

RESEARCH

Open Access



Explainable AI for Symptom-Based Detection of Monkeypox: a machine learning approach

Gizachew Mulu Setegn^{1*} and Belayneh Endalamaw Dejene²

Abstract

Background Monkeypox, a viral zoonotic disease, is an emerging global health concern, with rising incidence and outbreaks extending beyond its endemic regions in Central and, West Africa and the world. The disease transmits through contact with infected animals and humans, leading to fever, rash, and lymphadenopathy symptoms. Control efforts include surveillance, contact tracing, and vaccination campaigns; however, the increasing number of cases underscores the necessity for a coordinated global response to mitigate its impact. Since monkeypox has become a public health issue, new methods for efficiently identifying cases are required. The control of monkeypox infections depends on early detection and prediction. This study aimed to utilize Symptom-Based Detection of Monkeypox using a machine-learning approach.

Methods This research presents a machine learning approach that integrates various Explainable Artificial Intelligence (XAI) to enhance the detection of monkeypox cases based on clinical symptoms, addressing the limitations of image-based diagnostic systems. In this study, we used a publicly available dataset from GitHub containing clinical features about monkeypox disease. The data have been analysed using Random Forest, Bagging, Gradient Boosting, CatBoost, XGBoost, and LGBMClassifier to develop a robust predictive model.

Results The study shows that machine learning models can accurately diagnose monkeypox based on symptoms like fever, rash, lymphadenopathy and other clinical symptoms. By using XAI techniques for feature importance, the approach not only achieved high accuracy but also provided transparency in decision-making. This integration of explainable Artificial intelligence (AI) enhances trust and allows healthcare professionals to understand predictions, leading to timely interventions and improved public health responses to monkeypox outbreaks. All Machine learning methods have been compared with the evaluation matrix. The best performance was for the LGBMClassifier, with an accuracy of 89.3%. In addition, multiple Explainable Techniques tools were used to help in examining and explaining the output of the LGBMClassifier model.

Conclusions Our research shows that combining explainable techniques with AI models greatly enhances the accuracy of case detection and boosts the trust of medical professionals. These models result in directly involving the reader and health care professional in the decision-making process, making informed decisions, and efficiently allocating resources by providing insight into the decision-making process. In addition, this study underscores the potential of AI in public health surveillance, particularly in enhancing responses to emerging infectious diseases such as monkeypox.

*Correspondence:
Gizachew Mulu Setegn
Gizachew.Mulu@dku.edu.et

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Keywords Clinical symptoms, Health, Machine learning, Monkeypox, SHAP, XAI

Introduction

Monkeypox, as defined by the World Health Organization (WHO), is a viral infection caused by the monkeypox virus, a member of the orthopoxvirus genus and closely related to smallpox, spreading globally and raising public health concerns [1–3]. The disease is mainly transmitted from animals to humans via direct contact with infected animals, but it can also spread between humans through respiratory droplets, bodily fluids, or contaminated materials with symptoms including fever, headache, swollen lymph nodes, and skin lesions [4–6]. Severe cases can lead to pneumonia, sepsis, or encephalitis [1, 7]. Early monkeypox diagnosis is vital but symptoms overlap with other diseases complicate clinical assessments. Figure 1 illustrates monkeypox infection, which typically manifests within three days following the onset of fever. Initially, symptoms appear on the face and subsequently spread to other parts of the body. The World Health Organization (WHO) announced a significant increase in the number of infections of this virus, with over 318,000 cases reported in August 2022 [2] and more than 70,000 death cases began to appear. In addition, the WHO has declared the monkeypox epidemic a global health emergency due to the outbreak in the Democratic Republic of Congo (DRC) spreading to at least 13 other African countries, as well as potential cases in Europe and Asia on 14 August 2024 [8]. At the time of this writing, the World Health Organization (WHO) has reported over 14,000 cases and 524 deaths in Africa in 2024, surpassing the figures from 2023. The Democratic Republic of the Congo (DRC) accounts for nearly 96% of these cases and deaths. The potential case count now exceeds 17,000, marking a significant rise from the 7,146 cases in 2022 and almost 15,000 cases in 2023. The Africa Center for Diseases Control (CDC) has verified that monkeypox outbreaks have been reported in at least 12 African countries, including Burundi, Kenya, Rwanda, and Uganda. In 2024, these countries have recorded a total of 2,863 cases and 517 deaths, with the majority occurring in the Democratic Republic of the Congo (DRC). Suspected cases across the continent have risen to over 17,000, a notable increase from 7,146 cases in 2022 and 14,957 cases in 2023 [9]. Figure 1 illustrate the symptoms of skin infection in monkeypox disease.

Currently, there is no specific antiviral treatment for the monkeypox virus (MPXV). However, supportive care can help manage symptoms and reduce risks. Contact tracing, quarantine, and isolation are crucial for controlling the spread of the disease [1]. Monkeypox typically manifests with a rash within 1–3 days after the onset of fever, starting on the face and then spreading to other

parts of the body. The rash progresses through several stages, including macules, papules, vesicles, and pustules, before ultimately forming scabs [11]. Swollen Lymph Nodes: A notable feature of monkeypox is lymphadenopathy, or swollen lymph nodes, which usually appear early in the illness and are not a common characteristic of other poxviruses like smallpox [12].

However, the unexpected resurgence of monkeypox in new regions has raised alarms within the public health sector, highlighting deficiencies in outbreak response and healthcare systems. Issues such as delayed detection and a lack of vaccines have been identified as significant barriers. Current initiatives focus on enhancing surveillance and diagnostic capabilities, and the clinical presentation of monkeypox can be complex, often overlapping with other infectious diseases, making it challenging for healthcare professionals to accurately distinguish monkeypox from similar viral infections based solely on symptoms. This is particularly problematic given the potentially severe consequences of misdiagnosis, including delayed treatment, ineffective containment measures, and the risk of further disease transmission. The recent monkeypox outbreak has underscored the importance of timely and accurate diagnosis in controlling the spread of the disease. While laboratory testing remains the gold standard for monkeypox diagnosis, access to such tests is limited in many regions. Imaging techniques like X-rays and MRIs are crucial for diagnosing medical conditions, but they can sometimes be misinterpreted due to various factors. The subjective nature of image interpretation can complicate the diagnostic process, as it heavily relies on the radiologist's experience. The use of AI and ML has significantly enhanced the management of the pandemic (such as Covid-19) by improving treatment, medication, screening, prediction, forecasting, contact tracing, and drug/vaccine development, while also reducing the dependence on human intervention in medical practices [13]. The integration of AI in public health raises significant ethical concerns, particularly regarding privacy and fairness. The use of AI-driven data analysis may lead to unintentional biases and discrimination, potentially affecting marginalized groups disproportionately. Furthermore, the collection and utilization of personal health information pose risks to individual privacy if not sufficiently safeguarded [14]. In contrast, this study highlights the need for alternative diagnostic approaches, such as clinical symptom-based predictive models offer a more holistic approach by considering the patient's overall clinical picture, medical history, and reported symptoms. These assessments provide contextual clarity and can reveal important nuances that imaging alone may

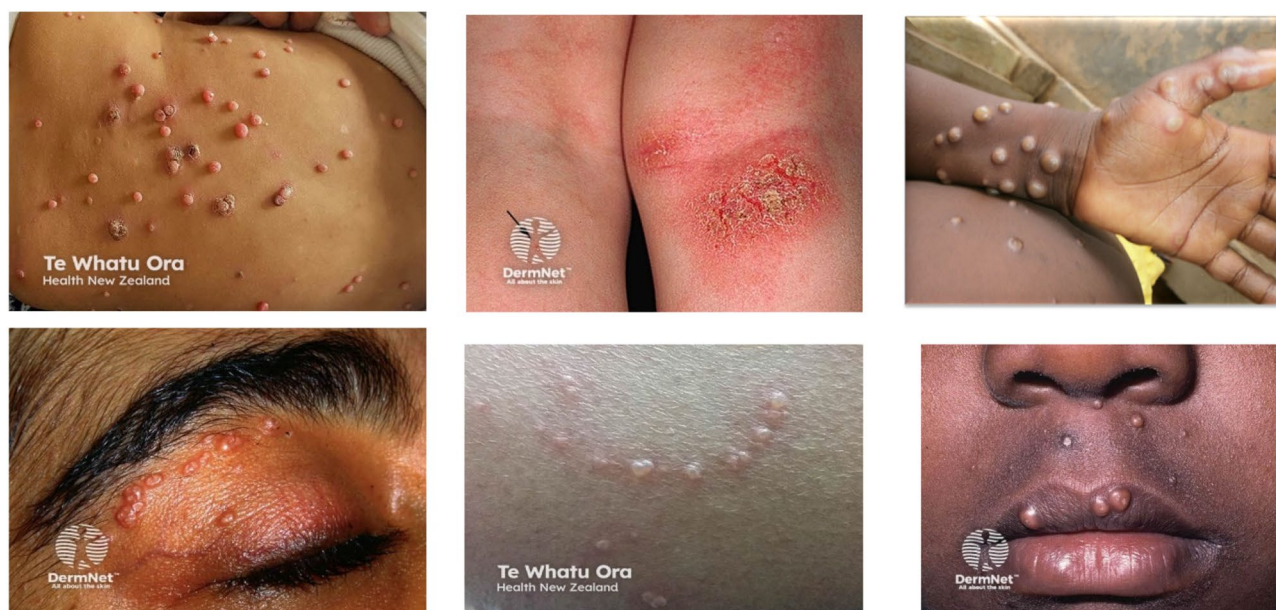


Fig. 1 The symptoms of skin infection are common in monkeypox disease [10]

miss, ultimately leading to a more accurate and comprehensive understanding of a patient's health status. Recent advances in machine learning (ML) have demonstrated the potential for precise disease prediction. Explainable Artificial Intelligence (XAI) is becoming increasingly crucial in the healthcare sector, particularly in the diagnosis and management of disease [15]. XAI plays a significant role in enhancing the interpretability of machine learning models, allowing healthcare professionals to understand the rationale behind predictions. This transparency is vital for establishing trust in AI-driven decisions, especially when using clinical data to evaluate disease progression and predict outcomes. Research has shown that XAI techniques can enhance the accuracy of disease predictions by providing clear insights into how clinical parameters affect the model's predictions, ultimately leading to more informed clinical decision-making and personalized treatment plans [16]. By enabling early detection and intervention, XAI has the potential to reduce mortality rates associated with diseases [17]. This study aims to fill this gap by developing a machine-learning model that predicts monkeypox based on clinical symptoms while leveraging Explainable AI (XAI) techniques to enhance trust and usability among healthcare professionals. This article adds to the current body of literature in the following ways.

- Develop a robust ensemble ML model that can accurately predict human monkeypox infection based on clinical symptoms.
- Enhancing the model's clarity by identifying the most influential clinical features through XAI.

- Assessing the minimal changes of the input data leads to the risk of monkeypox for model outcome through diverse counterfactual explanations.
- Offers a decision-support tool that can assist healthcare professionals in the early identification and management of monkeypox cases through artefacts.
- Assess how the explanations generated by Qlattice can enhance clinician decision-making in the diagnosis and treatment of monkeypox cases detection.

The rest of this paper is organized as follows: The first section is the related work, and the research materials and methods section discuss the methodology used in this study. The results section presents the findings of the research Comparison of Model Interpretation Methods. The paper concludes with the Conclusion section, summarizing the study's key findings and the last section demonstrates limitations and future recommendations for future researchers.

Related works

This part of the research involved analyzing studies that utilized artificial intelligence algorithms to predict the occurrence of monkeypox disease using data from skin imaging. In the study [18–21], researchers assess the effectiveness of machine and deep learning models in detecting monkeypox disease using publicly available datasets with monkeypox, chickenpox, measles, and normal/healthy patient classes. Along with data augmentation, preprocessing, and feature selection were utilized.

However, the study focuses on using AI to identify monkeypox based on images of the disease rather than its clinical symptoms. Image-based disease detection has its limitations and may not be effective in many real-life situations [22]. Furthermore, the study lacked explainability and interpretability, which made it challenging to grasp the rationale behind the decisions made. Therefore, utilizing clinical symptoms as the foundation for identification, and diagnostics can offer a more practical and effective solution. In [2], an artificial neural network (ANN) and an adaptive artificial bee colony algorithm were employed to enhance the early detection of monkeypox infections. The study involved 240 suspected cases, with the deep learning model attaining the highest accuracy at 75%, followed by the random forest model at 71.1%. However, the error rate remains significantly high, as models often struggle to generalize beyond the specific data on which they were trained. Moreover, the model cannot elucidate the decision-making process involved in detecting monkeypox infections. In [1], ML techniques were utilized to create models that forecast the spread and severity of monkeypox outbreaks in several countries (Portugal, Spain, the USA, and Canada) from June 3 to December 31, 2022. Various models were assessed, with ANN showing the best performance. In another study [23], a predictive tool was developed using five ML models to analyze monkeypox trends in the United States. This tool employed a dataset of 1,205 reported cases obtained from the CDC website to forecast future trends. The study revealed that the neural prophet model exhibited the best performance, accurately predicting cases with 95% accuracy based on historical data. However, the study did not address the model's explainability, and the dataset size was limited. In [5], a time series monkeypox dataset from five countries was analyzed using neural network models. Various architectures were tested with different numbers of hidden layer neurons, and the optimal structure was identified through K-fold cross-validation with early stopping. The ANN model had the highest R-value at nearly 99%, while the LSTM and GRU models both achieved around 98%. However, the study does not address the explainability of the model. Similarly [24], conducted a study on predicting the transmission rate of monkeypox using various ensemble ML techniques. The results indicated that stacking ensemble learning outperformed the other methods, achieving a root mean square error of 33.1075, a mean square error of 1096.1068, and a mean absolute error of 22.4214. However, the dataset for monkeypox was inadequate for training the model, so the COVID-19 dataset was used instead to train the machine-learning models. While the monkeypox dataset was utilized for testing. Additionally, the study fails to address the explainability of the model [22]. A gradient boosting method was created for diagnosing monkeypox

using Support Vector Machines (SVM) and Random Forest algorithms with data from the Kaggle dataset. XGBoost outperforms other models with a 1.0 accuracy and includes a total of 211 cases, while SHAP helps in analyzing and explaining its output. However, in ML models, small datasets can lead to overfitting, as there are insufficient examples for the model to learn generalizable patterns, resulting in a lack of real-time application to predict new, and unseen datasets. The literature often addresses the integration of AI tools into clinical workflows. Research shows that while ML models can improve diagnostic accuracy, their successful incorporation depends on addressing usability and interpretability issues to gain acceptance among healthcare professionals [25, 26]. A cutting-edge model for detecting monkeypox is introduced in this study. The model makes use of Residual Dilated Spatial Pyramid Integration (ResDSPI), Efficiency Channel Attention (ECA), and MobileNetV2 with progressive transfer learning. It performed exceptionally well in both multi-class as well as binary classification tasks, achieving an outstanding maximum accuracy of 99.34% [27]. However, this study lacks explainability and does not consider clinical symptoms to classify monkeypox disease.

Several studies have utilized various ML approaches to enhance accuracy. Due to the limitations of the previous researcher, this research focuses on expanding datasets, improving model interpretability, and developing accessible, practical ML tools for clinicians. This includes creating explainable AI techniques that clarify the reasoning behind model predictions, which is essential for building trust and promoting adoption in clinical settings [28]. XAI aids clinicians in identifying the most significant symptoms, such as fever, rash, and lymphadenopathy, for predicting monkeypox. This targeted approach can improve diagnostic accuracy and facilitate timely interventions during patient assessments [26]. The transparency provided by XAI fosters trust among healthcare professionals. When clinicians understand the reasoning behind a model's predictions, they are more likely to use AI tools for diagnostic support, which is a crucial step in integrating these technologies into clinical practice [29]. The proposed paper develops a robust ensemble ML model that can accurately predict human monkeypox infection based on clinical symptoms, enhancing the model's clarity by identifying the most influential clinical features that drive its predictions and understanding their contributions to the patient characteristics associated with monkeypox infection through XAI and assessing the minimal changes of the input data lead to the risk of monkeypox for model outcome through diverse counterfactual explanations. Furthermore, this research offers a decision-support tool that can assist healthcare professionals in the early identification and management of

monkeypox cases by developing artefacts and this study assesses how the explanations generated by Qlattice enhance clinician decision-making in the diagnosis and treatment of monkeypox cases detection.

Materials and methods

Materials

We developed using Jupyter Notebook with the Python programming language. Additionally, we utilized scikit-learn for machine learning tasks, Explainable AI libraries, and Matplotlib for data visualization. Our hardware specifications include an Intel(R) Core(TM) i7-8550U CPU @ 1.80 GHz, 2.00 GHz, and 12 GB of installed RAM, along with a 1 TB storage hard disk for handling datasets. Machine learning libraries are employed for data analysis and model development. These libraries also function as AI tools to enhance model interpretability. They play a vital role in data analysis and model development by offering tools and algorithms for constructing and evaluating models [30]. Moreover, these libraries facilitate the analysis and clarification of model predictions, enhancing understanding and credibility. Additionally, they aid in interpreting and elucidating model outcomes, promoting better insight and reliability [31].

Modeling approaches

In this study, an Ensemble ML model integrated with XAI techniques was utilized to detect monkeypox cases based on clinical symptoms. The approach is structured into data collection, preprocessing, model development, explainability (SHAP, LIME, Eli5, counterfactual and Qlattice)) and artifact development as illustrated in Fig. 2. The models' effectiveness is then assessed by validating them with performance metrics such as accuracy, precision, recall, and F1-score.

Data collection

The dataset utilized in our study is published on GitHub and published article by Ali Reza Farzi Pour titled "Detection of Monkeypox Cases Based on Symptoms [32]. This dataset encompasses clinical features and novel presentations of human monkeypox and the monkeypox dataset generally includes comprehensive clinical descriptions that provide essential information about the disease's characteristics in infected individuals. This typically encompasses symptoms like fever and the development of a rash. It includes only 211 identifiable cases remained in the dataset and 47 fields with a target variable indicating whether the patient has monkeypox or not. A comprehensive list of symptoms is utilized to

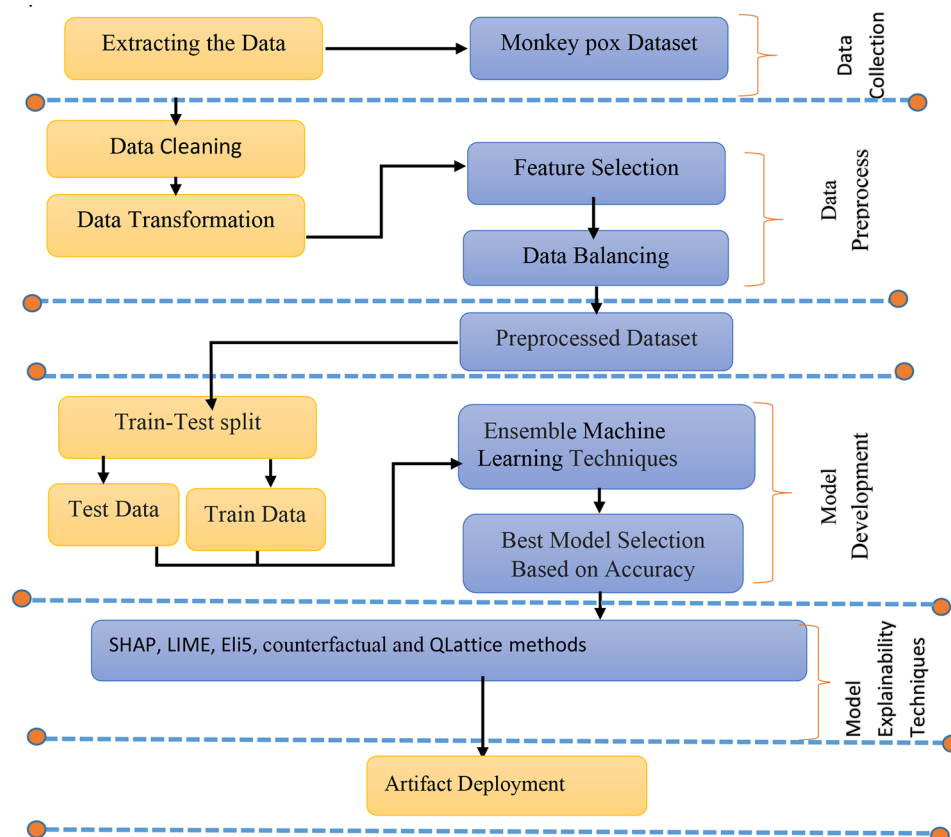


Fig. 2 The proposed methodology for the Monkeypox case detection

diagnose or identify monkeypox cases based on clinical presentation, including rash, skin lesions, headache, ulcerative lesions, oral and genital ulcers, fever, perianal papules, inguinal adenopathy, genital ulcer lesions, pustules, cough, blisters, erythema with vesicles and papules, difficulty breathing, fatigue, muscle pain, dysphagia, decreased physical strength, outbreak on the skin, hands, chest, chills, general weakness, general discomfort, adenomegaly, myalgia, itch, papules, swollen lymph nodes, sore throat, malaise, asthenia, diarrhea, Pain urinating, ulcers, loss of appetite, Vesicles, lymphadenopathy, myalgias, postules, encephalitis, blisters on limbs and genitals and monkeypox Status as a target variable. All the datasets are categorical and the result contains positive (1) and negative (0).

Data preprocessing

Data preprocessing is a crucial first step in converting raw data into valuable information. Raw data is frequently incomplete, redundant, or noisy. Data preprocessing can resolve these issues, allowing the data to be effectively used in creating ML models [33, 34]. The dataset underwent meticulous preprocessing to guarantee its high quality and suitability for training machine learning models. The dataset utilised in this study contains all cells or fields populated with valid data; no methods for handling missing values have been applied. Choosing only the necessary features is crucial for obtaining the best outcomes. Training the model on relevant data enhances the reliability of the predictions [35]. The features that maintain the same value across most of the dataset were removed; in this research, the quasi-constant feature adenomegaly was removed. In this research, we employed novel feature selection techniques from Featurewiz, including the Recursive XGBoost algorithm to select a reduced number of features in each iteration, and the SULOV method within the Featurewiz library that identifies pairs of variables that are highly correlated [36]. These methods utilize statistical analysis to assess

the strength and significance of these relationships and the result illustrated in Table 1. The performance of ML algorithms is negatively impacted by imbalanced datasets. Undersampling and oversampling are two techniques used to address class imbalance. While under-sampling can lead to the loss of crucial information and overfitting [37]. The application of synthetic data in health research can aid in developing and assessing new algorithms and methodologies by simulating various health scenarios, particularly for rare diseases or conditions with scarce real-world data. This can lead to more robust models capable of addressing a broad spectrum of health situations. Additionally, synthetic data can help mitigate bias in health datasets by including diverse demographic profiles, thereby improving the relevance of research findings and promoting equitable health interventions [38]. Synthetic Minority Oversampling Technique (SMOTE) and ADASYN are both effective techniques for handling imbalanced data in the current era [39]. SMOTE is widely used for creating synthetic samples to balance class distributions and improve classifier performance, but its effectiveness can vary in high-dimensional spaces [40]. ADASYN can elevate the risk of overfitting with small datasets by creating synthetic samples derived from existing minority class instances. This may lead to a model that learns noise rather than the underlying patterns in the data, resulting in poor generalization to unseen data [41]. This study implemented the SMOTE to balance the dataset and prevent improper model performance. Figure 3 shows the SMOTE algorithm's performance both before and after comparison. Correlation Analysis was also done to determine which symptoms are most strongly correlated with monkeypox infection. After implementing the SMOTE, the dataset now includes a total of 372 samples with 178 positive and 178 negative class. We utilized an 80/20 split for our dataset, meaning that 80% of the data was designated for training the model, while the remaining 20% was set aside for testing.

Table 1 Result of feature selection by using featurewiz

Featurewiz feature selection	'fever','skin lesions','headache','rash','muscle pain','itch','oraland-genitalulcers','adenomegaly','pustules','papules','fatigue','malaise','swollen-lymphnodes','Vesicles','myalgia','sorethroat'
------------------------------	---

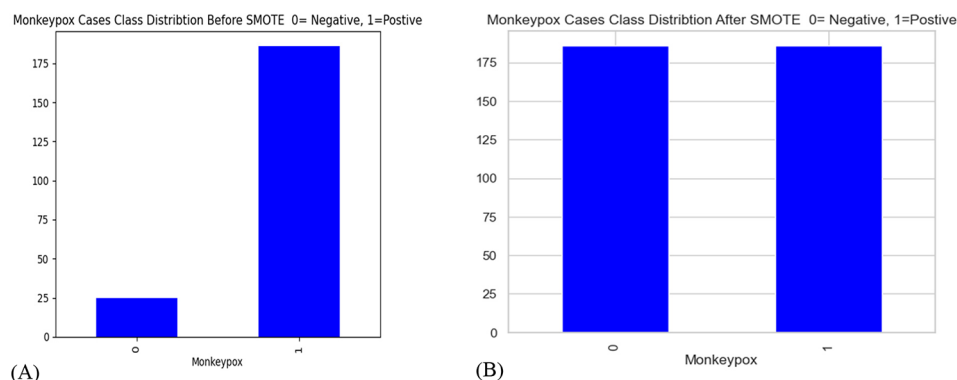


Fig. 3 (A) Unbalanced training dataset (B) Balanced training dataset using Borderline-SMOTE

Model development

This study explores Ensemble ML, which is highly effective in machine learning as it merges multiple models to enhance accuracy and perform reliably and effectively under various conditions [42]. Ensemble algorithms merge multiple individual models, such as decision trees (DT) or neural networks, to enhance overall performance and robustness.

Ensemble learning process

The training process entails selecting diverse base models and training them independently on the same dataset or different subsets of the data. Techniques like bagging and boosting are commonly employed, where models are trained on random samples of the data and their predictions are averaged, or models are trained sequentially to rectify errors. The predictions from these base models are combined using methods like voting or averaging to generate a final output, improving accuracy and minimizing overfitting. Ensemble machine learning algorithms leverage key principles such as diversity, balancing bias and variance, error correction, and aggregation to enhance predictive accuracy by combining multiple models. By integrating predictions from various models, ensemble techniques can improve accuracy by capturing distinct patterns in the data and lowering overall error rates. Strategies like error correction in boosting algorithms help rectify mistakes made by earlier models, while aggregation methods such as majority voting or averaging ensure more reliable and consistent predictions, minimizing the impact of noise and outliers. These principles collectively enable ensemble methods to outperform individual models across a range of machine-learning tasks [43]. By minimizing variance and bias, they produce more reliable predictions and superior performance on new data. Their adaptability enables them to be utilized with a range of base algorithms, and their capacity to adjust to different data traits and manage imbalanced datasets makes them a versatile option for a variety of applications.

Detailed techniques of machine learning algorithms

The Decision Tree, Support Vector Machine, Random Forest, Bagging, Gradient Boosting, CatBoost, XGBoost, and LGBM Classifier models were implemented. The model's performance is validated, and its generalizability is ensured through several steps, including splitting the dataset into training, and test sets via stratified sampling, evaluating performance metrics, tuning hyperparameters, and assessing the feature impact using XAI methods. In Fig. 4, we observe an outline of the ensemble learning process. In this process, the predictions from each classifier are combined to form a unified prediction. This combination enables the evaluation of ensemble models using various performance metrics such as accuracy, precision, recall, and F1-score.

In the following section, a comprehensive explanation of each ensemble learning technique employed in this study is presented.

Decision tree classifier

Decision tree is a successful classifier used in various domains, built using a recursive partition process. They are robust to parameter selection and perform well in imbalanced datasets. Ensembles of classifiers consisting of many members are combined for the final decision, which is generally better if the individual members are accurate and diverse [44]. But DT is prone to overfitting, especially when the tree is deep or when there are many feature and the model can be unstable when there is small change in the data [45].

Support vector machine

SVMs were developed in 1990 after being first presented in the 1960s. Compared to other machine learning algorithms, SVMs are implemented differently. Their ability to handle a large number of continuous and categorical variables has led to their recent rise in popularity. However, SVM can overfit the model and be computationally expensive [46].

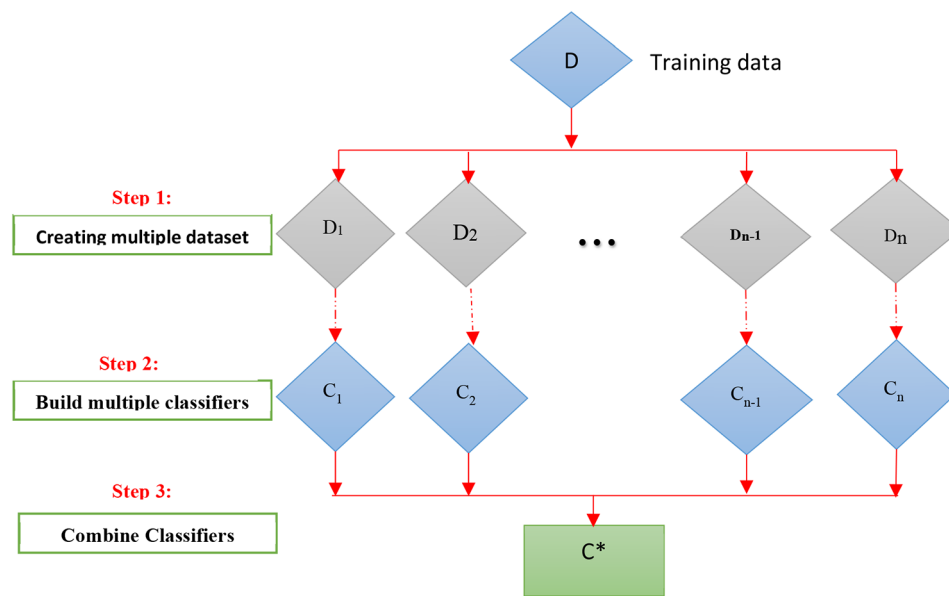


Fig. 4 Overview of the Ensemble Learning Procedure

Random forest classifier

Breiman introduced the Random Forest ensemble learning technique in 2001. This algorithm, based on classification and regression trees, is shaped by the number of decision trees and nodes [47]. Random Forest is a supervised machine learning algorithm that utilizes ensemble learning and consists of a collection of decision trees, with each tree trained on a randomly selected subset of the data [48]. Random forest can handle a wide variety of data including numerical and categorical data and also handle outlier and missing value. RF also reduce overfitting problem in decision tree and help to improve the accuracy. But RF can be computational expensive for large dataset and require a lot of memory which can be a constraint when working with limited resource [49].

Bagging classifier

Bagging allows multiple weak learners to work together to outperform a single strong learner, while also reducing variance and preventing overfitting in models. Bagging can result in a loss of model interpretability and introduce bias if appropriate procedures are not adhered to. Although it is highly accurate, the computational cost may restrict its application in certain situations [50].

Gradient boosting classifier

Gradient boosting is a highly effective predictive modelling technique that has proven successful in various tasks. The concept of boosting involves enhancing a weak learner to strengthen its performance. This method merges accurate prediction rules with less accurate ones to improve overall effectiveness. Gradient boosting algorithms necessitate the optimization of a loss function, a

weak learner for predictions, and an additive model for precise estimation. The initial success of boosting algorithms was observed with adaptive gradient boosting methods [51].

CatBoost classifier

In 2018, Dorogush et al. introduced CatBoost, an advanced GBDT toolkit akin to XGBoost [52]. CatBoost tackles challenges like gradient bias and prediction shift. It provides numerous benefits, such as automatically handling categorical features as numerical attributes, leveraging a combination of category features to improve feature dimensions, and employing a symmetrical tree model to minimize overfitting while boosting algorithm accuracy and generalizability [53].

XGBoost classifier

XGBoost, a boosting algorithm introduced by Chen et al. in 2016, is based on gradient-boosting decision trees and RF methods [54]. Compared to gradient-boosting decision trees, XGBoost offers improvements in multi-threaded processing, the classifier, and the optimization function. It also provides the following advantages: (1) the algorithm controls tree complexity to mitigate overfitting by incorporating a regularization term into the objective function. (2) A column sampling technique is employed to prevent overfitting, akin to the random forest algorithm. (3) The second-order Taylor expansion of the objective function is used to simplify and enhance the definition of the objective function when pursuing the optimal solution [53].

LGBM classifier

The high-performance gradient boosting algorithm, known as LGBM, is both fast and robust. It is employed for ranking and classification and is based on the decision tree algorithm. Unlike other algorithms that grow trees horizontally, LGBM grows trees vertically, specifically in a leaf-wise manner instead of level-wise [55]. Moreover, it incorporates techniques to prevent overfitting, such as early stopping and regularization, which have made it a preferred choice for data scientists and in competitive ML settings [56].

Model evaluation

The most crucial phase in the ML process is model evaluation, which involves assessing the model's performance through various metrics and techniques. This systematic overview of model assessment includes evaluating a trained model's effectiveness, comparing multiple models to identify the best performer, and verifying the model's ability to generalize to new data [57]. The researcher in this study utilized accuracy, precision, recall, F1-Score, and receiver operating characteristic (ROC) Curve for model evaluation.

Model explainability techniques

XAI is vital in healthcare for medical diagnoses, treatment plans, and predicting patient outcomes. It enhances understanding and trust between practitioners and patients, fostering more ethical AI applications [58, 59]. XAI illustrated how ML models, frequently criticized for their black-box nature, can be understood by revealing the specific reasoning behind their predictions. This has allowed medical professionals to gain a deeper understanding and effectively apply the findings of ML models for diagnosis [60]. Therefore, we employed several explainability techniques in our study, including QLattice, Eli5, counterfactual explanations, LIME (Local Interpretable Model-agnostic Explanations), and SHAP (SHapley Additive Explanations). To identify which symptoms or patient characteristics were most significant in diagnosing monkeypox, SHAP quantified the contribution of each feature to individual predictions, offering a consistent measure of feature importance. By highlighting the most relevant features for each case, LIME provided local approximations of the model's decision-making process

for specific instances, simplifying complex forecasts. Eli5 significantly streamlined model explanations by delivering metrics and visualizations that enhanced the understanding of model behaviour. Counterfactual explanations assisted clinicians in grasping the critical thresholds for specific symptoms by illustrating how minor adjustments in feature values could alter predictions.

Artifact deployment

Model deployment and monitoring are the last steps in the model creation process. In this section, we prepare the model for deployment as a web application featuring an interface that shows predicted probabilities and explanations for clinical use. developing and deploy a ML model on a local machine typically involves several steps [61]. The data flow after model deployment is illustrated in Fig. 5 below.

Results

Hyperparameter tuning

To select the best parameters for each algorithm, hyperparameter tuning using GridSearchCV was employed. The grid search algorithm can be utilized to identify the optimal hyperparameters in Sklearn [62] through "GridSearchCV." There is a reason grid search is regarded as the state of the art, even after years of scientific research into global optimization and various proposed hyperparameter optimization techniques. Its low dimensionality and reliability; its ease of implementation; it uncovers a significantly better alternative to manual sequential optimization [63]. The results of hyperparameter tuning using GridSearchCV are illustrated in Table 2.

Performance comparison

We assessed the predictive accuracy of ensemble machine learning models (Random Forest, Bagging, Gradient Boosting, CatBoost, XGBoost, and LGBM Classifier) for detecting MonkeyPox cases. Five key metrics accuracy, precision, recall, F1-Score, and ROC curve are utilized to assess the performance of each model as defined in the reference [64].

$$Accuracy = \frac{TN + TP}{TP + TN + FP + FN} \quad (1)$$

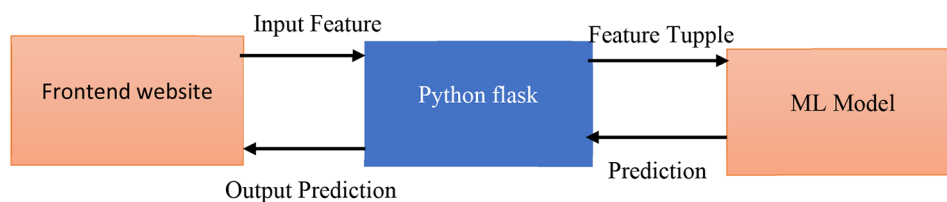


Fig. 5 Data flow after machine learning model deployment

Table 2 Result of GridSearchCV hyperparameter tuning

Model	Hyperparameter	Optimized Value
Decision Tree	max_depth	None
	min_samples_split	10
	min_samples_leaf	2
	max_features	sqrt
	Criterion	gini
Support Vector Machine (SVM)	C	50
	Kernel	rbf
	Gamma	auto
Random Forest	n_estimators	200
	max_depth	1
	min_samples_split	2
	max_features	auto
Bagging	Bootstrap	True
	max_features	sqrt
	n_estimators	50
	max_samples	1.0
	max_features	0.75
	Bootstrap	False
	bootstrap_features	False
	n_estimators	50
Gradient Boosting	learning_rate	0.1
	max_depth	5
	max_features	sqrt
	min_samples_leaf	1
	min_samples_split	5
	subsample	1.0
	learning_rate	0.5
	n_estimators	100
CatBoost	Depth	5
	learning_rate	0.5905175449124459
	Iterations	947
	l2_leaf_reg	7.628385045572297
	bagging_temperature	0.39025608827230857
	Subsample	0.9929766425198754
	max_depth	5
	Subsample	1
XGBoost	boosting_type	gbdt
	num_leaves	31
	learning_rate	1
	n_estimators	100
	max_depth	28
	learning_rate	0.05
	n_estimators	300
	Subsample	0.3
LGBM Classifier	colsample_bytree	0.7000000000000001
	min_child_sample	1
	n_jobs	1
	num_leaves	84
	objective	binary
	random_state	100

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$F1 - Score = 2x \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

Where (TP) is a true positive; (TN), is a true negative; (FP), is a false positive; and (FN), is a false negative. We commonly use these metrics in machine learning to evaluate the effectiveness of classification models. We calculated the values of these five metrics to compare the models' performance. For each model, we computed the values of these six metrics and compared them to identify which model demonstrated the highest performance. The optimal model for prediction accuracy is identified as the one exhibiting the highest performance. In addition to these metrics, we utilized various techniques to evaluate the models' performance. One such technique involves the use of confusion matrices, which offer a visual representation of the model's true positives, true negatives, false positives, and false negatives. Table 3 illustrate the Evaluation metrics of eight ML models with parameter tuning before Borderline-SMOTE analysis, Table 4 illustrate Evaluation metrics of eight ML models with Default parameters after SMOTE analysis and Table 5 presents the Evaluation metrics of eight ML models with parameter tuning after Borderline-SMOTE analysis. Here below are the table with result.

Before implementing Borderline-SMOTE with Parameter Tuning, all models demonstrated similar performance, with accuracy levels ranging from 86.0 to 88.4%. Random Forest, Bagging, Decision Tree, and Support Vector Machine attained the highest accuracy of 88.4%, while Gradient Boosting also reached 88.4%. In contrast, LGBMClassifier recorded the lowest accuracy of 78.5% after applying SMOTE with default parameters. After applying Borderline-SMOTE with Parameter Tuning, LGBMClassifier achieved the highest accuracy of 89.3%, whereas XGBoost had the lowest accuracy of 76.9%.

In the same train and test set split, as shown in Tables 3, 4 and 5, for our symptom-based monkeypox detection, LightGBM was selected as the top-performing model due to its higher accuracy compared to other machine learning models assessed during the study. The metrics for the Light Gradient Boosting Machine Classifier reach 89.3, surpassing the metrics of all other approaches. LightGBM emerged as the optimal choice due to its performance edge. Its superiority for our specific application was further confirmed by its efficiency in training and prediction. To achieve optimal accuracy with the LGBMClassifier, several key hyperparameters

Table 3 Evaluation metrics of eight ML models with parameter tuning before Borderline-SMOTE

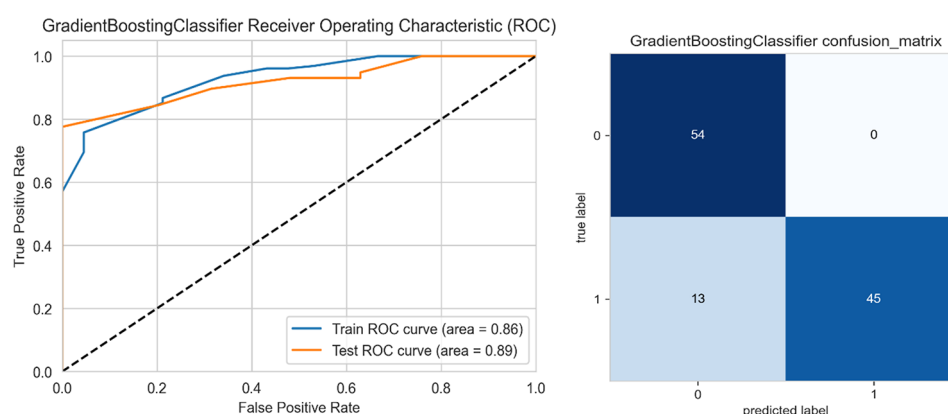
Metrics	Random Forest	Bagging	Gradient Boosting	CatBoost	XGBoost	LGBMClassifier	DT	SVM
Accuracy	88.4%	88.4%	86.0%	88.3%	88.3%	88.3%	88.4%	88.4%
Recall	50.0%	58.6%	97.3%	88.3%	97.3%	88.3%	94.7%	94.7%
Precision	88.3%	85.5%	88.8%	85.5%	90.2%	85.5%	92.3%	92.3%
F1_score	63.8%	69.5.1%	92.8%	86.9%	93.6%	86.9%	95.4%	87.6%
ROC	50.0%	72.2%	77.1%	75.0%	86.0%	75.0%	57.4%	63.4%

Table 4 Evaluation metrics of eight ML models with default parameters after SMOTE

Metrics	Random Forest	Bagging	Gradient Boosting	CatBoost	XGBoost	LGBMClassifier	DT	SVM
Accuracy	87.0%	87.5%	88.4%	87.5%	87.5%	78.5%	86.6%	79.4%
Recall	87.0%	87.9%	88.7%	87.5%	87.5%	78.5%	86.6%	79.4%
Precision	86.6%	89.7%	88.3%	90.0%	90.0%	82.9%	89.5%	79.5%
F1_score	86.8%	88.8%	88.4%	88.7%	88.7%	80.6%	88.0%	79.4%
ROC	87.0%	87.9%	89.0%	87.9%	87.9%	79.2%	87.1%	75.8%

Table 5 Evaluation metrics of eight ML models with parameter tuning after Borderline-SMOTE

Metrics	Random Forest	Bagging	Gradient Boosting	CatBoost	XGBoost	LGBMClassifier	DT	SVM
Accuracy	88.0%	88.4%	87.5%	85.1%	76.9%	89.3%	87.0%	86.6%
Recall	87.9%	88.7%	87.9%	85.1%	76.7%	89.2%	86.6%	87.0%
Precision	87.5%	88.3%	87.5%	85.7%	80.5%	91.2%	86.0%	86.6%
F1_score	87.7%	88.5%	87.7%	85.4%	78.6%	90.2%	86.3%	86.7%
ROC	87.9%	88.8%	88.0%	85.1%	76.5%	89.7%	90.1%	90.5%

**Fig. 6** Demonstrates the ROC curve of the Gradient Boosting model for monkeypox case detection with a Confusion matrix (The text describes a confusion matrix in which classes are designated as “0” for “no monkeypox disease” and “1” for “monkeypox disease.”)

were optimized during the model training process as illustrated in Table 2. To enhance our understanding of the models' effectiveness, we have included the confusion matrix for each method in Figs. 6 and 7. Among the eight machine learning approaches tested, Light Gradient Boosting exhibited superior performance and this research demonstrates that the Synthetic Minority Oversampling Technique (SMOTE) can enhance prediction performance on imbalanced datasets compared to analyses without SMOTE. By generating synthetic samples for the minority class, SMOTE effectively addresses the issue of class imbalance that often causes classifiers to favor the majority class in the original dataset. Studies have shown that SMOTE improves the model's

ability to enhance predictive performance and provides more reliable evaluation metrics across various classifiers. The boosting technique involves developing models that correct the errors of previous models, leading to improved accuracy with each iteration. LightGBM provides insights into feature importance, allowing users to focus on critical factors and improve model precision. By combining multiple weak learners, ensemble models combine multiple models to produce a unified prediction, often resulting in greater accuracy than individual models and tend to be more robust against overfitting compared to traditional models [43]. LightGBM creates a strong predictive model that enhances overall accuracy. Studies have shown that the LGBMClassifier

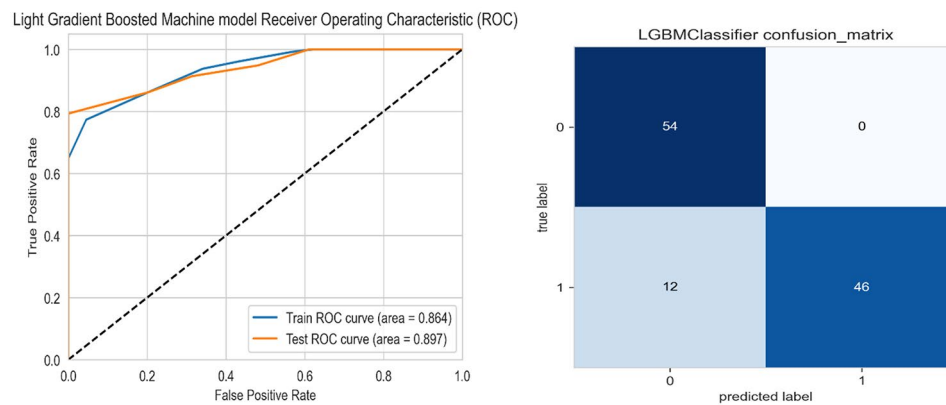


Fig. 7 Demonstrates the ROC curve of the Light Gradient Boosting model for monkeypox case detection with a Confusion matrix (The text describes a confusion matrix in which classes are designated as “0” for “no monkeypox disease” and “1” for “monkeypox disease”)

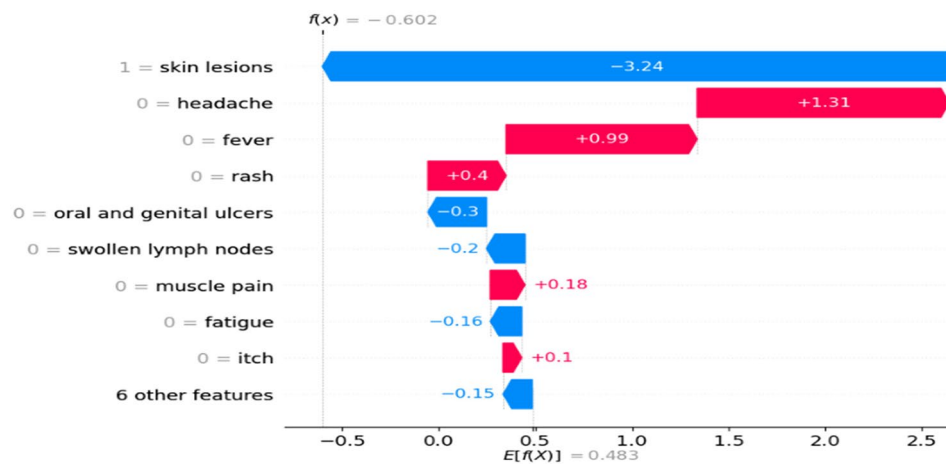


Fig. 8 Provides insights into the predictions made by the LightGBM (LGB) classifier model using Shap explanation

outperforms Random Forest, Bagging, Gradient Boosting, CatBoost, XGBoost, and Decision Trees (DT) on small datasets regarding training speed and scalability. This advantage stems from its leaf-wise tree growth and efficient histogram-based algorithms. Furthermore, the LGBMClassifier frequently exceeds SVM in both accuracy and speed, particularly on small datasets [65]. Also, LGBMClassifier computational efficiency and scalability render it ideal for real-time monitoring and public health interventions in low-resource settings. Its rapid training and inference, coupled with low memory consumption, have demonstrated an ability to effectively manage large-scale epidemiological data while preserving high accuracy and efficiency. Recent studies have underscored its effectiveness in predicting and responding to monkeypox outbreaks, establishing it as a viable option for public health interventions. The Light Gradient Boosting model is cost-effective and can be utilized in healthcare systems for monkeypox case detection along with explanation techniques like SHAP, LIME, Counterfactuals, and Eli5.

These techniques offer insights into the reasoning behind the model's predictions.

Explanation of ML models

To get an insight into features that are most important for the Light Gradient Boosting model prediction, the SHAP values plot of every monkeypox case feature was analyzed.

Waterfall plot Shap explanation for the LGBMClassifier model

The waterfall plot serves as a visual tool for comprehending ML models. It aids in grasping the importance of each feature in predicting the output for a particular instance. The initial prediction is displayed at the top of the plot, followed by the individual contributions of each feature. Each feature's impact is represented by a bar, with bars extending downwards for features that reduce the prediction and upwards for those that increase it [66]. In Fig. 8 $E[f(x)] = 0.483$ gives the average predicted number of monkeypox across all rows. $F(x) = -0.602$ is the predicted number of monkeypox for this particular record.



Fig. 9 a, b, c Represent SHAP force plot visualization represents the impact of each feature on the model's output, illustrating how each feature contributes to the prediction and aiding in understanding the importance of different features in the model

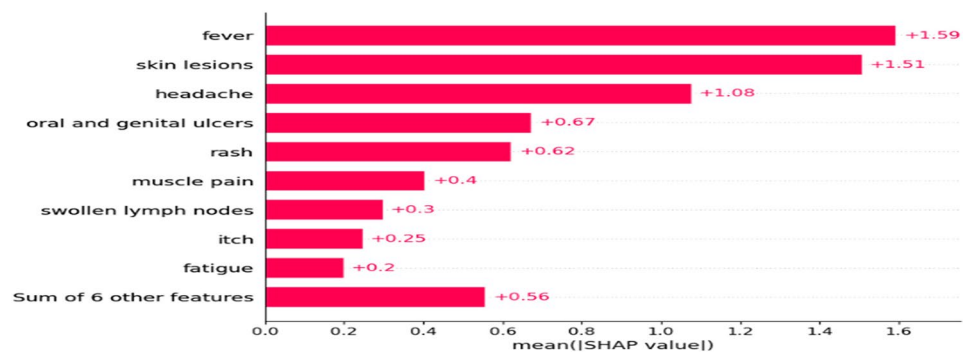


Fig. 10 Shape bar plot for Monkeypox cases Detection form clinical symptoms

The SHAP values are all the values in between. Ex: headache in the plot has increased the predicted number of monkeypox by 1.31. Features pushing the prediction to class 1 (Monkeypox positive) are shown in red, and those pushing the prediction to class 0 (Monkeypox negative)) are in blue.

Force plot

A force plot is another method for visualizing the impact of each attribute on the forecast for a certain observation. Figure 9 begins with the same base value of 0.4829 and examines how each attribute contributes to the final monkeypox case detection.

The plot above illustrates the impact of each feature on moving the model output from the baseline prediction (average predicted outcome across the testing dataset) to the actual model output (y_test value). Features that increase the prediction are highlighted in red, while those that decrease the prediction are highlighted in blue. In the bar plot, the features are arranged in order based

on their impact on the prediction. It considers the absolute SHAP values, so it makes no difference whether the feature influences the prediction positively or negatively. Figure 10 shows which clinical symptoms are most significant, by comparing the means SHAP values across all observations for each clinical feature. We can observe for example, that fever in the plot had the highest mean SHAP value.

In addition to the SHAP global explanation, local model explanation techniques with the help of ELI5 were performed to evaluate how the explainers provide interpretation to the LGBMClassifier model for monkeypox case detection. The ELI5 explanation enables researchers to comprehend black-box models and has capabilities for many platforms [67]. To make the LGBMClassifier model predictions more interpretable, the prediction of the model was presented as a sum of feature contributions and bias. Figure 10 shows how the monkeypox features lead to a particular prediction of the weights for each feature and their actual value. The weights depict

Table 6 Eli5 explains model predictions for Monkeypox case detection

Y=0 (probability 0.795, score -1.355) top feature	
Contribution	Feature
+2.178	Fever
+0.571	Oral and genital ulcers
+0.120	fatigue
+0.096	Sore throat
+0.085	rash
+0.056	malaise
+0.043	vesicles
+0.023	myalgia
-0.010	pustules
-0.117	headache
-0.138	itch
-0.296	papules
-0.308	Muscle pain
-0.483	BIAS
-0.616	Skin lesions

how influential each feature has been in contributing to the final prediction decision of the LGBMClassifier. As shown in Fig. 10, the top five influential features of the LGBMClassifier model prediction, are fever, oral and genital ulcers, swollen lymph nodes, fatigue, and sore throat. The ELI5 feature importance shown in Fig. 10 demonstrates that muscle pain and skin lesions have little contribution as compared to the other features for the LGBMClassifier prediction of the instance used in the simulation. The fever of the patient has the highest weight contributing the most to the prediction outcome of the model. We conclude that ELI5 is good at explaining the feature weight demonstrating the features contributing to the prediction outcome of the developed LGBMClassifier model monkeypox cases detection (Table 6).

we have also used the LIME explainable approaches to explain the predictive model. LIME is an approach that emphasizes the human aspect of connecting AI with humans. It specifically addresses model confidence and prediction confidence as its primary focus areas. LIME presents a unique explainable AI system that delivers explanations for predictions on a local scale [68]

As demonstrated in Fig. 11, the fever contributed much to the LGBMClassifier model predicting the given instance as monkeypox positive according to the explanation provided by LIME. Likewise, skin lesions and itch contribute to the model's prediction pushing the model to predict the instance of the monkeypox or positive class. In contrast, some of the features such as oral and genital ulcers, swollen lymph nodes, sore throat, fatigue, malaise, vesicles, and pustules pushed the model towards the negative or monkeypox-negative class.

Prediction involves mapping certain elements of the world (inputs) to other elements of the world (outputs) through the use of data. Another model explanation tool, Counterfactual prediction utilizes factual information to predict specific elements of the world that would have occurred in a different scenario [69]. The standard counterfactual approach aims to fulfill the objectives of XAI by providing counterfactuals. These counterfactuals provide insights into how the outcomes of a machine-learning model might have differed if modifications were made to the input variables [69]. Table 7 illustrates how the change in features affects the final result.

Qlattice employs a wide range of potential models to determine the most appropriate solution for a specific problem. Abzu employed QLattice to improve model transparency, questioning the core assumptions of black-box artificial intelligence. This simulation employs quantum-inspired assessments to enhance our understanding

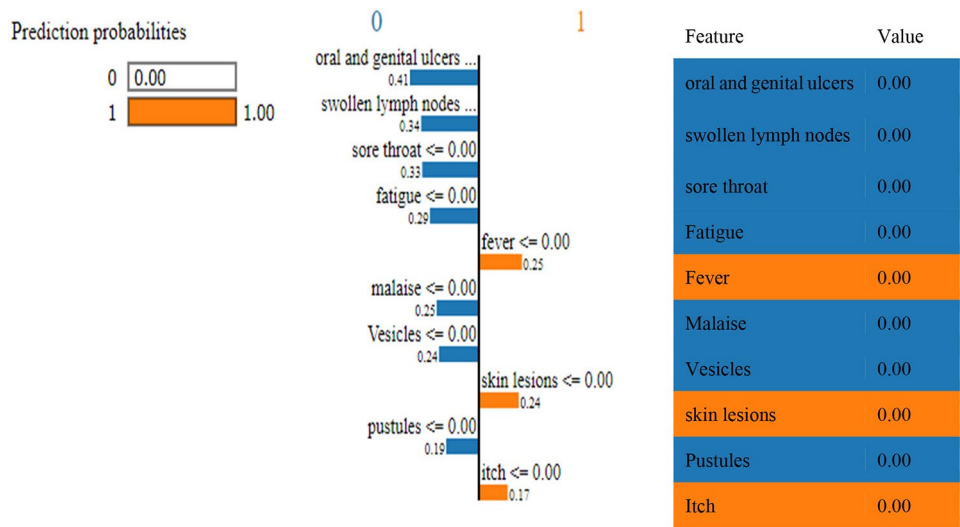


Fig. 11 Visualization of the local contribution of clinical features through the LIME model in classifying a single test instance using the LightGBM Model

Table 7 Counterfactual model explanation for Monkeypox cases detection

MonkeyPox	rash	skin lesions	Headache	oral and genital ulcers	Fever	pustules	fatigue	muscle pain	myalgia	itch	papules	swollen lymph nodes	sore throat	malaise	Vesicles
Original input (outcome: Negative)	No	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Counterfactuals (outcome: Positive)	Yes	-	-	-	Yes	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	Yes	-	-	-	-	-	-	-	-	-	Yes
	-	-	-	-	Yes	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	Yes	-	-	-	-	Yes	-	-	-	-	-
	-	-	-	-	Yes	-	-	-	-	Yes	-	-	-	-	Yes
	Yes	-	-	-	Yes	-	-	-	-	Yes	-	-	-	-	-
	Yes	-	No	-	-	-	-	-	-	-	-	-	-	Yes	Yes
	-	-	-	-	-	-	-	-	-	-	-	Yes	-	-	-
	-	-	-	-	Yes	Yes	-	-	-	-	-	-	-	-	-
	Yes	-	-	No	Yes	-	-	-	-	-	-	-	-	-	Yes

of data interactions. QGraphs are used for result interpretation, consisting of edges, registers, and activation functions. Each attribute is associated with a register, with connections formed through edges. Activation functions are utilized on the registers to generate meaningful insights [70, 71]. Users can adjust parameters such as input properties and variables. These parameters, known as registers, are utilized to construct a “QGraph” model, which consists of edges and nodes with weights and activation functions. Once training is complete, crucial attribute information is generated. QLattice is built using the “Feyn” library in Python. Figure 12 represents a QGraph. From the figure, it can be seen that the model considers fever, itch, headache, skin lesions, and oral and genital ulcers as the most important attributes. This model also uses the “multiply” and “addition” function to interpret results. To illustrate the relationships and importance of different clinical symptoms in monkeypox case detection, our study makes use of QGraph. The model’s emphasis on important symptoms such as fever, itching, headaches, skin lesions, and oral and vaginal ulcers is depicted in the diagram. Utilizing QLattice’s XAI tool to investigate how symptoms affect the model’s decision-making. Each attribute’s importance is determined by how well it helps the model forecast cases of monkeypox. The model’s behavior can be visualized to help physicians better comprehend its outcomes and identify important symptoms while evaluating patients. Health care professionals can use QGraph and QLattice to make well-informed decisions based on clear and intelligible model outputs, highlighting the significance of explainability in AI applications.

Comparison of model interpretation methods

The methodological difference between previous studies and this study is that this study investigates the performance of six ensemble ML methods for human monkeypox case detection using Featurewiz feature selection techniques, implementation of real-time monkeypox case detection using Ensemble ML model for clinicians to input patient data, and receive predictions output and compared model interpretation methods. Table 8 illustrates the methods used for interpreting model prediction for monkeypox case detection. The interpretation result of the three model explanation methods, namely, SHAP, LIME, and ELI5 is discussed in this section. The interpretation of the LIME, ELI5, and SHAP differs as the explanation of each method is different for similar test instances used in the experiment, as depicted in Table 8. The explanation generated by ELI5, LIME and Shap provides model interpretation showing that fever contributes to model prediction. However, the SHAP and LIME explanation interprets the model prediction output to be due to the skin lesions as the highest contributor or

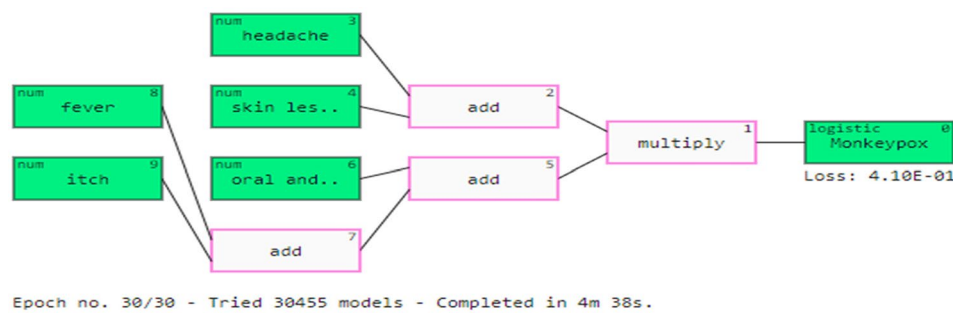


Fig. 12 the use of QGraph to explain model predictions for Monkeypox cases detection

Table 8 Comparison of model explanation methods

Feature Rank	ELI5	LIME	SHAP
1.	Fever	Fever	Fever
2.	oral and genital ulcers	skin lesions	skin lesions
3.	swollen lymph nodes	Itch	Headache
4.	Fatigue	oral and genital ulcers	oral and genital ulcers
5.	sore throat	swollen lymph nodes	Rash
6.	Rash	sore throat	Muscle pain
7.	Malaise	Fatigue	swollen lymph nodes
8.	Vesicles	Malaise	itch
9.	Myalgia	Vesicles	Fatigue
10.	Pustules	Pustules	others

reason for the model's prediction of the test instance as a monkeypox-positive case next to fever.

The differences in feature importance rankings among SHAP, LIME, and ELI5 underscore how these methods collaborate to provide physicians with a deeper understanding of monkeypox symptoms. Healthcare professionals can enhance diagnostic accuracy by leveraging these insights, especially in resource-limited settings where timely and accurate symptom-based screening is crucial. Alongside specialized training and flexible guidelines, incorporating these explainability techniques into clinical practices can boost their relevance and acceptance in real-world healthcare environments.

Model artifact

Real-time data ingestion and processing systems are essential for delivering data to analytics and machine learning platforms. These systems greatly influence the efficiency and effectiveness of data analysis and machine learning algorithms, enabling the end user to predict monkeypox cases to address the unique challenges posed by this dynamic data ingestion ecosystem. In this section, we used Flask API to design artefacts using the best-performed ML algorithm based on the important feature to

predict symptom-based monkeypox. To develop an artefact, we pickle the best-performed model and develop a flask API that connects the pickled model and the developed interfaces. As shown in Fig. 13 below the user inputs the relevant feature and the flask API give the result of the prediction based on the developed model. Early detection of monkeypox is essential for timely treatment and the prevention of further transmission. One method for detecting monkeypox is through symptom-based surveillance. This approach involves monitoring individuals who exhibit symptoms characteristic of monkeypox. It aids in the early identification and isolation of cases, thereby preventing the spread of the disease. By concentrating on the symptoms associated with monkeypox, healthcare providers can swiftly identify and diagnose cases, resulting in improved outcomes for both patients and the community.

Discussion

An explainable AI framework utilizing machine learning techniques proved effective in detecting monkeypox symptoms. The incorporation of explainable AI methods tackles the challenges of accurate detection and transparency in AI-driven healthcare tools. The LGBMClassifier, along with tools such as SHAP, LIME, ELI5, counterfactual, and QLattice, strikes a balance between predictive performance and interpretability. This approach offers clear insights into prediction features, aiding healthcare professionals in making confident diagnoses. The use of counterfactual explanations facilitates the exploration of alternative diagnoses, enhancing the tool's utility in clinical settings. Previous studies do not compare different model interpretation methods. This research utilized a dataset based on symptoms (211 cases, expanded to 372 using SMOTE) with clinical features such as fever, rash, and lymphadenopathy, prioritizing interpretability. However, current advancements are focused on image-based deep learning models (like CNNs and DenseNet) trained on skin lesion images, achieving higher accuracies of 97.63% [29]. These models typically depend on larger datasets but do not incorporate clinical symptoms. This study achieved an accuracy of 89.3% with

Fig. 13 Symptom-Based Detection of Monkeypox Interface

LGBMClassifier, which is lower than leading image-based models but competitive within symptom-based approaches. For instance [22], reported 100% accuracy on a similar small dataset but did not address class imbalance or real-world applicability. Our research employed various XAI techniques (SHAP, LIME, ELI5, and counterfactuals, Qlattice) to enhance transparency. We identified key symptoms (such as fever and swollen lymph nodes) that influence predictions, which is essential for clinical confidence. However, previous studies [18, 21, 71] lacked explainability. The incorporation of multiple XAI tools in this study is a unique contribution, and we also developed a Flask-based web interface for real-time symptom input and prediction, bridging AI research with clinical practice. Nonetheless, current research indicates that few studies implement practical tools; many remain theoretical or focus on retrospective analysis [23, 24]. Also Studies [1, 2, 5, 18, 21, 23], and [71] have evaluated ML/DL models for monkeypox detection; however, they often omitted explanations of the models. Although [22] improved accuracy using SHAP, it did not explore other interpretability tools such as LIME, ELI5, counterfactuals, or Qlattice, which offer actionable insights and visualize prediction criteria. This study addresses these gaps by comparing multiple XAI methods for monkeypox detection, emphasizing interpretability to enable faster and more precise diagnoses, timely interventions, and efficient resource allocation, ultimately enhancing public health outcomes during outbreaks. Table 9 shows a few new diagnostic techniques for monkeypox. The use of symptoms for diagnosis with the XAI tool is the primary difference between these approaches and the suggested strategy. The study identifies crucial clinical symptoms of monkeypox that can be utilized to create

a symptom-based diagnostic tool for healthcare professionals. The implementation of this tool in real-world settings may be obstructed by limited access to technology, variability in symptom presentation, and the necessity for training healthcare workers. Addressing these challenges may involve focused training programs, collaborations with local health organizations, and public awareness campaigns to diminish stigma and promote early reporting of symptoms. Meta-analyses aggregate data from multiple studies to quantitatively estimate effect sizes [72], while systematic reviews offer a structured approach to summarizing literature using predefined protocols [73]. By employing computational methods, machine learning can detect patterns in large datasets and reveal non-linear correlations. Merging the precision of systematic reviews and meta-analyses with the flexibility of machine learning enables a thorough understanding of research issues.

3. Conclusions

The rise of monkeypox as a significant public health issue highlights the need for fast and precise detection methods. Conventional diagnostic techniques often lack speed and efficiency, causing delays in response and potential outbreaks. The practical advantages of using XAI in the symptom-based detection of Monkeypox are substantial. Employing XAI offers numerous benefits, such as enhancing diagnostic accuracy, fostering trust in AI tools, and improving training for healthcare professionals. This transparency facilitates a deeper understanding of model predictions, helps identify key symptoms associated with Monkeypox, and leads to faster diagnoses, timely treatments, and improved patient outcomes.

Table 9 Comparative analysis of diagnostic methodologies used for MonkeyPox

Reference	Technique	Description	Limitation
[2]	ANN with the adaptive artificial bee colony algorithm	The deep learning model achieved the best result with an accuracy of 75% from clinical symptoms	The proposed model achieved an accuracy of 71%, yet it was outperformed by deep learning models, indicating that further efforts are required to enhance and refine the ANN.
[18]	A deep learning model was used, and a distinction was made between monkeypox and other diseases	Image-based dataset with A Siamese Deep Learning model and an accuracy of 91.09%	The model concentrates on analyzing skin lesions associated with monkeypox, potentially overlooking other clinical indicators and diminishing its effectiveness in real-world diagnostic environments. Although it demonstrates high accuracy, the resemblance of monkeypox lesions to those of other diseases may lead to misclassification.
[74]	Deep and machine learning.	By integrating multiple types of data, including clinical, laboratory, environmental, social media, demographic, and geospatial data ECNN prototype in training and testing. Extended Convolutional Technique attained an accuracy of 88.10%.	Many deep learning models function as “black boxes,” creating difficulties in understanding their decision-making processes. Hence, the concept of XAI emerges.
[29]	Deep Neural Networks (DNNs) for detecting Monkeypox disease with a dataset comprising skin images of three diseases: Monkeypox, Chickenpox, Measles, and Normal cases	DenseNet201-based architecture has the best performance, with Accuracy = 97.63%, F1-Score = 90.51%, and Area Under Curve (AUC) = 94.27	The dataset used for monkeypox diagnosis is clinically unapproved, limiting its representation of the full spectrum of symptoms and variants, and its resilience in clinical settings.
[1]	Machine learning techniques, namely ANN, Long Short-Term Memory, and Gated Recurrent Unit models	ANN models, particularly those with optimized Root Mean Squared Error, Mean Absolute Percentage Error, and the Coefficient of Determination values, are effective in infectious disease forecasting	The CNN model neglects clinical symptoms such as fever, lymphadenopathy, and muscle soreness, disregarding their significance for diagnosis. The variability of monkeypox lesions, particularly those resembling other diseases, heightens the risk of misdiagnosis.

Proposed Method: Ensemble ML techniques

Symptom-based dataset with an accuracy of 89.3%. Using LGBMClassifier and XAI (SHAP, LIME, eli5 counterfactual and Qlattice). Explainable AI makes symptom-based detection algorithms more transparent so that medical practitioners may comprehend its justification. This technique can improve public health interventions by more accurately and early detecting monkeypox cases. The model is a useful decision-support tool because of its practical use and resilience across a variety of datasets. By focusing on specific symptoms, XAI minimizes overlap among viral infections. Real-time feedback and explanations facilitate quicker clinical decision-making, facilitating earlier monkeypox detection and treatment, and improving patient outcomes

This study aims to create an explainable AI framework using machine learning algorithms for symptom-based monkeypox detection, to enhance diagnostic accuracy and enable timely interventions with a transparent and interpretable diagnostic tool that can be trusted by healthcare professionals. This research demonstrates that machine learning models can effectively analyze and identify symptom patterns linked to monkeypox. LGBMClassifier model shows high accuracy in distinguishing between monkeypox and similar conditions based on clinical symptoms with an accuracy of 89.3%. Incorporating explainable AI techniques enhances transparency, enabling healthcare professionals to understand AI-driven recommendations. By utilizing clinical patient information, the model improves its predictive abilities. A user-friendly interface ensures easy integration of this technology into healthcare providers’ workflows, bridging the gap between advanced AI tools and practical medical applications. This application could be used in healthcare settings to provide immediate risk assessments based on patient symptoms. By integrating the

application with existing health information systems and connecting it to a centralized database for continuous data gathering, the model can adapt to emerging trends. Implementing a feedback loop to learn from real-world outcomes and training healthcare professionals on how to use the tool effectively will be crucial for proactive disease surveillance and response. This study has significant implications for public health and disease management. The proposed AI framework can lead to quicker public health responses by enabling faster and more accurate monkeypox detection, potentially reducing transmission rates. Improved detection methods also allow for better allocation of healthcare resources, ensuring targeted interventions. Policymakers can use insights from this research to develop proactive strategies for managing outbreaks and enhancing epidemic preparedness. Using XAI for symptom-based detection of Monkeypox presents practical advantages. XAI improves diagnostic accuracy by offering clear explanations of model predictions, aiding healthcare professionals in understanding the underlying data and pinpointing key symptoms.

This transparency builds trust in AI-driven tools and facilitates more effective training for healthcare workers. Ultimately, XAI can result in earlier diagnoses, timely interventions, and enhanced patient outcomes during Monkeypox outbreaks. The clinical relevance of selected symptoms for Monkeypox matched well with the Centers for Disease Control and Prevention's (CDC's) Monkeypox Case Definition. According to the CDC's Case Definition for Monkeypox, which includes a distinctive rash or lesion, the selected symptoms meet the criteria: lymphadenopathy, fever, and other systemic symptoms, along with epidemiological connections to endemic regions or confirmed cases. The WHO also emphasizes the significance of rash development and lymphadenopathy in differentiating monkeypox from related illnesses such as varicella (chickenpox).

Research limitations and recommendations

The study acknowledges certain limitations, such as the model's performance being contingent on the quality and diversity of input data due to the study's dataset of 211 cases poses challenges for the model's generalizability due to overfitting and the difficulty of capturing various patterns. Despite efforts to address class imbalances, the small sample size may result in subpar outcomes with new data. We acknowledge this limitation and advocate for larger datasets to enhance reliability and real-world deployment. Collaboration, public datasets, and data augmentation can expand the dataset and improve the accuracy of conclusions. Future research should prioritize increasing the dataset size to enhance accuracy and pattern capture. Symptom-based detection may overlook asymptomatic cases or variations in symptoms, potentially resulting in missed diagnoses. Another potential limitation of this study is that the correlation analysis used to measure the relationship between the synthetic and original data produced by SMOTE required more research. Additionally, this study sets the stage for future research on AI applications in diagnosing other infectious diseases, promoting a transformative approach to utilizing technology in public health initiatives. Future studies should concentrate on integrating genomic data with environmental factors to enhance the predictive power of models. Longitudinal studies can evaluate performance over time and across various outbreak scenarios. Further studies are also necessary to evaluate the performance of the LGBMClassifier on various geographic and genomic datasets to confirm its effectiveness across different regions and healthcare settings. This will help establish its reliability and suitability for real-time monitoring of monkeypox in resource-limited areas, thereby enhancing public health intervention strategies. Furthermore, incorporating additional demographic and geographic characteristics could further enhance the

model's generalizability. Future research may also explore the use of real-time data sources such as social media or medical reports to develop a more dynamic detection system. It will be crucial to address ethical considerations when implementing AI systems in public health to ensure responsible use and minimize biases in symptom interpretation and disease prediction.

Abbreviations

AI	Artificial intelligence
ANN	Artificial Neural Network
ML	Machine learning
MPXV	Monkeypox virus
RF	Random Forest
ROC	Receiver Operating Characteristic curve
SMOTE	Borderline-Synthetic Minority Oversampling Technique
WHO	World Health Organization
XAI	Explainable artificial intelligence

Acknowledgements

We would like to extend our heartfelt gratitude to Ali Reza Farzi Pour for generously sharing the invaluable dataset titled "Detection of Monkeypox Cases Based on Symptoms" on GitHub.

Author contributions

Conceptualization: Gizachew, methodology: Gizachew, software: Belayneh; validation: Belayneh; formal analysis: Gizachew; investigation: Gizachew; resources: Belayneh; data curation: Gizachew; writing original draft preparation: Gizachew and Belayneh; writing review and editing: Gizachew and Belayneh; visualization: Gizachew; supervision: Belayneh; project administration: Gizachew and Belayneh; funding acquisition: Gizachew. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Data availability

<https://github.com/alirezafarzipour/MonkeyPoxDetection>

Declarations

Ethics approval and consent to participate

Not applicable.

Informed Consent

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Computer Science, Debark University, Debark, Ethiopian, Ethiopia

²Department of Information Science, University of Gondar, Gondar, Ethiopian, Ethiopia

Received: 23 October 2024 / Accepted: 3 March 2025

Published online: 26 March 2025

References

1. Alnaji L. Machine learning in epidemiology: neural networks forecasting of Monkeypox cases. *PLoS ONE*. 2024;19(5):0300216. <https://doi.org/10.1371/journal.pone.0300216>.

2. Muhammed Kalo Hamdan A, Ekmekci D. Prediction of Monkeypox infection from clinical symptoms with adaptive artificial bee colony-based artificial neural network. *Neural Comput Appl*. 2024;1–16. <https://doi.org/10.1007/s00521-024-09782-z>.
3. Bleichrodt A, Dahal S, Maloney K, Casanova L, Luo R, Chowell G. Real-time forecasting the trajectory of Monkeypox outbreaks at the National and global levels, July–October 2022. *BMC Med*. 2023;21(1):19. <https://doi.org/10.1186/s12916-022-02725-2>.
4. Patel A, Bilinska J, Tam JC, Fontoura DDS, Mason CY, Daunt A, Snell LB, Murphy J, Potter J, Tuudah C, Sundramoorthi R. 2022. Clinical features and novel presentations of human monkeypox in a central London centre during the 2022 outbreak: descriptive case series. *bmj*. 378.<https://doi.org/10.1136/bmj-2022-072410>
5. Manohar B, Das R. Artificial neural networks for the prediction of Monkeypox outbreak. *Trop Med Infect Disease*. 2022;7(12). <https://doi.org/10.3390/tropic120424>.
6. Devarajan D, Dhana lakshmi P, Krishnaveni S, Senthilkumar S. Human Monkeypox disease prediction using novel modified restricted Boltzmann machine-based equilibrium optimizer. *Sci Rep*. 2024;14(1):17612. <https://doi.org/10.1038/s41598-024-68836-3>.
7. Markewitz NF, DeLuca J, Kalejaiye A, Shidid S, Jain T, Parikh P, Norton BL. Mpox-Associated pneumonia: A case report. *Annals Intern Medicine: Clin Cases*. 2023;2(3):e220945. <https://doi.org/10.7326/aimcc.2022.0945>.
8. WHO. 2024. WHO Director-General declares mpox outbreak a public health emergency of international concern. [Online] Available at: <https://www.who.int/news/item/14-08-2024-who-director-general-declares-mpox-outbreak-a-public-health-emergency-of-international-concern> [Accessed 25 Aug. 2024].
9. Africa CDC. 2024. Africa CDC declares mpox a Public Health Emergency of Continental Security, mobilizing resources across the continent. Published online August 13, 2024.
10. DermNet NZ. 2024. Molluscum contagiosum images. [Online] Available at: <https://dermnetnz.org/images/molluscum-contagiosum-images> [Accessed 25 Aug. 2024].
11. Hussain A, Kaler J, Lau G, Maxwell T. Clinical conundrums: differentiating Monkeypox from similarly presenting infections. *Cureus*. 2022;14(10). <https://doi.org/10.7759/cureus.29929>.
12. Pattnaik H, Surani S, Goyal L, Kashyap R. Making sense of Monkeypox: a comparison of other poxviruses to the Monkeypox. *Cureus*. 2023;15(4). <https://doi.org/10.7759/cureus.38083>.
13. Lalmuanawma S, Hussain J, Chhakchhuak L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. *Chaos Solitons Fractals*. 2020;139:110059. <https://doi.org/10.1016/j.chaos.2020.110059>.
14. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. *J Global Health*. 2019;9(2):020318. <https://doi.org/10.7189/jogh.09.020318>.
15. Nilofer A, Sasikala S. A comparative study of machine learning algorithms using explainable artificial intelligence system for predicting liver disease. *Comput Open*. 2023;1:2350003. <https://doi.org/10.1142/S2972370123500034>.
16. Akpinar R, Panzeri D, De Carlo C, Belsito V, Durante B, Chirico G, Lombardi R, Fracanzani AL, Maggioni M, Arcari I, Roncalli M. Role of artificial intelligence in staging and assessing of treatment response in MASH patients. *Front Med*. 2024;11. <https://doi.org/10.3389/fmed.2024.1480866>.
17. Decharatanachart P, Chaiteerakij R, Tiyyaratnatchai T, Treeprasertsuk S. Application of artificial intelligence in chronic liver diseases: a systematic review and meta-analysis. *BMC Gastroenterol*. 2021;21:1–16. <https://doi.org/10.1186/s12876-020-01585-5>.
18. Alakus TB. Prediction of Monkeypox on the skin lesion with the Siamese deep learning model. *Balkan J Electr Comput Eng*. 11(3), pp.225–31<https://doi.org/10.17694/bajece.1255798>
19. Nayak T, Chadaga K, Sampathila N, Mayrose H, Muralidhar Bairy G, Prabhu S, Katta SS, Umakanth S. Detection of Monkeypox from skin lesion images using deep learning networks and explainable artificial intelligence. *Appl Math Sci Eng*. 2023;31(1):2225698. <https://doi.org/10.1080/27690911.2023.225698>.
20. Uysal F. Detection of Monkeypox disease from human skin images with a hybrid deep learning model. *Diagnostics*. 2023;13(10):1772. <https://doi.org/10.3390/diagnostics13101772>.
21. Akram A, Jamjoom AA, Innab N, Almujally NA, Umer M, Alsabai S, Fimiani G. SkinMarkNet: an automated approach for prediction of MonkeyPox using image data augmentation with deep ensemble learning models. *Multimedia Tools Appl*. 2024;1–17. <https://doi.org/10.1007/s11042-024-19862-w>.
22. Farzipour A, Elmi R, Nasiri H. Detection of Monkeypox cases based on symptoms using XGBoost and Shapley additive explanations methods. *Diagnostics*. 2023;13(14):2391. <https://doi.org/10.3390/diagnostics13142391>.
23. Long B, Tan F, Newman M. Forecasting the Monkeypox outbreak using ARIMA, prophet, neuralprophet, and LSTM models in the United States. *Forecasting*. 2023;5(1):127–37. <https://doi.org/10.3390/forecast5010005>.
24. Dada EG, Oyewola DO, Joseph SB, Emebo O, Oluwagbemi OO. Ensemble machine learning for Monkeypox transmission time series forecasting. *Appl Sci*. 2022;12(23):12128. <https://doi.org/10.3390/app122312128>.
25. Thieme AH, Zheng Y, Machiraju G, Sadee C, Mittermaier M, Gertler M, Salinas JL, Srinivasan K, Gyawali P, Carrillo-Perez F, Capodici A. A deep-learning algorithm to classify skin lesions from Mpox virus infection. *Nat Med*. 2023;29(3):738–47. <https://doi.org/10.1038/s41591-023-02225-7>.
26. Uzun Ozsahin D, Mustapha MT, Uzun B, Duwa B, Ozsahin I. Computer-aided detection and classification of Monkeypox and chickenpox lesion in human subjects using deep learning framework. *Diagnostics*. 2023;13(2):292. <https://doi.org/10.3390/diagnostics13020292>.
27. Al-Gaashani MS, Xu W, Obsie EY. MobileNetV2-based deep learning architecture with progressive transfer learning for accurate Monkeypox detection. *Appl Soft Comput*. 2025;169:112553. <https://doi.org/10.1016/j.asoc.2024.112553>.
28. Eliwa EHI, El Koshiry AM, Abd El-Hafeez T, Farghaly HM. Utilizing convolutional neural networks to classify Monkeypox skin lesions. *Sci Rep*. 2023;13(1):14495. <https://doi.org/10.1038/s41598-023-41545-z>.
29. Sorayaie Azar A, Naemi A, Babaei Rikan S, Bagherzadeh Mohasefi J, Pirnejad H, Wiil UK. Monkeypox detection using deep neural networks. *BMC Infect Dis*. 2023;23(1):438. <https://doi.org/10.1186/s12879-023-08408-4>.
30. Nguyen G, Dlugolinsky S, Bobák M, Tran V, López García Á, Heredia I, Malik P, Hluchý L. Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. *Artif Intell Rev*. 2019;52:77–124. <https://doi.org/10.1007/s10462-018-09679-z>.
31. Lu SC, Swisher CL, Chung C, Jaffray D, Sidey-Gibbons C. On the importance of interpretable machine learning predictions to inform clinical decision making in oncology. *Front Oncol*. 2023;13:1129380. <https://doi.org/10.3389/fonc.2023.1129380>.
32. Farzipour A. 2025. GitHub - alirezafarzipour/MonkeyPoxDetection. [Online] Available at: <https://github.com/alirezafarzipour/MonkeyPoxDetection> [Accessed 16 Jan. 2025].
33. Shiwangi K, Sandhu JK, Sahu R. 2023, August. Effective Heart-Disease Prediction by Using Hybrid Machine Learning Technique. In *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)* (pp. 1670–1675). IEEE. <https://doi.org/10.1109/ICCPCT58313.2023.10245785>
34. Setegn GM, Dejene BE. Explainable artificial intelligence models for predicting pregnancy termination among reproductive-aged women in six East African countries: machine learning approach. *BMC Pregnancy Childbirth*. 2024;24(1). <https://doi.org/10.1186/s12884-024-06773-9>.
35. Chadaga K, Prabhu S, Bhat V, Sampathila N, Umakanth S, Chadaga R, Kumar S, G.S. and, Swathi KS. An explainable multi-class decision support framework to predict COVID-19 prognosis utilizing biomarkers. *Cogent Eng*. 2023;10(2):2272361. <https://doi.org/10.1080/23311916.2023.2272361>.
36. featurewiz. 2025. *featurewiz* · PyPI. [online] Available at: <https://pypi.org/project/featurewiz/0.2.4/> [Accessed 17 January 2025].
37. Mooijman P, Catal C, Tekinerdogan B, Lommen A, Blokland M. The effects of data balancing approaches: A case study. *Appl Soft Comput*. 2023;132:109853. <https://doi.org/10.1016/j.asoc.2022.109853>.
38. Gonzales A, Guruswamy G, Smith SR. Synthetic data in health care: A narrative review. *PLOS Digit Health*. 2023;2(1). <https://doi.org/10.1371/journal.pdig.0000082>.
39. Gnip P, Vokorokos L, Drotár P. Selective oversampling approach for strongly imbalanced data. *PeerJ Comput Sci*. 2021;7:e604. <https://doi.org/10.7717/peerj-cs.604>.
40. Blagus R, Lusa L. 2013. SMOTE for high-dimensional class-imbalanced data. *BMC bioinformatics*, 14, pp.1–16. <https://doi.org/10.1186/1471-2105-14-106>
41. Khan TM, Xu S, Khan ZG, Uzair Chishti M. Implementing multilabeling, ADASYN, and relief techniques for classification of breast Cancer diagnostic through machine learning: efficient Computer-Aided diagnostic system. *J Healthc Eng*. 2021;2021(1):5577636. <https://doi.org/10.1155/2021/5577636>.
42. Rafid ARH, Azam S, Montaha S, Karim A, Fahim KU, Hasan MZ. An effective ensemble machine learning approach to classify breast cancer based on

- feature selection and lesion segmentation using preprocessed mammo-grams. *Biology*. 2022;11(11):1654. <https://doi.org/10.3390/biology11111654>.
43. Mahajan P, Uddin S, Hajati F, Moni MA. 2023, June. Ensemble learning for disease prediction: A review. In *Healthcare* (Vol. 11, No. 12, p. 1808). MDPI. <https://doi.org/10.3390/healthcare11121808>
 44. Ahmad A, Safi O, Malebary S, Alesawi S, Alkayal E. Decision tree ensembles to predict coronavirus disease 2019 infection: a comparative study. *Complexity*. 2021;2021(1):5550344. <https://doi.org/10.1155/2021/5550344>.
 45. AIML.com. 2024. Advantages and disadvantages of Decision Tree. [online] Available at: <https://aiml.com/what-are-the-advantages-and-disadvantages-of-using-a-decision-tree/> [Accessed 13 November 2024].
 46. Anguita D, Ghio A, Greco N, Oneto L, Ridella S. 2010, July. Model selection for support vector machines: Advantages and disadvantages of the machine learning theory. In *The 2010 international joint conference on neural networks (IJCNN)* (pp. 1–8). IEEE. <https://doi.org/10.1109/IJCNN.2010.5596450>
 47. Cutler A, Cutler DR, Stevens JR. Random forests. *Encyclopedia of machine learning*. Springer; 2011. <https://doi.org/10.1007/978-1-4419-9326-7>.
 48. AlmalBetter. 2024. Random Forest Algorithm in Machine Learning. [online] Available at: <https://www.almabetter.com/bytes/tutorials/data-science/random-forest> [Accessed 13 November 2024].
 49. AIML.com. 2024. Advantages and disadvantages of Random Forest. [online] Available at: <https://aiml.com/what-are-the-advantages-and-disadvantages-of-random-forest/> [Accessed 13 November 2024].
 50. Corporate Finance, Institute. 2024. Bagging (Bootstrap Aggregation) - Definition, How It Works. [online] Available at: <https://corporatefinanceinstitute.com/resources/data-science/bagging-bootstrap-aggregation/> [Accessed 13 November 2024].
 51. ScienceDirect. 2024. Gradient Boosting - an overview. [online] Available at: <https://www.sciencedirect.com/topics/computer-science/gradient-boosting> [Accessed 13 November 2024].
 52. Dorogush AV, Ershov V, Gulin A. CatBoost: gradient boosting with categorical features support. *ArXiv Preprint arXiv:1810.11363*. 2018. <https://doi.org/10.48550/arXiv.1810.11363>.
 53. Luo M, Wang Y, Xie Y, Zhou L, Qiao J, Qiu S, Sun Y. Combination of feature selection and catboost for prediction: the first application to the Estimation of aboveground biomass. *Forests*. 2021;12(2):216. <https://doi.org/10.3390/f12020216>.
 54. Chen T, Guestrin C. 2016, August. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785–794). <https://doi.org/10.1145/2939672.2939785>
 55. Keshri R. Prediction of employee turnover using light GBM algorithm. *Int J Innovative Sci Res Technol*. 2020;5(4):947–52. Available at: [Accessed 13 November 2024].
 56. Wang Y, Liu Y, Zhao J, Zhang Q. Low-Complexity fast CU classification decision method based on LGBM classifier. *Electronics*. 2023;12(11):2488. <https://doi.org/10.3390/electronics12112488>.
 57. Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. *ArXiv Preprint arXiv:1811.12808*. 2018. <https://doi.org/10.48550/arXiv.1811.12808>.
 58. Najim AH, Nasri N. Artificial intelligence for heart disease prediction and imputation of missing data in cardiovascular datasets. *Cogent Eng*. 2024;11(1):2325635. <https://doi.org/10.1080/23311916.2024.2325635>.
 59. Hulsén T. Explainable artificial intelligence (XAI): concepts and challenges in healthcare. *AI*. 2023;4(3):652–66. <https://doi.org/10.3390/ai4030034>.
 60. Islam MS, Hussain I, Rahman MM, Park SJ, Hossain MA. Explainable artificial intelligence model for stroke prediction using EEG signal. *Sensors*. 2022;22(24):9859. <https://doi.org/10.3390/s22249859>.
 61. Kale V. Cloud computing basics. *Creative Smart Enterprises*. 2017;June:141–71. <https://doi.org/10.1201/9781315152455-6>.
 62. Barupal DK, Fiehn O. Generating the blood exposome database using a comprehensive text mining and database fusion approach. *Environ Health Perspect*. 2019;127(9):097008. <https://doi.org/10.1289/EHP471>.
 63. Belete DM, Huchaiah MD. Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. *Int J Comput Appl*. 2022;44(9):875–86. <https://doi.org/10.1080/1206212X.2021.1974663>.
 64. Rainio O, Teuvo J, Klén R. Evaluation metrics and statistical tests for machine learning. *Sci Rep*. 2024;14(1):6086. <https://doi.org/10.1038/s41598-024-56706-x>.
 65. Machado MR, Karray S, De Sousa IT. 2019, August. LightGBM: An effective decision tree gradient boosting method to predict customer loyalty in the finance industry. In *2019 14th International Conference on Computer Science & Education (ICCSE)* (pp. 1111–1116). IEEE. <https://doi.org/10.1109/ICCSE.2019.8845529>
 66. Access O. Explainable AI in heart disease prediction. *Complexity*. 2024;04:11262–70. <https://doi.org/10.56726/IJMET554626>.
 67. Chadaga K, Sampathila N, Prabhu S, Chadaga R. Multiple explainable approaches to predict the risk of stroke using artificial intelligence. *Information*. 2023;14(8):435. <https://doi.org/10.3390/info14080435>.
 68. An J, Zhang Y, Joe I. Specific-Input LIME explanations for tabular data based on deep learning models. *Appl Sci*. 2023;13(15):8782. <https://doi.org/10.3390/app13158782>.
 69. Baron S. Explainable AI and causal understanding: counterfactual approaches considered. *Mind Mach*. 2023;33(2):347–77. <https://doi.org/10.1007/s11023-023-09637-x>.
 70. Chadaga K, Prabhu S, Bhat V, Sampathila N, Umakanth S. COVID-19 diagnosis using clinical markers and multiple explainable artificial intelligence approaches: a case study from Ecuador. *SLAS Technol*. 2023;28(6):393–410. <https://doi.org/10.1016/j.slast.2023.09.001>.
 71. Chadaga K, Prabhu S, Bhat V, Sampathila N, Umakanth S, Chadaga R. A decision support system for diagnosis of COVID-19 from non-COVID-19 influenza-like illness using explainable artificial intelligence. *Bioengineering*. 2023;10(4):439. <https://doi.org/10.3390/bioengineering10040439>.
 72. Li XM, Wan J, Xu CF, Zhang Y, Fang L, Shi ZJ, Li K. Misoprostol in labour induction of term pregnancy: a meta-analysis. *Chin Med J*. 2004;117(03):449–52. <https://pubmed.ncbi.nlm.nih.gov/15043790>.
 73. Shi Z, Luo K, Deol S, Tan S. A systematic review of noninflammatory cerebrospinal fluid biomarkers for clinical outcome in neonates with perinatal hypoxic brain injury that could be biologically significant. *J Neurosci Res*. 2022;100(12):2154–73. <https://doi.org/10.1002/jnr.24801>.
 74. Patel A, Srinivasulu K, Jani A, K. and, Sreenivasulu G. Enhancing Monkeypox detection through data analytics: A comparative study of machine and deep learning techniques. *Adv Eng Intell Syst*. 2023;2(04):68–80. <https://doi.org/10.22034/aeis.2023.415920.1131>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.