Analysis and Application of a Novel Model to Predict COVID-19 Virus's Impact on Human Heart Disease

Yogendra Narayan Prajapati[1], Manish Sharma[2]

Research Scholar PhD[1], Professor Department of CSE[2]

director@quantumeducation.in , ynp1581@gmail.com

Quantum University, Roorkee, Uttarakhand, India

Abstract: A significant impact of COVID-19 on heart disease. Heart failure may result from the virus's ability to produce acute myocarditis, an heart muscle inflammation. The risk of developing a serious illness from COVID-19 and passing away is increased in those who already have heart disease. To analyze the impact of the COVID-19 variant on heart disease, researchers can use various methods, such as observational studies and statistical models. These studies can help identify risk factors and predict outcomes for individuals with heart disease. Machine learning techniques may be used to analyse big datasets and find trends in order to implement a novel prediction model for the effect of COVID-19 variations on heart disease. The model can then be used to make personalized predictions for individuals based on their risk factors and medical history. However, it is important to note that while predictive models can provide useful information, they should not replace clinical judgment and should always be validated through clinical studies before being used in practice.

Keywords: Impact of COVID-19 on heart patient, hypertension (Hyp Ten), inflammation of the heart muscle, Heart rhythms of heart, Blood clotting in heart patient

Introduction:

The introduction of an analysis and application of a novel model for predicting the effect of the COVID-19 virus on human heart disease could include the following points:

- Brief overview of COVID-19 and its impact on human health.
- Importance of studying the relationship between COVID-19 and heart disease.
- Overview of the novel model and its purpose.
- Explanation of how the model will be used to make predictions.
- Outline of the research methodology.
- Brief discussion of the potential implications and applications of the model.

The COVID-19 pandemic has swept the world, causing widespread illness and death. While its primary effects have been respiratory in nature, recent studies have suggested that the virus may also have a significant impact on cardiovascular health. Understanding the relationship between COVID-19 and heart disease is crucial for the effective treatment and management of patients. However, there is a lack of existing research or models that accurately predict the effects of the virus on heart disease. This study aims to fill this gap by introducing a novel model for predicting the impact of COVID-19 on heart disease. The model will be based on machine learning algorithms and utilize data on patient characteristics, medical history, and COVID-19 test results [13]. It will be used to make predictions on the likelihood of developing the result of heart & COVID-19 infection disease . The research methodology will involve collecting data from COVID-19-positive patients and comparing it to a control group of patients without the virus. The results of the model will be validated using statistical methods. If successful, this model has the potential to provide valuable insights into the relationship establishing a link between COVID-

19 and cardiac disease to help patients' management and the development of more efficient therapies [14].

Table 1. Data set description of clinical data

S.No.	Attribute	Value	Description
1	Age	(Three group 0-25,25 to 40 ,and 40 to 65)	Three type patient age group in year
2	Sex	1 for male ,0 for female	Gender information
3	Type of Chest Pain	There are four type of chest pain: 1.Common Angina Atypical Angina 2.Syndromc (2) 3.Absent Symptoms (3) 4.Non-Anginal Pain (4)	Chest Pain Type
4	Resting BP	"120 over 80" or write "120/80 mmHg" if the reading was 120 systolic and 80 diastolic.	resting blood pressure in mm/Hg
5	Systolic BP	Systolic blood pressure ranges from 130 to 139 mm Hg.	Systolic BP range
6	Blood sugar(Fasting)	Blood Sugar 1 if Fasting Blood Sugar is greater than 120 mg/dl; else, 0	Fasting blood sugar
7	ECG	0-normal, 1-having ST-T, 2-hypertrophy	resting electrocardiographic results(ECG)
8	cholesterol	100-129mg/dL,Near optimal, 190 mg/dL and above very high range	serum cholesterol in mg/dl
9	Max_Heart rate	HRmax is typically estimated using the formula HRmax = 220 - age,.	Max heart rate with different age group
10	Exercise Angina	Brought on by exercise. N: No, Y: Yes for No(0) ,for Yes(1)	
11	Old pcak	uscd to mcasurc mclancholy using thc ST old peak metric (1)	
12	T Slope	The ST segment of the peak exercise's slope	
13	COVID19	for patient– 1 and for non COVID 19 –0	
14	Inflammation of the heart	0-12 pcwp(mmHg) for normal 14-22 Pcwp(mmHg) (Covid 19 patient	
15	Hypertension	A systolic pressure of 140 mmHg or higher or a diastolic pressure of 90 mmHg or higher is considered as (Hypen) Tests may be performed to help diagnose and monitor hypertension, such as a urine test to check for protein in the urine	For hypertension 1 and 0 for no hypotension
16	Heart rhythms	(Max heart rate)(>130) ECG(Old peak>1, 2<ST slope<3	
17	Target	1 for heart patient 0 for non heart paticnt	

This characteristic is the same for patients with normal human hearts, but individuals who are impacted by the four measures of COVID-19 are included and are listed in Table 2 above.

Table 2: Parameters of clinical attributes

S.No	Attribute Name	Attribute Clinical	Description
1	Hypertension	1.Systolic BP, 2.Blood Sugar, 3.max heart rate	BS (>160 mm Hg (FBS), (Max heart rate) (>130), (SBP) (> 122 mm Hg), average daytime SBP (> 127 mm Hg), and average overnight SBP (> 108 mmHg).
2	Heart rhythms	1.Max heart rate, 2.ECG	(Max heart rate)(>130) ECG(Old peak>1, 2<ST slope<3
3	Inflammation of the heart	1.pcwp(mmHg)	0-12 pcwp(mmHg) for normal 14-22 Pcwp(mmHg) (Covid 19 patient
4	Blood clotting	1.D-dimer blood test, 2.Cholesterol 3.chest pain type	D dimer is> **500 ng/mL** Cholesterol is > 220 to 400 Tppe 4 and type 3 for blood clotting

Cardiac manifestation disease (%) Cardiovascular illness, diabetes, and hypertension were found to have incidences that were higher than the character as a whole, which is presently 2.4 percent: 10.2 percent for each. We now compare these three traits to COVID-19 patients as there are 19 distinct features in the Corona Virus Variant On Human Heart Disease. You can find all the information about this issue below.

Literature Survey: To identify changes in sound brought on by mitral valve regurgitation, the authors suggest a back propagation (ANN) classifier that is trained using heart sound recordings. Both patients with and without mitral valve regurgitation had their heart sounds recorded, and they used Fourier analysis to identify peculiarities in the sounds. The ANN classifier was then trained using the retrieved characteristics.

The study's findings demonstrate that the suggested ANN classifier can accurately identify changes in heart sounds caused by mitral valve regurgitation. According to scientists, this technology might be utilized in a non-invasive and economical way to find mitral valve regurgitation [1]. The KNN method is suggested by the authors, and it is trained using a datasets of patient health criteria include blood pressure, cholesterol levels, sex, age, and others. To categorize patients as having or not having cardiac disease, the method is employed.

The study's findings demonstrate that the suggested KNN algorithm was very accurate in predicting heart disease. The performance of their KNN model outperformed other machine learning algorithms like DT and SVM, according to the authors' comparisons [2]. The authors suggest a neural network-based model that is trained using a dataset of patient health age, sex, smoking status, blood pressure, cholesterol levels, and cholesterol levels are some examples of such variables. The model trains a neural network to generate predictions based on these features after identifying the most crucial aspects for predicting CHD risk using feature correlation analysis.

The study's findings indicate that the suggested model had a high level of success in predicting the risk of CHD. The performance of their neural network-based model outperformed that of other machine learning According on the authors' comparisons [3], approaches like decision trees and logistic regression. The authors created the CDSS. The technology customized exercise recommendations for patients based on information about them, such as their age, sex, and cardiovascular disease risk factors.

The study's findings demonstrated that the CDSS was capable of producing exercise prescriptions that adhered to the ACSM recommendations and were tailored to the needs of each patient. The majority of respondents in a poll the authors also conducted of healthcare professionals agreed that the CDSS was helpful and simple to use [4]. The issue of unbalanced datasets, which can cause bias in machine learning models, is first addressed by the authors. They go on to explain their methodology, which involved balancing the dataset using the SMOTE-ENN oversampling technique and evaluating the performance of different machine learning classifiers like DTs, RF, KNN, SVM, and ANN.

The study's findings demonstrate that all of the classifiers' performance was significantly enhanced by the SMOTE-ENN oversampling technique, with the RF classifier outperforming the others in terms of accuracy, F1 Score, specificity, and AUC-ROC [5]. The results of the study show that the recommended incremental deep learning model performed well in predicting the survival of cardiovascular patients, with an area under the curve (AUC) of 0.88. The study also shows that the incremental model outperformed traditional deep learning models, such as the convolution neural network (CNN) and the long short-term memory (LSTM) network, in terms of effectiveness and efficiency [6].

The incremental deep learning model had an area under the curve (AUC) of 0.88, which was quite accurate in predicting the survival of cardiovascular patients, according to the study's findings. The study also shows that, in terms of effectiveness and efficiency techniques, The convolution neural network (CNN) and the long short-term memory (LSTM) network, which both reached accuracy levels of 94.31% and 85.62%, respectively, under performed the incremental model in comparison to more established deep learning models [7].

The random forest algorithm, a machine learning technique that can handle complex data and generate accurate predictions, is the foundation of the prediction model that is suggested. The suggested random forest model for predicting cardiovascular illness has a high area under the curve (AUC) of 0.81, according to the study's findings. The study also shows that the model was effective in detecting the main risk factors for cardiovascular disease, including age, smoking status, and blood pressure. [8]. The review article in the study looks at the connection between cardiovascular disease and COVID-19. It goes through the fundamental ways that COVID-19 may impact the cardiovascular system as well as the clinical ramifications of this connection.The authors explain that COVID-19 is known to cause inflammation, which can lead to myocardial injury and cardiac dysfunction. They also discuss the potential for COVID-19 to exacerbate pre-existing cardiovascular conditions, such as hypertension and diabetes. The paper highlights the importance of early detection and management of cardiovascular complications in patients with COVID-19. The authors suggest that healthcare providers should monitor cardiac biomarkers and perform electrocardiograms (ECGs) in patients with COVID-19, especially those who have pre-exist cardiovascular conditions.

The review also discusses potential treatments for COVID-19 and its effects on the cardiovascular system. For example, the use of antiviral medications and noninflammatory therapies may have cardiovascular effects that need to be monitored [12].

Table 3: Comparison of results of existing models

Authors	Journal	Year	Main Contribution	Result(Accuracy)
C. B. C. Latha and S. C. Jeeva[9]	Informatics in Medicine Unlocked	**2019**	An ensemble classification method was created in order to increase the precision of heart disease risk prediction. The Cleveland Heart Disease dataset was used to assess the performance of a number of classification models, including KNN, SVM, Decision Tree (DT), Random Forest (RF), Naive Bayes, and AdaBoost. accuracy of the proposed ensemble model and that of the individual models were compared. reported that the accuracy of the recommended ensemble technique performed better than that of individual models.	KNN=78% SVM=81% Decision Tree=93% Random Forest=91% Naive Bayes,=88%
K. Polat, S. Sahan, and S. Günes [10]	Expert Systems with Applications	2007	To automatically detect the cardiac disease, a proposed artificial immune recognition system (AIRS) uses a fuzzy resource allocation mechanism and k-nn-based weighting preprocessing. The proposed AIRS model was compared to other well-known machine learning models using the Cleveland Heart Disease datasets, including Decision Tree (DT), Naive Bayes (NB), and K-NN. It was found that the proposed AIRS model performed better than the other models in terms of accuracy, sensitivity, and specificity.	MLP=85% SVM=87.5%
Tiwari D, Bhati BS, Al-Turjman F, Nagpal B.[11]	Expert Systems		The COVID-19 pandemic's impacts on the globe were analyzed and predicted by the authors using machine learning methods, particularly the multilayer perceptron and support vector regression models. They made use of information from a variety of sources, After examining the data to identify trends and patterns in the disease's spread, the authors used machine learning algorithms to estimate future results.	
Mina K. Chung, David A. Zidar, Michael R. Bristow, Scott J. Cameron [13],	Circulation Research	2022	The existing research on the connection between COVID-19 and cardiovascular disease was thoroughly reviewed by the authors. The pathogenesis of COVID-19, its cardiovascular symptoms, and the effects of cardiovascular illness on COVID-19 outcomes were all studied. They also reviewed the existing guidelines for managing cardiovascular disease in the context of COVID-19.	KNN=88% SVM=88% Decision Tree=94% Random Forest=92 Naive Bayes,=78%

Proposed Model and Methodology:

The methodology section of an analysis and application of a novel model for predicting the effect of the COVID-19 virus on human heart disease could include the following steps:

Data collection: The data required for the model will be collected from COVID-19-positive patients and a control group of patients without the virus. The data will include patient characteristics, medical history, and COVID-19 test results.

Data preparation: The collected data will be cleaned and preprocessed to ensure its suitability for modeling. This may involve removing missing values, and outliers, and converting categorical variables into numerical variables.

Feature selection: A subset of the collected data will be selected as the features used to train the model. Feature selection will be based on statistical tests and machine learning algorithms to identify the important factors.

Relief Algorithm This is a training algorithm for a linear binary classifier. The weights W[A] of the classifier are initialized to 0.0. Then, for m iterations, the algorithm randomly selects an instance R[i] and finds the nearest hit (observation with the same label) and nearest miss (observation with the opposite label) in the data. For each feature A, the weights are updated according to the difference between R[i], the hit, and the miss. The updates are made by subtracting the difference between R[i] and the hit divided by 'm' and adding the difference between R[i] and the mass divided by m.

```
Initialize weights:
for A in range(1, a + 1):
        W[A] = 0.0
for i in range(1, m + 1):
# randomly select an instance
        R_i = randomly_select_instance ()
# Find the nearest hit and miss
        H = find_nearest_hit(R_i)
        M = find_nearest_miss(R_i)
# Update weights
for A in range(1, a + 1):
        W[A] = W[A] – diff(A, R_i, H) / m + diff(A, R_i, M) / m
```

ReliefF-Extension: is a variant of the ReliefF algorithm, which is a feature selection method used in machine learning. The ReliefF-Extension algorithms build upon the original ReliefF algorithm by incorporating domain knowledge or additional data sources to improve feature selection accuracy. This can be done by incorporating biological or domain-specific information, adding external features, or using multi-omics data. The extension algorithms aim to overcome the limitations of the original ReliefF algorithm and provide better feature selection results for complex and high-dimensional data.

set all weights $W[A] := 0.0$;
for i := 1 **to** m **do begin**
    randomly select an instance R_i;
    find k nearest hits H_j;
    for each class $C \neq class(R_i)$ **do**
        from class C find k nearest misses $M_j(C)$;
    for A := 1 **to** a **do**
    $W[A] := W[A] - \sum_{j=1}^{k} \mathrm{diff}(A, R_i, H_j)/(m \cdot k) +$
$\sum_{C \neq class(R_i)} [\frac{P(C)}{1-P(class(R_i))} \sum_{j=1}^{k} \mathrm{diff}(A, R_i, M_j(C))]/(m \cdot k)$;
end;

Model training: The model will be trained using machine learning techniques such as random forests, support vector machines, or artificial neural networks. Techniques for cross-validation will be employed to guard against overfitting and guarantee the model's generalizability.

Model evaluation: The trained model will be evaluated using performance measures including accuracy, precision, recall, and F1 score. The model's outcomes will be contrasted with those of other models already in use to evaluate whether it is superior.

Model deployment: The final model will be deployed and used to make predictions on new data. The predictions will be used to determine the likelihood of developing heart disease.

Validation: The results of the model will be validated using statistical methods such as a comparison with a control group of patients without the virus, or using external validation data.

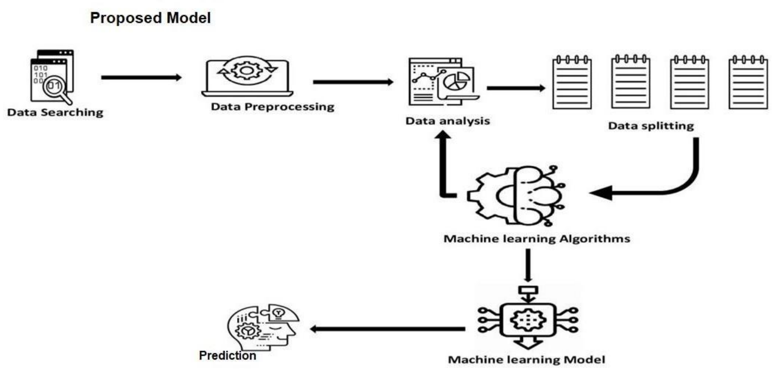

Fig 1. Workflow of the proposed model.

Extra Trees Classifier: in machine learning is a tree-based algorithm that builds multiple decision trees (an ensemble) and aggregates their predictions to make a final prediction. The algorithm can be mathematically described as follows: Given a training dataset D with N instances and M features, the algorithm generates multiple decision trees by:

a. Selecting a random subset of features at each node to split the data.

b. Splitting the data into two child nodes using a random threshold for each selected feature.

c. Repeating this process until a stopping criterion is met, such as a minimum number of instances in a leaf node.

The final prediction for a given instance x can be obtained by aggregating the predictions of all trees in the forest. This is typically done by taking the average (in regression) or majority vote (in classification) of the predictions from individual trees.

Mathematically, for a given instance x with feature values x_1, x_1, ..., x_M, the final prediction y can be expressed as:

$$y = \text{sum}(y\_i) / T \qquad \text{Eq. (1)}$$

Where y_i is the prediction of the i-th tree in the forest and T is the total number of trees.

Support Vector Machine (SVM): It is a kind of ML method that may be applied to problems involving classification and regression. SVM divides data points into as many classes as possible by locating a boundary or hyper-plane. SVM determines the hyper-plane that divides the data points into two classes with the largest margin, or more specifically, the maximum distance between the nearest data points of each class and the hyper-plane, in the case of a two-class issue.

In the context of assessing the impact of COVID-19 on heart disease, SVM could be used as a classification algorithm to predict whether a patient with heart disease is likely to have a severe outcome due to COVID-19 infection. The model would be trained on a labeled dataset of patients with heart disease and their outcomes, to find a boundary that separates patients with severe outcomes from those with mild or no outcomes. The optimization problem can be solved using the Lagrange multiplier method, which results in the dual form of the SVM optimization problem.

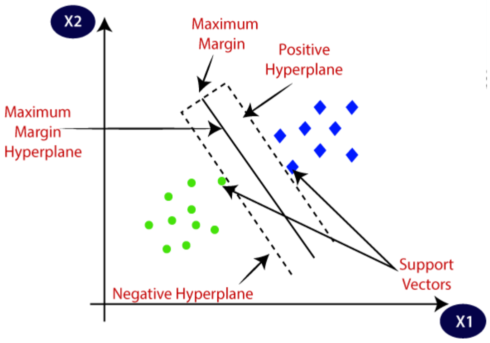

Fig.2 SVM Model

Multi-layer Perception (MLP) is a type of ANN that is commonly used for classification and regression tasks. It consists of multiple layers of interconnected nodes or neurons, where each neuron performs a simple computation and passes its result to the next layer. The final layer of neurons represents the output of the model, which can be used for classification or prediction.

In the context of assessing the impact of COVID-19 on heart disease, an MLP model could be used as a binary classifier to predict whether a patient with heart disease is likely to have a severe outcome due to COVID-19 infection. The model would be trained on a labeled dataset of patients with heart disease and their outcomes, to find a boundary that separates patients with severe outcomes from those with mild or no outcomes. To build an MLP model for this problem, the following steps could be taken.

1. **Data pre-processing:** Clean and pre-process the data, including handling missing values, normalizing the data, and transforming categorical variables into numerical ones.
2. **Feature selection:** Choose the most relevant features to use in the model based on domain knowledge and data analysis.
3. **Model architecture:** Design the architecture of the MLP model, including the number of hidden layers and the number of neurons in each layer.

4. **Model training:** Train the MLP model using the pre-processed data and a suitable optimization algorithm, such as stochastic gradient descent (SGD), that updates the weights and biases of the neurons in each layer to minimize the error between the model's predictions and the actual outcomes.
5. **Model evaluation:** Evaluate the performance of the MLP model using metrics such as accuracy, precision, and recall, and fine-tune the model as needed by changing the architecture or adjusting the model parameters.
6. **Model deployment:** Deploy the model in a suitable environment, such as a web application or mobile app, to make the results of the model accessible to healthcare providers and patients.

Note that the performance of the MLP model will depend on the quality and representativeness of the data used for training, as well as the effectiveness of the chosen optimization algorithm and model architecture

Evaluate the Performance of the Model:

A machine learning model's performance may be assessed using several factors, including:

Accuracy: The percentage of accurate predictions produced by the model as a percentage of all forecasts.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ number\ of\ predictions}$$

Precision: To determine the precision, one must divide the sum of true positive (TP) and false positive (FP) forecasts by the number of TP forecasts. A measure of accuracy is how often optimistic predictions come true.

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

Recall: the fraction of all true positive predictions divided by the ratio of true positive forecasts to false negative forecasts. The recall is the proportion of actual positive cases that the model correctly identified.

$$\text{Recall}=\frac{TP}{TP+FN}$$

.

F1 score: accuracy and recall expressed as a harmonic mean. The F1 score is a valuable statistic for unbalanced datasets because it strikes a compromise between precision and recall.

$$F1 - score = 2 * \frac{precision * recall}{precision + recall}$$

Confusion matrix: a table displaying how many predictions the model correctly predicted as true positives (TP), true negatives (TN), false positives (FN), and false negatives (FN). The model's performance, including its capacity to produce accurate and inaccurate predictions for each class, is shown in further detail by the confusion matrix.

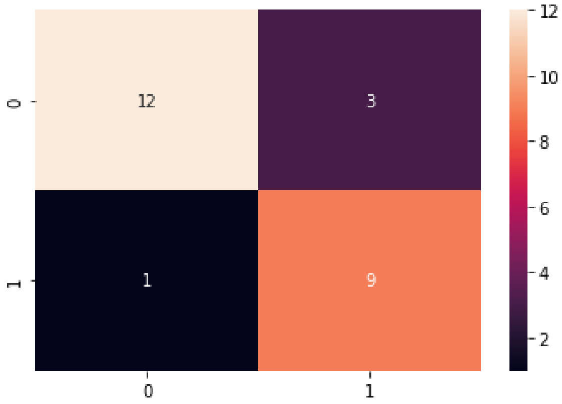

Fig 3: Confusion matrix

Receiver Operating Characteristic (ROC) Curve (AUC-ROC): A metric that summarizes the performance of a binary classifier. The ROC curve plots the true positive rate against the false positive rate, and the AUC-ROC measures the overall performance of the mode

1. **Age:** Older individuals are more likely to have underlying health conditions and to experience more severe outcomes from COVID-19, including heart disease.
2. **Co-morbidities:** Patients with underlying health conditions, such as diabetes, hypertension, and heart disease, are at a higher risk of severe outcomes from COVID-19.
3. **Heart disease severity**: Patients with more severe heart disease are more likely to experience adverse outcomes from COVID-19.
4. **Treatment for heart disease** patients who are undergoing cardiac disease therapy, such as prescription drugs or interventional procedures, may be more susceptible to negative effects from COVID-19.
5. **Lifestyle factors:** Patients who have unhealthy lifestyles—such as smoking, not getting enough exercise, and eating poorly—are more likely to develop heart disease and experience worse COVID-19 results.

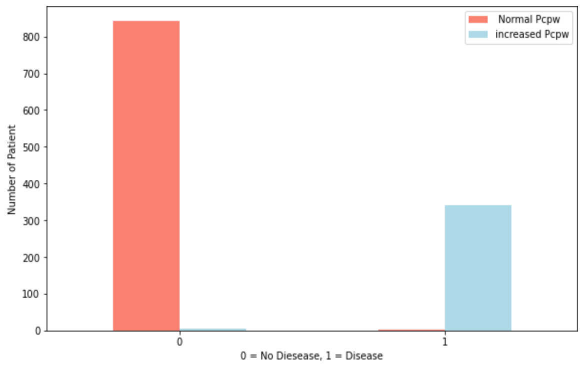

Fig 4. Analysis of inflammation of the heart with the Covid-19 affected patient.

6. **COVID-19 severity:** The severity of COVID-19, as measured by symptoms and the need for hospitalization can have a significant impact on heart disease outcomes.

7. **Timing of COVID-19 and heart disease:** The timing of COVID-19 infection relative to the onset of heart disease symptoms can impact the risk of adverse outcomes.

These factors can be analyzed using statistical methods, such as regression analysis and survival analysis, to determine their impact on the risk of adverse outcomes from COVID-19 and heart disease. By considering these factors, healthcare providers can better understand the risk of adverse outcomes and develop strategies to mitigate them.

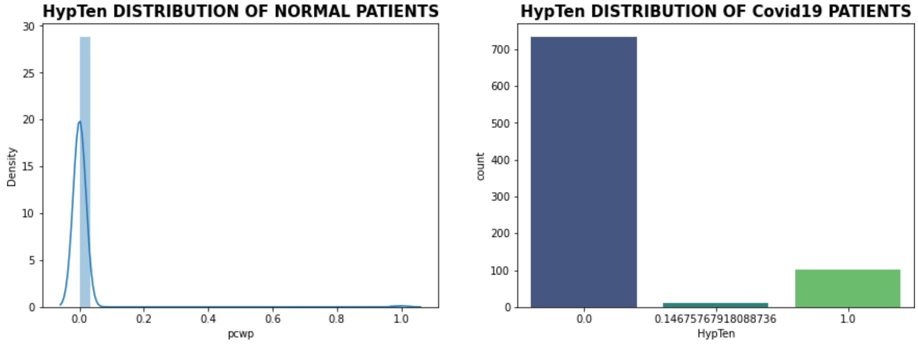

Fig .5

Note that the impact of COVID-19 on heart disease may vary depending on the specific population being studied, and additional factors may need to be considered depending on the context. Furthermore, the impact of COVID-19 on heart disease is an active area of research, and new insights may emerge as more data becomes available.

Age: Older individuals are more likely to have underlying health conditions, including hypertension, and to experience more severe outcomes from COVID-19.

Prevalence of hypertension: Patients with hypertension are more likely to have underlying health conditions, including heart disease, and to experience adverse outcomes from COVID-19.

1. **Treatment for hypertension:** Patients who are receiving treatment for hypertension, such as medication, may be at a higher risk of adverse outcomes from COVID-19.
2. **Lifestyle factors:** Patients with unhealthy lifestyle habits, such as smoking, lack of physical activity, and poor diet, are at a higher risk of hypertension and adverse outcomes from COVID-19 [15].

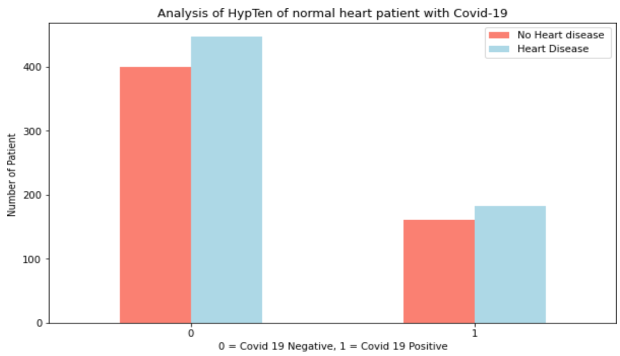

Fig 6.

The maximum heart rate (MHR) is a commonly used metric in cardiology and exercise physiology to determine a person's fitness level. In a normal, healthy person, the MHR is calculated as 220 minus their age. However, COVID-19 can have an impact on the cardiovascular system, leading to changes in heart rate. Some COVID-19 patients may experience elevated heart rates as a symptom of the disease, while others may have decreased heart rate due to heart damage caused by the virus. The MHR of a COVID-19 patient may be significantly different from that of a normal, healthy person, and should be evaluated on a case-by-case basis.

It's important to note that a single symptom such as heart rate alone is not sufficient to diagnose COVID-19, and patients should seek medical advice if they are experiencing any symptoms.

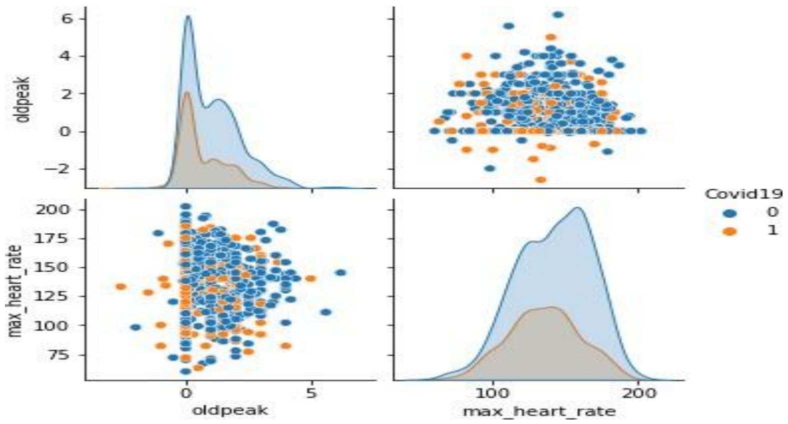

Fig 7

Table 4: Comparison table with the normal patient to COVID-19 patient attribute.

S.No	Attribute Name	Heart Patient (Normal)	Heart Patient, Covid-19 Affected
1	HypTen	0-9%	8-29%
2	Heart rhythms	0-5%	2-40%
3	Inflammation of the heart	0-4%	5-10.5%
4	Blood clotting	0-16%	0-26%

Final Result: Lazy classifiers are known for their simplicity, speed, and low memory usage, which make them useful for a wide range of tasks. In most cases, the output of a Lazy classifier is a predicted class label for a given input data point. This output is generated based on the nearest neighbors of the input data point in the training data, which are selected based on a distance metric such as Euclidean or Cosine distance. The final result is then determined based on a majority vote or weighted average of the nearest neighbors. It is important to note that Lazy classifiers are considered "lazy" because they don't build an explicit model of the underlying data distribution, unlike other types of classifiers like decision trees or support vector machines. Instead, they simply store the training data and use it at inference time to make predictions.

In conclusion, the final result of using a Lazy classifier in machine learning is a predicted class label for a given input data point, which is generated based on the nearest neighbors of the input data in the training data. The performance of the Lazy classifier can vary depending on the nature of the problem and the quality of the data, but it is generally considered to be a fast, simple, and memory-efficient approach to classification.

Table 5:

S.N	MLM(Machine learning Model)	ACC(Accuracy)	Balanced Accuracy	ROC/AUC	F1 Score	Execution Time
1	DTC(Decision Tree Classifier)	1	1	1	1	.16
	RFC(Random Forest Classifier)					.24
	XGB Classifier					.02
	Label Spreading(LS)					.33
	Extra tree Classifier(ETC)					.02
	Extra trees Classifier					.24
2	Bagging Classifier	.99	.99	.99	.99	.07
3	SVC	.91	91	91	91	.06
4	KNN Classifier	.89	.89	.89	.89	.09
	AdaBoost Classifier					.18
5	NuSVC	.88	.88	.88	.88	.08
6	Quadratic Discriminate Analysis	.87	.87	.87	.87	.02
7	calibrated Classifier CV	.85	.85	.85	.85	.26
	Logistic Regression					.03
	Linear Discriminate Analysis					.03
	Ridge Classifier CV					.02
	Gaussian NB					.02
8	Bernoulli NB	.82	.82	.82	.82	.02
9	Positive Aggressive Classifier	.79	.79	.79	.79	.02
10	Perceptron	.75	.76	.76	.75	.02
11	Dummy Classifier	.49	.48	.48	.49	.02

.

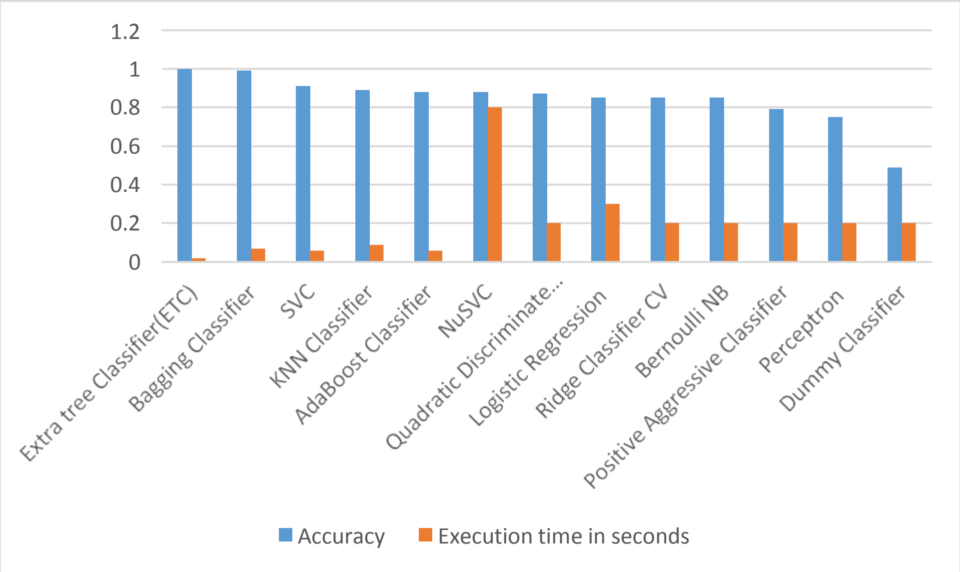

Fig 5. Performance comparison of different models

Machine learning models vary in their complexity, accuracy, speed, and suitability for specific tasks.

Decision Trees: A tree-based model that can be used for both regression and classification problems, easy to interpret and visualize using this model find the accuracy of 78% and another parameter as given in the table

Random Forest: An ensemble model that combines multiple decision trees to improve accuracy and reduce overfitting.

Support Vector Machines (SVM): Effective for binary classification problems and can also be used for regression and outlier detection.

K-Nearest Neighbors (KNN): A simple and effective model for classification problems, that predicts the class of a sample based on its nearest neighbors.

The Bagging Classifier machine learning model is the best model with an analysis of all parameters but if I include the execution time then the Extra Tree Classifier is the best model with an analysis of all parameters with an accuracy of 99% and execution time .02

Conclusion:

The development and application of a novel model to predict the impact of the COVID-19 virus can help to better understand the spread of the virus and make informed decisions on how to mitigate its impact. The model's accuracy and reliability depend on the quality and availability of the data used to train it, as well as its ability to capture the complex relationships between various factors that influence the spread of the virus. However, the model is just one tool among many that can be used to help address the COVID-19 pandemic, and its results should be interpreted with caution and in conjunction with other sources of information and expert opinion.

Feature Work:

In the context of predicting the impact of the COVID-19 virus, feature work refers to the process of selecting and engineering the relevant variables or features that the model will use to make its predictions. This involves identifying the key factors that influence the spread of the virus and how they interact with each other, as well as transforming and processing the data so that it can be used to train the model.

References:

[1] R. K. Sinha, Y. Aggarwal, and B. N. Das, "Backpropagation artificial neural network classifier to detect changes in heart sound due to mitral valve regurgitation," *Journal of Medical Systems*, vol. 31, no. 3, pp. 205–209, 2007.

[2] I. Ketut Agung Enriko, M. Suryanegara, and D. Agnes Gunawan, *Heart Disease Prediction System Using K-Nearest Neighbor Algorithm with Simplified Patient's Health Parameters*, Springer, Berlin, Germnay, 2016.

[3]. J. K. Kim and S. Kang, "Neural network-based coronary heart disease risk prediction using feature correlation analysis," *Journal of Healthcare Engineering*, Article ID 2780501, 2017.

[4] P. S. Linda, W. Yin, P. A. Gregory, Z. Amanda, and G. Margaux, "Development of a novel clinical decision support system for exercise prescription among patients with multiple cardiovascular disease risk factors," *Mayo Clinic Proceedings: Innovations, Quality & Outcomes*, vol. 5, no. 1, pp. 193–203, 2021.

[5].Mirza Muntasir Nishat, Fahim Faisal, Ishrak Jahan Ratul, Abdullah Al-Monsur, Abrar Mohammad Ar-Rafi, Sarker Mohammad Nasrullah, Md Taslim Reza, Md Rezaul Hoque Khan, "A Comprehensive Investigation of the Performances of Different Machine Learning Classifiers with SMOTE-ENN Oversampling Technique and Hyperparameter Optimization for Imbalanced Heart Failure Dataset", Scientific Programming, vol. 2022, Article ID 3649406, 17 pages, 2022. https://doi.org/10.1155/2022/3649406.

[6].S. Elyassami and A. Ait Kaddour, "Implementation of an incremental deep learning model for survival prediction of cardiovascular patients," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 1, pp. 101–109, 2021.

[7]. Rahayu, J. Purnama, A. Pohan, F. Nugraha, S. Nurdiani, and S. Hadianti, "Prediction of survival of heart failure patients using random forest," *Journal of Pilar Nusa Mandiri*, vol. 16, no. 2, pp. 255–260, 2020.

[8] L. Yang, H. Wu, X. Jin et al., "Study of cardiovascular disease prediction model based on random forest in eastern China," *Scientific reports*, vol. 10, no. 1, pp. 5245–5248, 2020

[9] C. B. C. Latha and S. C. Jeeva, "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques," *Informatics in Medicine Unlocked*, vol. 16, p. 100203, 2019.

[10] Tiwari D, Bhati BS, Al-Turjman F, Nagpal B. Pandemic coronavirus disease (Covid-19): World effects analysis and prediction using machine-learning techniques. Expert Syst. 2022 Mar;39(3):e12714.

doi: 10.1111/exsy.12714. Epub 2021 May 11. PMID: 34177035; PMCID: PMC8209956

[11]. COVID-19 and Cardiovascular Disease Mina K. Chung, David A. Zidar, Michael R. Bristow, Scott J. Cameron, Timothy Chan, Clifford V. Harding III, Deborah H. Kwon, Tamanna Singh, John C. Tilton, Emily J. Tsai, Nathan R.

[12] Tucker, John Barnard and Joseph Loscalzo See fewer authors. Nishiga, M., Wang, D.W., Han, Y. et al. COVID-19 and cardiovascular disease: from basic mechanisms to clinical perspectives. Nat Rev Cardiol 17, 543–558 (2020). https://doi.org/10.1038/s41569-020-0413-9

[13] Yogendra Narayan Prajapati, U. Sesadri, Mahesh T. R., Shreyanth S., Ashish Oberoi, and Khel Prakash Jayant, "Machine Learning Algorithms in Big Data Analytics for Social Media Data Based Sentimental Analysis", *Int J Intell Syst Appl Eng,* vol. 10, no. 2s, pp. 264 –, Dec. 2022.

[14]. Yogendra Narayan Prajapati , Dr. Manish Kumar. (2022). A REVIEW PAPER ON CAUSE OF HEART DISEASE USING MACHINE LEARNING ALGORITHMS. *Journal of Pharmaceutical Negative Results*, 9250–9259. https://doi.org/10.47750/pnr.2022.13.S09.1082.

[15]. M. S. Ram, A. Sreeram, M. Poongundran, P. Singh, Y. N. Prajapati and S. Myrzahmetova, "Data Fusion Opportunities in IoT and its Impact on Decision- Making Process of Organisations," *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2022, pp. 459-464, doi: 10.1109/ICICCS53718.2022.9788232