International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 02, ISSN No. 2455-2143, Pages 143-148
Published Online June 2024 in IJEAST (http://www.ijeast.com)

MACHINE LEARNING ALGORITHMS FOR FORECASTING MORAL BEHAVIOUR IN COVID-19 CARDIOVASCULAR PROBLEM

Mrs.M.Gandhimathi
Asst.professor and Head of BCA
G.Venkataswamy Naidu College, Kovilpatti, Tamilnadu, India

Abstract— **Patients with COVID-19 who are hospitalized are always at danger of passing away. There may be a chance to forecast the mortality of COVID-19 hospitalized patients by using machine learning (ML) methods. Therefore, the objective of our study was to evaluate a number of machine learning (ML) algorithms for predicting COVID-19 mortality based on the patient's initial admission data and Choose the most effective one to use as a decision-making predictor. Following feature selection based on validated predictions, approximately 10,000 patient sample data were retrieved from the registry of various Tamilnadu hospitals for this study. The Data Mining(DM) strategy utilized in this examination is the Random Forest (RF) and support Vector Machine (SVM) calculation with Disarray Framework and k-fold Cross Approval testing. The results showed that the machine learning model could accurately predict the COVID-19 patient's treatment status with up to 91% accuracy using the SVM algorithm and up to 90% accuracy using the RF algorithm. After that, a number of machine learning algorithms were taught to predict COVID-19 mortality.**

Keywords— **Data Mining(DM), Random Forest(RM), Support Vector Machine(SVM), prediction, COVID-19, Machine learning(ML).**

I. INTRODUCTION

In Wuhan District [1], Republic of China (ROC), the novel coronavirus disease (COVID-19) was discovered in December 2019. Since then, this virus has spread rapidly throughout the entire world. The World Health Organization (WHO) declared this outbreak a pandemic in January 2020 [2, 3]. The virus's clinical effects ranged from asymptomatic or mild symptoms to serious complications and, in some instances, death. COVID-19 is a highly contagious viral infection that has emerged as a significant threat to public health on a global scale and is currently continuing its rapid global spread. Quick spread of Coronavirus has brought about the extreme deficiency of clinical assets and depletion of bleeding edge medical care laborers [4,5,6,7,8,9].

Self-isolation for COVID-19 patients with non-severe illness and hospitalization for seriously ill patients with critical risk or death constitute the basic care status. Emergency clinics or clinical specialists take different ways of lessening the quantity of passing of Coronavirus patients, remembering by managing the situation with administrations for emergency clinics. After analyzing the patient's medical data, experienced doctors either advise the patient to self-isolate or to be admitted to the hospital. Deciding the degree of care for Coronavirus patients is a type of clinical treatment or treatment for Coronavirus patients to seek legitimate treatment or care.

Blunders in navigation frequently happen on the grounds that leaders consider a few standards as the reason for decision-production [10]. So it isn't is really to be expected that past specialists underlined that the blunders most frequently happen because of errors in decision making [11], and direction is a troublesome errand in view of the effect of the choices made [11]. In like manner, in suggesting whether a Coronavirus patient ought to be hospitalized or holed up, a few measures from the sickness side effects and the consequences of clinical trials are the reason for looking at whether as a patient ought to be hospitalized or holed up. Generally, it is challenging to decide the treatment status of Coronavirus patients, both long term and self-seclusion precisely.

In the meantime, AI is a quickly developing piece of software engineering today [12]. Although machine learning is used in the majority of scientific studies, it is only used in very few health studies [13]. AI helps mining information to foresee mining results precisely [14]. Predicting correlations based on case studies can be made easier with the help of machine learning [15]. Machine learning can be used to create a variety of solutions thanks to the availability of big data [16, 17]. Furthermore, thanks to advancements in information and communication technology [18, 19], it is simple to collect the necessary big data. Predictive modeling is one of the machine learning-based solutions [20–22]. In addition, machine learning can better identify patterns and uncover hidden patterns in big data (56 Journal of Advances in Information Technology, Vol. 14, No. accurately on February 1st, 2023 (doi: 10.12720/jait.14.1.56-65) and with high precision [23].

International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 02, ISSN No. 2455-2143, Pages 143-148
Published Online June 2024 in IJEAST (http://www.ijeast.com)

Using data mining techniques, the goal of this study is to propose a machine learning system model for making predictions (solutions) for COVID-19 patients' inpatient and self-isolation treatment status decisions.

II. PROPOSED ALGORITHM

A. Supervised model for machine learning

A decision is made by a supervised learning model when an unknown instance of feedback is provided [10]. In supervised learning, we practice the program on a classified dataset to show that this knowledge is associated with the correct response. A supervised learning algorithm learns and aids in the prediction of unanticipated data effects through labeled training results.[11].This paper employs regression models for the development of predictive models.

We have utilized six relapse models to gauge the investigation of Coronavirus:

- Least Absolute Shrinkage and Selection Operation (LASSO)
- Random forest
- Decision tree regressor
- Linear regression
- Support vector machine
- Polynomial regression

a) 1. LASSO

Lasso is a multivariate statistical regression analysis tool that not only improves prediction accuracy but also evaluates the capability of a mathematical models it generates [12]. It also works on attribute selection and validation. It was at first intended for straight relapse investigation and this fundamental case shows an impressive sum about the assessor's way of behaving, incorporates its relationship to edge relapse and best assortment of subsets, cooperation's between their coefficients thus called delicate limit. The "lasso" reduces the number of squares to the number of coefficients with low fixed values. [13]. Due to this restriction, it typically generates coefficients that are exactly zero, resulting in interpretable models. The shrinkage improves LASSO and reduces error, making it the most stable regression. That implies that LASSO regression works in order to minimize the following;

$\sum k=1n(ak-\sum mbkm\beta m)2+\lambda\sum m=1q|\beta m|(1)$

The coefficient is set where λ is a concept of penalty that is minimum squares of residuals sum), $(ak-\sum mbkm\beta m)2$ bkm $\beta m)2$ is the residual sum of squares, and $|\beta m|$ is the sum of absolute value.

c) Support Vector Machine (SVM)

Solution sparsity, the use of kernel functions, and the storage of decision function energy make up this particular subset of algorithms [14]. The fact that a system of machines is SVMs indicates that the function can solve a regression problem. The information vector, x, while managing non-direct relapse, is changed over into a high-layered highlight space by a nonlinear change, and afterward straight relapse is finished in that space [15]. Using multivariate training data containing N observations, we can define the term as "a series of observed responses" in the ML sense [16].

$g(x)=y'\beta+c(2)$ is the how far the value is, β is slope or gradient function, c is value of x when y' is zero.

d) Decision Tree Regressor

Over the past two decades, methods for creating discriminant analysis, also known as classification and regression problems trees, have been developed [17]. Methods for immediately stimulating decision trees from data sets are being developed by machine learning researchers [18]. A decision tree, which is a good way to solve problems with classification and regression. Binary recursive partitioning, which is an ongoing process that divides into partitions, is also the focus of building a regression tree. [19].

e) Linear Regression

The most practical measurable strategy is a type of relapse demonstrating, for the examination of AI. [20]. It links the independent and contingent causes in a causal way. The independent characteristics of regression modelling are predicated on a target class [21]. A task is carried out using a given independent variable (x) in linear regression in order to estimate a value for the dependent variable (y). As a result, this method finds a linear relationship between x (the input) and y (the output). Therefore, the equation shows how y and x are connected:

$y=\theta 1+\theta 2x+\varepsilon(3)$

Here, $\theta 1$ represents the intercept whereas $\theta 2$ represents the coefficient of x,ε is define as error term. The machine learning algorithm aims to find the right values to get the best-fit regression line for (intercept) and (coefficient). The discrepancy between the true values and expected values should be small to get the best match, so this issue of minimization can be expressed as:

$minimise1n\sum i=1n(predi-yi)2(4)$

$h=1n\sum i=1n(predj-yj)2(5)$

here, h is the cost function of linear regression in the root mean square (RMSE) between predicted value (predi) and real value (yi), the total number of data points is denoted by n.

f) Random forest

It is a comparison of multiple decision trees for numerous datasets, sub-samples, and the average, with the goal of improving prediction results while sparing unnecessary power. The calculation for outfit discovering that incorporates a wide assortment of relapse trees is the RF relapse calculation. A collection of hierarchical requirements and constraints that extend from the root to the leaf of a tree is referred to as a regression tree. [22].

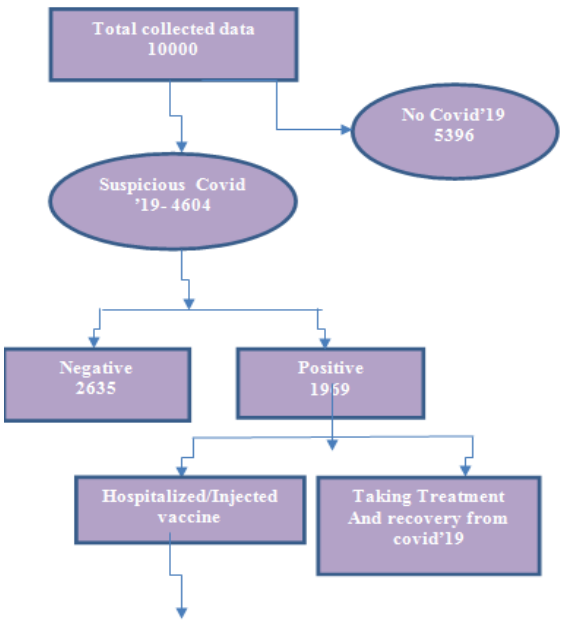
International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 02, ISSN No. 2455-2143, Pages 143-148
Published Online June 2024 in IJEAST (http://www.ijeast.com)

g) Polynomial regression
It is the part of man-made brainpower innovation, in various significance, it's a bunch of calculations for AI is relapse examination. The study includes a set of computer vision techniques that, in order to be predicted (X), a constant outcome variable (Y) must be present [23]. We have utilized the six-degree polynomial which is been utilized to foresee the better exactness of the model.

III. METHODOLOGY

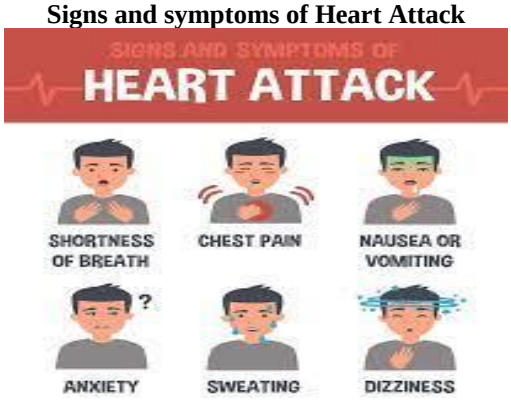
The COVID-19 forecast, also known as a novel coronavirus, is the subject of this study. A threat posed by COVID-19 to human life In India, the death rate rises each day, contributing to the deaths of thousands of people. This study aims to estimate death rates, the number of active cases, confirmed cases, and the total that will be recovered in the coming days in order to end this pandemic power, which continues to grow with each passing day. Prevision is delivered by the utilization of ways to deal with AI that are gainful in this regard. The day the pandemic began, confirmed cases, recovered cases, deaths, and active situations are all reported in the dataset. The data are first pre-processed in order to discover the specifics of known cases, cases that have been rescued, and deaths. The data are divided into two sets following the completion of the pre-processing step: a training set for the training model and a test set for the testing model. LASSO, SVM, decision tree, random forest regression, and linear regression are the learning models utilized. The methods are taught at random. The results show that models are then evaluated using the parameters R2 score, mean absolute error, root mean square error, and mean square error.
The proposed work is depicted in Fig. 1.

Fig-1



Some of them attack by Heart Problem

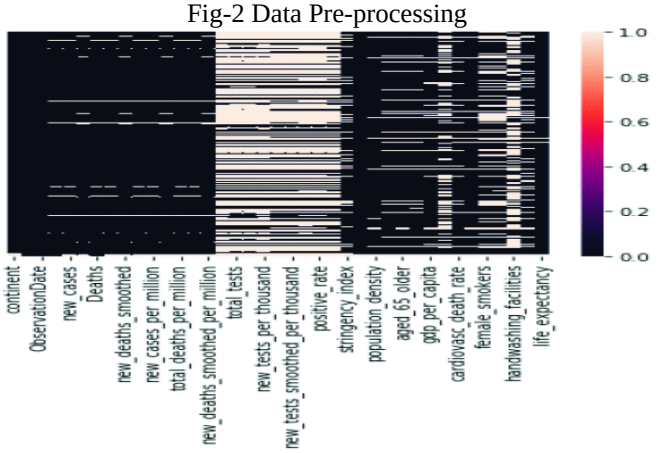
Signs and symptoms of Heart Attack



IV.ALGORITHMS COMPARISON

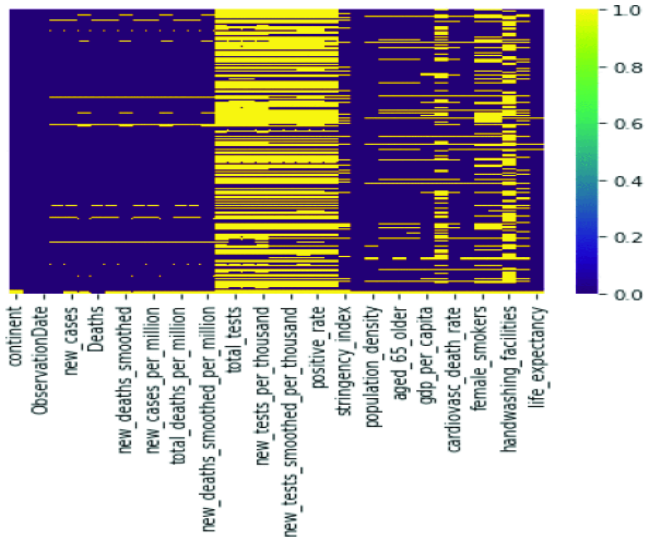
In this paper examination reason scene and python dissected and execution section two arrangement of information utilized. One bunch of information is covid19 information and one more arrangement of information cardiovascular information from UCI AI archive. Regression analysis is made possible by two-dimensional scatter plots of the dependent and independent variables for each data set. The best data fit is found in linear regression analysis's straight line. The next step in collecting and analyzing data was data wrangling. The most common way of fighting the information is eliminating invalid.
Coding:
values.coviddt.isnull(),coviddt.isnull.sum(),,sns.heatmap(co viddt.isnull(), yticklabels == false, cmap="viridis")
The Support Vector Machine, the polynomial regression algorithm, and the linear regression algorithm were utilized. Existing data are used to predict future deaths, cardiovascular deaths, and confirmed deaths. Table2 is the After effect of Affirmed cases forecast in direct, Polynomial and SVM Calculations and Table3 is the Consequence of Cardiovascular Passing cases expectation in straight polynomial and SVM Calculations. Figure 2 shows the Covid Data Results.

Fig-2 Data Pre-processing



International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 02, ISSN No. 2455-2143, Pages 143-148
Published Online June 2024 in IJEAST (http://www.ijeast.com)

Processed Data



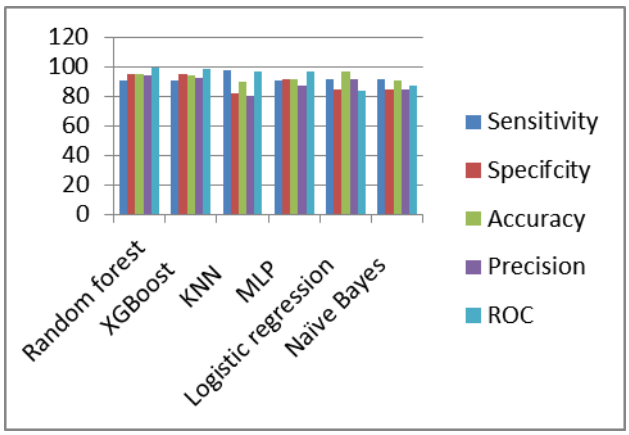
V. DISCUSSION

A comprehensive literature review and a two-round Delphi survey were used to retrospectively develop and validate ML models based on the most relevant characteristics for determining the risk of COVID-19 mortality. Using a dataset of COVID-19 hospitalized patients with laboratory confirmation, the J48 decision tree, RF, k-NN, MLP, NB, XG-Boost, and LR models were developed for this purpose. With an accuracy of 95.03%, sensitivity of 90.70 %, precision of 94.23%, specificity of 95.10 %, and ROC of around 99.02 %, RF outperformed the other seven ML methods. The ROCs of the RF, XG-Boost, KNN, and MLP models all exceeded 96.49%, and their diagnostic efficiency was superior to that of the LR model trained with the same parameters, as demonstrated by our findings.

Table-1

A review of the effectiveness of the selected machine learning algorithms in forecasting COVID-19 fatalities

Algorithms	Precision	Accuracy	Sensitivity	Specificity	**R**OC
Random forest	91.7	94.1	93.03	92.23	97.02
XGBoost	92.89	93.01	95.25	91.43	96.18
KNN	95.38	83.15	90.56	81.11	95.78
MLP	91.81	92.07	90.25	86.19	94.49
Logistic regression	91.45	84.47	96.49	91.23	83.94
Naïve Bayes	91.07	84.47	90.44	84.31	87.47



Dyspnoea, admission to the ICU, oxygen therapy (intubation), age, fever, and cough were deemed to be of the utmost importance in the current study. On the other hand, alcohol/addiction, platelet count, alanine aminotransferase (ALT), and smoking were deemed to be of the utmost significance in predicting COVID-19 mortality. However, from the perspective of physicians, being aware of these aspects may be essential for the success of drug therapy and the prediction of mortality. However, mortality can be predicted with fewer factors using ML techniques, and many of these factors can be excluded from analysis

VI.CONCLUSION

Using the most important clinical features, we developed and evaluated ML-based prediction models for in-hospital mortality in this study. In terms of classification accuracy, the RF model outperformed the other four ML algorithms. Predicting COVID-19 patient mortality risk and maximizing utilization of limited hospital resources can be accomplished using the proposed model. Patients at high risk could be identified automatically as soon as they were admitted or were in the hospital. In conclusion, COVID-19 patients' mortality risk can be accurately and promptly classified using ML algorithms and qualitative and extensive hospital databases like patient registries. If we test more classification methods on larger, qualitative, multi centre datasets in the future, our model's performance will improve.

VII. REFERENCE

[1]. Kazemi-Arpanahi H, Moulaei K, Shanbehzadeh M. Design and development of a web-based registry for Coronavirus (COVID-19) disease. Med J Islam Repub Iran. 2020;34:68–68.

[2]. Peeri NC, Shrestha N, Rahman MS, Zaki R, Tan Z, Bibi S, Baghbanzadeh M, Aghamohammadi N, Zhang W, Haque U. The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? Int J Epidemiol. 2020;2020:1–10.

[3]. Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, Fan Y, Zheng C. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. Lancet Infect Dis. 2020;20:425–34.

[4]. Liu Y, Wang Z, Ren J, Tian Y, Zhou M, Zhou T, Ye K, Zhao Y, Qiu Y, Li J. A COVID-19 risk assessment decision support system for general practitioners: design and development study. J Med Internet Res. 2020;22(6): e19786.

[5]. Alom MZ, Rahman M, Nasrin MS, Taha TM, Asari VK: COVID_MTNet: COVID-19 detection with multi-task deep learning approaches. 2020. arXiv preprint

[6]. Bansal A, Padappayil RP, Garg C, Singal A, Gupta M, Klein A. Utility of artificial intelligence amidst the COVID 19 pandemic: a review. J Med Syst. 2020;44(9):156.

[7]. Lai C-C, Shih T-P, Ko W-C, Tang H-J, Hsueh P-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. Int J Antimicrob Agents. 2020;55(3): 105924.

[8]. Hussain A, Bhowmik B, do Vale Moreira NC. COVID-19 and diabetes: knowledge in progress. Diabetes Res Clin Practice. 2020;162: 108142.

[9]. Moujaess E, Kourie HR, Ghosn M. Cancer patients and research during COVID-19 pandemic: a systematic review of current evidence. Criti Rev Oncol/Hematol. 2020;150: 102972.

[10]. P. H. D. Santos, S. M. Neves, D. O. Sant'Anna, et al., "The analytic hierarchy process supporting decision making for sustainable development: An overview of applications," J. Clean. Prod., vol. 212, pp. 119-138, 2019.

[11]. D. Laureiro-Martínez and S. Brusoni, "Cognitive flexibility and adaptive decision-making: Evidence from a laboratory study of expert decision-makers," Strateg. Manag. J., vol. 39, no. 4, pp. 1031-1058, 2018.

[12]. D. Prasad, S. K. Goyal, A. Sharma, et al., "System model for prediction analytics using k-nearest neighbors algorithm," J. Comput. Theor. Nanosci., vol. 16, no. 10, pp. 4425-4430, 2019.

[13]. Z. Zhang, "Introduction to machine learning: K-nearest neighbors," Ann. Transl. Med., vol. 4, no. 11, pp. 1-7, 2016.

[14]. P. Guleria, S. Ahmed, A. Alhumam, et al., "Empirical study on classifiers for earlier prediction of COVID-19 infection cure and death rate in the Indian states," Healthc., vol. 10, no. 1, 2022.

[15]. A.Yosipof, R. C. Guedes, and A. T. García-Sosa, "Data mining and machine learning models for predicting drug likeness and their disease or organ category," Front. Chem., vol. 6, pp. 1-11, May 2018.

[16]. J. Bullock, A. Luccioni, K. H. Pham, et al., "Mapping the landscape of artificial intelligence applications against COVID19," J. Artif. Intell. Res., vol. 69, pp. 807-845, 2020.

[17]. L. Wynants, B. V. Calster, G. S. Collins, et al., "Prediction models for diagnosis and prognosis of COVID-19: Systematic review and critical appraisal," BMJ, vol. 369, no. m1328, pp. 1-16, 2020.

[18]. A.Anggrawan, A. H. Yassi, C. Satria, et al., "Comparison of online learning versus face to face learning in english grammar learning," in Proc. the 5th International Conference on Computing Engineering and Design, 2018, pp. 1-4.

[19]. Anggrawan, "Interaction between learning preferences and methods in face-to-face and online

learning," ICIC Express Lett., vol. 15, no. 4, pp. 319-326, 2021.

[20]. H. Zhang, L. L. Wang, Y. Y. Chen, et al., "A tool to early predict severe corona virus disease 2019 (COVID-19): A multicenter study using the risk nomogram in Wuhan and Guangdong, China," Cancer, vol. 46, pp. 1-17, May 2020.

[21]. X. Jiang, M. Coffee, A. Bari, et al., "Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity," Comput. Mater. Contin., vol. 63, no. 1, pp. 537- 551, 2020.

[22]. J. Liu, Y. Liu, P. Xiang, et al., "Neutrophil-to-lymphocyte ratio predicts critical illness patients with 2019 coronavirus disease in the early stage," J. Transl. Med., vol. 18, no. 1, pp. 1-12, 2020.