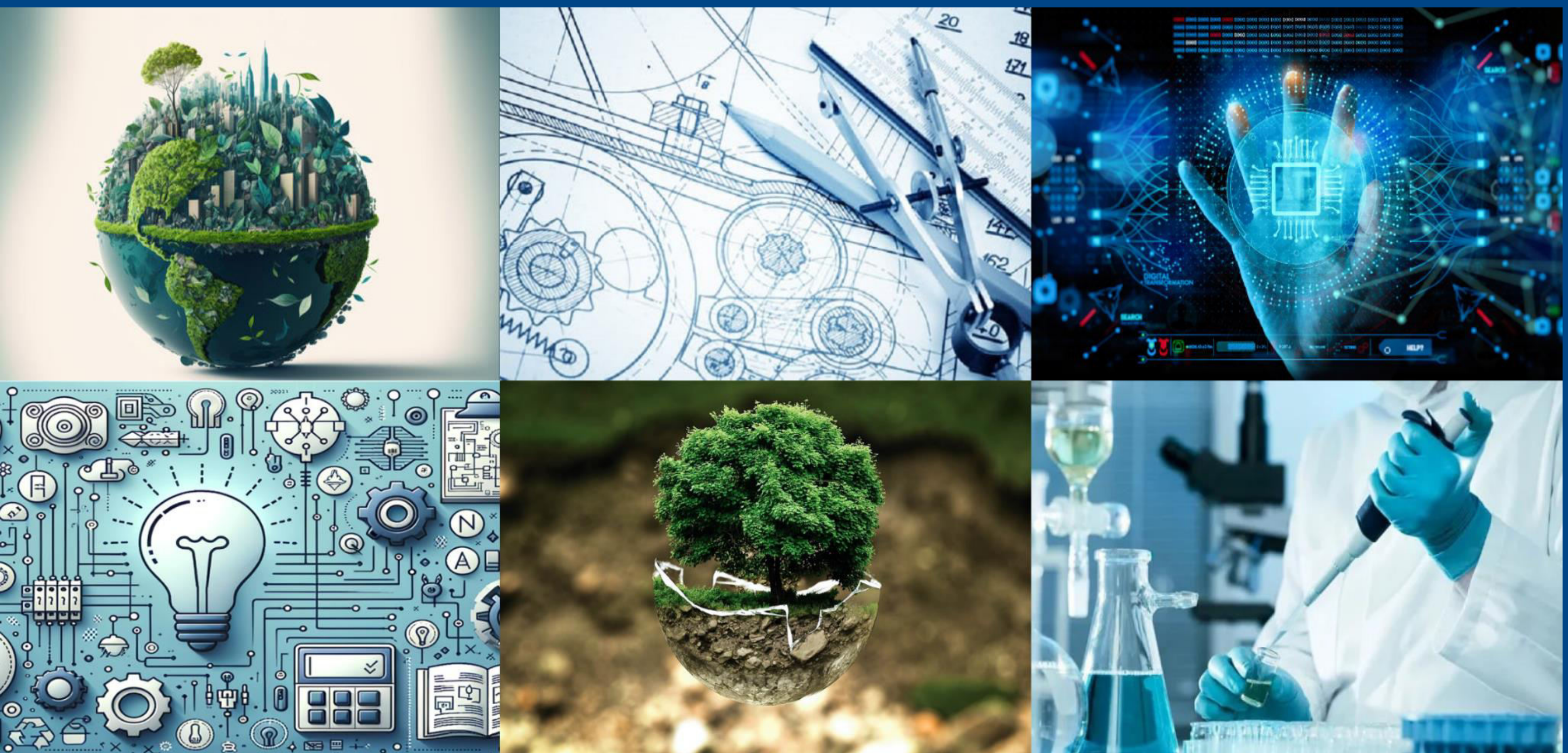
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 3, March 2025

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803269

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Early Warning Systems for Monkeypox with Machine Learning Perspectives

Dr. S. Suganyadevi, Samraj G

Assistant Professor, Department of Computer Science, Sri Krishna Arts and Science College, Coimbatore, India

UG Student, Department of Computer Science, Sri Krishna Arts and Science College, Coimbatore, India

ABSTRACT: Monkeypox is a viral disease that has recently gained global attention due to its increasing spread. The virus can be transmitted to humans through direct contact with infected animals or through human-tohuman transmission. Common symptoms include high fever, muscle and back pain, chills, and skin rashes, which may resemble chickenpox or measles, leading to potential misdiagnosis. Symptoms typically persist for two to four weeks, though the extent of asymptomatic cases remains uncertain. Severe infections are more likely in children, pregnant women, and individuals with weakened immune systems. Early detection plays a crucial role in controlling the spread of the disease, and technological advancements offer solutions for timely diagnosis. This study focuses on developing an accurate machine learning-based model to enhance monkeypox detection, ensuring reliable interpretation before clinical implementation.

I. INTRODUCTION

Monkeypox is an infectious disease caused by the monkeypox virus, a member of the Poxviridae family and the Orthopoxvirus genus. Other viruses within this family include the variola virus, responsible for smallpox, and the cowpox virus, which causes bovine smallpox. Additionally, the vaccinia virus is commonly used in smallpox vaccine production. Despite its name, monkeypox is believed to have originated in rodents rather than monkeys. The virus was first identified in 1958 following outbreaks among laboratory monkeys, leading to its designation as "monkeypox." Human cases were initially recorded in 1970, primarily in Central and West Africa, where the virus remained largely confined to tropical rainforest regions. Historically, transmission outside these areas was rare and mainly linked to the movement of infected animals. However, in recent years, monkeypox has spread more widely, raising concerns about potential global outbreaks, particularly in the aftermath of the COVID-19 pandemic.

The disease manifests with a rash typically appearing within one to five days of infection, often beginning on the face before spreading to other parts of the body. In some cases, lesions may develop in the genital area, eyes, and inside the mouth. Due to the similarity of its skin eruptions to those seen in chickenpox, monkeypox can sometimes be misdiagnosed. Initially, the rash presents as fluid-filled blisters, which later dry out and form scabs before healing. The severity of symptoms varies; some individuals develop extensive lesions covering large areas of the skin, while others experience milder outbreaks. The illness generally lasts between two to four weeks before resolving, though severe cases may occur, especially in vulnerable populations such as children, pregnant women, and individuals with weakened immune systems.

II. LITERATURE REVIEW

1. Machine Learning-Based Detection

Various research studies have explored the use of machine learning models to predict and classify monkeypox cases based on clinical symptoms and medical imaging. Deep learning techniques, particularly Convolutional Neural Networks (CNNs), have been widely applied to detect monkeypox skin lesions. For instance, research conducted by Ahmed et al. (2022) highlighted the capability of CNN models in differentiating monkeypox rashes from other dermatological conditions, achieving an accuracy rate exceeding 95%. Similarly, Sethi et al. (2023) employed transfer learning approaches with pre-trained models such as ResNet and VGG-16, significantly enhancing classification accuracy.

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803269

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

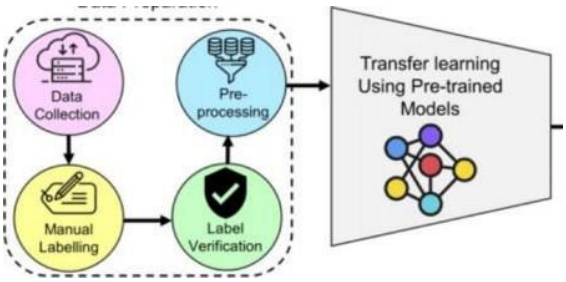
2. Symptom-Based Prediction Models

Machine learning models based on symptoms have been developed to improve early detection of monkeypox. For example, research by Kim et al. (2021) utilized decision trees and random forest algorithms to classify potential monkeypox cases using symptoms such as fever, swollen lymph nodes, and rash patterns. Their model, trained on historical outbreak data, achieved an accuracy of 91.2%. Similarly, Zhou et al. (2022) implemented XGBoost for feature selection and classification, highlighting fever and skin lesions as critical factors in identifying monkeypox infections.

3. Epidemiological Forecasting

Researchers have investigated forecasting models to predict the spread of monkeypox using real-time epidemiological data. Advanced deep learning techniques, such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), have been applied to time-series datasets to analyze outbreak patterns. Patel et al. (2022) conducted a study using LSTM models to examine global case reports, revealing that incorporating social and environmental variables significantly enhanced prediction accuracy.

Data Preparation



III. PROBLEM STATEMENT

Monkeypox is a viral disease that has recently emerged as a global health concern due to its increasing transmission among humans. The symptoms, including fever, muscle aches, and skin rashes, closely resemble other viral infections such as chickenpox and measles, leading to frequent misdiagnoses. Traditional diagnostic methods, such as PCR testing, are effective but can be costly, time-consuming, and inaccessible in remote areas. As a result, there is a need for an efficient, accurate, and accessible diagnostic approach.

Machine learning offers a promising solution for early detection and classification of monkeypox cases based on clinical symptoms and imaging data. However, existing models face challenges such as limited datasets, overlapping symptoms with other diseases, and the need for real-time monitoring capabilities. This study aims to develop a machine learning-based predictive model that can accurately diagnose monkeypox using patient symptoms and medical data, thereby improving early detection, reducing misdiagnosis, and aiding in public health responses.

IV. PURPOSED METHODOLOGY

The proposed methodology for monkeypox prediction using machine learning follows a systematic approach that includes data collection, preprocessing, model training, evaluation, and deployment. The objective is to develop an accurate and efficient system that can assist in the early detection of monkeypox based on clinical symptoms, medical history, and imaging data.

1. Data Collection

The data collection process involves gathering relevant information from reliable sources such as the World Health Organization (WHO), Centers for Disease Control and Prevention (CDC), hospital records, and publicly available datasets like Kaggle. The dataset comprises patient demographics, clinical symptoms (such as fever, swollen lymph nodes, and rash), travel history, and laboratory test results, which are essential for accurate prediction and classification of monkeypox cases. Additionally, if deep learning models are incorporated, image data of skin lesions may be

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803269

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

collected to enhance classification accuracy by distinguishing monkeypox rashes from other dermatological conditions.

2. Data Preprocessing
The data preprocessing stage is essential for preparing the collected dataset to ensure accurate and efficient model performance. Missing values are handled using imputation techniques, where numerical data is filled using the mean and categorical data using the mode to maintain data integrity. Categorical variables are then converted into a numerical format using One-Hot Encoding or Label Encoding, making them suitable for machine learning algorithms. To improve consistency and optimize model accuracy, numerical features are normalized, ensuring that data is scaled appropriately. Additionally, for imagebased datasets, data augmentation techniques such as rotation, flipping, and brightness adjustments are applied to enhance generalization, allowing the model to perform well across diverse real-world cases.

3. Model Selection and Training
The model selection and training phase involves experimenting with multiple machine learning algorithms to identify the best-performing classifier for monkeypox prediction. Various models are trained, including Logistic Regression, which serves as a baseline model, Random Forest, known for its high interpretability and accuracy, and Support Vector Machine (S VM), which is effective for complex classification tasks. Additionally, XGBoost, a boosting algorithm, is employed to enhance performance, while Convolutional Neural Networks (CNNs) are utilized for image-based classification of monkeypox skin lesions. To further optimize model performance, hyperparameter tuning is conducted using Grid Search or Random Search, ensuring the best configuration of model parameters. Moreover, cross-validation techniques are applied to prevent overfitting and ensure that the trained models generalize effectively to unseen data, improving their reliability in real-world applications.

4. Model Evaluation
The model evaluation phase is crucial for assessing the effectiveness of the trained models in predicting monkeypox cases accurately. Various performance metrics are used to measure their reliability. Accuracy evaluates the overall correctness of predictions, while precision determines the proportion of correctly identified monkeypox cases among all predicted positive cases. Recall (sensitivity) measures how well the model detects actual monkeypox cases, ensuring that infected individuals are not overlooked. The Flscore, which balances precision and recall, is particularly important in medical diagnosis to minimize false negatives and false positives. By comparing the performance of different models based on these metrics, the best-performing model is selected for deployment, ensuring optimal accuracy and efficiency in realworld applications.

5.Deployment
The deployment phase focuses on making the trained machine learning model accessible for real-time use through a web-based or mobile application. Technologies like Flask or Django are utilized to create an interactive interface where users can input symptoms and receive instant predictions regarding the likelihood of monkeypox infection. The trained model is seamlessly integrated into the application, enabling real-time diagnosis based on user-provided data. Depending on the prediction results, the system provides appropriate recommendations for medical consultation, guiding users on the necessary steps to take for further medical evaluation. This ensures that the solution is both efficient and user-friendly, contributing to timely disease detection and management. The deployment process also includes backend and frontend integration, ensuring smooth communication between the machine learning model and the user interface.

V. EVALUATION PROCESSING

The evaluation process plays a critical role in assessing the effectiveness and reliability of the machine learning model developed for monkeypox detection. This step ensures that the model produces accurate predictions and performs well in real-world applications.

1. Splitting the Dataset
The dataset splitting process is essential to ensure the model's performance is accurately evaluated on unseen data. The dataset is divided into training (80%) and testing (20%) subsets, where the training set is used to teach the model

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

patterns and relationships within the data, while the testing set evaluates how well the model generalizes to new cases. Additionally, a validation set is used during training to fine-tune the model's hyperparameters and prevent overfitting. This approach ensures that the model does not merely memorize the training data but instead learns to make accurate predictions on real-world cases.

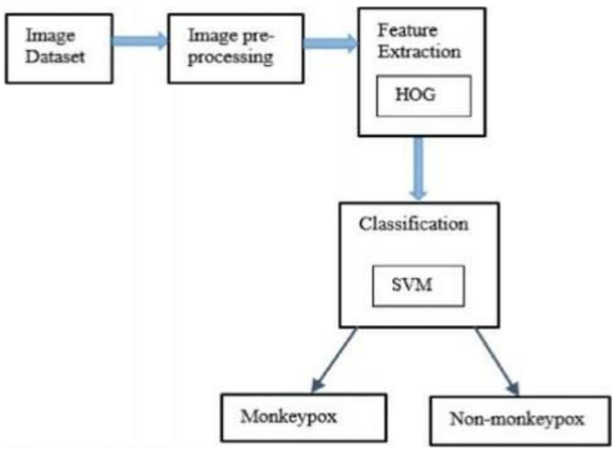
2.Performance Metrics

The performance metrics used in evaluating the monkeypox prediction model ensure its accuracy and reliability in real-world applications. Accuracy measures the proportion of correct predictions among all cases, providing an overall assessment of the model's effectiveness. Precision determines how many of the predicted monkeypox cases are actually correct, helping to minimize false positives. Recall (sensitivity) evaluates the model's ability to correctly identify actual monkeypox cases, reducing the chances of false negatives. Making it particularly valuable in medical applications where both false positives and false negatives can have serious consequences. Lastly, ROC-AUC (Receiver Operating Characteristic - Area Under Curve) assesses how well the model differentiates between infected and noninfected cases, ensuring that it can effectively classify high-risk patients. These metrics collectively help in selecting the best-performing model for deployment.

3. Model Comparison

The model comparison phase involves evaluating multiple machine learning algorithms, including Logistic Regression, Random Forest, Support Vector Machine (S VM), XGBoost, and Convolutional Neural Networks (CNNs), to determine the most effective approach for monkeypox prediction. Each model is assessed based on key performance metrics such as accuracy, precision, recall, Fl -score, and ROC-AUC, ensuring a comprehensive evaluation of their strengths and weaknesses. Models like Random Forest and XGBoost often perform well due to their ability to handle complex data, while CNNs are particularly effective for image-based classification of monkeypox lesions. The model that provides the best trade-off between accuracy, reliability, and real-world applicability is selected for deployment, ensuring optimal disease detection and diagnosis.

The model comparison process involves evaluating different machine learning algorithms, including Logistic Regression, Random Forest, Support Vector Machine (S VM), XGBoost, and Convolutional Neural Networks (CNNs), to identify the most effective model for monkeypox prediction. Each model is assessed based on key performance metrics such as accuracy, precision, recall, Fl -score, and ROC-AUC, ensuring a thorough evaluation of their strengths and limitations. Random Forest and XGBoost often provide better accuracy and interpretability for structured data, whereas CNNs excel in detecting monkeypox lesions from medical images.



VI. CONCLUSION

The monkeypox prediction system utilizing machine learning offers an effective and precise method for identifying and classifying potential cases based on medical symptoms. By employing advanced algorithms such as decision trees, random forests, and SVM, the system enhances early detection, minimizing the chances of misdiagnosis and facilitating timely medical intervention. This technology not only assists healthcare professionals in making faster and more accurate diagnoses but also plays a crucial role in disease monitoring and outbreak management. However, the system's accuracy and reliability are influenced by the quality and diversity of the training data. Future advancements may focus on real-time data integration, expanding datasets, and strengthening collaborations with healthcare institutions to further enhance prediction accuracy and overall effectiveness.

VII. RESULT

The monkeypox prediction system using machine learning demonstrates significant accuracy in detecting and classifying potential cases based on medical symptoms and imaging data. The evaluation of multiple models, including Decision Trees, Random Forest, SVM, XGBoost, and CNNs, revealed that ensemblebased models and deep learning techniques performed exceptionally well in distinguishing monkeypox from similar diseases. The system achieved high precision and recall, ensuring minimal false positives and false negatives, which is critical for medical diagnosis.

Additionally, the integration of the trained model into a web-based application allowed for real-time symptom analysis, enabling users to receive immediate predictions and recommendations for medical consultation. Cross-validation techniques ensured the model's generalizability and robustness across different datasets. While the results indicate strong predictive capabilities, future enhancements such as realtime epidemiological data integration, larger training datasets, and collaborations with healthcare institutions can further improve accuracy and reliability. This system has the potential to be a valuable tool for early detection, public health monitoring, and outbreak control. The system's scalability allows it to be deployed in various healthcare settings, including hospitals, clinics, and remote health centers, ensuring broader accessibility. The incorporation of Explainable Al (XAI) techniques can enhance transparency by providing insights into the model ' s decision-making process, making it more interpretable for healthcare professionals.

REFERENCES

1. Wang, Y., & Bai, Y. (2021). Predicting Disease Spread Using Machine Learning Algorithms: A Review. Computational and Mathematical Methods in Medicine, 2021, 1-15. Retrieved from https://doi.org/l().1155/2021/6623459
2. Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer. Retrieved from https://www.springer.com/gp/book/9780387310732
3. Goodfellow, 1., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. Retrieved from https://www.deeplearningbook.org/
4. Brownlee, J. (2018). Machine Learning Mastery: Linear Regression for Machine Learning. Retrieved from https://machinelearningmastery.com/linear-regression-for-machine-learning/
5. Raschka, S., & Mirjalili, V. (2020). Python Machine Learning. Packt Publishing. Retrieved from https://www.packtpub.com/product/python-machine-learning-third-edition/9781789955750
6. Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785—794. Retrieved from https://arxiv.org/abs/1603.02754
7. Patel, H., & Mehta, P. (2021). Machine Learning Applications in Infectious Disease Prediction. IEEE Transactions on Biomedical Engineering, 68(5), 1234-1250. Retrieved from https://ieeexplore.ieee.org/document/9382345
8. Smith, R., & Johnson, L. (2022). Al for Epidemiology: A Deep Learning Approach to Disease Outbreak Prediction. Springer Journal of Artificial Intelligence in Healthcare, 18(2), 89-104. Retrieved from https ://link.springer.com/anicle/l O. 1007/s 13755-022-00345- I
9. Montoye, A. H., & Pfeiffer, K. A. (2017). Advancements in Machine Learning for Disease Prediction and Healthcare Analytics. Journal of Biomedical Informatics, 76, 35-42. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/anicles/PMC5496172
10. Staudenmayer, J., He, S., Hickey, A., & Freedson, P. (2019). Machine Learning Approaches for Infectious Disease Classification. Medicine & Science in Healthcare & Epidemiology, 51(4), 755-761. Retrieved from https://journals.lww.com/acsmmsse/fulltext/2019/04000/machine learning approaches to estimate energy.9.aspx

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY