International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

FORECASTING COVID-19 AND MANAGING RESOURCES IN INDIA THROUGH MACHINE LEARNING APPROACHES

Dr. Vidyashree L[1*], Vinay Bharadwaj K[2], Shruthi N[3], Anusha R S[4]

[1*,2,3,4]Department of Information Science & Engineering, JSS Science and Technology University, Mysuru, Karnataka, 570006, India

Abstract

The COVID-19 pandemic has presented unparalleled obstacles to healthcare systems across the globe, with India being no different. Precise predictions of the disease's development and efficient management of hospital beds are essential for maximizing healthcare resources and guaranteeing prompt care for patients. This research introduces a novel strategy for managing COVID-19 in India by utilizing machine learning methods for forecasting and optimizing bed usage. Machine learning models, including time series analysis and epidemiological frameworks, are used to anticipate trends in COVID-19 cases, hospital admissions, and resource needs at both national and local scales. These models are regularly updated with real-time information, enabling flexible responses to the changing circumstances. Additionally, data on demographics, geography, and healthcare facilities are incorporated to offer a thorough understanding of the pandemic's effects on different regions within India. Strategies for managing resources are essential to ensure that healthcare institutions are equipped to handle increases in COVID-19 cases. Algorithms powered by machine learning are employed to allocate beds in an efficient manner, taking into account elements such as the severity of patients, availability of resources, and the geographical distribution of cases. This adaptive method allows healthcare officials to shift resources to areas of greatest need and to respond proactively to developing hotspots. The findings of this research emphasize the success of machine learning-driven forecasting and resource management in enhancing the distribution of healthcare resources during the COVID- 19 crisis. By leveraging data-centered methods, India can more effectively predict and respond to the hurdles presented by the virus, ultimately leading to improved patient care and a decrease in pressure on the healthcare system. This study offers significant insights into the realm of pandemic management and showcases the capabilities of machine learning in tackling healthcare emergencies, not just in India but also in other regions encountering similar issues.
Keywords: COVID-19, Prediction, Logistic Regression, Linear Regression, Random forest (RF), Decision Tree (DT), Gaussian Naive Bayes.

I. OVERVIEW

Strategies for forecasting COVID-19 and managing resources using machine learning in India have become crucial tools in tackling the difficulties brought on by the pandemic. In this scenario, the application of machine learning algorithms, especially decision trees, has shown considerable potential in delivering precise predictions and effective resource management. The COVID-19 pandemic has exhibited an alarming capacity for rapid transmission and placing immense pressure on healthcare systems. Precise forecasting of case counts, hospitalizations, and resource needs is vital for proactive and effective distribution of resources. Machine learning algorithms offer a robust method for analyzing large datasets to produce well-informed predictions. A commonly employed machine learning algorithm for forecasting COVID-19 is the decision tree. Decision trees are adaptable, easy to interpret, and proficient in managing both classification and regression tasks. Using previous data and pertinent attributes, decision trees can be used in forecasting to anticipate the pandemic's course. Recursively dividing the dataset into subsets according to the most important features is how decision trees operate. This results in a structure resembling a tree, with each internal node denoting a decision based on a particular feature and each leaf node denoting a

Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

classification or prediction. Decision trees can be used to forecast the number of COVID- 19 cases in the future by analyzing epidemiological variables such mask mandates, vaccination rates, population density, and others.

These models' decision-making process is simple for healthcare administrators and policymakers to comprehend and interpret, which is essential for making well-informed choices about the distribution of resources and intervention tactics. When assessing machine learning models for COVID-19 predictions, accuracy is a crucial parameter. When properly trained and optimized, decision trees have demonstrated remarkable levels of accuracy. Decision trees' interpretability further improves the model's accuracy by making it simpler to locate and correct possible sources of bias or inaccuracy.

Decision trees can be used in resource management plans in addition to forecasting. In order to guarantee that medical facilities can handle the influx of COVID-19 patients during spikes in cases, effective bed allocation is essential. Decision trees can help in this process by taking into account a number of variables, including the geographic distribution of cases, patient severity, and resource availability. Depending on their condition and the resources available, decision trees can rapidly decide which patients should be admitted, transferred to intensive care facilities, or sent home for quarantine. Healthcare institutions can adjust to the changing circumstances and distribute resources as efficiently as possible thanks to this dynamic strategy.

This initiative's main objective is to perform a thorough analysis and predictive assessment of the COVID-19 situation in India in order to forecast and analyze the impact of the pandemic. Its main goal is to provide vital support to India's efforts to mitigate the negative public health effects of the pandemic. The creation of an extremely effective and well-planned system for the fair distribution of medical supplies, including medications and critical medical equipment, to COVID-19 patients nationwide is at the heart of our objective. This comprehensive strategy includes careful resource allocation, logistical, and procurement tactics meant to guarantee the prompt and efficient delivery of these life-saving supplies to people and communities in urgent need. The program aims to obtain a thorough grasp of the changing COVID-19 situation in India by employing data-driven research and predictive modeling. This involves keeping an eye on and evaluating death, hospitalization, and infection rates. These findings are crucial for helping authorities prioritize high-need regions for focused support and for directing evidence-based decision-making. Ensure timely delivery to high-need locations and establish an effective mechanism for allocating vital medical resources, such as medications and medical equipment (such as oxygen concentrators and ventilators), to COVID-19 patients nationwide. Additionally, this initiative's success depends on a dedication to continuous monitoring and assessment, which will allow it to modify and improve its operations in response to shifting conditions. In conclusion, this all-encompassing strategy is an attempt to strengthen India's ability to control the COVID-19 pandemic. In the end, it protects the health of the populace across the country by fusing data-driven intelligence with an effective and responsive medical resource distribution system.

In light of India's varied geographies and dense population, it entails the creation of sophisticated machine learning models that can produce accurate, localized predictions of COVID-19 cases and hospital admissions. The study also intends to develop resource allocation plans that adjust to regional variations in healthcare infrastructure and changing demand. To provide administrators and healthcare professionals with up-to-date forecasts and resource recommendations, an intuitive platform will be developed. The success of the project will depend heavily on cooperation with data scientists and local health authorities, as well as respect to legal and ethical requirements. Sustained efficacy in tackling the changing COVID-19 situation in India requires constant observation and flexibility.

The pressing need to effectively combat the ongoing epidemic is the driving force behind the use of machine learning techniques, in particular decision trees, for COVID-19 forecasting and resource management in India. Our goal is to give healthcare authorities precise insights into illness patterns

Received: August 10, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

and resource needs by utilizing data-driven predictive models. This will help them make well-informed decisions and allocate resources effectively. In addition to improving patient treatment, this proactive strategy lessens the burden on healthcare systems, which eventually benefits public health and safety in these trying times.
The contribution to this paper is as follows:
1. In order to facilitate prompt and well-informed decision-making, machine learning models that can produce highly localized COVID-19 forecasts at the state, municipal, and district levels must be developed.
2. To create and put into practice dynamic resource allocation plans that, in accordance with anticipated demand, maximize the distribution of hospital beds, ventilators, and medical personnel, guaranteeing effective use of healthcare resources.
3. To provide an intuitive platform that gives administrators and medical experts instant access to forecasts and suggestions for resource allocation, enabling timely and efficient pandemic responses.
4. To work together with data scientists and local health authorities to guarantee the security, privacy, and accuracy of healthcare data while closely adhering to legal and ethical requirements.
5. In order to handle the changing dynamics of the COVID-19 pandemic in India, methods for ongoing monitoring, assessment, and system adaptation must be established. The main goals are life preservation and reducing the strain on healthcare systems.

II. RELATED WORKS

Sujath et al [1] presented a model that can forecast how COVID-19 will spread. To forecast the incidence and rate of COVID-19, we used linear, multilayer perceptron, and vector autoregression techniques on COVID-19 Kaggle data.
Mbunge et al [2] analysed to slow the spread of the coronavirus, legislators, healthcare providers, and IT specialists should think about incorporating new technology into COVID-19 contact tracing. In order to combat the COVID-19 pandemic, more research is required on how to employ new technology to increase the efficacy and efficiency of contact tracing while protecting people's privacy and safety.
Iqra Sardar et al [3] analysed the effect on PSX returns of the initial wave of COVID-19. According to earlier research, the pandemic had a significant impact on the economy. During the final phase of COVID-19, the market experienced numerous interruptions. Beyond this, developing nations like Pakistan present a distinct image. The market saw a steep drop in March when the government declared a total lockdown due to the rise in Covid-19 cases, but after precautions were taken, things got better. PSX has benefited from the reduction in public, economic, small business, and discount programs. Through the distribution of resources to the unemployed, the government has managed economic demand. I continue to anticipate that the economy will recover steadily and make both modest and significant strides during the COVID- 19 pandemic. It demonstrates that the contagion has had a greater impact on stock markets in developing nations than in developed ones. In this instance, in order to regularly measure protection, authorities must forecast COVID-19 developments, the market, and the economy. Financial support for businesses and the general public is required in these situations. By enabling local governments to satisfy consumer demand, these financial services promote economic growth and draw in investment.
Angira Katyayan et al [4] presented In the Covid era, computerization of healthcare is the only way to ensure health security, particularly in a nation like India with a sizable population and few providers or systems, according to International Journal for Multidisciplinary Research (IJFMR). Thus, the combination of AI and computer-based algorithms appears to be a strong and useful paradigm that can assist physicians. There is discussion of the benefits of computer algorithms in creating improved tracking, monitoring, and diagnosis as well as the application of artificial intelligence (AI) to aid in disease diagnosis.

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Yu -Yuan Wang et al [5] proposed Researchers frequently employ artificial models to perform research when data is limited and simulation time is short; when data is still adequate and simulation time is lengthy, researchers opt to use machine learning techniques or differential equations. Furthermore, our study demonstrates that management performance is a crucial metric for assessing how well epidemic preventive and control efforts are working. Optimizing the distribution of vaccines and medical professionals can lead to better prevention and control.

The data gathered by Kaggle is used to evaluate the possible COVID-19 impact patterns in India. It aids in future prediction and forecasting by using reliable data on confirmed cases, fatalities, and relapses throughout India over time. Regular maintenance of definitions and documentation is necessary for future research or theories. We can determine that the MLP technique produces superior anticipated results than the LR and VAR methods utilizing WEKA and Orange by examining the predicted results and contrasting them with cases from the Johns Hopkins University11 data.

Studies on the health implications of COVID-19 have relied heavily on Electronic Health Records (EHRs), which are a collection of data pertaining to patient interactions with a health system in India, including lab test results. Lessons learned during the pandemic are essential to supporting a successful and robust analytics ecosystem that can meet operational and research data needs for electronic medical records. Reliable reporting using electronic health records will become even more crucial as the nation's healthcare system grows. However, this also causes a greater gap between the clinicians and personnel who create the data and the people who use it, such as analysts, researchers, and data scientists. The methods used by hospitals, clinics, and other care facilities to gather patient data may vary, and if data creators and users are not in communication, this could result in inaccurate EHR data interpretation.

Because information must be gathered from various health systems and viewpoints, this effect is exacerbated during a public health emergency. In particular, we discuss some of the obstacles of generating high-quality, real-time data from an EHR data source and emphasize the value of the HER in assisting with pandemic response. We think that these issues can be resolved by forming a cooperative team that integrates data and analysis with clinical studies early on in treatment planning, which will lessen the workload for medical staff and produce better information.

The current system has a number of drawbacks, including:

- Financial concerns include the costs of adoption and use, ongoing maintenance, incomes related to temporary unemployment, and decreased income, as well as the ban on doctors and patients using EHRs to make home modifications. These expenses have been reported in numerous research conducted in both inpatient and outpatient settings. Maintaining electronic medical records can potentially be costly. Both software and hardware need to be updated and changed on a regular basis. End users of the HER must also receive assistance and training from providers.
- Workflow disruptions cause short-term productivity losses for medical staff and providers. End users picking up new skills leads to this product churn, which can cost businesses money.
- Patients are becoming more concerned about the possibility of patient privacy intrusions as a result of the growing number of electronically. Health information exchanged Legislators have taken action to protect patient privacy and security in an effort to allay some of these worries. The strict guidelines established by the new law make it considerably more difficult for electronic data to be accessed improperly, even though few electronic data are completely safe.
- Bad effects of electronic health information include increased medical attention, bad emotions, energy level changes, and technological intrusion. As was already indicated, because of inadequately planned interventions or a lack of end-user training, researchers have discovered a link between the usage of CPOE and an increase in medical errors.

The suggested system was created to address the shortcomings of the current one. The suggested solution has generated a lot of interest in fusing machine learning techniques with medical applications, including medical imaging and patient monitoring. In this instance, this hybrid

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

application is a good substitute for any type of treatment, including disease control or cancer diagnosis. To the best of our knowledge, there isn't yet a thorough and in-depth explanation of the machine learning process, including all of its components and applications, in the current pandemic. With an emphasis on the significance of advancements in surveillance, tracking, image analysis applications, social issues, and more, our research attempts to give a thorough review of the contemporary applications planned employing this ML technique. Examining the most fascinating topic of study while taking into account recent research on the many ML implementations set up for every part of the outbreak is where our work has the biggest impact. Modeling, survival analysis, forecasting, business and complex domains, analytical techniques, medicine discovery, and hybrid applications are some of the strategies employed in the current pandemic, according to the report. The majority of these applications make use of machine learning techniques including Gaussian Naive Bayes, Random Forest (RF), Logistic Regression, Linear Regression, and Decision Trees (DT). We looked at practically all of the characteristics for each classification and utilized ML algorithms to do so. Machine learning algorithms have been employed to develop resource management systems that optimize the allocation of resources, including ICU beds and ventilators. Based on this, several advantages they are,

- Using a data-driven approach with machine learning classifiers, we can easily analyse and predict COVID-19 in India.
- Provides valuable insights into the dynamics of the COVID-19 outbreak in India.
- Time consumption is reduced.
- Reduced financial losses.
- Reduce the risk of patient security violation.

III. DATA SET

Dataset Name: The dataset we are referring to might have a specific name on Kaggle, such as "Covid- 19 in India".

Size: This dataset is separated into four smaller datasets: Individual Details, HospitalBedsIndia, Covid_19_India, and AgeGroupDetails. As seen in Table 1, Figures 1 and 2, each subdataset has four rows with four columns, nine rows with nine columns, ten rows with thirteen columns, and ten rows with twelve columns, respectively.

Division into Training and Testing Sets: In machine learning, splitting a dataset into two subsets—a training set and a testing set—is standard procedure. Machine learning models are trained on the training set, and their performance is assessed on the testing set. This aids in assessing the model's fit to fresh, untested data. The goal of the Covid-19 forecasting and bed management strategy is to anticipate Covid-19 and distribute medical resources to patients by utilizing machine learning and data analysis techniques. For competitors in various competitions, Kaggle offers extensive information and resources.

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

1	Sno	Date	Time	State/Unio	ConfirmedIndianNational	ConfirmedForeignNational	Cured	Deaths	Confirmed
2	1	30-01-2020	6:00 PM	Kerala	1	0	0	0	1
3	2	31-01-2020	6:00 PM	Kerala	1	0	0	0	1
4	3	01-02-2020	6:00 PM	Kerala	2	0	0	0	2
5	4	02-02-2020	6:00 PM	Kerala	3	0	0	0	3
6	5	03-02-2020	6:00 PM	Kerala	3	0	0	0	3
7	6	04-02-2020	6:00 PM	Kerala	3	0	0	0	3
8	7	05-02-2020	6:00 PM	Kerala	3	0	0	0	3
9	8	06-02-2020	6:00 PM	Kerala	3	0	0	0	3
10	9	07-02-2020	6:00 PM	Kerala	3	0	0	0	3
11	10	08-02-2020	6:00 PM	Kerala	3	0	0	0	3
12	11	09-02-2020	6:00 PM	Kerala	3	0	0	0	3
13	12	10-02-2020	6:00 PM	Kerala	3	0	0	0	3
14	13	11-02-2020	6:00 PM	Kerala	3	0	0	0	3
15	14	12-02-2020	6:00 PM	Kerala	3	0	0	0	3
16	15	13-02-2020	6:00 PM	Kerala	3	0	0	0	3
17	16	14-02-2020	6:00 PM	Kerala	3	0	0	0	3
18	17	15-02-2020	6:00 PM	Kerala	3	0	0	0	3
19	18	16-02-2020	6:00 PM	Kerala	3	0	0	0	3
20	19	17-02-2020	6:00 PM	Kerala	3	0	0	0	3
21	20	18-02-2020	6:00 PM	Kerala	3	0	0	0	3
22	21	19-02-2020	6:00 PM	Kerala	3	0	0	0	3
23	22	20-02-2020	6:00 PM	Kerala	3	0	0	0	3

Figure 1: Covid-19 cases details

Age Group Details:

Table 1: Age group Details

Sl No	Age Group	Total Cases	Percentage
1	0-9	22	3.18%
2	0-10	27	3.90%
3	20-29	172	24.86%
4	30-39	146	21.10%
5	40-49	112	16.18%
6	50-59	77	11.13%
7	60-69	89	12.86%
8	70-79	28	4.05%
9	9>=80	10	1.45%
10	10 Missing	9	1.30%

Received: August 10, 2025

Covid_19_India:

Hospital Beds India:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Sno	State/UT	NumPrimai	NumComm	NumSubDi:	NumDistric	TotalPublic	NumPublic	NumRuralH	NumRuralB	NumUrbanI	NumUrbanB
2	1	Andaman 8	27	4		3	34	1246	27	575	3	500
3	2	Andhra Pra	1417	198	31	20	1666	60799	193	6480	65	16658
4	3	Arunachal F	122	62		15	199	2320	208	2136	10	268
5	4	Assam	1007	166	14	33	1220	19115	1176	10944	50	6198
6	5	Bihar	2007	63	33	43	2146	17796	930	6083	103	5936
7	6	Chandigarh	40	2	1	4	47	3756	0	0	4	778
8	7	Chhattisgar	813	166	12	32	1023	14354	169	5070	45	4342
9	8	Dadra & Na	9	2	1	1	13	568	10	273	1	316
10	9	Daman & D	4	2		2	8	298	5	240	0	0
11	10	Delhi	534	25	9	47	615	20572	0	0	109	24383
12	11	Goa	31	4	2	3	40	2666	17	1405	25	1608
13	12	Gujarat	1770	385	44	37	2236	41129	364	11715	122	20565
14	13	Haryana	500	131	24	28	683	13841	609	6690	59	4550
15	14	Himachal P	516	79	61	15	671	8706	705	5665	96	6734
16	15	Jammu & K	702	87		29	818	11342	56	7234	76	4417
17	16	Jharkhand	343	179	13	23	558	7404	519	5842	36	4942
18	17	Karnataka	2547	207	147	42	2943	56333	2471	21072	374	49093
19	18	Kerala	933	229	82	53	1297	39511	981	16865	299	21139
20	19	Lakshadwe	4	3	2	1	10	250	9	300	0	0
21	20	Madhya Pra	1420	324	72	51	1867	38140	334	10020	117	18819
22	21	Maharashtr	2638	430	101	70	3239	68998	273	12398	438	39048
23	22	Manipur	87	17	1	9	114	2562	23	730	7	697

Figure 2: Hospital Beds Details

IV. PROPOSED METHODOLOGY

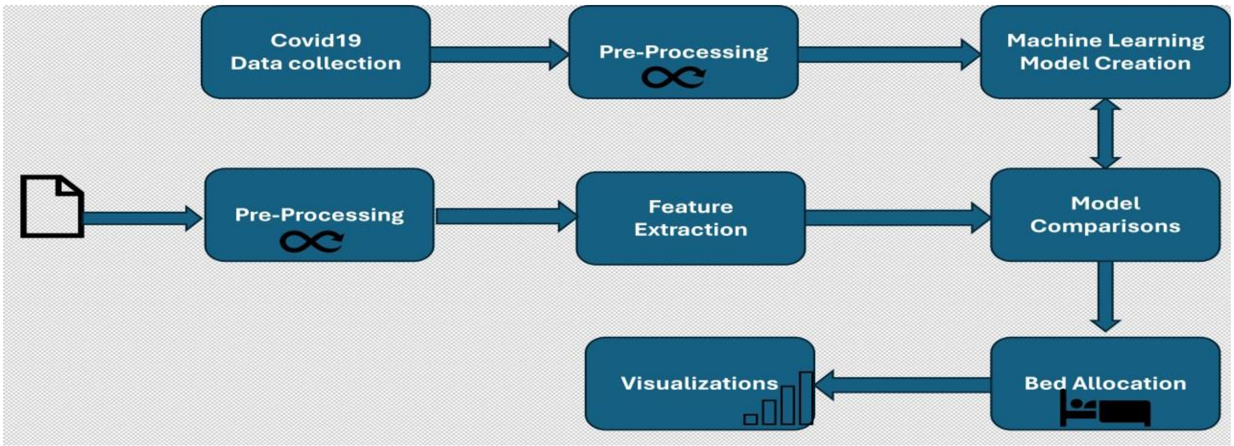

Figure 3: Proposed System

In the larger framework of public health management, the proposed machine learning-based COVID-19 forecasting and resource management system in India is intended to offer a responsive and comprehensive solution. The system's fundamental components are bed management, modeling, pre-processing, data gathering, and visualization. The data collection module ensures accuracy and completeness by obtaining hospital data and real-time COVID-19 statistics. To get data ready for modeling, pre-processing entails feature extraction and data cleansing. The

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

machine learning models that predict COVID-19 cases and bed demand are at the core of the system. These models make use of both recent and historical data to forecast future patterns. In order to optimize resource allocation, bed management strategies assign hospital beds based on forecasts and real-time updates. Effective hospital resource management depends on the system's responsiveness and agility. Stakeholders can monitor hospital locations and bed availability thanks to interactive maps and dashboards created with visualization tools. The resource status may be quickly understood thanks to color-coded indicators. All things considered, this design facilitates data-driven decision-making, improves resource allocation, and helps respond quickly to India's changing COVID-19 situation, all of which eventually lead to improved healthcare management and results (Figure 3).

a. Context Diagram

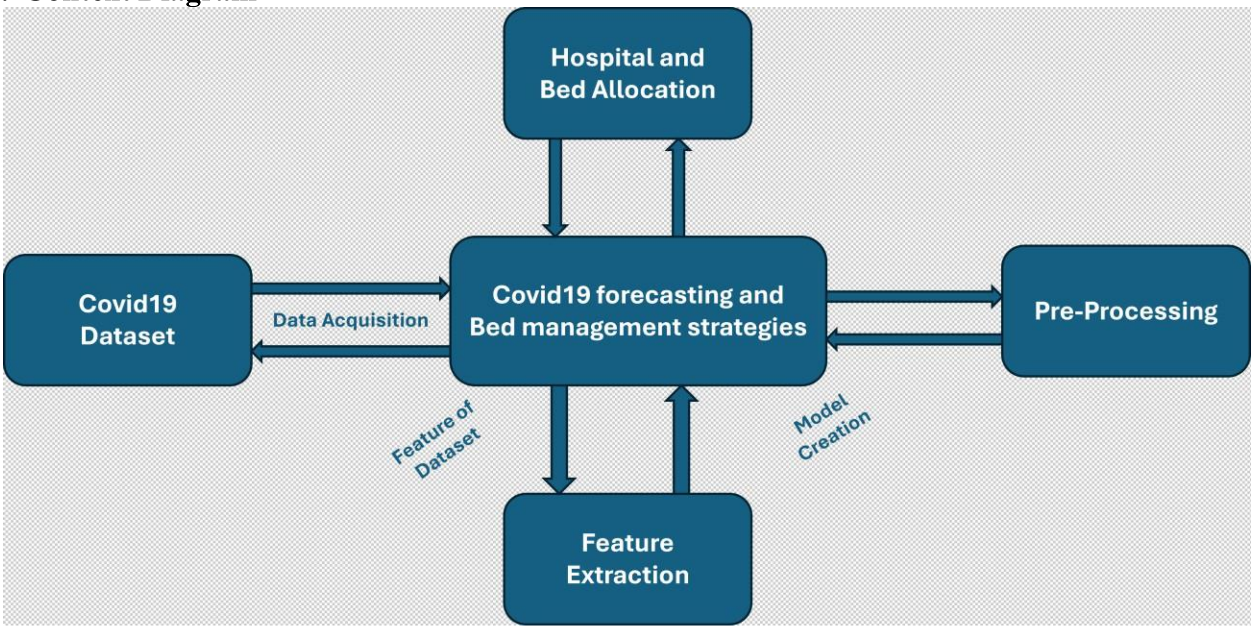

Figure 4: Context Diagram

A machine learning algorithm that forecasts the number of COVID-19 patients and the demand for beds is at the heart of the system. These models offer insights into upcoming patterns by utilizing both historical and contemporary data. By distributing hospital beds according to projections and in-the-moment modifications, bed management techniques enhance resource allocation. By extracting features from datasets, hospital and bed allocation can be based on this. To get the data ready for modeling, preprocessing entails data cleaning and specific removal procedures (Figure 4).

V. RESULTS AND DISCUSSION

This system makes extensive use of data collection in addition to the data pre-processing technique. Additionally, a number of metrics were employed to analyze the research's performance. Eventually, within our model and that of multiple classifiers, comparative analysis is performed.

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

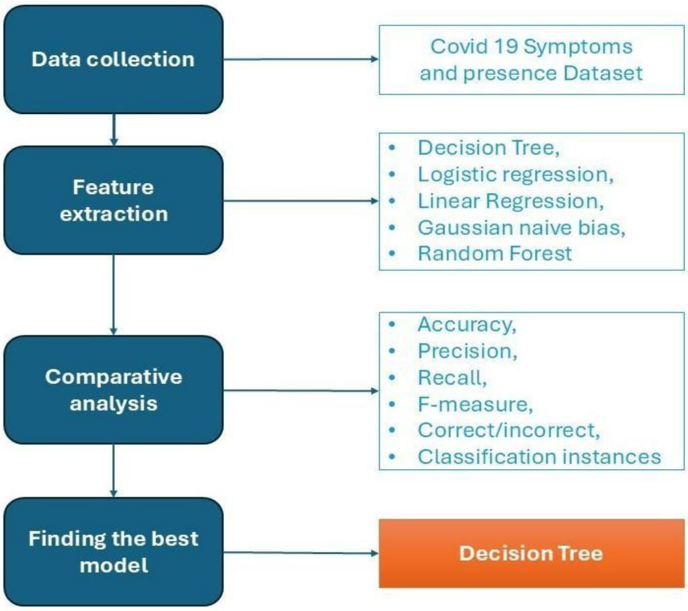

Figure 5: System Architecture

a. Data Collection

i. This step involves gathering data from an outside source, in this case Kaggle.

ii. The dataset has a large number of rows and columns, indicating that each instance has multiple data points (instances) and distinct qualities (features).

iii. The data may contain patient and hospital information as well as other characteristics that may be important for predicting COVID-19 and allocating medical resources to those who need them, among other things.

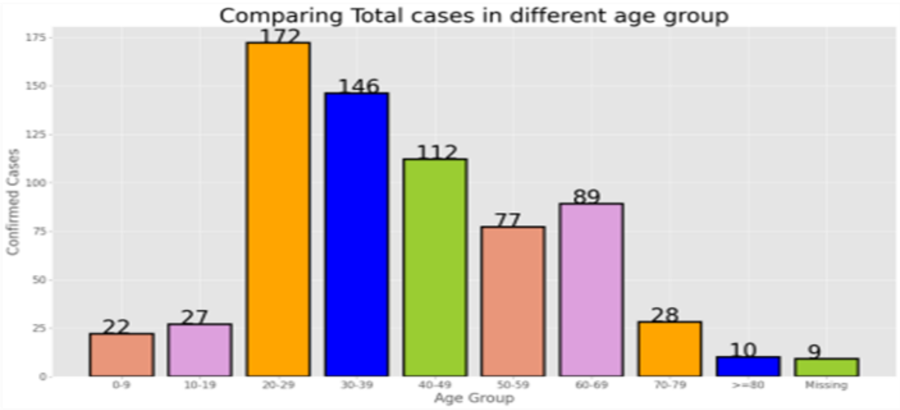

Figure 5.2: Comparing total cases

X-Axis (Horizontal Line): The dataset's "Agegroup" column provides the labels or values on the chart's horizontal line. A category on the x-axis is represented by each distinct number in the "Agegroup" column.

Y-Axis (Vertical Line): The values in the dataset's "confirmed cases" column dictate the heights of the bars on the chart's vertical line. The height of each bar represents a certain "confirmed case" value.

Bar Colors: The values in the dataset df's "age" column determine the color of the bars in the chart. Each color in the "age" column denotes a distinct category. In order to illustrate the relationship between various "agegroups" and "confirmed cases," this bar chart use color to depict the distribution or relationship between these parameters for various "age" categories.

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

b. Feature Extraction

Finding and choosing the most pertinent characteristics (or traits) that can help create precise resource allocation models is the goal of feature extraction, a crucial stage in the Covid-19 forecasting and bed management process. In this situation, we frequently use libraries such as Scikit-Learn for feature selection and dataset partitioning, and Seaborn for data visualization.

Dataset Splitting: The training set and the testing set are the two sets into which the dataset is normally separated. The train_test_split function in Scikit-Learn is used to accomplish this divide.
The resource allocation model is trained using the training approach, and its performance is assessed using the testing method.
This division guarantees that the model is evaluated on unobserved data, which is essential for determining how well it can generalize.
Feature Selection: One of the most important steps in the process is choosing the best features. Methods for choosing features, such as decision tree feature importance ratings. Retaining only those characteristics that offer useful information for differentiating between cured and confirmed instances is the aim.
We may improve the efficacy and efficiency of our Covid-19 forecasting and bed management system by utilizing Scikit-Learn for dataset splitting and feature selection methods, as well as Seaborn for data visualization. With the use of these tools, we can extract valuable insights from the data, build strong training and testing datasets, and pinpoint the most important characteristics, which eventually results in a more precise and dependable COVID detective and resource allocation model.
Patient information, age, hospital information, the number of cured deaths and confirmed cases in hospitals for each state, and other factors are important determinants of the Covid-19 forecasting and bed management system.

c. Model Training

i. Decision trees are a type of classification technique that works with binary results, which makes it appropriate for detecting viruses (whether or not COVID is present).
ii. The model learns the best weights for each feature by minimizing a cost function, usually using iterative optimization techniques like gradient descent.
iii. During training, the algorithm fits a sigmoid curve to the training data, which represents the probability of an instance belonging to the positive class (Covid is confirmed) as a function of its features.

```
model = DecisionTreeRegression() model.fit(X_train, y_train)
```

We import the DecisionTreeRegression class from scikit-learn.
We have already split our dataset into X_train (features) and y_train (target variable). We create a decision tree regression model using DecisionTreeRegression()
We fit (train) the model using the fit method, providing the training features X_train and their corresponding labels y_train.
After this code is executed, our decision tree regression model is trained and ready for making predictions on new data.

d. Model Evaluation

To determine how effectively the trained decision tree regression model operates, model evaluation is essential.

Accuracy Score: The ratio of accurately predicted instances to all instances in the dataset is determined by the accuracy score. However, when dealing with contradicting information from one class to another, facts can be deceptive.

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

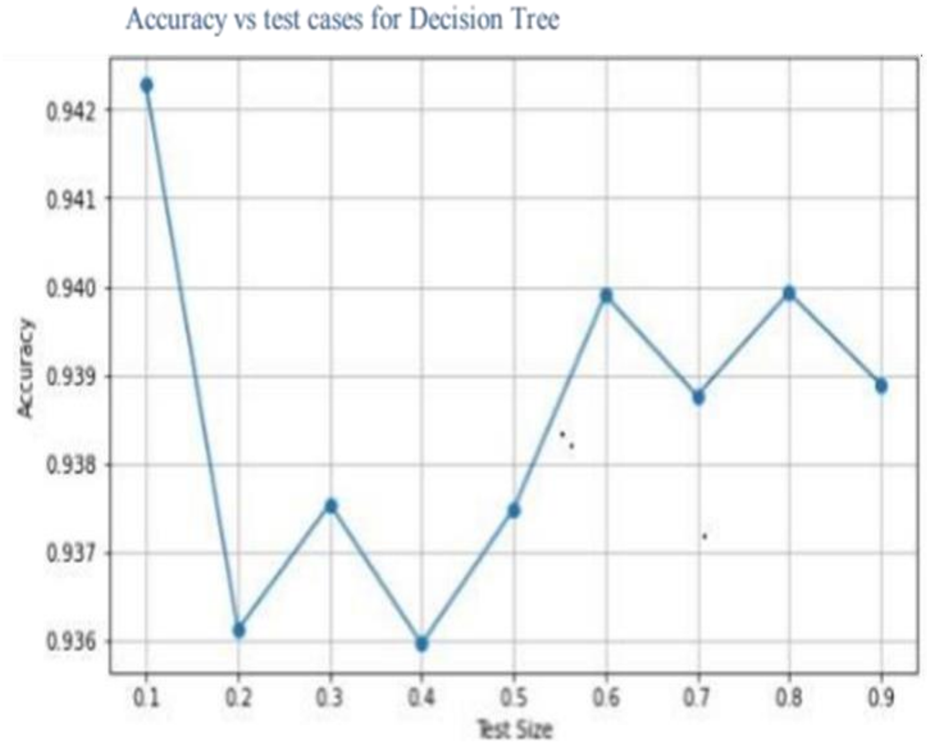

Figure 5.3: Accuracy Vs Test Cases for Decision Tree

In this case, I have made an accuracy graph in which the y-axis shows the matching accuracy scores and the x-axis shows various test sizes, from 10% to 100%.

In particular, a 93% accuracy was achieved by dividing the data into 80% for training and 20% for testing. For this particular instance, let's explain:

X-Axis (Test Size): The test size, which shows the percentage of my dataset used for testing, is represented by the x-axis. 20% is the test size I selected in this instance, meaning that 80% of the data is used for training and the remaining 20% is set aside for testing.

Y-Axis (Accuracy): The machine learning model's accuracy is shown on the y-axis. A popular evaluation statistic for classification tasks is accuracy, which calculates the proportion of properly predicted cases in the test set relative to all occurrences.

Result (96% Accuracy): Using 80% of the data for training and 20% for testing, it produced a high accuracy of 96.06%. This suggests that our decision tree regression model did a good job of identifying whether or not COVID was present in test instances of patients.

With an accuracy of 96.06%, our model was able to accurately predict the results of almost 93% of the test cases involving COVID patients in the test set. This is a good result that suggests your model can effectively determine whether COVID-19 is present or not.

e. Virus Detection

The decision tree regression model can be used to forecast new, unseen transactions once it has been trained and assessed. As seen in Figure 5.4, the model determines a probability score for every transaction that ranges from 0 to 1, representing the possibility that the transaction is infected with a virus.

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

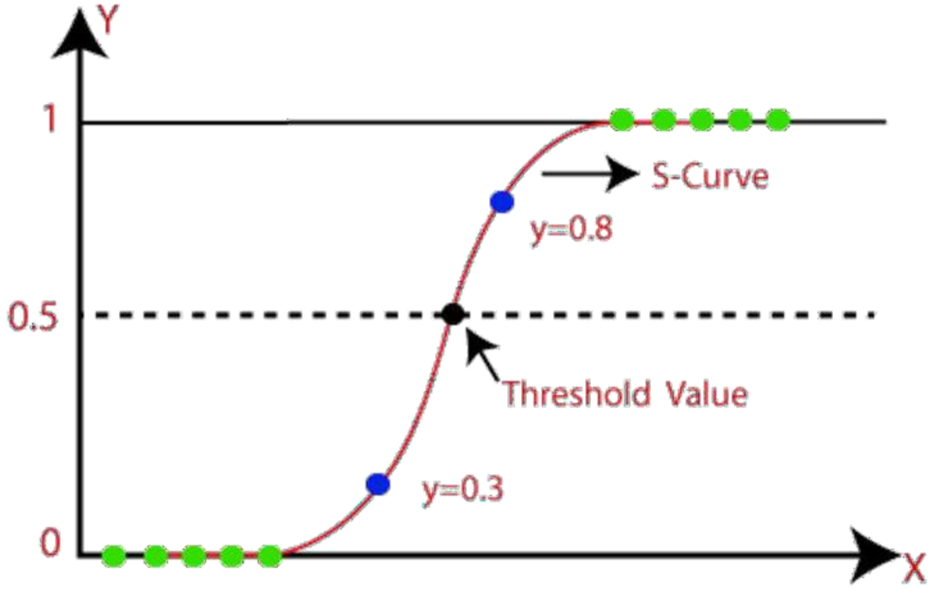

Figure 5.4: Decision Tree Regression Model Evaluation

A threshold value, typically 0.5, is chosen, and if the probability score is higher than this threshold, the transaction is categorized as indicating that the patient has COVID; if not, it is deemed to be free of COVID.

The estimated probabilities produced by a decision tree regression model are plotted against the pertinent dataset features in the graph we are discussing. In this case, the y-axis shows the anticipated viral probability, and the x-axis shows the pertinent attributes. The aforementioned threshold, which is usually set at 0.5, acts as the decision boundary: predictions with probabilities higher than 0.5 are categorized as having COVID-19, whereas predictions with probabilities lower than 0.5 are categorized as not having COVID-19. This graph's operation is explained as follows:

X-Axis (Relevant Features): The x-axis shows the pertinent characteristics or features you have chosen for your model. Our decision tree regression algorithm uses these characteristics as inputs to forecast each patient's chance of confirmation. Every point on the x-axis represents a distinct quality or characteristic. **Y-Axis (Predicted Probabilities):** The decision tree regression model's anticipated probabilities are shown on the y-axis. The model determines a probability score between 0 and 1 for each patient's details in our dataset, which represents the patient's likelihood of having COVID-19. The expected chance of confirmation for each test scenario is represented by a point on the y-axis.

Threshold (0.5): The classification of insurance claims is based on a criterion of 0.5. The model classifies a claim as a confirmed case (positive class) if the projected probability for that claim is greater than or equal to 0.5. The model classifies it as not confirmed (negative class) if the predicted probability is less than 0.5.

For instance, the model defines a test case as a confirmed case if the anticipated probability is 0.7, which is greater than the 0.5 threshold. On the other hand, the model classifies a test case as not confirmed if the predicted probability is 0.3, which is below the threshold.

VI. EXPERIMENTAL ANALYSIS

The term "experimental analysis" is a methodical, controlled inquiry or research carried out to collect empirical data, better comprehend a certain occurrence, or test theories. It entails changing one or more variables while maintaining the same values for the others in order to see how they affect and relate to one another.

Scientific study frequently uses experimental analysis, particularly in disciplines like physics,

Received: August 10, 2025

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

chemistry, biology, psychology, and the social sciences. Establishing causal linkages, identifying cause-and-effect patterns, and offering evidence in favor of or against theories or hypotheses are the objectives.
A key component of the scientific method, experimental analysis advances knowledge across a range of domains by offering empirical support and a foundation for deriving findings or making well-informed judgments.

Table 6.1 Experimental Results

Classifiers	Accuracy
Decision Tree	96.06%
Logistic Regression	86.49%
Linear Regression	85.07%
Random Forest Classifier	94%
Gaussian Naive Bayes	31.09%

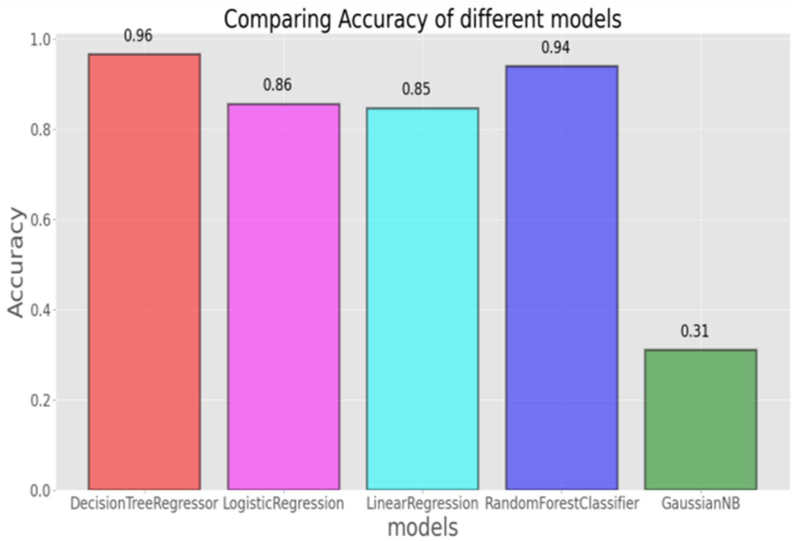

Figure 5.5: Comparative Analysis of different Models

One non-parametric supervised learning approach used for both classification and regression problems is the decision tree. A root node, branches, internal nodes, and leaf nodes make up its hierarchical tree structure.
By dividing the source set into smaller groups according to attribute selection measures, a tree can be learned. One method for assessing the value of various qualities for data classification in decision tree algorithms is attribute selection analysis. Recursive partitioning is an iterative procedure that repeats this operation in each partition. When the target variable is the same in every subset of a node or when the partition does not improve the prediction, recursion is successful. The decision tree's structure is appropriate for the information search because it doesn't require any registration or configuration data. Large amounts of data can be processed via decision trees. Segmentation and so expanding the data is the aim of ASM.

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

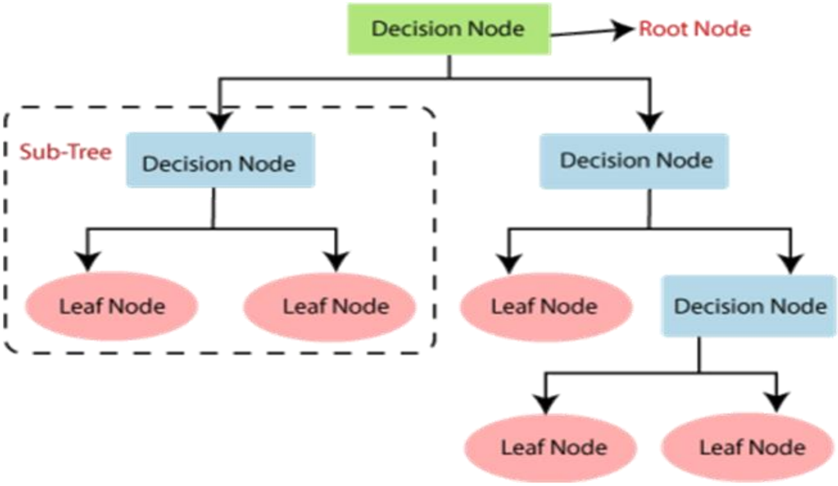

Figure 5.6: Decision Tree

With a high accuracy of 96.06%, Decision Tree (DT) is used to forecast Covid-19 and allocate resources, making it a useful model for determining whether Covid is present or not.

A supervised machine learning technique called logistic regression is mostly applied to classification tasks in which estimating the likelihood that an instance will belong to a particular class is the main objective. It is known as logistic regression and is utilized in classification algorithms. Because it utilizes a sigmoid function to estimate the probability for the specified class and takes the output of the linear regression function as input, it is known as regression.

The distinction between logistic regression and linear regression is that the former predicts the likelihood of a class occurring, whilst the latter provides a fixed value that can be any value.

- The results of categorical dependent variables are predicted by logistic regression. As a result, the outcomes need to be definitive or distinct.
- It may be true or false, 0 or 1, Yes or No, etc.
- With the exception of their respective applications, logistic regression and linear regression are very similar.
- In logistic regression, we fit a "S" shaped logistic function that predicts two maximum values (0 or 1) rather than a regression line.
- The logistic function curve shows the probability of things like whether or not the cells are malignant, whether or not a mouse is obese based on its weight, etc.
- Because it can categorize new data using continuous and discrete datasets and provide probabilities, logistic regression is an important machine learning technique.

Logistic Regression (LR) is applied in Covid-19 detection, achieving a accuracy of 86.49%, making it an effective model for classifying Covid or not cases.

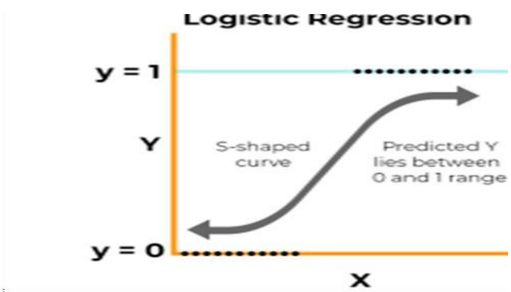

Figure 5.7: Logistic Regression

One kind of supervised machine learning algorithm is called linear regression, which determines the linear connection between one or more independent features and a dependent variable. It is referred

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

to as univariate linear regression when the number of independent features is one, and multivariate linear regression when there are several features. The algorithm's objective is to identify the optimal linear equation that, given the independent variables, can forecast the value of the dependent variable. The relationship between the independent and dependent variables is shown by the equation as a straight line. The dependent variable's change for every unit change in the independent variables is indicated by the line's slope.

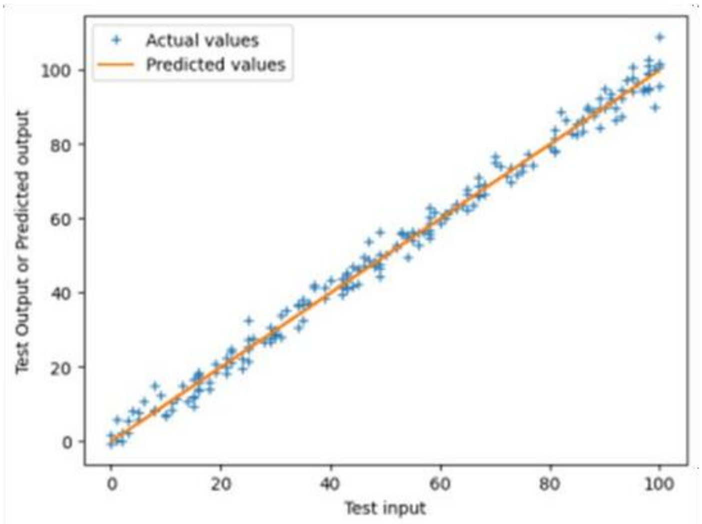

Figure 5.8: Linear Regression

Leo Breiman and Adele Cutler developed the machine learning algorithm Random Forest, which aggregates the output of several decision trees to arrive at a single conclusion. Because it can handle both regression and classification issues, it is simple to use and implement. The two major procedures that Random Forest alludes to are: "Random observations to grow each tree." To keep them apart, pick a different size. Every decision tree in a random forest generates forecasts on its own, and the values are either averaged (regression) or removed (maximum vote) to arrive at the final price. This model's benefit is that the features are used to create various trees with various sub-features. Each tree's features are chosen at random, therefore the trees.

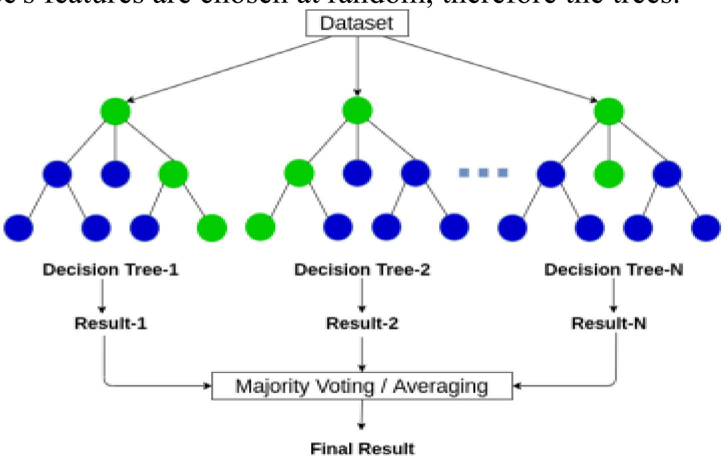

Figure 5.9: Random Forest

Random Forest is a well-liked option because of its capacity to manage intricate, high- dimensional data and its resistance to over-fitting. Covid-19 detection tasks, with claimed accuracy rates as high as 94%. A quick and easy machine learning approach for classifying datasets is called Naive Bayes. It can be applied to both multi-format and binary classification. It outperforms other algorithms on a wide range of prediction classes. The Naive Bayes model will give your test dataset a probability

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

of zero and be unable to predict anything if it comprises distinct classes that are absent from the training data.

CONCLUSION

An important role is played by resource allocation, which includes the distribution of material and human resources. It is crucial to properly match resources with the scope, budget, and schedule in order to ensure effective task execution and efficient use. Because it can significantly affect the efficiency and cost-effectiveness, resource allocation is not just a practical decision but also a strategic one. Robust project management, informed by best practices and established procedures, is essential to the implementation phase. The key players in this phase ensure that predetermined milestones are met. Their capacity to handle complexity, foresee difficulties, and act quickly is essential to maintaining the research direction and guaranteeing its successful completion. Experimental analysis, a fundamental part of the scientific process, provides empirical support and a basis for drawing inferences or reaching well-informed conclusions, advancing knowledge in a variety of fields.

Decision Tree (DT) is an effective technique for detecting the presence or nonexistence of Covid-19 because it forecasts the virus and allocates resources with a high accuracy of 96.06%.

FUTURE WORK

In the context of an Indian machine learning-based COVID-19 forecasting and bed management system, future research promises ongoing innovation, adaptation, and improvement to meet changing difficulties. The development of machine learning algorithms for COVID-19 predictions should be the main focus of future research. Predictions may be more accurate if additional data sources are included, such as environmental factors, community migration, and genomic data. Model accuracy and the capacity to identify new trends and variations can be further enhanced by utilizing developments in artificial intelligence, deep learning, and natural language processing. The ability of the system to react quickly to changing conditions can be improved by creating automated systems for real-time data integration from several testing facilities and healthcare facilities. Using Internet of Things (IoT) devices to collect and transmit data continuously can give decision-makers access to a constant flow of pertinent data. When it comes to resource allocation and limited outbreaks, geographic analysis can offer important information. Geospatial modeling approaches should be investigated in future studies in order to identify high-risk regions and direct focused interventions. Decision-makers can improve their reactions by using Geographic Information Systems (GIS) to display and evaluate the geographic distribution of resources and the virus's transmission. To further optimize resource allocation, advanced optimization methods should be investigated, taking into account not only bed availability but also variables like patient transit logistics and healthcare professional availability.

REFERENCES

1. Sujath, R. A. A., Jyotir Moy Chatterjee, and Aboul Ella Hassanien."A machine learning forecasting model for COVID-19 pandemic in India." Stochastic Environmental Research and Risk Assessment 34 (2020): 959-972.
2. Mbunge, Elliot."Integrating emerging technologies into COVID-19 contact tracing: Opportunities, challenges and pitfalls." Diabetes & Metabolic Syndrome: Clinical Research & Reviews 14.6 (2020): 1631-1636.
3. Sardar, Iqra, et al. "MACHINE LEARNING-BASED COVID-19 FORECASTING: IMPACT

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

ON PAKISTAN STOCK EXCHANGE." International Journal of Agricultural & Statistical Sciences 17.1 (2021).

4. Katyayan, Angira, Adri Katyayan, and Anupam Mishra. "Enhancing India's health care during COVID era: role of artificial intelligence and algorithms." Indian Journal of Otolaryngology and Head & Neck Surgery 74.Suppl 2 (2022): 2712-2713.
5. Wang, Yu-Yuan, et al. "Optimal resource allocation model for COVID-19: a systematic review and meta- a nalysis." BMC Infectious Diseases 24.1 (2024): 200.
6. "Machine learning for COVID-19 prediction and patient management: A review of current applications and future prospects" Rajaraman, V., et al. 2021.
7. "COVID-19 outbreak prediction with machine learning" Devi, S., et al. 2021.
8. "Machine learning-based forecasting of COVID-19 spread in India" Chimmula, V. K. R. 2020.
9. "Machine learning approaches for predicting and mapping COVID-19 in India: A review" Arora, P., et al. 2020".
10. A review of machine learning-based forecasting models for COVID-19 and comparisons with the traditional models" Hossain, M. S., et al. 2021.
11. "A review of machine learning techniques for COVID-19 vaccine development" Thakur, V., et al. 2021.
12. "Machine learning-based bed management in Indian hospitals during the COVID-19 pandemic" Jain, S., et al. 2021.
13. "Machine learning-based COVID-19 severity prediction in India" Kumar, A., et al. 2021.
14. "Machine learning-driven vaccine distribution strategies for COVID-19 in India"Reddy, S., et al. 2021.
15. "Integration of contact tracing and machine learning for COVID-19 management in India" Sharma, S., et al.2020.
16. "AI-enhanced diagnosis of COVID-19 in Indian healthcare settings" Verma, S., et al. 2020.
17. "Machine learning for early warning systems in COVID-19 management in India"Gupta, R., et al. 2020.
18. "Machine learning-based resource allocation for COVID-19 patients in Indian hospitals" Tiwari, S., et al. 2021.
19. "Optimizing COVID-19 testing strategies in India using machine learning" Singh, P., et al. 2020.
20. "Machine learning for modeling the impact of COVID-19 in India" Mishra, D., et al. 2021.

Received: August 10, 2025