East Journal of Computer Science
ISSN: 3079-9406

Volume 1, Issue 3
Publisher: East Publication & Technology
DOI: https://doi.org/10.63496/ejcs.Vol1.Iss3.96

Predict Covid-19 Pandemic Phases using Several Machine Learning Algorithms

Yasir Hussein Shakir[1*], Eshaq Aziz Awadh AL Mandhari [2], Issa Nasser Salim Al-ghatrifi[3], Ahmed Yahya Awad [4]

[1] College of Graduate Studies (COGS), University Tenaga Nasional (UNITEIN), Kajang, Malaysia, yasserhessein19855@gmail.com.
[2] Graduate School of Technology at Asia Pacific University of Technology and Innovation (APU) in Malaysia, eshaq.almandhari@utas.edu.om.
[3] Department of Information Systems, University of Nizwa, Oman, issa.ghatrifi@unizwa.edu.om.
[3] Department of Engineering and information technology, Al Zahraa University for Women, Iraq, ahmed.yahya.awad@alzahraa.edu.iq.
*Corresponding Author.

Received: 23/04/2025, **Revised**: 09/05/2025, **Accepted**: 13/05/2025, **Published**: 27/05/2025

***Abstract*:**

COVID-19, despite the worldwide reduced mortality and severity, is still a menace all over the world. This paper also provides the outline of a classification of pandemic stages and describes the methods for using machine learning to evaluate present-day pandemic intensity. Some aspects that have been targeted are economic and educational status and health, both physical and mental, concerning different aspects of life in Malaysia. The chosen algorithms for our analysis include Support Vector Machine (SVM), Passive Aggressive (PA), k-Nearest Neighbors (KNN), AdaBoost, Decision Tree (DT), Extra Trees (ET), Random Forest (RF), Gradient Boosting (GB), Histogram-based Gradient Boosting (HGB), and Multilayer Perceptron (MLP). Among them the MLP model achieved the best performance, with an accuracy of 94.50%, precision of 95.22%, recall of 94.50%, and an F1-score of 93.50%. These results suggest that machine learning can be effectively leveraged to identify pandemic stages and inform public health strategies. The proposed application is adaptable to future pandemics and can support data-driven decision-making for mitigating disease spread.

***Keywords*: COVID-19, Machine Learning, Prediction, Healthcare.**

1. Introduction

The coronavirus also known as COVID-19 became a global pandemic on the 11th of March 2020 according to World Health Organization (WHO). COVID-19 is an airborne disease that transmits from one person to another through the respiratory tract by breathing, speaking, laughing, singing, coughing, or sneezing; it has been contagious to date, and it is fought and contained to date through tests, quarantines, and perhaps the most extensive use of hygiene practices and measures such as wearing masks face masks, social distancing, frequent washing of hands, and using sanitizer among others. Among the countries, Middle-Income Countries have been significantly impacted by COVID-19. For instance, overpopulation coupled with a huge number of refugees in Malaysia has been counterproductive due to inadequacy of equipment, human skills, facilities, funds, and awareness [1]. The initial case of contamination by COVID-19 in Malaysia was confirmed on 25th January 2020. Since then, the global figure has risen to 4,848,314 for infected cases, and 36,387 for deaths [2]. Besides, it has engulfed many aspects of people's lives such as socio-economic, mental, job, and study [3,4,5,6,7]. Foreign trade among other economic activities has also felt the heat through a slowdown that is attributable to the crisis [8]. The COVID-19 pandemic quick spread has severely harmed people and caused financial losses. It has also had a big influence on the healthcare system. Early COVID-19 analysis is critically necessary for efficient pandemic response policies and future prediction, which will lessen the burden on the healthcare system and governments [9].

When Three phases were involved in the pandemic:

1. Movement Control Order by Phase MCO (2020): The first case of COVID-19 this year was detected on March 18, 2020. After that, the federal government of Malaysia introduced a sequence of national quarantine and cordon sanitaire measures in response to the COVID-19 pandemic. Restriction orders were primarily called 'lockdowns' in local and global newspapers and television.

2. Recovery Movement Control Order (RMCO) by Phase (2021) 31 March 2021. This year, the government continues to allow activities to be normal in all public and private offices and courts, restaurants, shopping malls, recreation centers, religious places, etc., but with supervision, restrictions, and instructions to follow.

3. National Recovery Plan (NRP) by Phase Terminal (2022) 3 January 2022 Most people returned to their normal lives as they ensured that they took the vaccine. The infection rate is significantly lower than in the previous reviews. and the government has been removing all the restrictions in all sectors, including areas of the economy, future education, and individuals' physical and mental condition.

This study provides various contributions to the problem of pandemic Predict and healthcare analytics:

According to the World Health Organization, COVID-19 vaccination has helped to reduce the severity of the pandemic in Malaysia. This paper identifies five stages of the pandemic in Malaysia, and, using machine learning, identifies which stage of the pandemic is being faced based on personal and national factors in four sectors: literacy level, employment, and both physical and mental well-being. Ten machine learning algorithms were used namely Support vector machine, Passive Aggressive, k-nearest neighbors, AdaBoost, Decision Tree, Extra Trees, Random Forest, Gradient boosting, Histogram Gradient boosting and Multilayer Perceptron.

The remainder of this paper is structured as follows: Section 2 Previous works related to this topic. Section 3 Methodology expounds the dataset and method of this research. Section 4 describes the results and discussion Section. 5 concludes the paper.

2. Related Work

Machine learning applies in various ways to the data from the COVID-19 pandemic. For example, univariate logistic regression (LG) was used to determine the factors affecting youth mental health during the pandemic [10]. A study [11] using DASS-21 and a multilayer perceptron neural network on 339,781 respondents showed high prediction accuracy for depression (76.4%), anxiety (76.3%), and stress (87.4%). Stress and anxiety were key factors highlighting the potential of machine learning for early mental health issue detection. Therefore, they are likely to be similarly affected by the pandemic psychological effects of COVID-19 during lockdowns using the K-means algorithm, demonstrating effectiveness in clustering accuracy and understanding COVID-19's psychological impact [12]. A boosted random forest algorithm has been used to predict the riskiness of single patients' infections [13]. In contrast, this research seeks to predict the severity of an infection. Adaboost algorithms were utilized to classify pandemic growth rates in various regions of Italy, considering geographical differences and containment measures [14]. This study examined COVID-19 topics on Facebook and Twitter employing a phenomenological method (January 1–March 17, 2020). Through an examination of the top 10 influencers and pages, which included the Ministry of Health, the study garnered 6,068 mentions and 51 million interactions. Prior to the MCO, health, political, economic, and policy challenges were identified using thematic analysis. These results provide perceptions of public opinion and attitude, helping health organizations interact with the public and formulate plans to track and stop the spread of COVID-19 [15]. Another [16] study utilized COVID-19 data from the Ministry of Health (MOH) Malaysia (25/1/2020 to 17/6/2022) to applied regression models describing the trends of COVID-19 cases, considering their unpredictable nature and they are using Weka software three techniques were exercised (60:40 / 70:30) split ratios (10 and 20-fold cross-validation) Support Vector Regression (SVR), Multi Linear Regression (MLR), and Random Forest (RF). The findings reveal that RF best the strongest coefficient correlation and the lowest Root Mean Square Error (22.7611) predicted new COVID-19 deaths. In [17], the authors generate rankings of 35 countries from all continents (except Africa) on performance against COVID-19. These indices correctly assess a country's risk by quantifying both propagation and estimated instances. They tracked the evolution of these indicators across numerous countries. In [18], Gompertz, Logistic, and Artificial Neural Networks models are shown for predicted results of COVID-19 infection cases. For the model, some of these were used to make guesses around how many new COVID-19 infections may appear over time. Based on the results, trends could become projected and extrapolated until the ending the epidemic in Mexico. However, their models were limited to predicting the new positive case of COVID-19. In another work [19], the approach is based on decision tree (DT) machine learning to predict the gene sequence of the COVID-19 virus using intrinsic genomic signatures specific to infectious characteristics. In another study [20], the use of machine learning models to predict outcomes among COVID-19 patients. Four common models LASSO, exponential smoothing (ES, linear regression (LR), and (SVM) were evaluated to predict the risk factors of interest for COVID-19. For every single model, we forecast the new infection cases and deaths over the subsequent 10 days by current COVID-19 data. The ES model outperforms all the other models in predicting new cases of infection and it is best according to this comparison. Nowadays, machine learning is a trending area of the fourth industrial revolution [21] that broadly used in real-world problems, data prediction [22], computer vision [23] and this trending field has been awaited in IoT [24], cloud platform [25], and Cyber Security [26].

3. Methodology

This study uses supervised machine learning models to classify the phases of the COVID 19 pandemic. The process involves the acquisition of dataset, cleaning and preparation of the same, carrying out exploration data analysis (EDA), choosing and applying a number of classification algorithms and evaluating them based on standard metrics and the whole procedure is depicted in Figure 1.

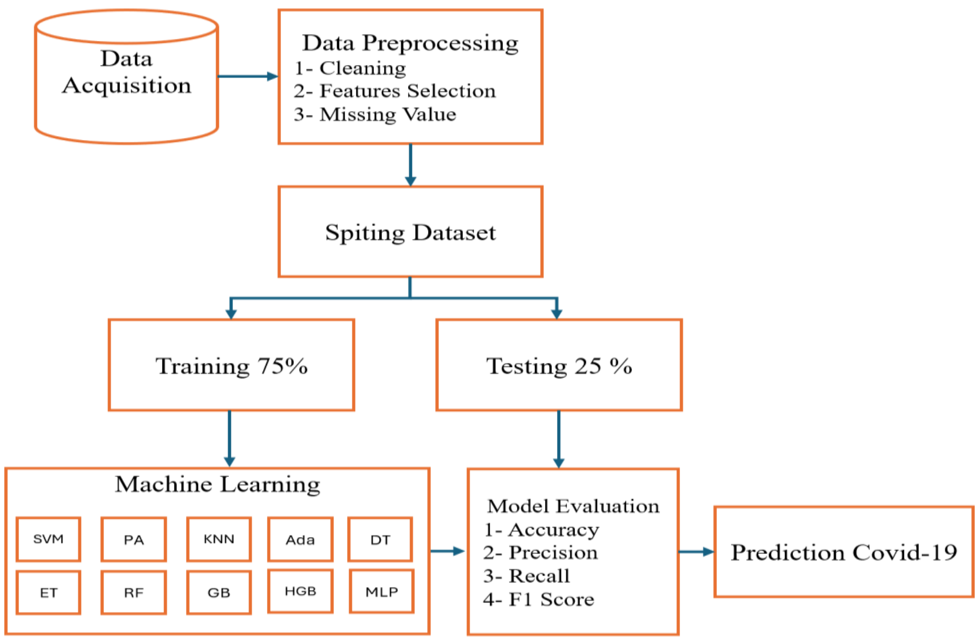

Figure 1: Overall procedure of pandemic Phases.

3.1 Acquisition of Dataset

This paper also uses some data collected from the available sources, Ministry of Health Malaysia (MoH) [13] and additional surveys were conducted particularly for this research. A total of 32 features that were impacted on by the pandemic in Malaysia were enumerated in Table1. the dataset contains various columns related to COVID-19 cases in Malaysia, including sample 1611 of form 2020 to 2024 data instances along with 31 features which are date, Cases new, Cases import, Cases recovered, Cases active, Cases cluster, Cases unvax, Cases pvax, Cases fvax, Cases boost, Cases child, Cases adolescent, Cases adult, Cases elderly, cases_0_4, cases_5_11, cases_12_17, cases_18_29, cases_30_39, cases_40_49, cases_50_59, cases_60_69, cases_70_79, cases_80, Cluster import, Cluster religious, Cluster community, Cluster high Risk, Cluster education, Cluster detention Centre, Cluster workplace as shown Table 1.

Table 1: Feature pandemic in Malaysia

Feature	Description
Date	The date of the recorded data
Cases new	Number of new cases
Cases import	Number of imported cases
Cases recovered	Number of recovered cases

Cases active	Number of active cases
Cases cluster	Number of cluster cases
Cases unvax	Number of unvaccinated cases
Cases pvax	Number of partially vaccinated cases
Cases fvax	Number of fully vaccinated cases
Cases boost	Number of cases with booster shots
Cases child	Number of child cases
Cases adolescent	Number of adolescent cases
Cases adult	Number of adult cases
Cases elderly	Number of elderly cases
cases_0_4	Number of cases in age group 0-4 years
cases_5_11	Number of cases in age group 5-11 years
cases_12_17	Number of cases in age group 12-17 years
cases_18_29	Number of cases in age group 18-29 years
cases_30_39	Number of cases in age group 30-39 years
cases_40_49	Number of cases in age group 40-49 years
cases_50_59	Number of cases in age group 50-59 years
cases_60_69	Number of cases in age group 60-69 years
cases_70_79	Number of cases in age group 70-79 years
cases_80	Number of cases in age group 80+ years
Cluster import	Number of import clusters
Cluster religious	Number of religious clusters
Cluster community	Number of community clusters
Cluster high Risk	Number of high-risk clusters
Cluster education	Number of education clusters
Cluster detention Centre	Number of detention Centre clusters
Cluster workplace	Number of workplace clusters

3.2 Dataset Preprocessing

1. Cleaning

Unnecessary and malformed data entries were removed. The date column was processed to extract the year information, The mentioned above attributes were utilized to predict the pandemic year (2020 to 2024), which was retrieved from the date field. The year of the epidemic is used here to identify the pandemic phase. Which was added as a new feature as shown in sample dataset Table 2.

Table 2. Sample dataset

Index	Date	Year
1	2020-01-25	2020
2	2020-01-26	2020
-	-	-
-	-	-
-	-	-
1609	2024-06-04	2024
1610	2024-06-22	2024

2. Handling Missing Values

In order to guarantee the quality of data, missing values have been defined and deleted as shown in Table 3 after and before handling missing data. An equation that can represent such an equation is the following (1):

$$D' = \{x \in D \mid \forall xi \neq NaN\} \quad (1)$$

Table 3. Handling missed data Before and After

Feature	Missing Values (Before)	Missing Values (After)
cluster_import	449	0
cluster_religious	449	0
cluster_community	449	0
cluster_highRisk	449	0
cluster_education	449	0
cluster_detentionCentre	449	0
cluster_workplace	449	0
All Other Columns	0	0

3.3 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is essential for finding the patterns identifying anomalies and understanding relationships within the dataset. In this study in EDA, the age distribution of cases of COVID-19, the effect of vaccination status (unvaccinated, partially, fully, and boosted), and the incidence of cases of different clusters (workplace, community and education) was analyzed. Visualizations indicated that 18–29 age group had the highest incidence of infections and that workplace clusters dominated case count. These led to the feature selection and design of classification models.

3.4 Spitting Dataset

The dataset, consisting of 1,611 samples, was divided into 75% training and 25% testing parts by using the train_test_split function from Scikit-learn. This strategy creates an opportunity to use much of the data to train the model and still reserve a good number of unseen samples to test model performance, as shown in Table 4.

Table 4.The Split Summary

Dataset Portion	Number of Samples	Percentage
Training Set	1,208	75%
Testing Set	403	25%
Total	1,611	100%

3.5 Machine Learning Algorithms (MLA)

We selected ten machine learning algorithms (SVM, PA, KNN, AdaBoost, DT, ET, RF, GB, HGB, and MLP) to facilitate a comprehensive comparison across a wide range of models including linear models (SVM, PA), tree-based methods (DT, RF, GB, HGB, ET), ensemble methods (AdaBoost); instance-based learning (KNN), and neural networks (MLP). This wide range enables us to assess performance according to various principle of algorithms and to determine the best model applicable for hypothesis about pandemic phases as shown in Table 5 parameters algorithms.

3.4.1. Support Vector Machine

SVM is a form of a supervised learning algorithm that is best used for classification with some form of regression. It is used for identifying the best hyperplane that provides the maximum margin between the various classes within feature space [27].

3.4.2. Passive Aggressive

In the large ranges of algorithms of machine learning, the Passive Aggressive algorithm can be initially classified as an online learning algorithm. It only alters the model when a particular input would have been classified incorrectly, it is passive when it gets it right and aggressive when it is wrong [28].

3.4.3. k-Nearest Neighbors

K-NN is a straightforward, instance based learning technique which decides the class of a sample, based on the most frequent class among its ‘K nearest neighbors in the feature space [29].

3.4.4. Adaptive Boosting

AdaBoost is the abbreviation of Adaptive Boosting which is an instance of the boosting algorithms to build a strong classifier from a lot of weak classifiers. It modifies the weights of misclassified instances to boost the performance of the subsequent classifiers [30].

3.4.5. Decision Tree

Decision Tree is a type of supervised learning technique and is also classified non-parametric in nature used for both classification and regression techniques. They divide the data into several subsets depending on the result of computations on the input variables and construct a tree like structure of decisions [31].

3.4.6. Extra Trees

Extra Trees or Extremely Randomized Trees is a type of ensemble, which creates more trees using random splits of features. It shrinks the variance by taking the average of all the trees [32].

3.4.7. Random Forest

Random forest is a type of ensemble learning technique which builds several numbers of decision trees during the training process and result of each tree the mode of the classes for classification or mean for regression [19].

3.4.8. Gradient Boosting

It is modeled based on boosting and builds sequentially so that each model, in turn, learns from the previous model’s mistakes. It uses one of the operations of gradient descant to reduce the amount reflected in the loss function [34].

3.4.9. Histogram Gradient Boosting

Histogram Gradient Boosting is an improvement of Gradient Boosting, where instead of using the exact computation of split points, histograms are used to improve the performance. It is efficient for big data especially where a result from an analysis of the dataset is required [35].

3.4.10. Multilayer Perceptron

MLP can be described as feedforward artificial neural networks belonging to a specific class. This comprises many layers of the nodes where each layer is a fully connected node with the nodes of the next layer. The former is used for classification whereas the latter for regression analysis is termed as [36].

Table 5. The Parameters machine learning algorithms

Classifier	Parameters
Support Vector Machine	kernel="linear", gamma='auto', random_state=10
Passive Aggressive	random_state=10
K-Nearest Neighbors	n_neighbors=36
Adaptive Boosting	n_estimators=500, learning_rate=0.1
Decision Tree	criterion="entropy", max_depth=3
Extra Trees	n_estimators=100, random_state=10
Random Forest	n_estimators=1000, max_depth=2, random_state=5
Gradient Boosting	n_estimators=100, learning_rate=1.0, max_depth=1, random_state=10
HistGradient Boosting	random_state=10
Multilayer Perceptron	random_state=10, max_iter=3000

3.6 Model Evaluation

Different classifier performance assessments occurred during this phase utilizing the parameters described below parameters. The performance metrics provided an overview of model capabilities so users can select an appropriate model for their needs while evaluating the metrics include:

1- The accuracy measure defines the proportion between correct classed instances and the full data set illustration.

$$Accracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

2- The rate of positive instances between the total positive and the true positive instance among the positive ones.

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

3- The recall represents sensitivity by providing information about the number of correct true positive cases among all actual positive cases.

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

4. The F1 Score provides an equilibrium between precision and recall measurements through their single 'F-measure'.

$$F1 - score = \frac{2 \times Precision \ \times Recall}{Precision + Recall} \quad (5)$$

5. Results and Discussion

5.1 Exploratory Data Analysis (EDA)

Figure 2 illustrates the case distribution based on different age groups for COVID 19 cases. The visualization gives insight on how age brackets were affected during the pandemic. The 18–29 group recorded the most cases, a fact that suggests that it was the most affected part of the population. On the other hand, the elderly individuals aged 80 years and over had the fewest case reports which could be explained by stringent health regimes or less exposure.

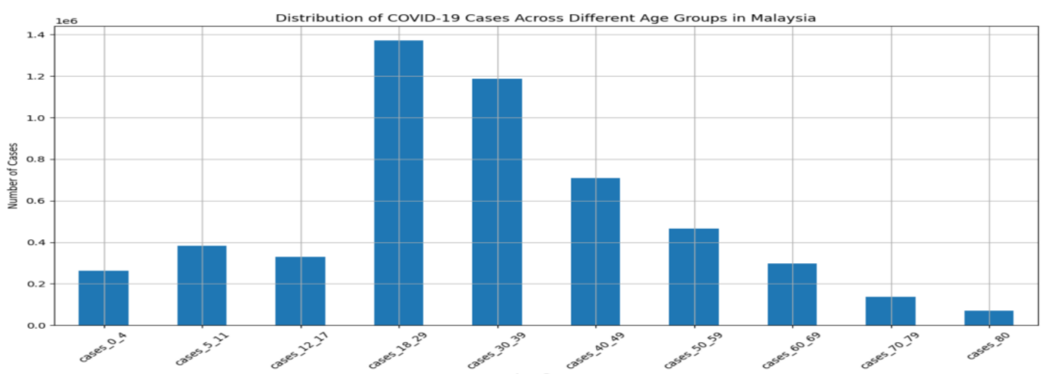

Figure 2: Visualize the distribution of cases across different age groups

Figure 3 compares new confirmed COVID-19 cases compared to total confirmed cases. The data pointed to a worrisome trend in 2022, during which new infections skyrocketed above confirmed cases. This means that 2022 was especially harsh with respect to COVID-19 spread.

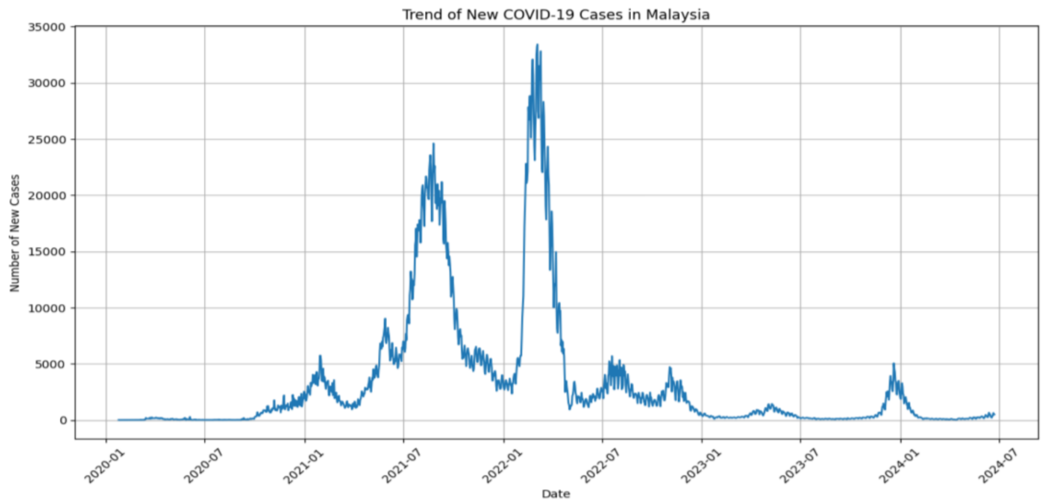

Figure 3: Trend of New COVID-19

Figure 4 shows the distribution of individuals based on their vaccination status in Malaysia. The data classifies people as: Unvaccinated, Fully Vaccinated and Boosted; Of these, there were greater increases, among those given booster doses, followed by those who were fully vaccinated. This tendency shows the increasing public health movement seeking to improve immunity via booster initiatives.

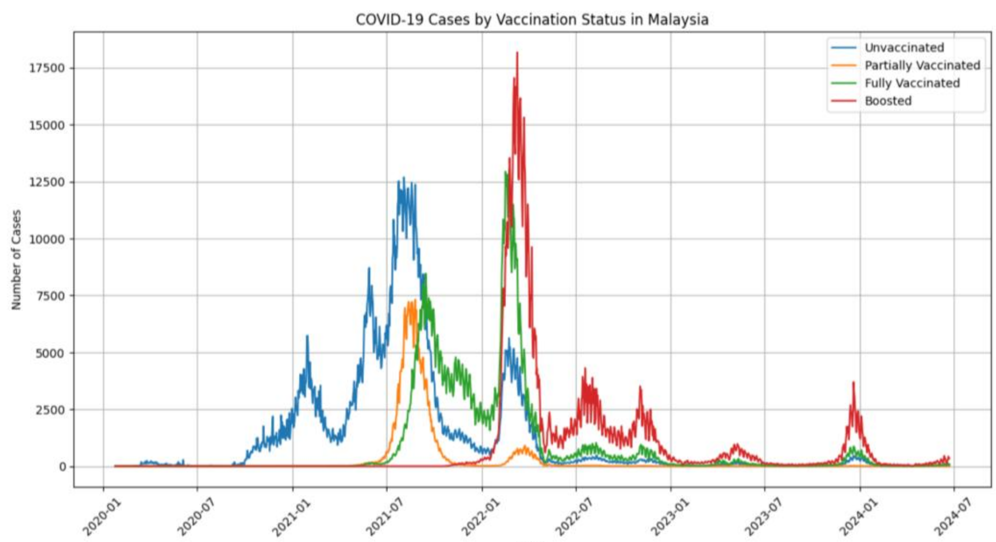

Figure 4: Vaccination Status in Malaysia

In Figure 5 illustrates the division of the COVID-19 cases by the cluster types. Workplace clusters developed as the leading perspective of the infections, highlighting the importance of strict health means at the workplace settings. Community clusters were also a significant contributor to the number of cases, which supported the role of community level intervention. On the other hand, clusters associated with religious events and international travel reported relatively few cases, which could have been caused by effective control policies. Educational and detention center clusters recorded moderate infection rates, which represent a need for enhanced preventive measures for such environments.

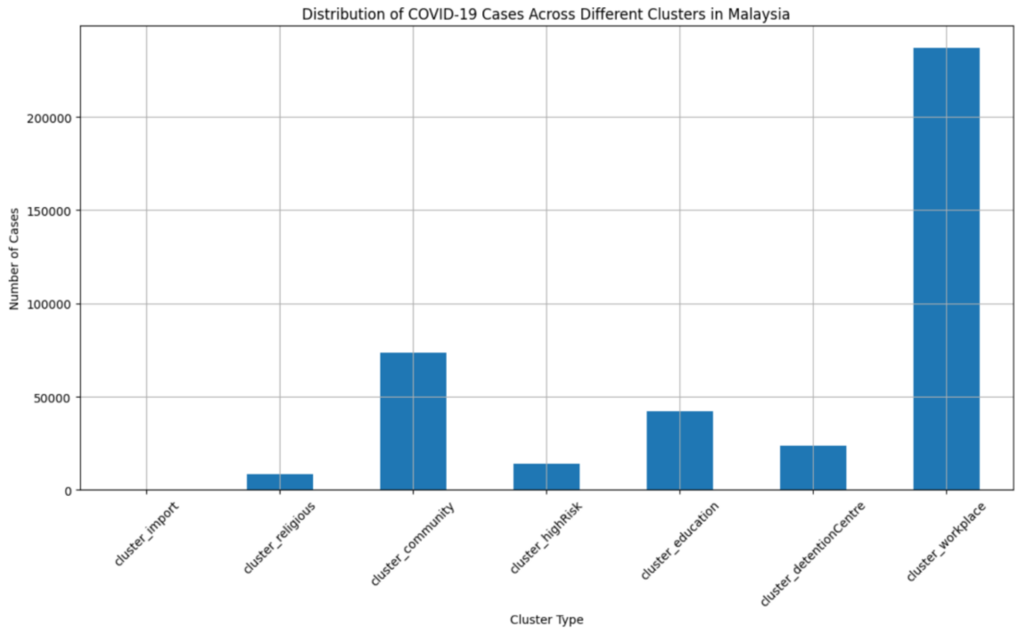

Figure 5: Distribution of COVID-19 cases across different cluster types in Malaysia

5.2 Comparative Analysis of Classification Models

Based In this part, a comparative analysis of ten diverse classification models is carried out in order to assess their ability to predict outcomes using the given dataset. The measures of evaluation used are Precision, Recall, F1-Score and Accuracy, which provide a broad picture of every model's predictive performance. The Figure 6 shows the precision performance of all ten classifiers. MLP classifier observed the best precision i.e., 95.22% followed closely by HistGradientBoosting 94.92% and Gradient Boosting 93.33%. The Passive Aggressive classifier also did well with a score of 93.78%. On the other hand, AdaBoost recorded the least precision of 82.18%, meaning more magnitudes of false positives than other models.

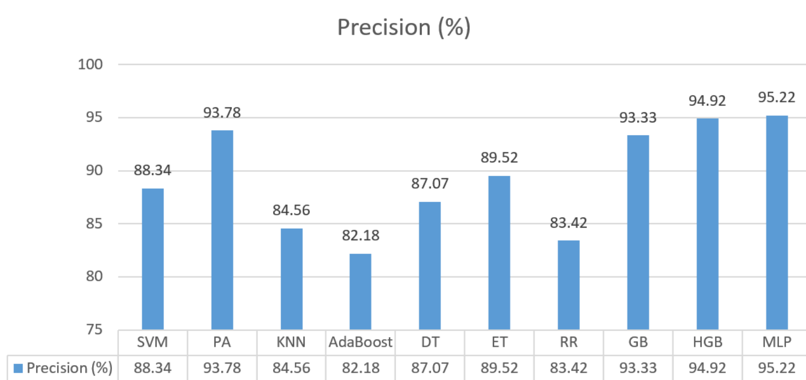

Figure 6: The Precision Performance

Synchronization depicted in the Figure 7 also showed the MLP model achieving the highest recall at 94.5%, showing that it was quite able to identify the real positives. HistGradientBoosting and Extra Trees came close at

94.16% while Gradient Boosting succeeded at 93.47%. The lowest recall was reported by AdaBoost 81.79% showing its failure to positively identify a good number of positive cases.

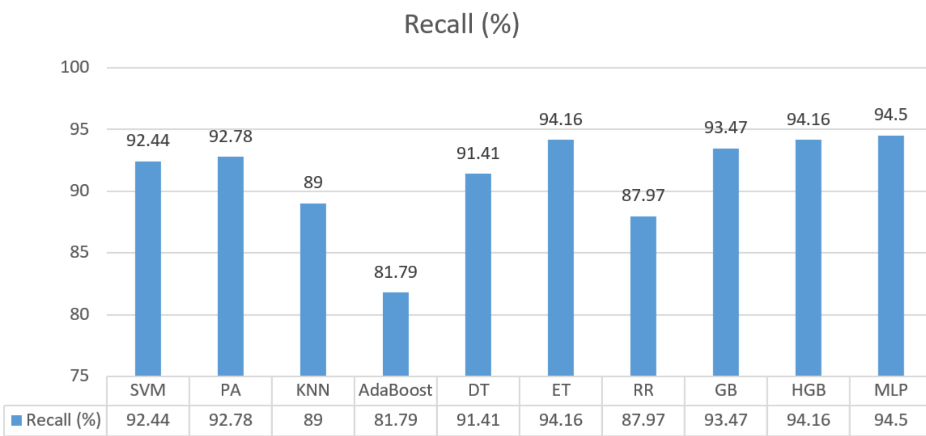

Figure 7: The Recall Performance

The F1 Score that addresses the balance of precision and recall is shown in the following Figure 8. As expected, the MLP classifier takes top spot with an F1 score of 93.46%, by a very small margin, Gradient Boosting 93.36% then HistGradientBoosting 92.82%. AdaBoost meanwhile, has the least F1 score at 80.49%, hence providing further evidence for its poorer performance on all metrics.

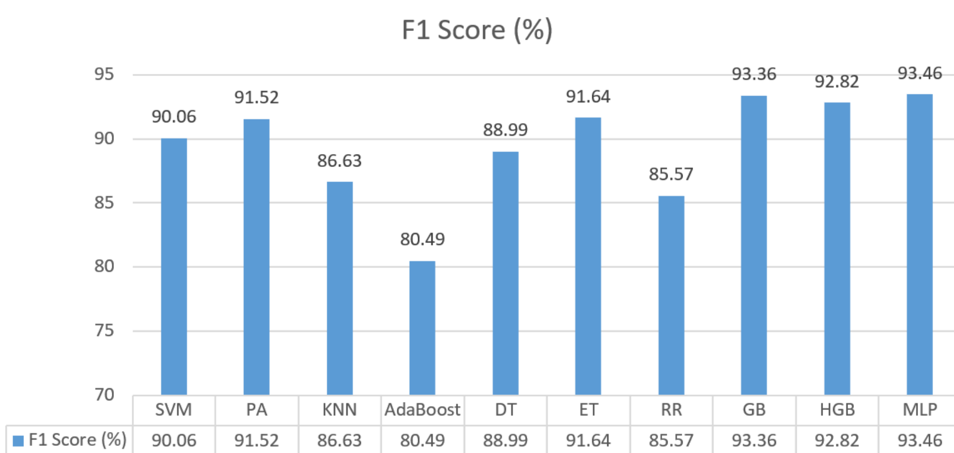

Figure 8: The F1 score Performance

Also, Figure 9 shows the accuracy in general for each classifier. MLP classifier proved to be the most accurate classifier 94.5% followed by Gradient Boosting 93. 47%, HistGradientBoosting 94.16% and Extra Trees 94.16%. AdaBoost had the lowest accuracy again at 81.79%, which again was an indication of poor performance relative to all other models.

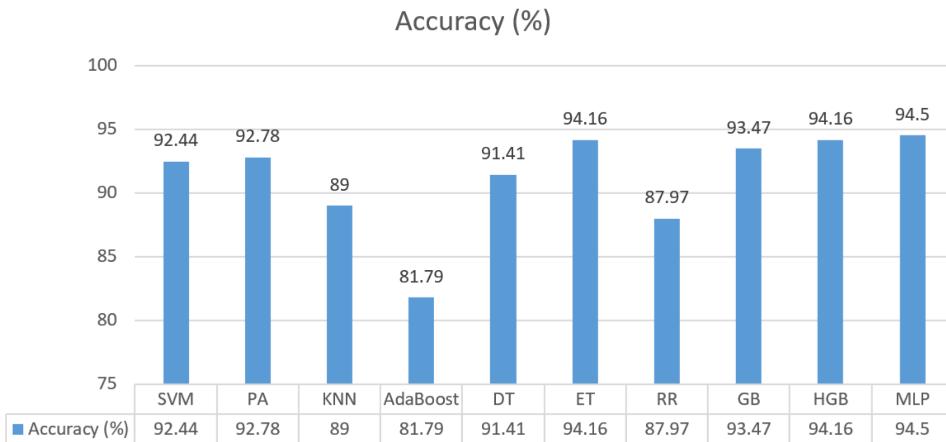

Figure 9: The F1 score Performance

5.3 Confusion Matrix Analysis of Classification Models

The confusion matrices in Figure 10 illustrate the classification performance of ten machine learning models on COVID-19 data by year. The MLP model proved the most effective; it classified the best majority of cases throughout all years, with closest accuracy in the years 2021 and 2022. Extra trees and gradient boosting also gave good results – both precision and recall were high. Models such as AdaBoost, KNN and Decision Tree had greater misclassifications especially in 2023 and 2024. These findings validate the fact that MLP is the best model for this dataset indeed, it is capable of accurately determining COVID-19 trends in the years, superior accuracy, and consistency.

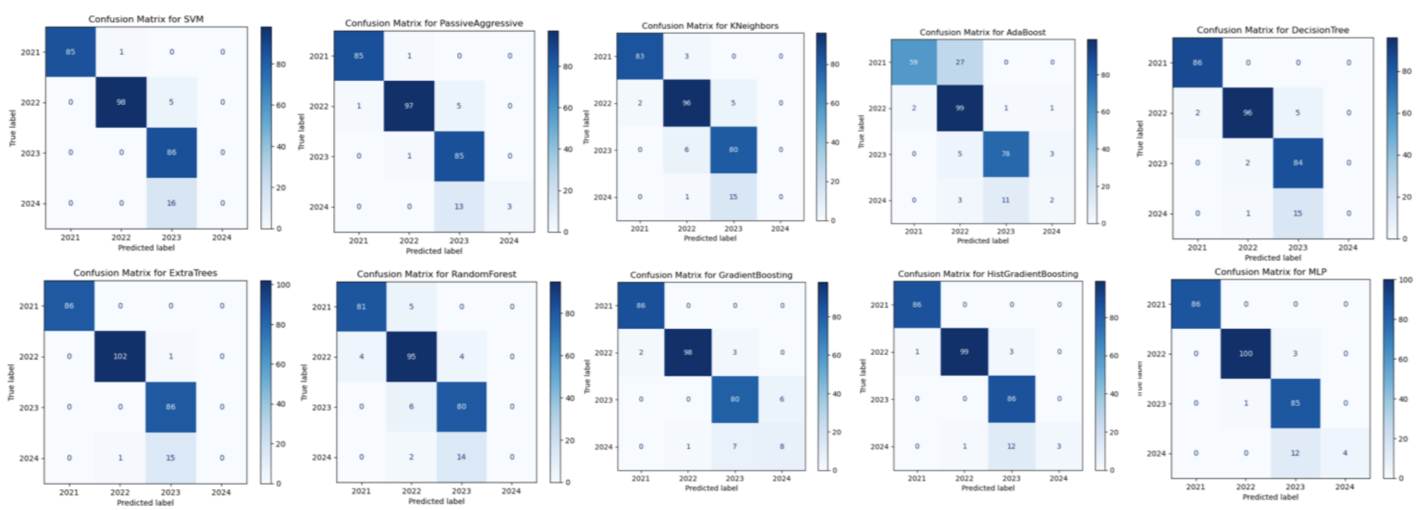

Figure 10: Confusion Matrix Performance

6. Conclusion

This study has discovered that COVID-19 has brought about the loss of the source of income, affected Vaccination Status, cases based on age, workplace clusters are the most important source of disruptions in learning, and changed lifestyle among other effects in the lives of humans. If the country could prognosticate the trend of a disease, it would be able to allocate resources where it was most needed in the following difficulties. This paper is an extension to earlier papers where a new method of analyzing the Phases of a pandemic by using historical data is put forward. This proves that it is quite possible to pinpoint the timeline of the pandemic and the outcome of various statistics with a fair degree of accuracy. Therefore, it can be smooth to establish the dynamics of going from one phase to another just by using the current data. The approach targeted increasing the countries' readiness for potential new waves of COVID-19 and other pandemic threats.

Acknowledgement

This project was created under the context of the Laboratory of Information System and Educational Technology Center, IT Students Course Project Laboratory at the University of Technology and Applied Sciences-Nizwa (APU) in Oman. You can get the dataset and code pretrained models from GitHub at: https://github.com/yasserhessein/Coivd19-Malaysia.

References

[1] A. U. M. Shah, S. N. A. Safri, R. Thevadas, N. K. Noordin, A. Abd Rahman, Z. Sekawi, and M. T. H. Sultan, "COVID-19 outbreak in Malaysia: Actions taken by the Malaysian government," *International Journal of Infectious Diseases*, vol. 97, pp. 108-116, 2020.

[2] N. S. Othman, Z. Mohamed Isa, N. Ali, D. Ibrahim, A. Anuar, and S. Abu Bakar, "Socio-economic differences in response to the COVID-19 pandemic: a case in Malaysia," in *Proceedings*, vol. 82, no. 1, p. 32, Sep. 2022.

[3] M. Irfan, F. Shahudin, V. J. Hooper, W. Akram, and R. B. Abdul Ghani, "The psychological impact of coronavirus on university students and its socio-economic determinants in Malaysia," *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, vol. 58, 00469580211056217, 2021.

[4] H. Shanmugam, J. A. Juhari, P. Nair, C. S. Ken, and N. C. Guan, "Impacts of COVID-19 pandemic on mental health in Malaysia: A single thread of hope," *Malaysian Journal of Psychiatry*, vol. 29, no. 1, pp. 78-84, 2020.

[5] W. M. A. W. M. Khalik, "Status kualiti air Tasik Temenggor, Perak," Master's thesis, Universiti Kebangsaan Malaysia, 2012.

[6] M. A. Khalid, "Covid-19: Malaysia experience and key lessons," *Asian Economic Papers*, vol. 20, no. 2, pp. 73-95, 2021.

[7] K. G. Kadhim, A. Harun, N. M. Hamawandy, A. N. Qader, H. A. Khudhur, R. M. Kakil, and R. M. Ali, "The Measures to Overcome the Impact of Covid-19 On Malaysia Economy," *Turkish Online Journal of Qualitative Inquiry*, vol. 12, no. 8, 2021.

[8] S. Agarwal, Z. He, and B. Yeung, Eds., "Impact of COVID-19 on Asian economies and policy responses," World Scientific, 2020

[9] S. Amaran, A. Z. M. Kamaruzaman, N. Y. M. Esa, and Z. Sulaiman, "Malaysia healthcare early response in combatting COVID-19 pandemic in 2020," *Korean Journal of Family Medicine*, vol. 42, no. 6, p. 425, 2021.

[10] C. Takolia, D. Priftis, M. Charakopoulou-Travlou, I. Rannou, K. Magklara, I. Giannopoulou, and K. Kotsis , "An explainable machine learning approach for COVID-19's impact on mood states of children and adolescents during the first lockdown in Greece," *Healthcare*, vol. 10, no. 1, p. 149, 2022..

[11] E. M. Aidid and R. Musa, "Accuracy of supervised machine learning in predicting depression anxiety and stress using web-based big data: Preserving the humanistic intellect," *Malaysian Journal of Medicine and Health Sciences*, vol. 18, no. 19, pp. 87-92, 2022.

[12] A. Arora, P. Chakraborty, and M. P. S. Bhatia, "Problematic use of digital technologies and its impact on mental health during COVID-19 pandemic: assessment using machine learning," in *Emerging Technologies during the Era of COVID-19 Pandemic*, 2021, pp. 197-221.

[13] C. Iwendi, A. K. Bashir, A. Peshkar, R. Sujatha, J. M. Chatterjee, S. Pasupuleti, and O. Jo, "COVID-19 patient health prediction using boosted random forest algorithm," *Frontiers in Public Health*, vol. 8, p. 357, 2020.

[14] B. Niu, R. Liang, S. Zhang, H. Zhang, X. Qu, Q. Su, and Q. Chen, "Epidemic analysis of COVID-19 in Italy based on spatiotemporal geographic information and Google Trends," *Transboundary and Emerging Diseases*, vol. 68, no. 4, pp. 2384-2400, 2021.

[15] A. K. Khairulnissa, L. Krisnan, M. K. Kaundan, and A. A. Aziz, "Emerging themes in Facebook and twitter before movement control order (MCO) in Malaysia," *SEARCH Journal of Media and Communication Research*, vol. 13, no. 1, pp. 47-60, 2021.

[16] S. Mutalib, S. N. M. Pungut, A. W. Z. Abidin, S. A. Halim, and I. S. M. Zawawi, "Development of Regression Models for COVID-19 Trends in Malaysia," *WSEAS Transactions on Information Science and Applications*, vol. 20, pp. 398-408, 2023.

[17] D. T. Jamison, L. J. Lau, K. B. Wu, and Y. Xiong, "Country performance against COVID-19: rankings for 35 countries," *BMJ Global Health*, vol. 5, no. 12, e003047, 2020.

[18] O. Torrealba-Rodriguez, R. A. Conde-Gutiérrez, and A. L. Hernández Javier, "Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models," *Chaos, Solitons and Fractals*, vol. 138, 2020, doi: 10.1016/j.chaos.2020.109946.

[19] G. S. Randhawa, M. P. Soltysiak, H. El Roz, C. P. de Souza, K. A. Hill, and L. Kari, "Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study," *PLOS One*, vol. 15, no. 4, e0232391, 2020.

[20] F. Rustam, A. A. Reshi, A. Mehmood, S. Ullah, B. W. On, W. Aslam, and G. S. Choi, "COVID-19 future forecasting using supervised machine learning models," *IEEE Access*, vol. 8, pp. 101489-101499, 2020.

[21] H. Nozari, J. Ghahremani-Nahr, and A. Szmelter-Jarosz, "AI and machine learning for real-world problems," in *Advances in Computers*, vol. 134, pp. 1-12, 2024.

[22] C. Iwendi, C. G. Y. Huescas, C. Chakraborty, and S. Mohan, "COVID-19 health analysis and prediction using machine learning algorithms for Mexico and Brazil patients," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 36, no. 3, pp. 315-335, 2024.

[23] A. S. Varde, D. Karthikeyan, and W. Wang, "Facilitating COVID recognition from X-rays with computer vision models and transfer learning," *Multimedia Tools and Applications*, vol. 83, no. 1, pp. 807-838, 2024.

[24] H. E. Saroğlu, I. Shayea, B. Saoud, M. H. Azmi, A. A. El-Saleh, S. A. Saad, and M. Alnakhli, "Machine

learning, IoT and 5G technologies for breast cancer studies: A review," *Alexandria Engineering Journal*, vol. 89, pp. 210-223, 2024.

[25] Y. Zhang, B. Liu, Y. Gong, J. Huang, J. Xu, and W. Wan, "Application of Machine Learning Optimization in Cloud Computing Resource Scheduling and Management," *arXiv preprint arXiv:2402.17216*, 2024.

[26] F. Nabi and X. Zhou, "Enhancing intrusion detection systems through dimensionality reduction: A comparative study of machine learning techniques for cyber security," *Cyber Security and Applications*, vol. 100033, 2024.

[27] V. Singh, R. C. Poonia, S. Kumar, P. Dass, P. Agarwal, V. Bhatnagar, and L. Raja, "Prediction of COVID-19 corona virus pandemic based on time series data using support vector machine," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 23, no. 8, pp. 1583-1597, 2020

[28] L. Abreu, A. Koebach, O. Díaz, S. Carleial, A. Hoeffler, W. Stojetz, and T. Brück, "Life with corona: increased gender differences in aggression and depression symptoms due to the COVID-19 pandemic burden in Germany," *Frontiers in Psychology*, vol. 12, 689396, 2021.

[29] F. Fitriyadi and M. Muqorobin, "Prediction System for the Spread of Corona Virus in Central Java with K-Nearest Neighbor (KNN) Method," *International Journal of Computer and Information System (IJCIS)*, vol. 2, no. 3, pp. 80-85, 2021.

[30] J. K. Tsai and C. H. Hung, "Improving AdaBoost classifier to predict enterprise performance after COVID-19," *Mathematics*, vol. 9, no. 18, p. 2215, 2021.

[31] M. D. Guerrero, L. M. Vanderloo, R. E. Rhodes, G. Faulkner, S. A. Moore, and M. S. Tremblay, "Canadian children's and youth's adherence to the 24-h movement guidelines during the COVID-19 pandemic: a decision tree analysis," *Journal of Sport and Health Science*, vol. 9, no. 4, pp. 313-321, 2020.

[32] S. A. F. Sayed, A. M. Elkorany, and S. S. Mohammad, "Applying different machine learning techniques for prediction of COVID-19 severity," *IEEE Access*, vol. 9, pp. 135697-135707, 2021.

[33] J. Wang, H. Yu, Q. Hua, S. Jing, Z. Liu, X. Peng, and Y. Luo, "A descriptive study of random forest algorithm for predicting COVID-19 patients outcome," *PeerJ*, vol. 8, e9945, 2020.

[34] T. K. Dash, C. Chakraborty, S. Mahapatra, and G. Panda, "Gradient boosting machine and efficient combination of features for speech-based detection of COVID-19," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 11, pp. 5364-5371, 2022..

[35] M. Tamim Kashifi and I. Ahmad, "Efficient histogram-based gradient boosting approach for accident severity prediction with multisource data," *Transportation Research Record*, vol. 2676, no. 6, pp. 236-258, 2022.

[36] R. Zhang, L. Wang, S. Cheng, and S. Song, "MLP-based classification of COVID-19 and skin diseases," *Expert Systems with Applications*, vol. 228, p. 120389, 2023.