applied sciences

MDPI

Article

Ensemble Machine Learning for Monkeypox Transmission Time Series Forecasting

Emmanuel Gbenga Dada [1], David Opeoluwa Oyewola [2], Stephen Bassi Joseph [3], Onyeka Emebo [4] and Olugbenga Oluseun Oluwagbemi [5,6,*]

1 Department of Mathematical Sciences, University of Maiduguri, Maiduguri P.M.B. 1069, Nigeria
2 Department of Mathematics and Statistics, Federal University Kashere, Gombe P.M.B. 0182, Nigeria
3 Department of Computer Engineering, University of Maiduguri, Maiduguri P.M.B. 1069, Nigeria
4 Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA
5 Department of Computer and Information Technology, Sol Plaatje University, Kimberley 8301, South Africa
6 National Institute for Theoretical and Computational Sciences (NiTheCS), Stellenbosch 7602, South Africa
* Correspondence: olugbenga.oluwagbemi@fulbrightmail.org; Tel.: +27-663926506

Abstract: Public health is now in danger because of the current monkeypox outbreak, which has spread rapidly to more than 40 countries outside of Africa. The growing monkeypox epidemic has been classified as a "public health emergency of international concern" (PHEIC) by the World Health Organization (WHO). Infection outcomes, risk factors, clinical presentation, and transmission are all poorly understood. Computer- and machine-learning-assisted prediction and forecasting will be useful for controlling its spread. The objective of this research is to use the historical data of all reported human monkey pox cases to predict the transmission rate of the disease. This paper proposed stacking ensemble learning and machine learning techniques to forecast the rate of transmission of monkeypox. In this work, adaptive boosting regression (Adaboost), gradient boosting regression (GBOOST), random forest regression (RFR), ordinary least square regression (OLS), least absolute shrinkage selection operator regression (LASSO), and ridge regression (RIDGE) were applied for time series forecasting of monkeypox transmission. Performance metrics considered in this study are root mean square (RMSE), mean absolute error (MAE), and mean square error (MSE), which were used to evaluate the performance of the machine learning and the proposed Stacking Ensemble Learning (SEL) technique. Additionally, the monkey pox dataset was used as test data for this investigation. Experimental results revealed that SEL outperformed other machine learning approaches considered in this work with an RMSE of 33.1075; a MSE of 1096.1068; and a MAE of 22.4214. This is an indication that SEL is a better predictor than all the other models used in this study. It is hoped that this research will help government officials understand the threat of monkey pox and take the necessary mitigation actions.

Keywords: monkeypox; machine learning; time series; forecasting; stacking ensemble learning

Citation: Dada, E.G.; Oyewola, D.O.; Joseph, S.B.; Emebo, O.; Oluwagbemi, O.O. Ensemble Machine Learning for Monkeypox Transmission Time Series Forecasting. *Appl. Sci.* **2022**, *12*, 12128. https://doi.org/10.3390/app122312128

Academic Editors: William Yu Chung Wang, Yung-Chun Chang, Jheng-Long Wu and Hong-Jie Dai

Received: 24 October 2022
Accepted: 25 November 2022
Published: 27 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Public health, peace, and safety are being threatened by a number of diseases all over the world today. Over the past three (3) years, infectious diseases like COVID-19 (SARS-CoV-2) have spread throughout the world [1]. Different diseases are transmitted in different ways. For instance, SARS-CoV-2 is typically spread by contact with bodily fluids and human touch, among other things [2]. Millions of individuals have died as a result of SARS-CoV-2 around the world [3,4]. HIV can be passed from one person to another through sexual contact, the use of unsterilized, contaminated needles, and blood transfusions, among other methods [5]. Tuberculosis can spread from person to person when an infected person coughs into the air and an uninfected person breathes in the contaminated air [6–8]. Ebola Virus Disease (EVD) is an infectious disease that is transmitted or spread at alarming rates through body fluid contact and human body contact, among others [9–11]. Malaria

is another disease that can be transmitted by vectors such as *Anopheles* mosquitoes, by sucking human blood infected with Plasmodium falciparum [12,13].

Vaccine candidates have been developed to deal with the spread of the COVID-19 virus, and some of the outcomes of the vaccination effort have been impressive [14,15]. Computational health applications and software have been developed and applied to support the diagnosis and self-management of HIV disease and reduce the transmission of the disease [16,17]. Docking-based molecular simulation has been applied towards simulating possible drug targets to combat the transmission of Ebola virus disease (EVD) [18]. In another study, informatics tools such as the open data kit and the innovation hub were used to reduce the spread of an Ebola pandemic [19]. Bioinformatics approaches have been used to control Ebola virus disease (EVD) transmission [20]. Biological methods such as polymerase chain reaction (PCR) have been applied toward the diagnosis of malaria. This strategy can be employed as a viable tool for minimizing malaria transmission [21]. Other methods, such as computational informatics approaches, have been implemented for the correct dosage prescription of anti-malaria drugs [22], reducing the transmission of the disease. Recently, vaccine development approaches have been used to reduce malaria transmissions [23]. Magagula and colleagues [24] have modeled clinical immunity on malaria infection to reduce the transmission of the disease [24]. Malaria transmission has been controlled using modeling and simulation approaches [25–28].

Computational frameworks have also been developed to reduce the transmission of different diseases [29–31]. One of the methods other scientists have utilized for controlling and reducing the transmission of diseases is the development of health-related information systems [32–35]. Different medical and computational approaches have been applied towards combating hereditary diseases [36,37]. Biological and bioinformatics methods have been applied to control and reduce the transmission rates of tuberculosis [38–42]. Mathematical modeling techniques have been utilized and applied to control other types of diseases to reduce their transmissions [43–47]. Phylogenomics, mathematical modeling, and bioinformatics methods have been applied towards controlling the monkeypox virus [48–51]. However, recently, the monkeypox (MPX) virus was identified and listed as a disease of global emergency and international concern by the World Health Organization (WHO) [52]. Computational approaches such as machine learning can be used to forecast time-series transmission of the monkeypox virus to improve preparedness and reduce the transmission of the disease.

The monkeypox (MPX) virus was discovered in orangutans in a zoo in Indonesia in 1949 by Rijk Gispen [53]. It was not until about a decade later, in 1958, that it was recognized as a member of the poxvirus family, and Magnus et al. [54,55] reported their findings in 1959. The monkeypox virus is a double-stranded DNA virus that belongs to the orthopoxvirus taxonomy, which also includes human variola (VAR), cowpox (CPV), and vaccinia (VAC) viruses. The monkeypox virus is a zoonotic infection, which means it can be transmitted from one animal to another. Monkeypox virus has been detected in rodents, Gambian poached rats, dormice, and non-human primates, but further research is needed to determine the true natural repository (s). The monkeypox virus is divided into two genomic lineages: Central African (Congo Basin) and West African. The Congo Basin clade is expected to become a hotspot and cause more serious infections [56,57].

Whenever a person comes into indirect or direct contact with infected animals' or humans' secretions, injured substances, or polluted materials, the monkeypox virus can easily be transmitted. Transmission from animal to human can happen via bites, scratches, or cooking of animal parts. The monkeypox virus is assumed to be transmitted among people mostly through respiratory secretions during prolonged face-to-face interaction, although it can also be spread sexually through sex with an infectious individual's bodily fluids [58,59]. The virus can enter the body through cuts or wounds (even if they are not apparent), the respiratory tract, or the mucous membranes of the eyes, nose, and mouth. Whereas the MPX virus was first detected in monkeys, animals including squirrels and striped grass mice are thought to represent the virus's initial reservoir [60,61]. The virus's

propagation to infected humans was first documented in the Democratic Republic of the Congo in 1970 [56,61,62]. It produces a pathogenic infection in humans [63] with clinical manifestations comparable to smallpox. Repeated epidemics in human populations have been recorded primarily in the Congo Basin since the first transmission [61].

The virus, meanwhile, expanded beyond Africa, where an increase in incidence is thought to be linked to the termination of the smallpox immunization campaign [64,65]. Since natural and foreseeably modified variants [66] of MPX are apparently being studied in biowarfare programs [67,68], countries like the United States (US) have already been amassing smallpox vaccines as a safety measure against biological terrorism and biowar [69,70] and as an antiretroviral therapy that is said to be resistant to various varieties of previously discovered smallpox [71].

In view of the current COVID-19 pandemic and its repercussions, the current epidemic of MPX has now spread to 19 countries across five continents (by May 2022) [72] and has therefore become a new center of interest. An epidemic is characterized by a rapid spike in the transmission of a disease when the number of cases surpasses the estimated number for the place or period [72]. As a result, it is critical to respond rapidly to new diseases so that therapeutic and non-therapeutic countermeasures can be implemented promptly. Cases have been reported in individuals in western and central African countries ever since the revelation of the monkeypox virus in 1958: Cameroon, the Central African Republic, Côte d'Ivoire, the Democratic Republic of the Congo, Gabon, Liberia, Nigeria, the Republic of the Congo, and Sierra Leone, with many people infected in the Democratic Republic of the Congo. The majority of monkeypox cases in people beyond Africa have been connected to overseas travel or contact with exotic animals, with cases reported in the United States, Israel, Singapore, and the United Kingdom [73]. Most new occurrences of monkeypox in Sweden, Italy, Belgium, the United States, and Canada have not been related to travel to endemic areas in West and Central Africa, implying that local transmission occurs within societies. Monkeypox incidents have been reported in the UK, Spain, and Portugal among males who have sex with men, although not exclusively [73]. As of 31 May, 321 cases had been confirmed from 17 EU/EEA member countries, although there had been no fatalities. The majority of cases present with lacerations in the region of the sexual organ, implying that spread occurred most likely through direct physical contact during sexual intercourse. Cases have been recorded in a number of countries that appear to be linked to occurrences in Spain (Madrid and the Canary Islands) and Belgium (Antwerp). Several countries, on the other hand, record instances with little or no established pathogenic connection to travel, contact with other patients or animals, or participation in particular activities. There have been 557 confirmed cases worldwide (including the EU and EEA) [74].

The symptoms of monkeypox infection are less severe than those of smallpox, although they are identical. The fact that monkeypox causes enlarged lymph nodes (lymphadenopathy) but smallpox does not is a massive distinction. Inflammation can occur in many different places on the body, particularly in the neck and armpits, and it can be restricted. Monkeypox has a 7–14-day gestation period (from infection to development of symptoms), although it can be as short as 521 days. Fever, headache, muscle pain, tiredness, stiffness, and shivering are all early indications and symptoms, in addition to swollen lymph nodes [75]. Within 1 to 3 days (occasionally more) following the onset of fever, a rash appears on the body, probably beginning on the face and expanding to other areas of the body. Prior to coming off, the infections advance through the phases of red bumps, enlarged pores, blisters, abscesses, and eventually open sores. Biomedical predictions are commonly inaccurate, and their inaccuracy is frequently overlooked [76]. Because the number of incidents to be analyzed cannot be approximated by a single person, forecasting the future of outbreaks and epidemics is difficult [77]. Irrespective of the flaws involved in medical forecasting, it is nonetheless vital to give a summary of the situation to the public. This will enable extensive preparation for future problems to be undertaken.

Monkeypox research has been published in recent years, but it lacked exact measurements because most of the studies were country-specific rather than performance

comparisons using machine learning techniques in some locations. When it comes to monkeypox data, they also lacked the time variable. Much of the research is based on linear knowledge, whereas such problems are not linear in nature [77]. Observational study models, which can only be initiated based on hypotheses, have been examined in some investigations. The bulk of monkeypox research does not take advantage of artificial intelligence or machine learning techniques. This is despite their benefits in diagnosing and preventing the disease.

This paper seeks to apply different machine learning algorithms to forecast the transmission rate of the monkeypox endemic. Machine learning has proven over time to be a very effective and robust algorithm that can handle large volumes of data effectively. It can therefore be employed to prudently forecast the spread of the monkeypox pandemic. In specific terms, the main contributions of this work include:

i. A survey of machine learning algorithms that can be used for the prediction and control of monkeypox was presented.
ii. The study determined the best predictive model for confirmed monkey pox cases in various continents around the world that have become monkey pox hotspots.
iii. A forecasting model for monkeypox outbreaks in countries around the world, focusing on Africa, Europe, and the Americas, was developed using different machine learning models. The models include adaptive boosting regression, gradient boosting regression, random forest regression, the least absolute shrinkage selection operator, ridge regression, ordinary least squares regression, and proposed stacking ensemble learning methods for predicting monkeypox transmission rate.
iv. An evaluation of the performance of the proposed stacking ensemble machine learning technique for monkeypox transmission time series forecasting was done.

The rest of the paper is organized as follows: The review of the related together with the summary of contributions table is presented in Section 1. Section 2 is the materials and methods. Section 3 discussed the machine learning models that were used for forecasting. The simulation and statistical results of the experiment analysis were presented in Section 4. The conclusion of the paper is in Section 4.

Related Works

Much work has been done in monkeypox detection and diagnosis, as shown in Table 1. Hughes et al. [78] carried out research to see if coinfections are common and to define the clinical characteristics of these infections. The MPX/VZV coinfections were investigated using clinical, epidemiological, and analytical outcomes. The findings suggest that diseases with both MPX and VZV may modify the extent of infection. The ability to detect coinfections is enhanced by collecting numerous lesion samples. As the program progresses, it will be critical to have these methods in place to track changes in the particulate matter of situations in which many pathogenic strains simultaneously infect a victim over time. In that same time span, nevertheless, the percentage of coinfection cases has been steady. Many factors may contribute to the lack of spatial or temporal connections between coinfections. The number of reported cases examined may fluctuate because of community engagement operations and climatic conditions like the rainy season. While the monitoring police are gone from their positions, there could be a delay in the reporting of suspected cases, followed by an upsurge in cases after personnel come back for inspections. Any links between case categorization and season or location was not reported in the work. These relatively brief gaps in examinations may lessen as greater monitoring keeps going, leading to a more thorough study. One of the drawbacks of the work is the risk of cross-contamination between different samples from the same case. Another limitation of the study is that artificial intelligence techniques such as machine learning, with all their benefits, are not taken advantage of.

Lash et al. [79] explored the implications of georeferencing attempts on modeling monkeypox case data distribution and propagation risk. They evaluated the amount of time and effort put into transforming explanations of sickness prevalence sites into quantitative matrix dimensions (latitude and longitude). The researchers created three

datasets from the same original monkeypox ailment occurrence data, each with a different standard of treatment and commitment: the first based on an interactive website; the second enhancing the first by referencing supplementary maps and digital guidebooks; and the third strengthened even more by consulting legacy monitoring records that give excellent extra details of each case. They created environmental habitat models and made predictions of monkeypox propagation risk based on each of the three incidence data sets to demonstrate the ramifications of these apparently minor enhancements in data quality. Subject to the type of georeferencing approach utilized, they discovered macrogeographic variances in environmental niche predictions. In the Africa region and along the periphery of the Congo Basin, less careful georeferencing indicated significantly smaller areas as having potential for monkeypox transmission. These findings have consequences for mapping operations, as each increased degree of georeferencing precision necessitates a significant increase in time investment. The limitation of the study is that it is only limited to the Congo Basin. Additionally, the performance of the proposed technique is poor because of the high error rate.

Nolen et al. [80] investigated the prolonged human-to-human transmission during a monkeypox epidemic in the Democratic Republic of the Congo. The population attack rate (the percentage of people who live with an infected person and develop MPXV symptoms) was 50%. Nine families within this health zone experienced multiple transmission incidents, totaling more than six transmission episodes. An average of 8 days (between 4 and 14 days) passed during incubation. The increasing incidence and spread seen in this study highlight the need for monkeypox surveillance and early detection. The study has some flaws that need to be pointed out. First, in 48% of the cases, PCR verified the presence of the MPXV virus; the other cases were discovered through the symptoms of the victims. Because patients were checked after their symptoms had subsided, an accurate diagnosis was not always possible. Confirmation of MPXV infection in many patients was not achievable during the pandemic because of a lack of local resources for obtaining samples. The lack of specimen collection in the current monitoring program has been discovered as a defect. Second, it is difficult to determine the incubation time for many patients. This is because they were unable to identify a specific cause of the illness or the date of infection.

Liu et al. [81] applied deep learning as an alternative diagnosis of skin disorders. A deep learning system (DLS) is utilized to provide a diagnostic evaluation of skin problems in clinical situations. The DLS differentiates between 26 of the most prevalent skin disorders, which account for around 80% of all skin issues seen in the healthcare system. The initial 14,021 cases from a tele dermatology practice that served 17 clinical sites were used for creation, and the final 3756 cases were used for validation. These de-identified cases were used for both the development and assessment of the DLS. The findings show that the DLS has the capacity to boost health providers' (GPs') capacity to effectively detect skin problems without additional specialty training by proposing disparity diagnoses that may not have been considered. The work's limitation is that even if the performance was satisfactory, more could have been done. Additionally, the proposed system was not evaluated against any existing deep learning systems. Finally, monkeypox is not one of the skin diseases that the suggested system is designed to identify.

Tom and Anebo [82] created a neuro-fuzzy-based model for detecting monkeypox virus variants in other pox families. The authors considered 18 symptoms linked to monkeypox, as reported by the medical professionals interviewed, as well as symptoms from the Nigeria Centre for Disease Control and other publications about monkeypox diagnosis. However, only three symptoms were chosen for the simulation when the model was put into practice. The proposed system has the drawback of not being able to employ the 18 inputs that are theoretically analyzed during analysis. Additionally, the study's dataset was kept secret. The system's performance is also mediocre. Moreover, not all the monkeypox symptoms are present in the input set.

Bunge et al. [83] performed a comprehensive analysis of the peer-reviewed and unpublished literature on the evolution of monkeypox epidemiology, focusing particularly

on the quantity of verified, suspected, and/or potential cases, age at presentation, death, and regional distribution. For data extraction, they found 48 peer-reviewed articles and 18 pieces of gray literature. Since the 1970s, there have been more instances of human monkeypox, with the DRC experiencing the largest rise. The analysis reveals an increase in instances of monkeypox, particularly in the DRC, where it is extremely endemic; a spread to other nations; and a rising median age from young children to young adults. These findings could be explained by the rise in human-to-human transmission following the end of smallpox immunization, which offered some cross-protection against monkeypox. The emergence of outbreaks outside of Africa emphasizes the disease's global significance. Increased surveillance and case detection are crucial for comprehending the epidemiology of this resurgent disease, which is changing continuously. The study was limited in that it was not possible to conduct a thorough investigation of the percentage of instances that were transmitted from person to person. Additionally, because data quantity and quality differed by jurisdiction, the writers did not give a clear view of the number of reported, suspected, and/or possible cases.

Table 1. Summary of Related Works.

Author and Year	Technique	Contributions	Research Gap
Hughes et al. [78].	Define the clinical characteristics of MPX/VZV.	MPX/VZV coinfections were investigated using clinical, epidemiological, and analytical outcomes.	Lack of spatial or temporal connections between coinfections. Risk of cross contamination between different samples from the same case.
Lash et al. [79].	Georeferencing attempt on modeling monkeypox case data distribution and propagation risk.	Discovered macrogeographic variances in environmental niche predictions.	Study is limited to Congo Basin. Additionally, the performance of the proposed technique is poor because of high error rate.
Nolen et al. [80].	Human-to-human transmission during a monkeypox using laboratory analysis.	In 48% of the cases, PCR verified the presence of the MPXV virus.	Lack of specimen collection in the current monitoring program. Inability to determine the incubation time for many sufferers since they failed to properly identify a specific cause of the illness or a date of infection.
Liu et al. [81].	Deep learning for diagnosis of skin disorders.	DLS distinguishes between 26 of the most common skin disorders, which account for around 80% of all skin issues seen in healthcare system.	Proposed system was not evaluated against existing deep learning systems. Additionally, monkeypox is not one of the skin diseases that the proposed system is designed to identify
Tom and Anebo [82]	Neuro-fussy based model for diagnosis of monkeypox virus.	Was able to differentiate Monkeypox from other pox families Authors took into account 18 symptoms linked to it.	The system only use 3 out of the 18 monkeypox symptoms as inputs. Additionally, the dataset used was kept a secret. The system's performance is also mediocre.
Bunge et al. [83].	Analysis of the peer-reviewed and unpublished literature on the evolution of the monkeypox epidemiology	Proposed increased surveillance and case detection are crucial for comprehending the epidemiology of this resurgent disease, which is changing continuously.	Inability to conduct a thorough investigation of the percentage of instances that were transmitted from person to person. Additionally, data quantity and quality differed by jurisdiction, No information on the number of reported, suspected, and/or possible cases.

The motivation for using stacking ensemble learning for this work is that machine learning models have historically been constructed with the presumption that a model will operate effectively if the training and test data are derived from the same feature space and distribution. However, if there is any alteration in the feature space or the distribution of data changes, then there will be a need to create a new model. It is costly to build a new model from scratch each time, as well as gather fresh training data each time. Stacking ensemble learning allows easy retrieval of the enormous volumes of training data with less time and effort. The motivation behind using "stacked" ensemble learning is that the

algorithm has the capacity to intelligently use knowledge acquired earlier for a different task or area to tackle new problems more quickly or effectively.

2. Materials and Methods

2.1. Adaptive Boosting Regression (Adaboost)

The boosting technique known as the AdaBoost algorithm, sometimes known as adaptive boosting, is used as an ensemble method in machine learning. The weights are redistributed to each instance, with higher weights being given to instances that were mistakenly classified, hence the name "adaptive boosting." For supervised learning, boosting is used to lower bias and variance. It operates under the premise that learners improve incrementally. The meta-estimator adaptive boosting regression [84] starts by fitting a regressor to the initial dataset. It then fits additional copies of the regressor onto the same dataset. The only exception is that the weights of the instances are changed in accordance with the error of the most recent prediction. Consider a dataset $S = (x_1, y_1), \ldots, (x_n, y_n)$ that is derived from a time series. The dataset is composed of n pairs of observations, and each observation is given a weight w_i. The likelihood that an observation will be included in the training set at iteration k is then determined for each observation i based on the weight it is given. The weighted sum of the probabilities is then used to get the average loss (l_k) for the model k over all the observations i. The average loss (l_k) and probability p^k mathematical formulae are as shown in Equations (1)–(3):

$$p_k = \frac{w_i}{\sum w_i} \tag{1}$$

$$l_k = \sum_{i=1}^{n} l_k p_k \tag{2}$$

$$w_i^{k+1} = w_i^k \beta_k (1 - l_k) \tag{3}$$

where p_k is the probability at iteration k, average loss at iteration k, w_i^{k+1} is the updated weight at iteration i, w_i^k is the prior weight and β_k is the model loss.

2.2. Gradient Boosting Regression (GBOOST)

Gradient boosting is a machine learning method that is used, among other things, for regression and classification problems. It provides a prediction model in the form of an ensemble of weak, decision-tree-like prediction models [85]. Gradient boosting repeatedly selects a function that points in the opposite direction of the gradient to maximize a cost function across the function space. Decision trees are frequently used as weak predictors for gradient boosting. Weakly learned models are ones with low variance and regularization, a large bias toward the training dataset, and outputs that are just marginally better than random guesses. An additive model, weak learners, and a loss function are the three main components of boosting techniques. Gradient boosting machines work by using gradients to spot the weaknesses in poor models. This is done by using an iterative strategy where the goal is to eventually merge base learners to reduce prediction errors, whereby decision trees are joined using an additive model; and the loss function is reduced using gradient descent. The definition of the gradient boosting tree (g) is presented in Equations (4) and (5):

$$g = \sum_{i=1}^{n} f_i x_t \tag{4}$$

$$argmin \sum_{t} L(y_t, g) + f_{n+1}\, x_t \tag{5}$$

where g is the gradient boosting tree, $L()$ is the loss function and $f_{n+1} x_t$ is the new decision tree.

2.3. Random Forest Regression (RFR)

Random forest regression is a well-known decision tree approach [86] that creates numerous decision trees from an input dataset. The technique splits the dataset randomly into various sub-parts before creating several decision trees for each sub-part. The anticipated output of each decision tree is then combined to create a forecast that is more reliable and accurate. The output value of each input or subset in random forest regression is the mean of the values predicted by several decision trees. An n-tree bootstrap sample is generated from the real input dataset in random forest regression [87]. Expanding an unpruned regression tree from the various bootstrap samples is the next step. However, by averaging the predictions of the decision trees, a new data value is calculated. The error rate may be predicted using average out-of-bag predictions based on the training data.

2.4. Ordinary Least Square Regression (OLS)

Ordinal least squares regression is a well-liked technique for finding the coefficients of linear regression equations that describe the relationship between one or more individual quantitative variables and a dependent variable [88]. The errors of the smallest squares are represented by those of the least squares. OLS has two alternatives: maximum likelihood and the expanded technique of the moment's estimator. The OLS equation is depicted in (6).

$$Y = \beta_o + \sum_{j=1} \beta_j x_j + \varepsilon \tag{6}$$

where Y is the dependent variable, β, is the intercept of the model, x_j corresponds to the jth explanatory variable of the model and ε is the random error.

2.5. Least Absolute Shrinkage Selection Operator Regression (LASSO)

A linear regression analysis technique that lowers both the total absolute values of regression coefficients and the sum of squared errors is called the Least Absolute Shrinkage Selection Operation [89]. The following objective function is minimized to get the regression coefficient as shown in Equations (7) and (8).

$$lasso = \left(\sum_{i=1}^{n} y_i - \beta_o - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \gamma \sum_{j=1}^{p} |\beta_j| \tag{7}$$

$$\sum_{j=1}^{p} |\beta_j| \leq \gamma \tag{8}$$

where β_j is the regression coefficient operating on the standardized covariate j, β_o is the intercept and γ is a penalty term which controls the value of shrinkage.

2.6. Ridge Regression (RIDGE)

Ridge regression is a technique for determining the coefficients of multiple-regression models in situations where the autonomous variables are very interrelated. Data from multiple regression that exhibits multicollinearity can be analyzed using the Ridge Regression method [90]. Although least squares estimates are unbiased when multicollinearity occurs, their enormous variances make it possible that they are far from the true value. Ridge regression lowers the standard errors by introducing some bias into the regression estimates. Overall, it is hoped that this will result in more accurate estimations.

Consider a regression Equation (9):

$$Y = X^{-1}\beta + \varepsilon \tag{9}$$

Regression coefficients of ordinary least square is depicted by (10):

$$\hat{\beta} = \left(X^{-1}X\right)^{-1}X'Y \tag{10}$$

The variance covariance matrix of the estimate is represented by (11):

$$V\left(\hat{\beta}\right) = \sigma^2 R^{-1} \tag{11}$$

From the above, we find that Equation (12):

$$V\left(\hat{\beta}_j\right) = r^{ij} = \frac{1}{1-R^2} \tag{12}$$

The amount of bias in this estimator is given by Equation (13):

$$E\left(\widetilde{\beta} - \beta\right) = \left[X'X + kI\right)^{-1}X'X - I]\beta \tag{13}$$

The covariance matrix is given by Equation (14):

$$V\left(\widetilde{\beta}\right) = \left(X'X + kI\right)^{-1}X'X\left(X'X + kI\right) \tag{14}$$

where Y is the dependent variable, X represents the independent variables, β is the regression coefficients to be estimated, and ε represents the errors are residuals.

2.7. Proposed Model

2.7.1. Stacking Ensemble Learning (SEL)

Ensemble learning is a very powerful machine learning technique [91–93] that uses the combined output of two or more models or weak learners to address a specific computational intelligence issue. The main goal of ensemble learning is to boost model performance in areas like classification and prediction. A machine learning model called an ensemble model integrates the forecasts from two or more models.

Stacking is an ensemble learning technique that aggregates several machine learning algorithms through meta-learning. The algorithms at the base level are trained using a comprehensive training dataset; afterwards, a meta model is trained on the final results of the total base level model as features. Bagging, boosting, and stacking are the three most used ensemble learning techniques in machine learning. Significant differences between comparable models can be averaged using bagging to reduce discrepancies. Boosting constructs numerous sequential models to reduce the bias, while minimizing variation. Stacking is an augmentation of the voting classifier or voting regressor by an upper level (amalgamation level) that learns from the superior combination of the distinctive outputs. A different classifier or regressor is located at the top of the stack. Stacking is exceptionally beneficial when the outputs of different algorithms are likely not to be the same, and this is practically always the situation with regression. Stacking algorithms have the capacity to extend over a number of layers. This makes them very strenuous to train. The stacking approach is distinct. Stacking is used to investigate the space of many models for the same problem. According to the concept, it is possible to approach a learning problem using several kinds of models that can only learn a portion of the problem space. It will build multiple distinct learners and use them to create a transitional prediction, one for each learned model. Subsequently, a new model that gains knowledge from earlier predictions for a similar goal is added.

The ultimate model (meta model) is referred to as being stacked on top of all the others, hence the appellation. As a result, the general performance is enhanced. The stacking technique usually produces a model that performs better than any of the individual transitional models.

To forecast monkeypox transmission rate, this study used stacked ensemble learning (SEL) models. Stacking is a well-known ensemble modeling approach in machine learning. It entails merging several weak learners in parallel such that, by combining them with meta learners, we may anticipate better future predictions [94]. This ensemble strategy works by combining the predictions of numerous weak learners and meta learners to produce a superior output prediction model. Stacking is a technique in which an algorithm takes the outputs of sub-models as input and attempts to learn how to optimally combine the input predictions to get a superior output prediction. The model is called "stacking" since it is placed on top of the others.

Architecture of Stacking

The stacking model architecture is designed to include six base learner models: adaptive boosting regression (Adaboost), gradient boosting regression (GBOOST), random forest regression, ordinary least square regression (OLS), least absolute shrinkage selection operator regression (LASSO), and ridge regression (RIDGE), as well as a random forest meta-model that combines the predictions of the base models. The level 0 models are the base models, while the level 1 model is the meta-model. As a result, the stacking ensemble approach consists of original (training) data, primary level models, primary level prediction, secondary level models, and final prediction. The fundamental stacking architecture is as follows:

- Original data: The dataset is divided into training data and test data.
- Base models: Level-0 models include adaptive boosting regression (Adaboost), gradient boosting regression (GBOOST), random forest regression (RFR), ordinary least square regression (OLS), least absolute shrinkage selection operator regression (LASSO), and ridge regression (RIDGE). These models employ training data to provide assembled predictions (level 0).
- Level-0 Predictions: Each base model produces various level-0 predictions when it is activated on a set of training data.
- Meta Model: To aggregate the predictions of the base models as effectively as possible, the stacking model's architecture consists of a single meta-model that uses random forest regression. An alternate name for the meta-model is the level-1 model.
- Level-1 Prediction: The meta-model learns how to combine the predictions of the base models in the best way possible and is trained on the various predictions made by individual base models. For instance, data that was not used to train the base models is fed to the meta-model, predictions are made, and these predictions, along with the expected outputs, provide the input and output pairs of the training dataset that was used to fit the meta-model. See Figure 1 (The architecture of the proposed system).

The proposed system makes use of stacking ensemble learning (SEL) techniques for time series forecasting. Experiments were conducted using COVID-19 datasets [95] as the training sets, while as the test sets for each of the datasets, we employed the monkey pox dataset. The extracted features that are used for the prediction are confirmed_cases, suspected_cases, hospitalized, travel_history_yes and travel_history_no. COVID-19 was employed as the training set, while monkey pox was used as the test set. Six models were used in the architecture of the proposed stacking ensemble learning system. The primary goal of this research is to predict monkey pox transmission rates on different continents around the world using stacking ensemble machine learning. This is achieved by stacking the six models (adaptive boosting regression, gradient boosting regression, random forest regression, regularized regression (the least absolute shrinkage selection operator, ridge regression), and ordinary least squares stacking ensemble learning methods).

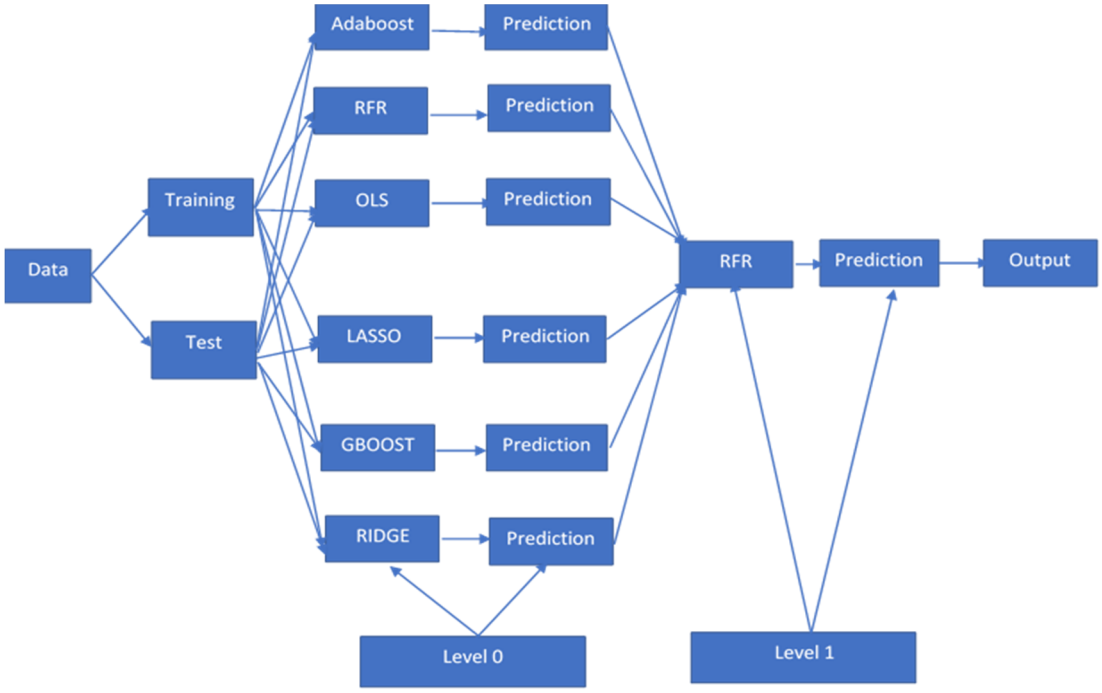

Figure 1. Architecture of the Proposed Stacking Ensemble Learning (SEL) Systems.

2.7.2. Dataset Description

For both training and testing, a diverse range of datasets were used. A total of 1836 COVID-19 datasets [96] were used in this study, with a sample size of 1836. Because COVID-19 is an epidemic that has spread globally, our study used it as a training set. The severe acute respiratory syndrome coronavirus (SARS-CoV-2) virus is the source of the COVID-19 pandemic, a coronavirus infection that has spread globally. COVID-19 has caused a sharp decline in the price of the stock market. Financial markets throughout the world responded negatively and behaved in a way that had not been seen prior to the 2019 economic crunch because of the worldwide spread of the COVID-19 virus [97]. The test dataset is the monkey pox outbreak dataset.

The theory holds that after pre-training, machine learning algorithms will have already discovered a solution that is close to being optimal. This works well if the solution spaces in the two domains are comparable. However, by locating a different (i.e., better) optimum in the solution space, this method may be able to reveal intricate correlations that exist between the two datasets. Machine learning algorithms were used on the COVID-19 and monkey pox datasets to evaluate the performance of the proposed stacking ensemble learning methodologies. The same experimental design and dataset are used for the comparisons. The collected data, comprising the COVID-19 and monkey pox datasets, is used for the experiments in this paper. Additionally, the monkey pox dataset was used as test data for this paper [98]. The COVID-19 dataset used for training our models was collected between 22 January 2020, and 7 August 2022, while the monkey pox dataset used for testing the models was collected between 6 May 2022, and 24 June 2022.

2.7.3. Experimental Configurations

The experimental configuration and parameter tuning for the work are depicted in Table 2. Training a model involves choosing suitable values for each weight and bias from labeled samples. Tuning parameters is one of the most important steps in the training of machine learning models. The parameters used to control the COVID-19 dataset for training and the monkey pox dataset for testing the models are all represented in Table 2 and are used to fine-tune the model's performance.

Table 2. Experimental configuration and parameters tuning of Adaboost, GBOOST, RF, OLS, LASSO, RIDGE and SEL.

All Models Were Trained Using Scikit Learn Package Machine Learning Model	Hyperparameter	Values
Adaboost	n_estimators	50
	learning_rate	0.2
	Loss	Exponential
RFR	n_estimators	400
	Random_state	0
OLS	Alpha	0.1
LASSO	Alpha	0.1
GBOOST	n_estimators	400
	max_depth	5
	Loss	Squared_error
	min_samples_split	2
	learning_rate = 0.1	0.1
RIDGE	Alpha	0.1
SEL	n_estimators	400
	Random_state	0

2.7.4. Performance Metrics

By computing the root mean square error, mean square error, and mean absolute error from the predictions, all techniques are ultimately assessed on the same test set. The following are the root mean square error (RMSE), mean square error (MSE), and mean absolute error (MAE). The RMSE, MSE and MAE are represented in Equations (15)–(17):

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(o-f)^2}{n}} \tag{15}$$

$$\text{MSE} = \frac{\sum_{i=1}^{n}(o-f)^2}{n} \tag{16}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|o-f| \tag{17}$$

where n is the number of observations, o is the observed values and f is forecasted values.

3. Results and Discussion

The section presents the results and discusses the major findings from our experiments. Experiments were performed using the Jupiter notebook (Python 3.9 version) programming environment. The Python programming language was used for training and testing all the models used in this work. The Python libraries used in this paper are Numpy, Matplotlib, Sklearn, Pandas, Math, Seaborn, Scipy, and Time.

The COVID-19 dataset was used for training the machine learning models, while the monkey pox dataset was used for testing the models. Experimental results of the original COVID-19 dataset that was used for training our model is already reported in our previous work [99]. The reason for using the COVID-19 dataset for training is that it is also an epidemic, just like monkey pox. COVID-19 has a large dataset, which the machine learning algorithms can easily learn from. However, the dataset for monkey pox is not so big, and is therefore more suitable for testing the trained models. The virus that causes monkey pox is a newly discovered infectious illness that affects people and is typically spread by rodents.

Although it can spread to other individuals, an epidemic cannot be readily sustained by person-to-person transmission alone. The clinical appearance is less severe but similar to that of smallpox in the past. Although smallpox was eliminated from the world in 1980, monkey pox still periodically arises in areas of Central and West Africa that are close to tropical rainforests. Monkey pox epidemics have historically had a 1–10% case fatality rate, but with the right care, many patients will recover. Health professionals in charge of monkey pox prevention and control are the target audience for this research, which gives a general introduction to the disease.

This paper considered the different continents that have experienced outbreaks of monkey pox. Figure 2 displays hospitalized cases of monkey pox in Africa, Asia, Europe, North America, South America, and Oceania. Africa and North America had the lowest number of people that were hospitalized (0.0%), while Oceania came in second with 2.6%. The European continent has the highest number of people hospitalized with cases of monkey pox, at 89.7%. Figure 3 shows the confirmed cases of monkey pox, with North America having the lowest with 0.2%, followed by Oceania with 0.3%, while the European continent has the highest number of confirmed cases with 97.6%. Figure 4 shows the suspected cases of monkey pox, with Africa, Asia, North America, and Oceania having the lowest with 1.6%. The European continent has the highest, with 75.8%. Figure 5 is the pie chart of the suspected travel history of monkey pox patients. Europe has the highest rate with 72.2%, followed by South America with 12.7%. The African continent had the lowest percentage, at 0.8%. Figure 6 is the pie chart of suspected cases of monkey pox with no travel history. The result showed that Europe is leading with 87.9%, while Asia, Oceania, and North America are the lowest with 1.5%. Figure 7 is the time series of monkey pox from 6 May 2022 to 24 June 2022. It displays the downtrend of the monkey pox, which was followed by the horizontal trend and the downtrend. This shows that there is a gradual reduction in monkey pox transmission in Europe.

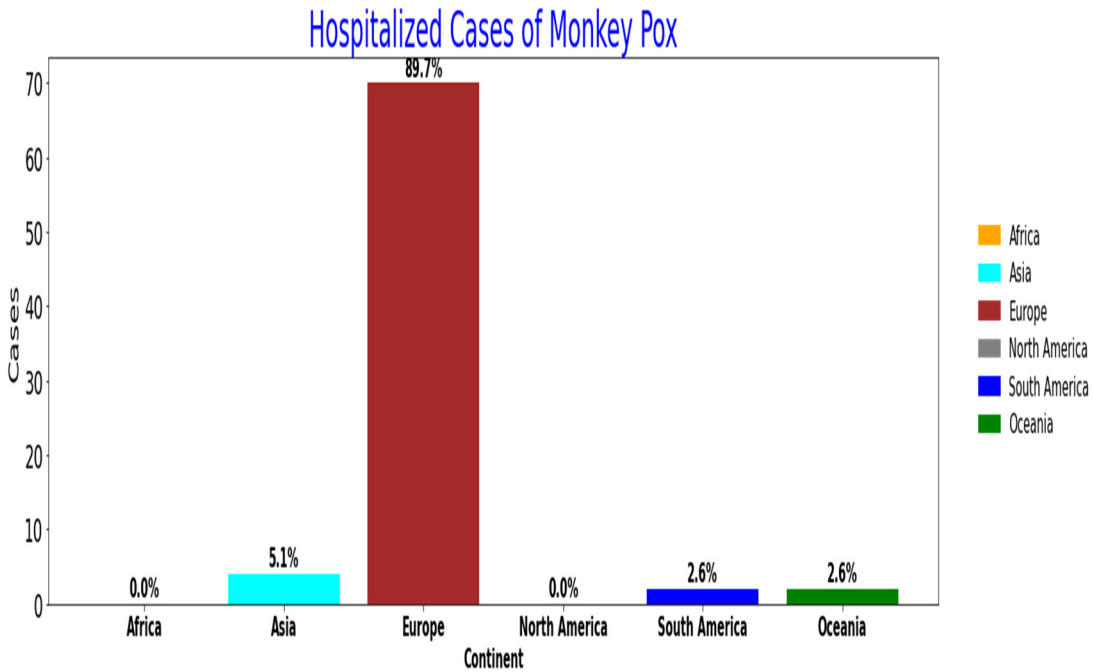

Figure 2. Hospitalized Cases of Monkey Pox.

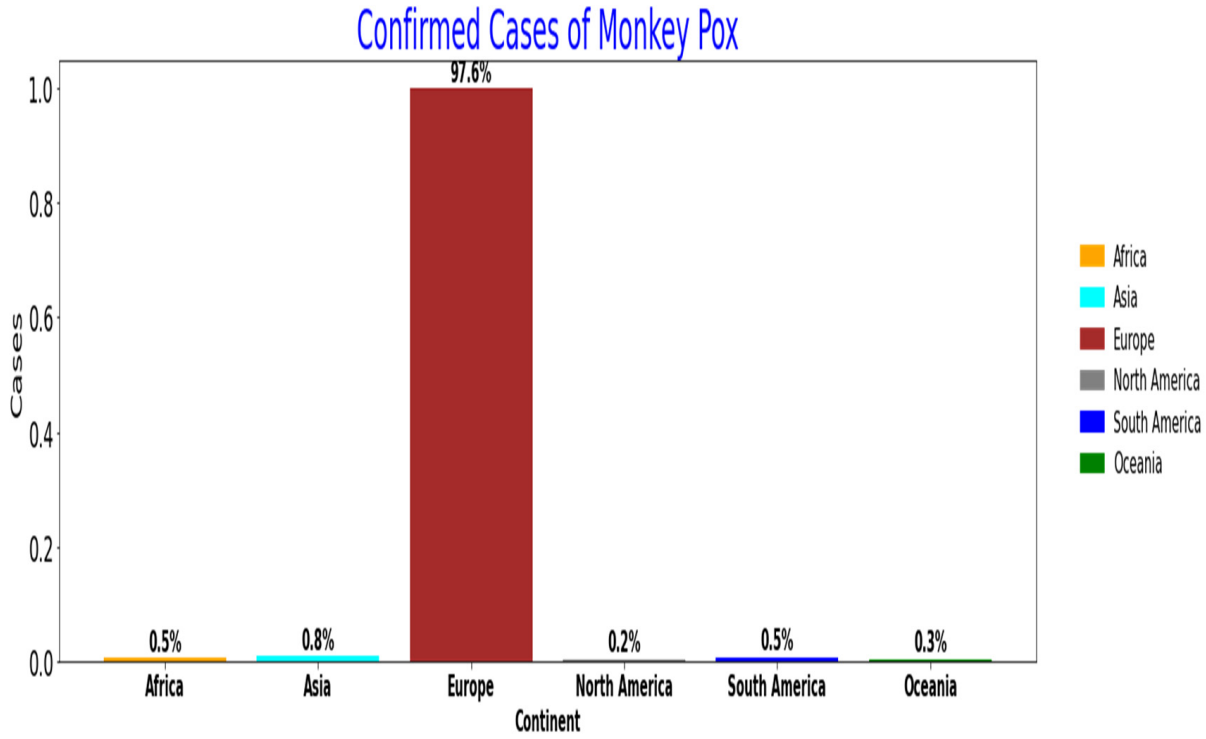

Figure 3. Confirmed Cases of Monkey pox.

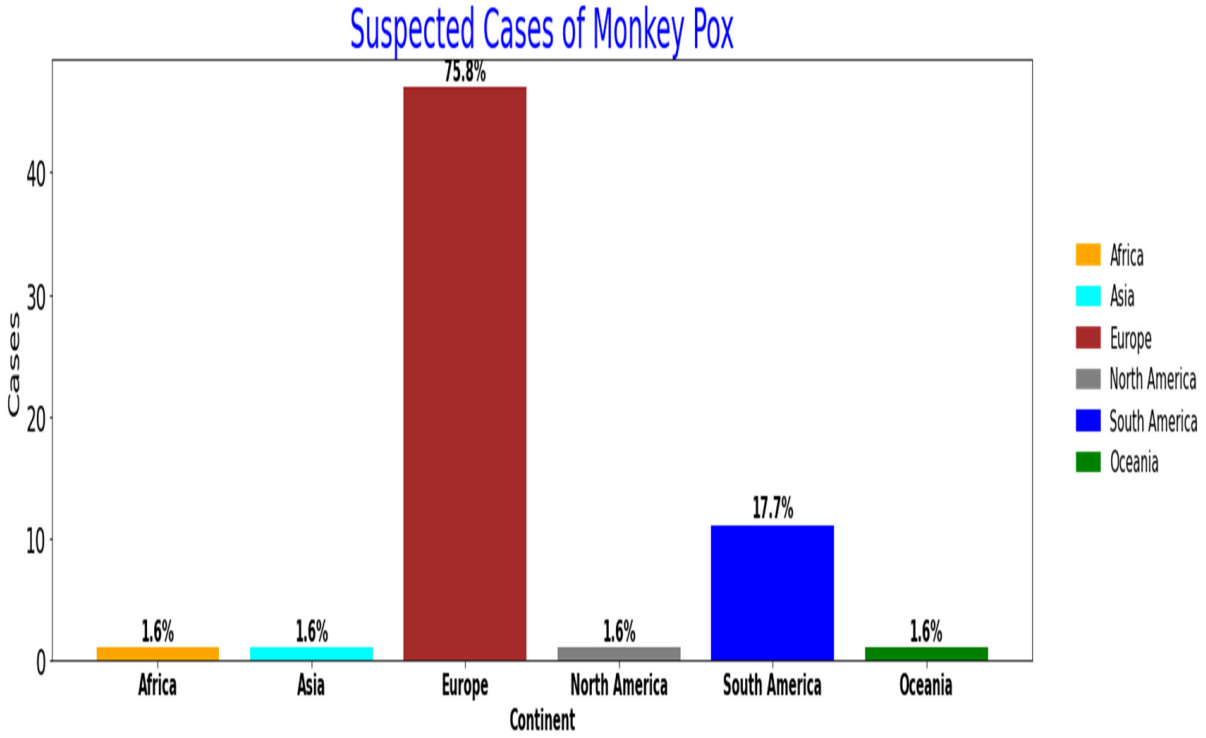

Figure 4. Suspected Cases of Monkey Pox.

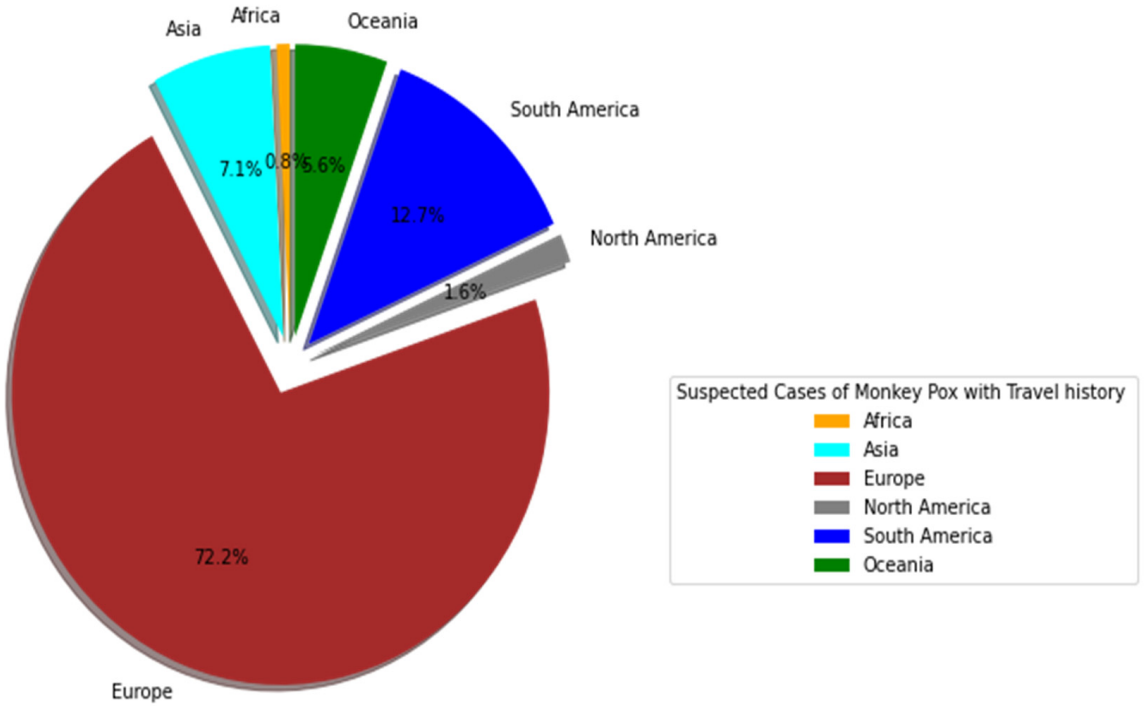

Figure 5. Suspected Travel History of Monkey Pox.

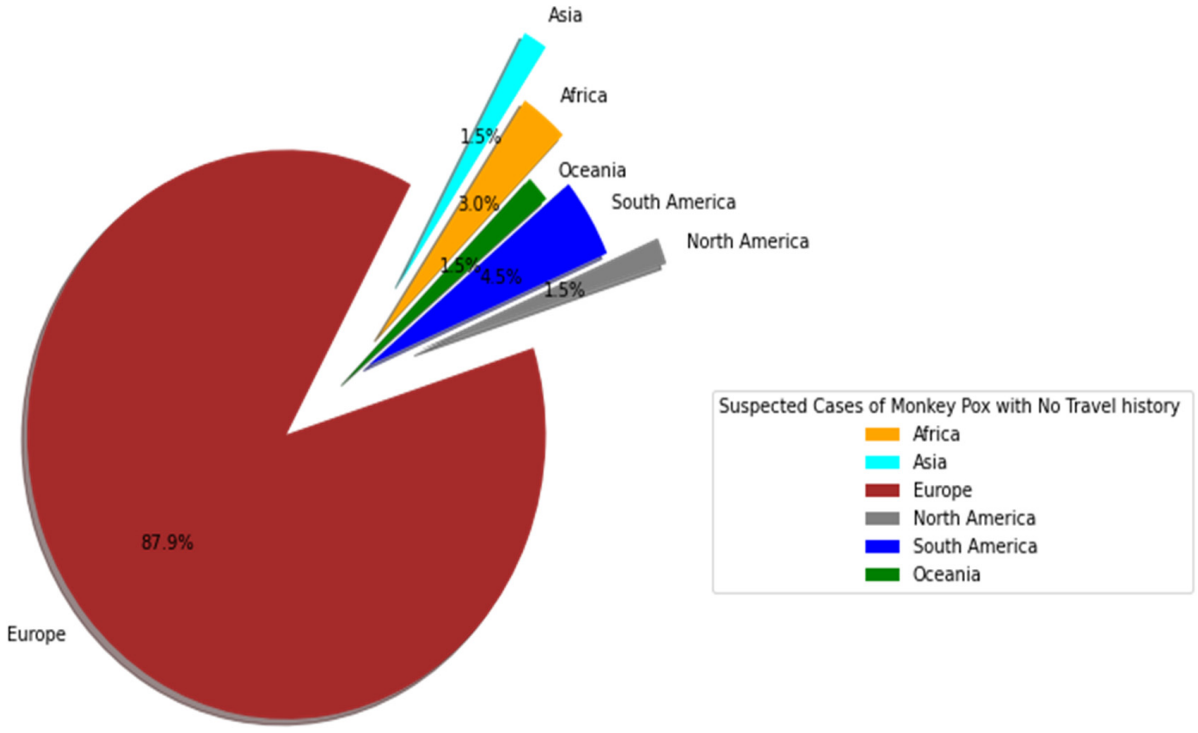

Figure 6. Suspected Cases of Monkey Pox with No Travel History.

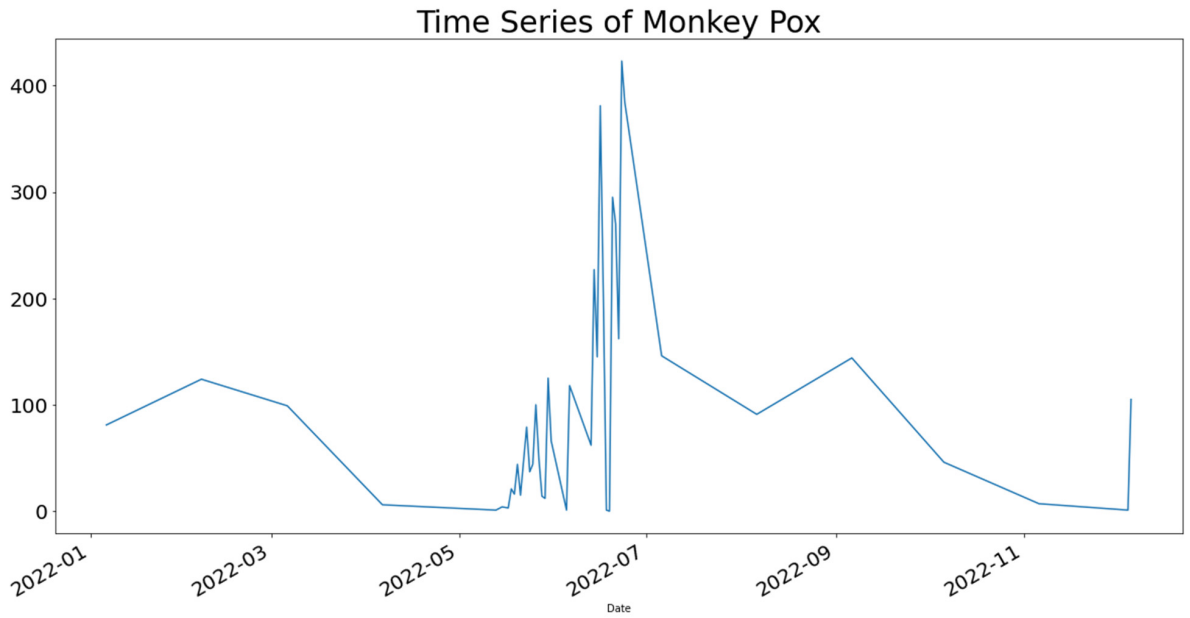

Figure 7. Time Series of Monkey pox.

In this study, the computation times of the six machine learning algorithms were analyzed. The time lapses adopted in this paper for each machine learning algorithm start at the beginning and end of each machine learning algorithm used in this research. The time lapse was obtained by subtracting the start time from the end time of the machine learning model. Figure 8 displays the training time for all the machine learning. The results show that ordinary least square has the least computational training time, followed by LASSO and RIDGE. This implies that ensemble machine learning algorithms, such as adaptive boost (Adaboost), gradient boost (GBOOST), and random forest, have a higher computational time than regularized regression algorithms, such as LASSO and RIDGE. Figures 9–14 is the time series of all the machine learning compared in this paper, which consists of actual monkey pox and predicted monkey pox transmission rates.

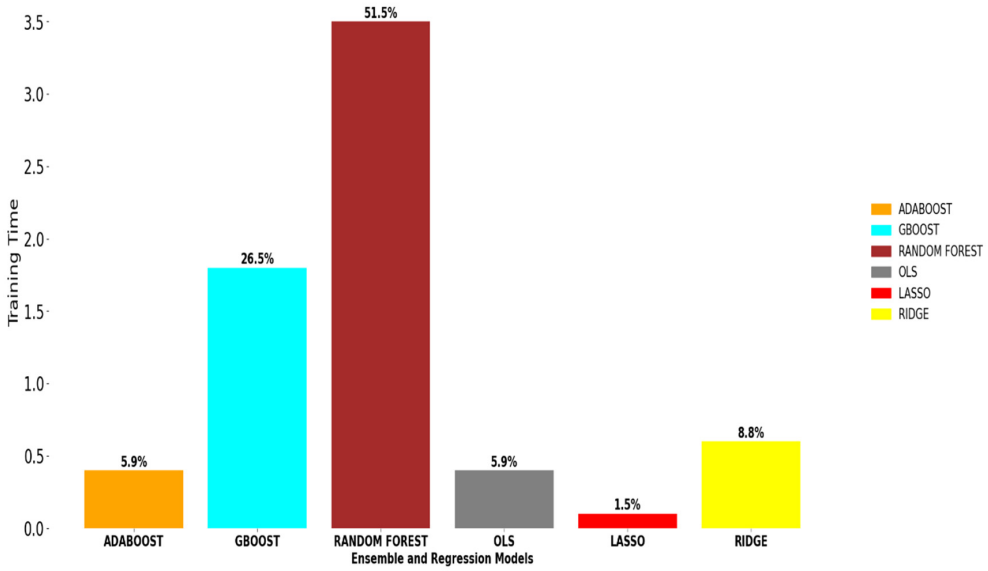

Figure 8. Comparison of Training time of machine learning.

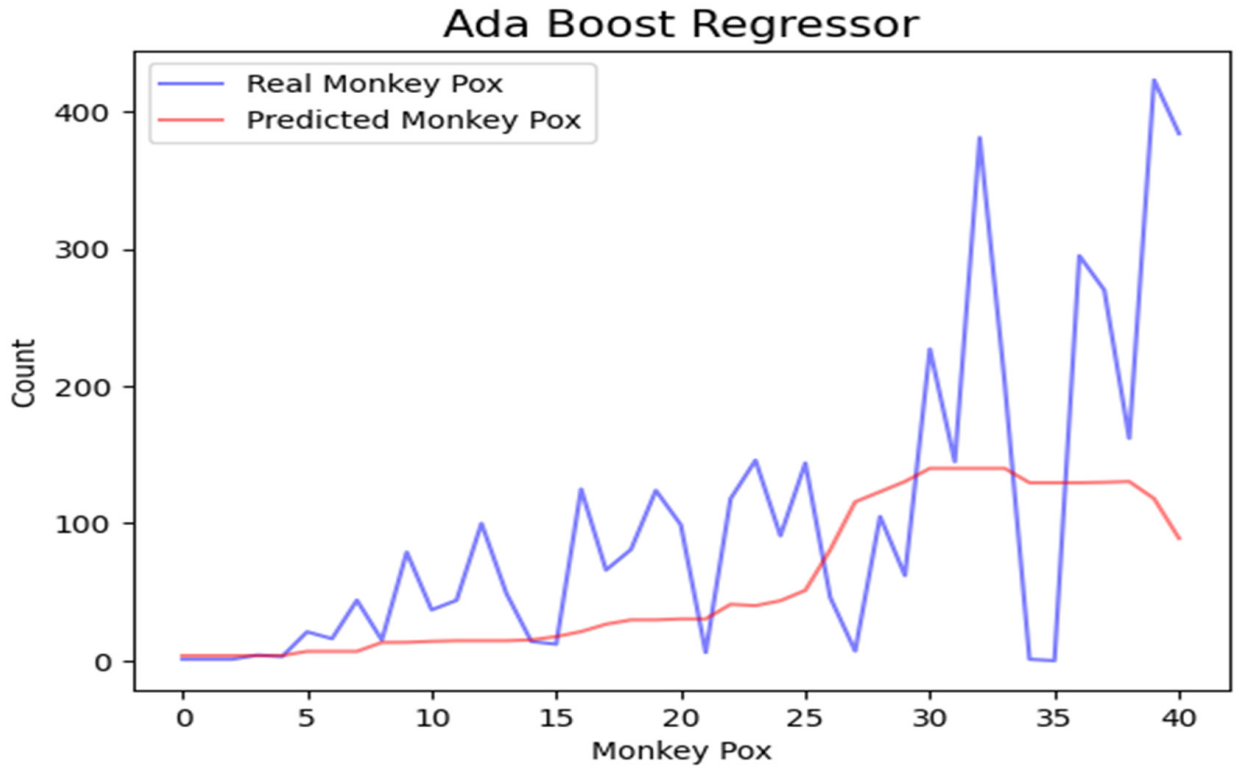

Figure 9. Time Series of Adaboost machine learning.

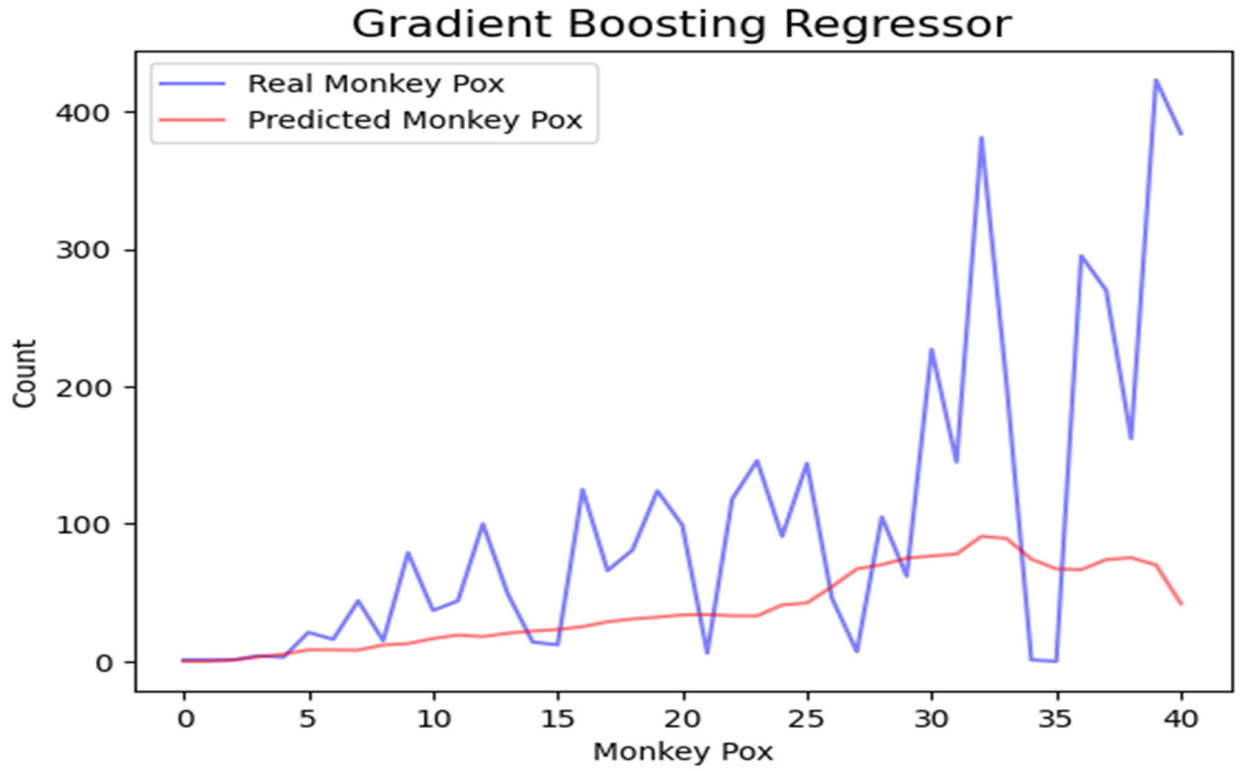

Figure 10. Time Series of Gradient Boost Algorithm.

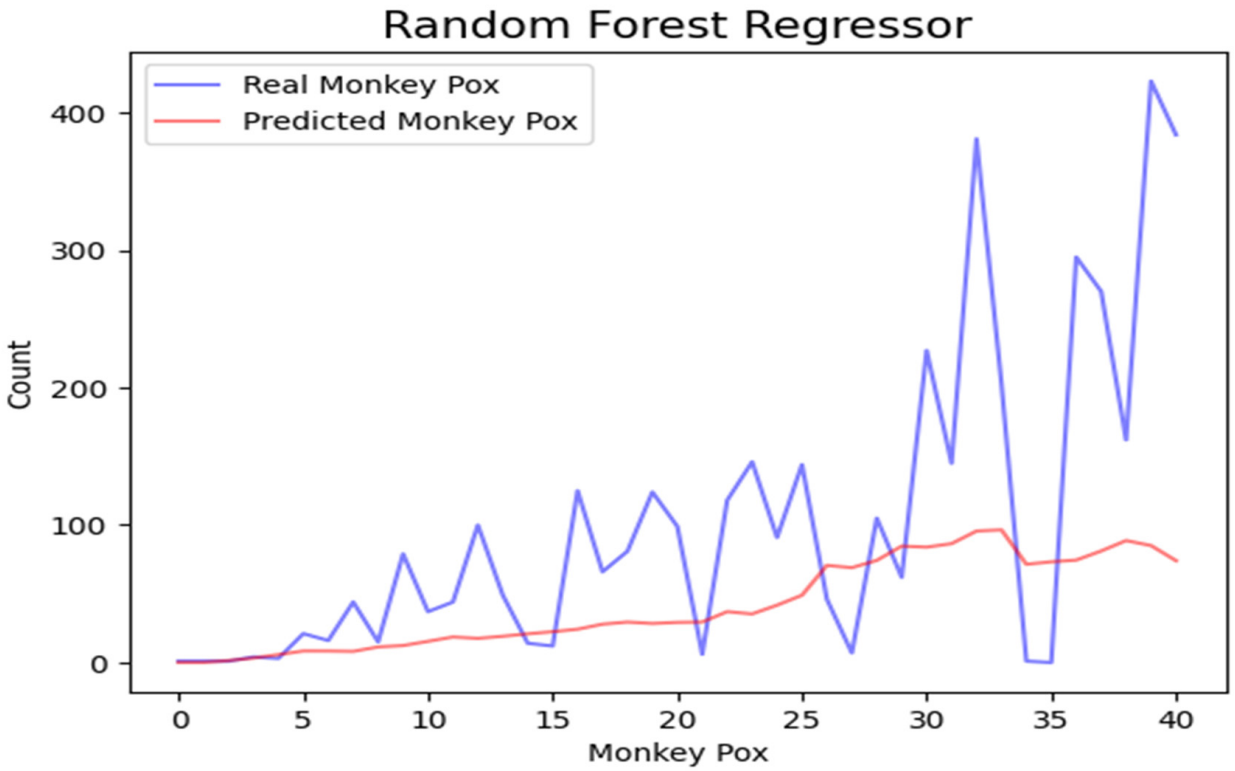

Figure 11. Time Series of Random Forest Algorithm.

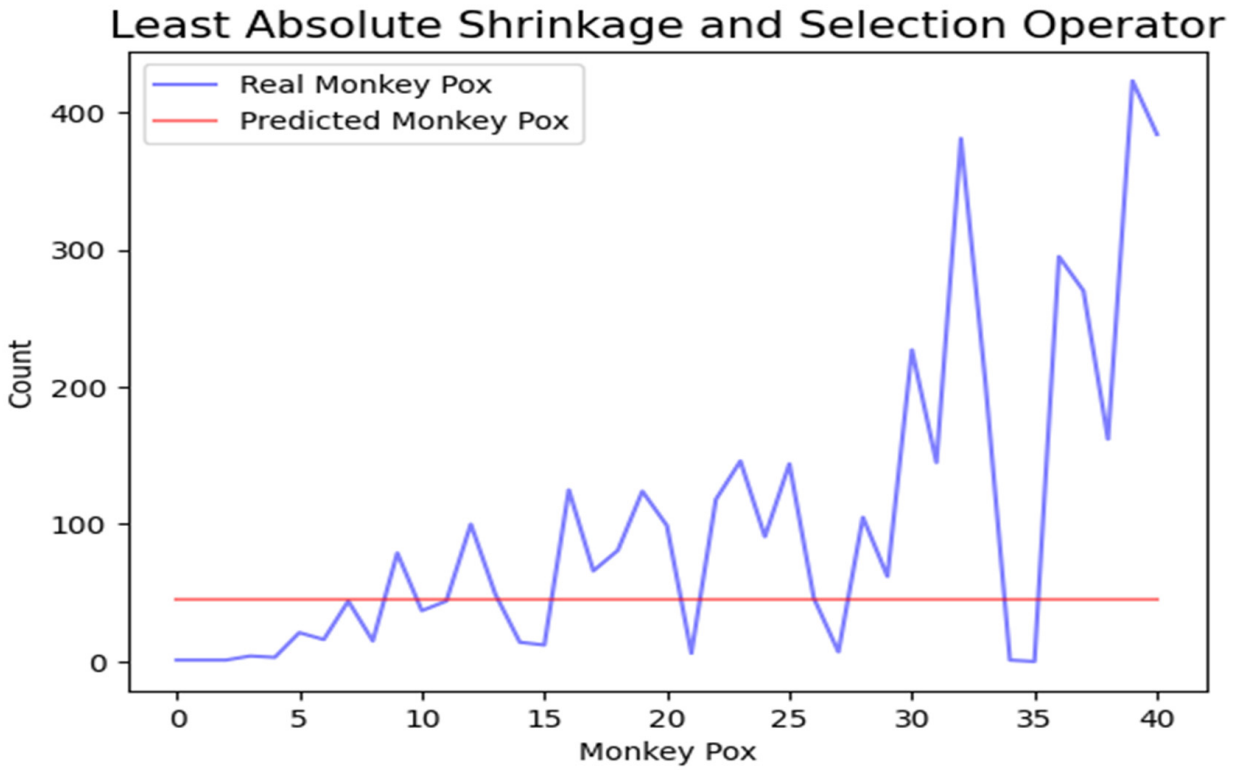

Figure 12. Time Series of Least Absolute Shrinkage and Selection Operator Algorithm.

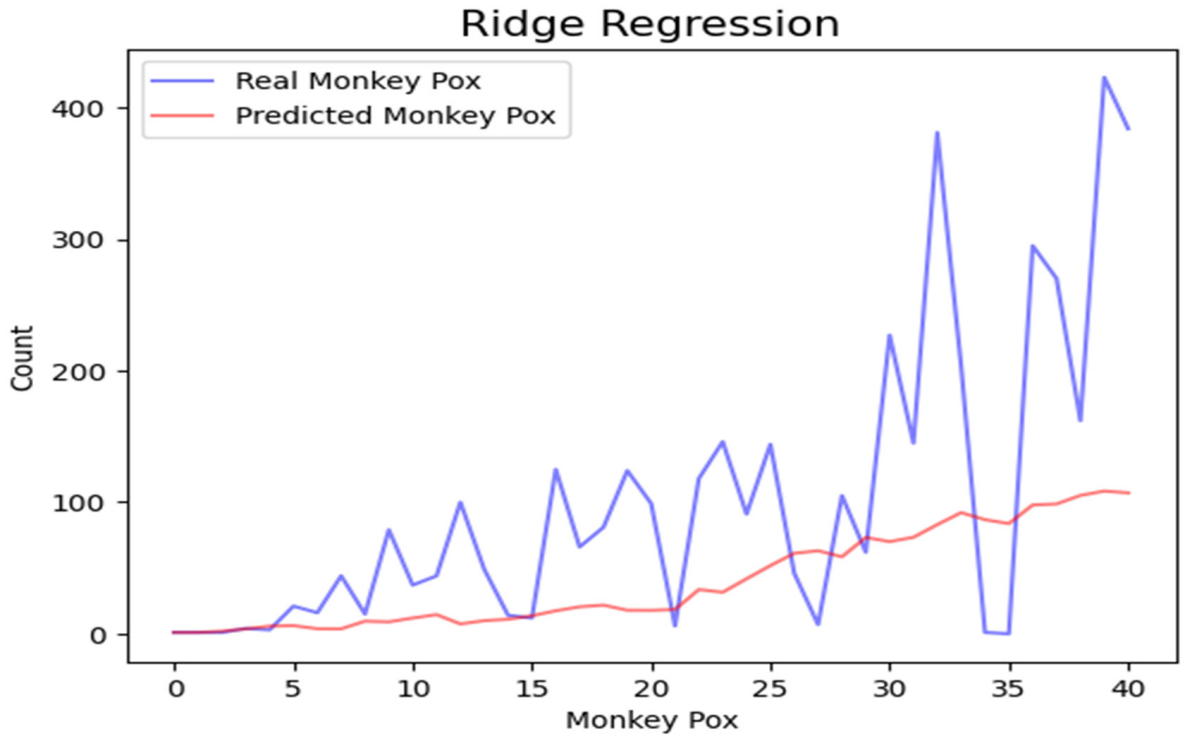

Figure 13. Time Series of Ridge Regression Algorithm.

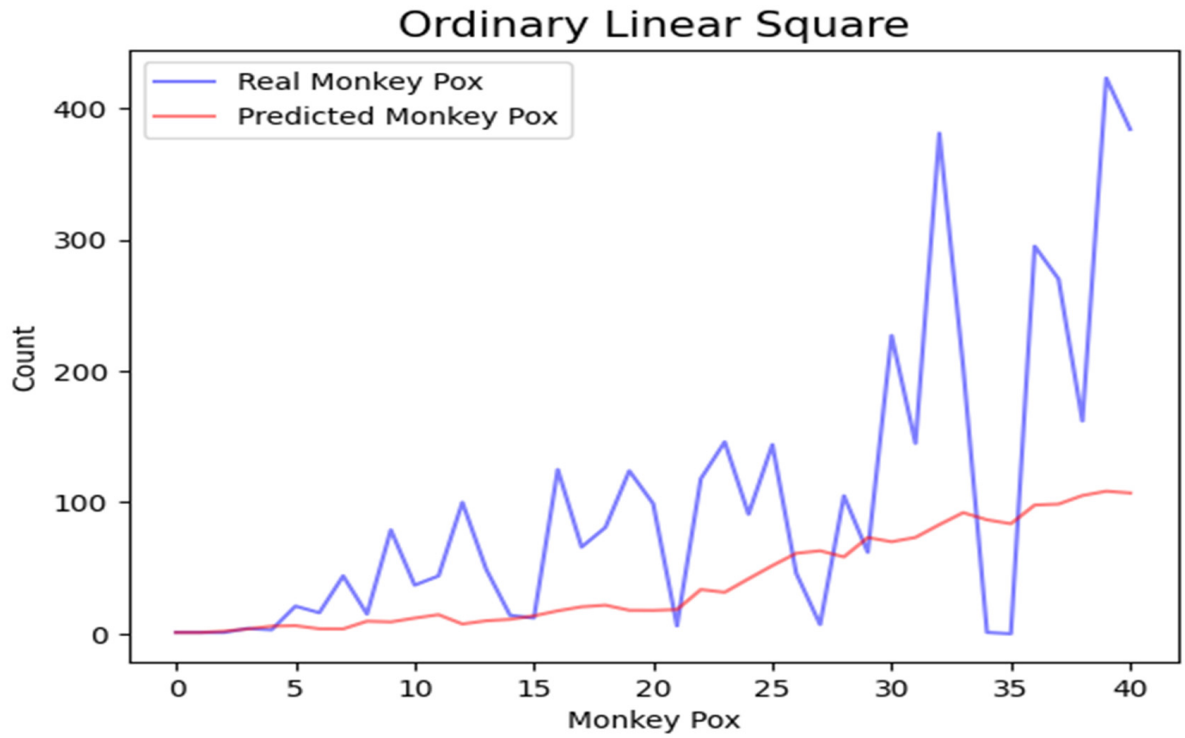

Figure 14. Time Series of Ordinary Least Square.

The COVID-19 dataset was used for training the models since COVID-19 was also declared an epidemic by the WHO in 2020, while the test dataset remained monkey pox, which was also declared an epidemic by the WHO in 2022. Three performance metrics were adopted in this paper, which were RMSE, MSE, and MAE. Figure 15 shows the actual and predicted monkey pox transmission rate obtained from the Stacking Ensemble Learning (SEL) model proposed in this paper. Table 3 displays the statistical results of all the machine

learning methods considered in this paper. SEL performed excellently with a RMSE of 33.1075, followed by Adaboost with 100.7981 when considering the RMSE.

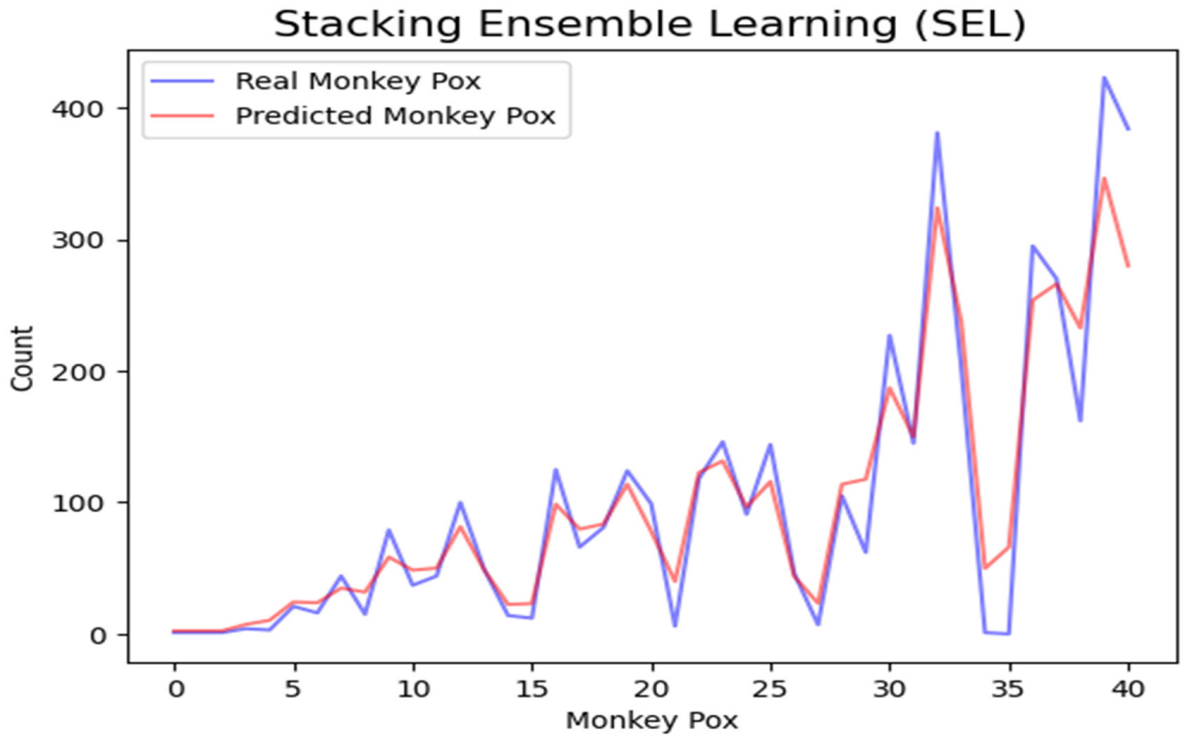

Figure 15. Time Series of Stacking Ensemble Learning (SEL).

Table 3. Performance Metrics of Machine Learning and Proposed Stacking Ensemble Learning (SEL).

Algorithm	RMSE	MSE	MAE
Adaboost	100.7981	10,160.2726	68.5727
GBOOST	115.8856	13,429.4795	75.7660
Random Forest	111.1798	12,360.9503	73.9469
OLS	108.4988	11,772.0099	74.2072
LASSO	124.5257	15,506.6696	82.3904
RIDGE	113.4759	12,876.7957	70.2174
SEL	33.1075	1096.1068	22.4214

4. Conclusions

This research presents a novel technique that uses stacking ensemble learning for monkey-pox time series forecasting. The COVID-19 and monkey-pox datasets were used for training and testing the proposed method, respectively. Experiment results show that the proposed stacking ensemble learning outperformed other machine learning techniques considered in the paper. In addition, the three proposed training strategies—OLS, LASSO, and RIDGE—achieved the lowest computational training time. Despite only being tested with two source datasets, the proposed method can also be used to train more datasets. According to experiment findings, compared to other continents, Europe has the highest rate of monkey pox, and, unless necessary measures are taken, there is a probability that the pandemic will spread faster. The stacking ensemble learning method was adopted for this paper because it saves us the trouble of having to train several machine learning models from scratch to fulfill similar tasks, therefore saving time and resources. SEL also serves as a cost-cutting measure in areas of machine learning that need a lot of resources,

such as image classification or natural language processing. Moreover, it is very useful in compensating for a shortage of labeled training data maintained by an organization, as pre-trained models are used. Stacking ensemble learning makes use of minimal computational resources and helps attain enhanced results using a smaller dataset. Furthermore, stacking ensemble learning models attains optimal performance quickly compared to conventional ML models. The reason for this is that the models leverage knowledge from base models and meta models.

Our findings from experimental results indicated that Europe has the highest monkey pox transmission rate with 72.2%, followed by South America with 12.7%. The continent of Africa now has the lowest transmission rate at 0.8%. Based on the time series prediction of the monkey pox dataset used for this study, the transmission rate will decrease. This indicates that monkey pox transmission is steadily declining in Europe and other parts of the world. In the future, deep neural network algorithms such as deep belief networks, convolutional neural networks, generative adversarial networks, and autoencoders will be applied to detecting and diagnosing monkey pox infections using image datasets.

Author Contributions: Conceptualization, E.G.D., D.O.O., S.B.J. and O.O.O.; methodology, E.G.D., D.O.O., S.B.J., O.O.O.; software, E.G.D., D.O.O. and S.B.J.; validation, E.G.D., D.O.O., S.B.J., O.O.O. and O.E.; formal analysis, E.G.D., D.O.O., S.B.J. and O.O.O.; investigation, E.G.D., D.O.O., S.B.J. and O.O.O.; resources, E.G.D., D.O.O., S.B.J., O.O.O. and O.E.; data curation, E.G.D., D.O.O., S.B.J. and O.O.O.; writing—original draft preparation, E.G.D., D.O.O., S.B.J. and O.O.O.; writing—review and editing, E.G.D., D.O.O., S.B.J., O.O.O. and O.E.; visualization, E.G.D., D.O.O., S.B.J., O.O.O. and O.E.; supervision, E.G.D. and O.O.O.; project administration, E.G.D. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Virginia Tech University's research support to O.E. O.O.O. received research funding from the Oppenheimer Memorial Trust (OMT) Foundation with grant number: OMT Ref. 21563/01); O.O.O. was also supported with DAAD ClimapAfrica grant with grant number: ST32/ 91769426).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Monkeypox Dataset was gotten from Kaggle open access online database: https://www.kaggle.com/datasets/deepcontractor/monkeypox-dataset-daily-updated (accessed on 28 June 2022). The data is freely available, and no consent is needed.

Acknowledgments: We thank all co-authors for their various support and contributions. The APC was funded by Virginia Tech University's research support to O.E. O.O.O. received research funding from the Oppenheimer Memorial Trust (OMT) Foundation with grant number: OMT Ref. 21563/01); O.O.O. was also supported with DAAD ClimapAfrica grant with grant number: ST32/ 91769426).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Petersen, E.; Ntoumi, F.; Hui, D.S.; Abubakar, A.; Kramer, L.D.; Obiero, C.; Tambyah, P.A.; Blumberg, L.; Yapi, R.; Al-Abri, S.; et al. Emergence of new SARS-CoV-2 Variant of Concern Omicron (B.1.1.529)-highlights Africa's research capabilities, but exposes major knowledge gaps, inequities of vaccine distribution, inadequacies in global COVID-19 response and control efforts. *Int. J. Infect. Dis.* **2022**, *114*, 268–272. [CrossRef] [PubMed]
2. Patel, K.P.; Vunnam, S.R.; Patel, P.A.; Krill, K.L.; Korbitz, P.M.; Gallagher, J.P.; Suh, J.E.; Vunnam, R.R. Transmission of SARS-CoV-2: An update of current literature. *Eur. J. Clin. Microbiol. Infect. Dis.* **2022**, *39*, 2005–2011. [CrossRef] [PubMed]
3. Smith, C.; Odd, D.; Harwood, R.; Ward, J.; Linney, M.; Clark, M.; Hargreaves, D.; Ladhani, S.N.; Draper, E.; Davis, P.J.; et al. Deaths in children and young people in England after SARS-CoV-2 infection during the first pandemic year. *Nat. Med.* **2022**, *28*, 185–192. [CrossRef] [PubMed]
4. Sigal, A.; Milo, R.; Jassat, W. Estimating disease severity of Omicron and Delta SARS-CoV-2 infections. *Nat. Rev. Immunol.* **2022**, *22*, 267–269. [CrossRef]

5. Gisselquist, D.; Upham, G.; Potterat, J. Efficiency of Human Immunodeficiency Virus Transmission Through Injections and Other Medical Procedures Evidence, Estimates, and Unfinished Business. *Infect. Control Hosp. Epidemiol.* **2006**, *27*, 944–952. [CrossRef]
6. Edward, A. Nardel. Transmission and Institutional Infection Control of Tuberculosis. *Cold Spring Harb. Perspect. Med.* **2016**, *6*, a018192.
7. Nazaroff, W.W.; Nicas, M.; Miller, S.L. Framework for Evaluating Measures to Control Nosocomial Tuberculosis Transmission. *IndoorAir* **1998**, *8*, 205–218. [CrossRef]
8. Cleveland, J.L.; Robison, V.A.; Panlilio, A.L. Tuberculosis Epidemiology, Diagnosis and Infection Control Recommendations for Dental Settings: An Update on the Centers for Disease Control and Prevention Guidelines. *J. Am. Dent. Assoc.* **2009**, *140*, 1092–1099. [CrossRef]
9. Inungu, J.; Iheduru-Anderson, K.; Odio, O.J. Recurrent Ebolavirus disease in the Democratic Republic of Congo: Update and challenges. *AIMS Public Health* **2019**, *6*, 502–513. [CrossRef]
10. Cenciarelli, O.; Pietropaoli, S.; Malizia, A.; Carestia, M.; D'Amico, F.; Sassolini, A.; di Giovanni, D.; Rea, S.; Gabbarini, V.; Tamburrini, A.; et al. Ebola Virus Disease 2013-2014 Outbreak in West Africa: An Analysis of the Epidemic Spread and Response. *Int. J. Microbiol.* **2015**, *2015*, 769121. [CrossRef]
11. Oluwagbemi, O.O.; Oluwagbemi, F.; Abimbola, O. Ebinformatics: Ebola Fuzzy Informatics Systems on the diagnosis, prediction, and recommendation of appropriate treatments for Ebola Virus Disease (EVD). *Inform. Med. Unlocked* **2016**, *2*, 12–37. [CrossRef]
12. Oluwagbemi, O.O.; Fornadel, C.M.; Adebiyi, E.F.; Norris, D.E.; Rasgon, J. AnoSpEx: A Stochastic, Spatially-Explicit Computational Model for Studying *Anopheles* Metapopulation Dynamics. *PLoS ONE* **2013**, *8*, e68040. [CrossRef] [PubMed]
13. Oluwagbemi, O. *A Stochastic Computational Model for Anopheles Metapopulation Dynamics: Towards Malaria Control and Insight for Possible Eradication*; VDM Publishing Group: Riga, Latvia, 2013; pp. 1–184. ISBN 978-3-659-41990-4.
14. Oluwagbemi, O.O.; Oladipo, E.K.; Kolawole, O.M.; Oloke, J.K.; Adelusi, T.I.; Irewolede, B.A.; Dairo, E.O.; Ayeni, A.E.; Kolapo, K.T.; Akindiya, O.E.; et al. Bioinformatics, Computational Informatics and Modeling Approaches to the Design of mRNA COVID-19 vaccine candidates, the Special Issue Computation to Fight SARS-CoV-2 (COVID-19). *Computation* **2022**, *10*, 117. [CrossRef]
15. Oluwagbemi, O.O.; Oladipo, E.K.; Dairo, E.O.; Ayeni, A.E.; Irewolede, B.A.; Oyewole, M.P.; Ogunrinde, B.H.; Olawale, B.M.; Adegoke, H.M.; Jimah, E.M.; et al. Computational construction of a glycoprotein multi-epitope subunit vaccine candidate for old and new South-African SARS-CoV-2 virus strains. *Inform. Med. Unlocked J.* **2022**, *28*, 100845. [CrossRef] [PubMed]
16. Stonbraker, S.; Cho, H.; Hermosi, G.; Pichon, A.; Schnall, R. Usability Testing of a mHealth App to Support Self-Management of HIV-Associated Non-AIDS Related Symptoms. *Stud. Health Technol Inform.* **2018**, *250*, 106–110.
17. Oluwagbemi, O.O.; Oluwagbemi, F.E.; Jatto, A.; Hui, C. MAVSCOT: A fuzzy logic-based HIV diagnostic system with indigenous multi-lingual interfaces for rural Africa. *PLoS ONE* **2020**, *15*, e0241864. [CrossRef]
18. Nasution, M.A.F.; Toepak, E.P.; Alkaff, A.H.; Tambunan, U.S.F. Flexible docking-based molecular dynamics simulation of natural product compounds and Ebola virus Nucleocapsid (EBOV NP): A computational approach to discover new drug for combating Ebola. *BMC Bioinform.* **2018**, *19*, 419. [CrossRef]
19. Tom-Aba, D.; Olaleye, A.; Olayinka, A.T.; Nguku, P.; Waziri, N.; Adewuyi, P.; Adeoye, O.; Oladele, S.; Adeseye, A.; Oguntimehin, O.; et al. Innovative Technological Approach to Ebola Virus Disease Outbreak Response in Nigeria Using the Open Data Kit and Form Hub Technology. *PLoS ONE* **2015**, *10*, e0131000. [CrossRef]
20. Oluwagbemi, O.; Olaitan, A. A Comparative Computational Genomics of Ebola Virus Disease Strains: In-silico Insight for Ebola Control. *Inform. Med. Unlocked* **2018**, *12*, 106–119. [CrossRef]
21. Rahi, M.; Sharma, R.; Saroha, P.; Chaturvedi, R.; Bharti, P.K.; Sharma, A. Polymerase Chain Reaction-Based Malaria Diagnosis Can Be Increasingly Adopted during Current Phase of Malaria Elimination in India. *Am. J. Trop. Med. Hyg.* **2022**, tpmd210966. [CrossRef]
22. Oluwagbemi, O.; Oluwagbemi, F.; Fagbore, O. Malavefes A computational fuzzy voice-enabled anti-malarial drug informatics software for correct dosage prescription of anti-malaria drugs. *J. King Saud Univ. Comput. Inf. Sci.* **2017**, *30*, 185–197. [CrossRef]
23. Kuehn, B.M. Vaccine Development Is Charting a New Path in Malaria Control. *JAMA* **2022**, *327*, 510–512. [CrossRef] [PubMed]
24. Magagula, V.; Odhiambo, J.N.; Oluwagbemi, O.; Pandey, S.; Rerolle, F.; Van Ness, S. Modeling the impact of clinical Immunity on Malaria Infection: Insight from the Garki Project. In *A Technical Report for the (ICI3D) of the International Clinic on the Meaningful Modeling of Epidemiological Data (ICI3D)*; 2018 Workshop; African Institute of Mathematical Sciences (AIMS): Cape Town, South Africa, 2018.
25. Owen, B.N.; Winkel, M.; Bonnington, C.; Nuwa, A.; Achan, J.; Opigo, J.; Tibenderana, J.K.; Pothin, E. Dynamical malaria modeling as a tool for bold policymaking. *Nat. Med.* **2022**, *28*, 610–611. [CrossRef] [PubMed]
26. Li, Y.; Stewart, K.; Han, K.T.; Han, Z.Y.; Aung, P.P.; Thein, Z.W.; Htay, T.; Chen, D.; Nyunt, M.M.; Plowe, C.V. Understanding spatio-temporal human mobility patterns for malaria control using a multi-agent mobility simulation model. *Clin. Infect. Dis.* **2022**, ciac568. [CrossRef]
27. Burgert, L.; Reiker, T.; Golumbeanu, M.; Möhrle, J.J.; Penny, M.A. Model-informed target product profiles of long-acting-injectables for use as seasonal malaria prevention. *PLoS Glob. Public Health* **2022**, *2*, e0000211. [CrossRef]
28. Ngowo, H.S.; Okumu, F.O.; Hape, E.E.; Mshani, I.H.; Ferguson, H.M.; Matthiopoulos, J. Using Bayesian state-space models to understand the population dynamics of the dominant malaria vector, *Anopheles* funestus in rural Tanzania. *Malar. J.* **2022**, *21*, 161. [CrossRef]

29. Oluwagbemi, O.; Clarence, S.Y. Computational Predictive Framework towards the Control and Reduction of Malaria incidences in Africa. *Egypt. Comput. Sci. J.* **2012**, *36*, 1–17.
30. Oluwagbemi, O.; Adewumi, A.; Misra, S.; Mafodkm, L.M. Mobile Application Framework for the management of Omics Data and Knowledge Mining. *J. Phys. Conf. Ser.* **1566**, *2020*, 012132. [CrossRef]
31. Daramola, J.O.; Osamor, V.C.; Oluwagbemi, O.O. A Grid based framework For Pervasive HealthCare Using Wireless Sensors Networks: A Case for Developing Nations. *Asian J. Inf. Technol.* **2008**, *7*, 260–267.
32. Oluwagbemi, O.; Ojutalayo, T.; Nwinyi, O. Development of a Secured Information System to Manage Malaria Related Cases in Southwestern region of Nigeria. *Egypt. Comput. Sci. J.* **2010**, *34*, 23–34.
33. Oluwagbemi, O.; Ofoezie, U.; Nwinyi, O. A Knowledge-based data mining system for diagnosing malaria related cases in Healthcare Management. *Egypt. Comput. Sci. J.* **2010**, *34*, 23–34.
34. Oluwagbemi, O.; Adeoye, E.T.; Fatumo, S. Building a Computer-Based Expert System for Malaria Environmental Diagnosis: An Alternative Malaria Control Strategy. *Egypt. Comput. Sci. J.* **2009**, *33*, 55–69.
35. Oluwagbemi, O.; Oladunni, B. Diagnosis and Recommender System for Some Neglected Tropical Diseases. *Int. J. Nat. Appl. Sci.* **2010**, *6*, 181–188.
36. Fad, S.A.; Revels, J.W.; Gharai, L.R.; Hanneman, K.; Dana, F.; Proffitt, E.K.; Grizzard, J.D. Cardiac MRI of Hereditary Cardiomyopathy. *RadioGraphics* **2022**. [CrossRef]
37. Oluwagbemi, O.; Oluwagbemi, F.; Ughamadu, C. Android Mobile Informatics Application for some Hereditary Diseases and Disorders (AMAHD): A Complementary framework for medical practitioners and patients. *Inform. Med. Unlocked* **2016**, *2*, 38–69. [CrossRef]
38. Qin Yuan, Q.; Wen, Z.; Yang, K.; Zhang, S.; Zhang, N.; Song, Y.; Chen, F. Identification of Key CircRNAs Related to Pulmonary Tuberculosis Based on Bioinformatics Analysis. *BioMed Res. Int.* **2022**, *2022*, 1717784. [CrossRef]
39. Peter van Heusden, P.; Mashologu, Z.; Lose, T.; Warren, R.; Christoffels, A. The COMBAT-TB Workbench: Making Powerful Mycobacterium tuberculosis Bioinformatics Accessible. *mSphere* **2022**, *7*, e00991-21. [CrossRef]
40. Harouna Hamidou, Z.; Morsli, M.; Mamadou, S.; Drancourt, M.; Saad, J. Emergence of multi-drug-resistant Mycobacterium tuberculosis in Niger: A snapshot based on whole-genome sequencing. *PLoS Negl. Trop. Dis.* **2022**, *16*, e0010443. [CrossRef]
41. Tomas, N.; Leonelli, D.; Campoy, M.; Marthey, S.; Le, N.H.; Rengel, D.; Martin, V.; Pál, A.; Korduláková, J.; Eynard, N.; et al. Bioinformatic Mining and Structure-Activity Profiling of Baeyer-Villiger Monooxygenases from Mycobacterium tuberculosis. *Msphere* **2022**, *7*, e00482-21. [CrossRef]
42. Ishikawa, S.; Ozeki, Y.; Suga, S.; Mukai, Y.; Kobayashi, H.; Inouchi, E.; Kaboso, S.A.; Gebretsadik, G.; Dewi, D.N.S.S.; Nishiyama, A.; et al. Monitoring IgG against Mycobacterium tuberculosis proteins in an Asian elephant cured of tuberculosis that developed from long-term latency. *Sci. Rep.* **2022**, *12*, 4310. [CrossRef]
43. Sweileh, W.M. Global research activity on mathematical modeling of transmission and control of 23 selected infectious disease outbreak. *Glob. Health* **2022**, *18*, 4. [CrossRef] [PubMed]
44. Oluwagbemi, O.; Ogeh, D.; Adewumi, A.; Fatumo, S. Computational and Mathematical Modeling: Applicability to Infectious Disease Control in Africa. *Asian J. Sci. Res.* **2016**, *9*, 88–105. [CrossRef]
45. Ojo, M.M.; Gbadamosi, B.; Benson, T.O.; Adebimpe, O.; Georgina, A.L. Modeling the dynamics of Lassa fever in Nigeria. *J. Egypt Math. Soc.* **2021**, *29*, 16. [CrossRef]
46. Birger, R.; Antillón, M.; Bilcke, J.; Dolecek, C.; Dougan, G.; Pollard, A.J.; Neuzil, K.M.; Frost, I.; Laxminarayan, R.; Pitzer, V.E. Estimating the effect of vaccination on antimicrobial-resistant typhoid fever in 73 countries supported by Gavi: A mathematical modelling study. *Lancet Infect. Dis.* **2022**, *22*, 679–691. [CrossRef] [PubMed]
47. Peter, O.J.; Abioye, A.I.; Oguntolu, F.A.; Owolabi, T.A.; Ajisope, M.O.; Zakari, A.G.; Shaba, T.G. Modelling and optimal control analysis of Lassa fever disease. *Inform. Med. Unlocked* **2020**, *20*, 100419. [CrossRef]
48. Grant, R.; Nguyen, L.L.; Breban, R. Modelling human-to-human transmission of monkeypox. *Bull. World Health Organ.* **2020**, *98*, 638–640. [CrossRef]
49. Isidro, J.; Borges, V.; Pinto, M.; Sobral, D.; Santos, J.D.; Nunes, A.; Mixão, V.; Ferreira, R.; Santos, D.; Duarte, S.; et al. Phylogenomic characterization and signs of microevolution in the 2022 multi-country outbreak of monkeypox virus. *Nat. Med.* **2022**. [CrossRef]
50. Antinori, A.; Mazzotta, V.; Vita, S.; Carletti, F.; Tacconi, D.; Lapini, L.E.; D'Abramo, A.; Cicalini, S.; Lapa, D.; Pittalis, S.; et al. Epidemiological, clinical and virological characteristics of four cases of monkeypox support transmission through sexual contact, Italy. *Eurosurveillance* **2022**, *27*, 2200421. [CrossRef] [PubMed]
51. Vandenbogaert, M.; Kwasiborski, A.; Gonofio, E.; Descorps-Declère, S.; Selekon, B.; Nkili Meyong, A.A.; Ouilibona, R.S.; Gessain, A.; Manuguerra, J.C.; Caro, V.; et al. Nanopore sequencing of a monkeypox virus strain isolated from a pustular lesion in the Central African Republic. *Sci. Rep.* **2022**, *12*, 10768. [CrossRef]
52. WHO. 2022. Available online: https://www.who.int/europe/news/item/23-07-2022-who-director-general-declares-the-ongoing-monkeypox-outbreak-a-public-health-event-of-international-concern (accessed on 30 July 2022).
53. Gispen, R. Smallpox reinfections in Indonesia. *Ned T Geneesk* **1949**, *93*, 3686–3695.
54. Arita, I.; Henderson, D.A. Smallpox and monkeypox in non-human primates. *Bull. World Health Organ.* **1968**, *39*, 277. [PubMed]
55. Magnus, P.V.; Andersen, E.K.; Petersen, K.B.; Birch-Andersen, A. A pox-like disease in cynomolgus monkeys. *Acta Pathol. Microbiol. Scand.* **1959**, *46*, 156–176. [CrossRef]

56. Shchelkunov, S.N.; Totmenin, A.V.; Babkin, I.V.; Safronov, P.F.; Ryazankina, O.I.; Petrov, N.A.; Gutorov, V.V.; Uvarova, E.A.; Mikheev, M.V.; Sisler, J.R.; et al. Human monkeypox and smallpox viruses: Genomic comparison. *FEBS Lett.* **2001**, *509*, 66–70. [CrossRef] [PubMed]
57. Alakunle, E.; Moens, U.; Nchinda, G.; Okeke, M.I. Monkeypox virus in Nigeria: Infection biology, epidemiology, and evolution. *Viruses* **2020**, *12*, 1257. [CrossRef] [PubMed]
58. Weaver, J.R.; Isaacs, S.N. Monkeypox virus and insights into its immunomodulatory proteins. *Immunol. Rev.* **2008**, *225*, 96–113. [CrossRef]
59. Ježek, Z.; Fenner, F. *Hum. Monkeypox*; S. Karger Ag: Basel, Switzerland, 1988; Volume 17.
60. Ngambika, G.K.; Mbembo-wa-Mbembo, B.; Séraphin, K.P.; Fabrice, K.K.; Bongo, G.N.; Colette, M.A.; Ruphin, D.D. First Report on Three Cases of Monkey pox in Nord Ubangi Province (Democratic Republic of the Congo). *Br. Int. Exact Sci. (BIoEx) J.* **2020**, *2*, 120–125.
61. Damon, I.K. Status of human monkeypox: Clinical disease, epidemiology and research. *Vaccine* **2011**, *29*, D54–D59. [CrossRef]
62. Marennikova, S.S.; Šeluhina, E.M.; Mal'Ceva, N.N.; Čimiškjan, K.L.; Macevič, G.R. Isolation and properties of the causal agent of a new variola-like disease (monkeypox) in man. *Bull. World Health Organ.* **1972**, *46*, 599.
63. Di Giulio, D.B.; Eckburg, P.B. Human monkeypox: An emerging zoonosis. *Lancet Infect. Dis.* **2004**, *4*, 15–25. [CrossRef]
64. Bunge, E.M.; Hoet, B.; Chen, L.; Lienert, F.; Weidenthaler, H.; Baer, L.R.; Steffen, R. The changing epidemiology of human monkeypox—A potential threat? A systematic review. *PLoS Negl. Trop. Dis.* **2022**, *16*, e0010141. [CrossRef]
65. Rimoin, A.W.; Mulembakani, P.M.; Johnston, S.C.; Lloyd Smith, J.O.; Kisalu, N.K.; Kinkela, T.L.; Blumberg, S.; Thomassen, H.A.; Pike, B.L.; Fair, J.N.; et al. Major increase in human monkeypox incidence 30 years after smallpox vaccination campaigns cease in the Democratic Republic of Congo. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16262–16267. [CrossRef] [PubMed]
66. Gilsdorf, J.R.; Zilinskas, R.A. New considerations in infectious disease outbreaks: The threat of genetically modified microbes. *Clin. Infect. Dis.* **2005**, *40*, 1160–1165. [CrossRef] [PubMed]
67. Kuhn, J.H.; Leitenberg, M. The Soviet biological warfare program. In *Biological Threats in the 21st Century: The Politics, People, Science and Historical Roots*; Harvard University Press: Cambridge, MA, USA; London, UK, 2016; pp. 79–102.
68. Tucker, J.B. Biological weapons in the former Soviet Union: An interview with Dr. Kenneth Alibek. *Nonproliferation Rev.* **1999**, *6*, 1–10. [CrossRef]
69. Henderson, D.A.; Inglesby, T.V.; Bartlett, J.G.; Ascher, M.S.; Eitzen, E.; Jahrling, P.B.; Hauer, J.; Layton, M.; McDade, J.; Osterholm, M.T.; et al. Smallpox as a biological weapon: Medical and public health management. *JAMA* **1999**, *281*, 2127–2137. [CrossRef]
70. Cieslak, T.J.; Kortepeter, M.G.; Wojtyk, R.J.; Jansen, H.J.; Reyes, R.A.; Smith, J.O.; NATO Biological Medical Advisory Panel. Beyond the dirty dozen: A proposed methodology for assessing future bioweapon threats. *Mil. Med.* **2018**, *183*, e59–e65. [CrossRef] [PubMed]
71. Kozlov, M. Monkeypox goes global: Why scientists are on alert. *Nature* **2022**, *606*, 15–16. [CrossRef]
72. Green, M.S.; Swartz, T.; Mayshar, E.; Lev, B.; Leventhal, A.; Slater, P.E.; Shemer, J. When is an epidemic an epidemic? *Isr. Med. Assoc. J. IMAJ* **2002**, *4*, 3–6.
73. Assessment, R.R. Monkeypox Multi-Country Outbreak. 2022. Available online: https://www.ecdc.europa.eu/en/publications-data/risk-assessment-monkeypox-multi-country-outbreak (accessed on 30 June 2022).
74. World Health Organization. *Surveillance, Case Investigation and Contact Tracing for Monkeypox: Interim Guidance*; World Health Organization: Geneva, Switzerland, 22 May 2022; No. WHO/MPX/Surveillance/2022.1.
75. Makridakis, S.; Wakefield, A.; Kirkham, R. Predicting medical risks and appreciating uncertainty. *Foresight Int. J. Appl. Forecast.* **2019**, *1*, 28–35.
76. Meraj, T.; Rauf, H.T.; Zahoor, S.; Hassan, A.; Lali, M.I.; Ali, L.; Bukhari, S.A.C.; Shoaib, U. Lung nodules detection using semantic segmentation and classification with optimal features. *Neural Comput. Appl.* **2021**, *33*, 10737–10750. [CrossRef]
77. Roda, W.C.; Varughese, M.B.; Han, D.; Li, M.Y. Why is it difficult to accurately predict the COVID-19 epidemic? *Infect. Dis. Model.* **2020**, *5*, 271–281. [CrossRef] [PubMed]
78. Hughes, C.M.; Liu, L.; Davidson, W.B.; Radford, K.W.; Wilkins, K.; Monroe, B.; Metcalfe, M.G.; Likafi, T.; Lushima, R.S.; Kabamba, J.; et al. A Tale of Two Viruses: Coinfections of Monkeypox and Varicella Zoster Virus in the Democratic Republic of Congo. *Am. J. Trop. Med. Hyg.* **2021**, *104*, 604. [CrossRef]
79. Lash, R.R.; Carroll, D.S.; Hughes, C.M.; Nakazawa, Y.; Karem, K.; Damon, I.K.; Peterson, A.T. Effects of georeferencing effort on mapping monkeypox case distributions and transmission risk. *Int. J. Health Geogr.* **2012**, *11*, 23. [CrossRef] [PubMed]
80. Nolen, L.D.; Osadebe, L.; Katomba, J.; Likofata, J.; Mukadi, D.; Monroe, B.; Doty, J.; Hughes, C.M.; Kabamba, J.; Malekani, J.; et al. Extended human-to-human transmission during a monkeypox outbreak in the Democratic Republic of the Congo. *Emerg. Infect. Dis.* **2016**, *22*, 1014. [CrossRef] [PubMed]
81. Liu, Y.; Jain, A.; Eng, C.; Way, D.H.; Lee, K.; Bui, P.; Kanada, K.; de Oliveira Marinho, G.; Gallegos, J.; Gabriele, S.; et al. A deep learning system for differential diagnosis of skin diseases. *Nat. Med.* **2020**, *26*, 900–908. [CrossRef] [PubMed]
82. Tom, J.J.; Anebo, N.P. A Neuro-Fussy Based Model for Diagnosis of Monkeypox Diseases. *Int. J. Comput. Sci. Trends Technol. (IJCST)* **2018**, *6*, 143–153.
83. Beer, E.M.; Rao, V.B. A systematic review of the epidemiology of human monkeypox outbreaks and implications for outbreak strategy. *PLoS Neglect. Trop. Dis.* **2019**, *13*, e0007791. [CrossRef]

84. Barrow, D.K.; Crone, S.F. A comparison of AdaBoost algorithms for time series forecast combination. *Int. J. Forecast.* **2016**, *32*, 1103–1119. [CrossRef]
85. Oyewola, D.O.; Dada, E.G.; Misra, S.; Damaševičius, R. Predicting COVID-19 Cases in South Korea with All K-Edited Nearest Neighbors Noise Filter and Machine Learning Techniques. *Information* **2021**, *12*, 528. [CrossRef]
86. Oyewola, D.O.; Dada, E.G.G.; Olaoluwa, O.E.; Al-Mustapha, K.A. Predicting Nigerian stock returns using technical analysis and machine learning. *Eur. J. Electr. Comput. Eng.* **2019**, *3*, 1–8. [CrossRef]
87. Oyewola, D.; Hakimi, D.; Adeboye, K.; Shehu, M.D. Using five machine learning for breast cancer biopsy predictions based on mammographic diagnosis. *Int. J. Eng. Technol. IJET* **2016**, *2*, 142–145. [CrossRef]
88. Bilginol, K.; Denli, H.H.; Şeker, D.Z. Ordinary least squares regression method approach for site selection of automated teller machines (ATMs). *Procedia Environ. Sci.* **2015**, *26*, 66–69. [CrossRef]
89. Oyewola, D.O.; Dada, E.G.; Omotehinwa, O.O.; Ibrahim, I.A. Comparative Analysis of Linear, Non-Linear and Ensemble Machine Learning Algorithms for Credit Worthiness of Consumers. *Comput. Intell. Wirel. Sens. Netw.* **2019**, *1*, 1–11.
90. Al-Hassan, Y.M. Performance of a new ridge regression estimator. *J. Assoc. Arab Univ. Basic Appl. Sci.* **2010**, *9*, 23–26. [CrossRef]
91. Oyewola, D.O.; Dada, E.G. Exploring machine learning: A scientometrics approach using bibliometrix and VOSviewer. *SN Appl. Sci.* **2022**, *4*, 143. [CrossRef] [PubMed]
92. Dada, E.G.; Yakubu, H.J.; Oyewola, D.O. Artificial Neural Network Models for Rainfall Prediction. *Eur. J. Electr. Eng. Comput. Sci.* **2021**, *5*, 30–35. [CrossRef]
93. Dada, E.G.; Oyewola, D.O.; Joseph, S.B.; Duada, A.B. Ensemble Machine Learning Model for Software Defect Prediction. *Adv. Mach. Learn. Artif. Intell.* **2021**, *2*, 11–21. [CrossRef]
94. Oyewola, D.O.; Ibrahim, A.; Kwanamu, J.A.; Dada, E.G. A new auditory algorithm in stock market prediction on oil and gas sector in Nigerian stock exchange. *Soft Comput. Lett.* **2021**, *3*, 100013. [CrossRef]
95. Oyewola, D.O.; Al-Mustapha, K.A.; Ibrahim, A.; Dada, E.G. High-Performing Machine Learning Algorithms for Predicting the Spread of COVID-19. In *Socioeconomic Dynamics of the COVID-19 Crisis*; Springer: Cham, Switzerland, 2022; pp. 371–401. [CrossRef]
96. Oyewola, D.O.; Dada, E.G.; Ndunagu, J.N.; Umar, T.A.; Akinwunmi, S.A. COVID-19 Risk Factors, Economic Factors, and Epidemiological Factors nexus on Economic Impact: Machine Learning and Structural Equation Modelling Approaches. *J. Niger. Soc. Phys. Sci.* **2021**, *3*, 395–405. [CrossRef]
97. Cryptocurrencies. 2022. Available online: https://finance.yahoo.com/cryptocurrencies (accessed on 10 May 2022).
98. Monkeypox Dataset (Daily Updated). 2022. Available online: https://www.kaggle.com/datasets/deepcontractor/monkeypox-dataset-daily-updated (accessed on 25 June 2022).
99. Oyewola, D.O.; Dada, E.G.; Al-Mustapha, K.A.; Fadele, A.A.; Joseph, S.B.; Ibrahim, A. Predicting Transmission Rate of Coronavirus (COVID-19) Pandemic Using Machine Learning Techniques. In *Computational Intelligence for COVID-19 and Future Pandemics*; Springer: Singapore, 2022; pp. 45–69. [CrossRef]