European Journal of Molecular & Clinical Medicine

ISSN 2515-8260 Volume 10, Issue 3, Winter 2023

PROSPECTIVE PROJECTION ON COVID-19 UTILISING INTEGRATED MACHINE LEARNING MODELS

Panjala Sravani[1], V. Rama Krishna[2]

[1]PG Student Department of CSE, Anurag University, Hyderabad, Telangana

[2]Assistant professor, Department of CSE, Anurag University, Hyderabad, Telangana [1]

panjalasravani7@gmail.com , [2] ramakrishnacse@cvsr.ac.in

ABSTRACT:

ML-based simulations have been shown to be helpful in predicting intraoperative outcomes in order to enhance judgment on the course of action. ML strategies are being used in this research to predict future COVID cases, and they are being evaluated to identify whichever algorithm is most appropriate for the COVID sample. This research confirms how ML algorithms can anticipate the proportion of upcoming Covid-19 individuals who will be harmed. Forecasting models such as LR, Random Forest, SVR, KNN, DT, and Elastic net were used to make the projections. For the upcoming 10 days, the number of newly infected cases is predicted by each model. The consequences display that the Decision Tree achieves top amongst those simulations, trailed by Linear Regression and K-NN, which are good at predicting new confirmed cases. Whereas SVR performs worst among those models using the dataset that is currently available.

Keywords: *Linear regression, Random Forest, Support vector regression, KNN, DT, Elastic net, and COVID-19.*

INTRODUCTION:

Machine Learning models are used for future projections in order to find the risk and to precautions based on that risk that may occur. Based on these machine learning models so many predictions are done previously like Stock market forecasting, weather forecasting, and disease prognosis, and lots of predictions are made for future forecasting. Now we are performing the future projections on the Covid-19 dataset to find the future risk and making comparisons of the models that we used. In this study, the predictions were done using forecasting models, including Linear regression(LR), Random Forest, SVR, KNN, Decision Tree , and Elastic net. Each model foresees the quantity of newly ill cases for the ensuing 10 days. Every model foresees the decrease in the quantity of newly definite circumstances and the Decision Tree implements best amongst the prototypes trailed by LR and KNN. Whereas SVR performed poorly among the models. The dataset is taken from https://covid19.who.int/data up to 25th November 2022.

LITERATURE SURVEY:

It is necessary to take appropriate management measures now that Covid19 has reached the pandemic status. Therefore, the authorities have adopted common models that were widely used in the media, such as simple medical specialization and applied mathematics models. However, normal models

have demonstrated poor accuracy for semi-permanent calculation in line for high level of ambiguity and a nonexistence of vital information. To reduce uncertainty, the author SinaF. Ardabili[1] contrasts machine learning and computational intelligence algorithms with especially vulnerable (SEIR) and vulnerable (SIR) approaches to estimate the prevalence of COVID-19. Organic process algorithms (EA), Genetic formula (GA), PSO, GWO, MLP, and Conciliating Network-Based Fuzzy Reasoning Systems are some of the techniques and models utilized to produce accurate results (ANFIS). An inference system that relies on a reconciled net (ANFIS) and inter perceptron's (MLP) both exhibited positive outcomes. The report provides preliminary testing to highlight the potential of learning algorithms for additional investigation.

The prognostication techniques and any forecasts of COVID-19 for recently diseased circumstances, quantity of deceases, and thus the quantity of reclamations within the following ten existences have been introduced by the author Furqan Rustam[2]. According to the author, this method is improbably useful in wishing for appropriate steps to counteract the COVID-19 threat components. The model was created using ML models such as regression toward the mean (LR), slightest out-and-out disappearance and choice operator (LASSO), SVM, and ES. Of all the simulations used, metallic element performs the best, trailed by LR and LASSO which perform fit in prognosticating the newly inveterate circumstances, decease ratio, as well as rescue frequency, whereas SVM implements hostile overall in the estimation.

In order to estimate new cases or the overall quantity of ill circumstances in a reasonable situation, the author Arti M.K[3] has projected a replacement mathematical model. This forecast is crucial in order to modify medical infrastructure and move forward with future plans of action. To produce the predictions, a decision making model is used, which assumes that certain individuals are isolated and a small number are missing unrecognized for a variety of causes, such as not showing symptoms etc. These concealed nodules reveal the illness in society. A simple application of the projected model will be used to make an approximation of recent cases. The findings make it clear that lockdown is a particularly important tool for controlling how a sickness develops. The accuracy of the generated model is demonstrated by a near correspondence between the diagnostic and obtainable results.

The author Rishi raj Sharma[4] proposed a method that is used to anticipate the regular new individuals suffering from renewed coronavirus virus in Brazil, the United States, and India. Predicting the everyday innovative instances for current days is beneficial for developing time strategies. Therefore, necessary steps for preventing the symptoms of Underlying disorders are no longer active. To break down a datum into its component parts and reduce the no stationary, EVDHM is applied over the datum. The long-term temporal values for each subcomponent are predicted by an ARIMA-based model. The final word output values are also pushed by the prediction values of each subcomponent.

In India, the coronavirus epidemic is spreading rapidly and infecting the infrastructure. It might be anything from a severe infection to a respiratory condition. The patient wants a ventilator and hospital beds. It is impossible to assign beds for a wide variety of patients, nevertheless. The author Dr. Gayathri[5] has since created a formula for increased resource allocation that is frequently used in hospitals to allocate beds. Hospital bed allocation essentially eliminated the perceptions of queueing concept and utilised the principle of first in, first aided. The models and algorithms used to address the problem of inefficient resource allocation include organic process algorithms, the skeeter hawk formula, and queueing theory.

LITERATURE REVIEW TABLE:

Author and Year	Model Utilize	Performance metrics	Demerits
Aayush Jain(2021)	ARIMA and SARIMA	RMSE	Applicable only on univariate time- series.
Kolla Bhanu(2020)	Decision Tree Classifier, Gaussian Naïve Bayesian Classifier	-	It mostly depends on the dataset.
Arti M.K.(2020)	Treebased model	-	We may not accurate results for hidden nodes.
Rishi raj Sharma(2021)	ARIMA, EVDHM, Time series forecasting.	-	The segment is comprised is employed to anticipate outcomes for the coming years.
William Darmawan (2020)	ARIMA model and PROPHET	R2, MS, MAE	Applicable only on univariate data.

ARCHITECTURE:

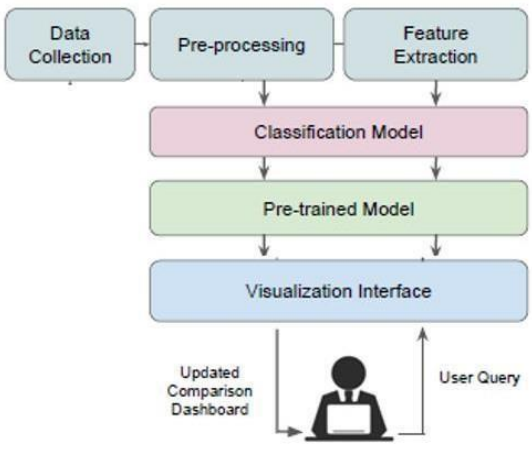

Fig: Architecture system.

ISSN 2515-8260 Volume 10, Issue 3, Winter 2023

EXISTING SYSTEM:

One of the most substantial uses of ML is anticipating. A few widely used ML procedures have existed and implemented inside the research part to advise subsequent action steps appropriate in a wide range of domains, such as meteorology, illness trying to predict, share price forecasting, as well as illness prognosis. There are numerous regress and models of neural networks that can be employed to foresee clinical symptoms for a specific illness.

Numerous research have been conducted utilising machine learning approaches to predict various diseases, including tumors, artery disease, and heart failure.

When making decisions to handle the current situation and direct early actions to manage certain diseases very successfully, these prediction algorithms can be quite helpful**.**

DISADVANTAGES OF EXISTING SYSTEM:

- Limited to COVID-19 forecasts
- Fewer MLS were taken into account**.**

PROPOSED SYSTEM:

As a contribution to the present humanitarian crisis, this work aims to create a COVID- 19 forecasting system. For the next 10 days, forecasts are made for the following three crucial diseaserelated variables:

1) The total quantity of newly confirmed instances.

2) The quantity of fatalities quantity of rescues.

The predicting problem is handled as a base classifier in this work , therefore it is created using state-of-the-art classifier path coefficients like regression analysis, Stochastic, fully convolutional, KNN, Decision Tree, and Elastic net. The COVID-19 individuals statistics data analysis, which was delivered by WHO, was used to train the learning models.

ADVANTAGES OF THE PROPOSED SYSTEM:

- Different ML techniques seem to accomplish enhanced class predictions.
- DT executes best when the time-series of data analyzed with only few entries.
- To forecast future, the majority of ML algorithms need a sample size of data; as the dataset size grows, so do the model performances.
- Forecasting using ML models can help decision-makers control pandemics like COVID- 19.

METHODOLOGY:

Decision Tree:

Using logistic regression as regression, the decision tree goes from making interpretations about an entry to making suppositions about the product's goal importance . It is one of the foretelling strategies used in stats, information retrieval, and algorithms. Decision trees are models with a goal variable that really can take a finite range of values; within that dendrogram, all edges symbolize the target class and the branching depict the convergences of traits that give rise to such classifiers. Decision trees with a linear attribute value are called extrapolation graphs. A clustering algorithm may be employed in expert judgment to depict issues aesthetically and explicitly**.**

ISSN 2515-8260 Volume 10, Issue 3, Winter 2023

Random Forest Algorithm:

An established machine learning method utilized in the continuous process is random forest. It can be applied to resolve ML issues that call for both classifiers. Its foundation is the idea of classifiers, which combines a number of categories to solve complex problems and enhance prediction results**.**

KNN Algorithm:

The K-NN technique assigns the training examples to the category that is most similar to the previous categories on the assumption that the prospective and current examples are better compared**. Linear Regression:**

A process that shows a relationship between a dependent (y) and first or perhaps more unbiased (y) parameters is alluded to as "regression analysis."

Support Vector Regression:

To forecast discrete values, Support vector regression is a trained discrete event simulator. Outcome the finest acceptable mark is the ultimate tenet of SVR**.**

Elastic Net:

Elastic Net is a linear regression extension that increases the loss function's regularisation penalty during training.

FLOW DIAGRAM:

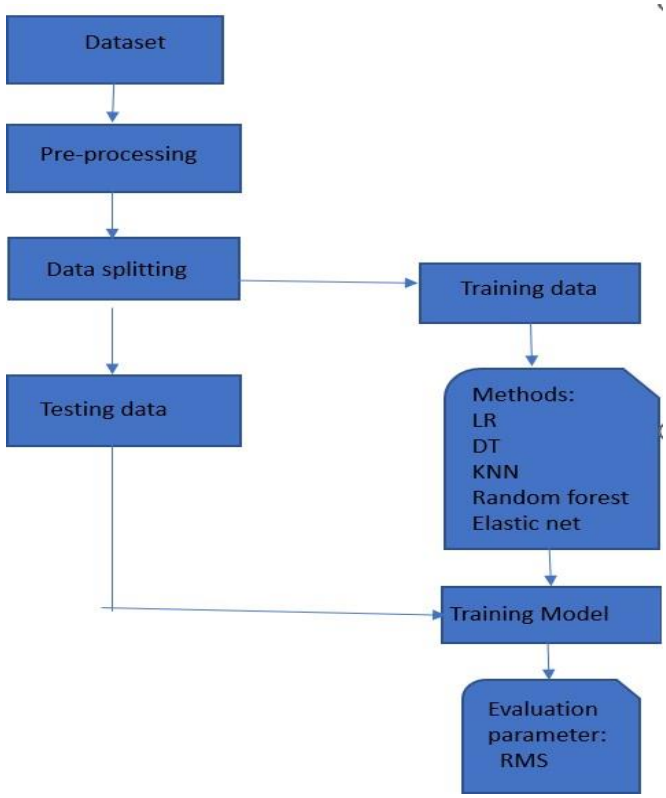

Fig: Flow chart

EXPERIMENTAL RESULTS:

Models applied on covid India and USA dataset that forecasts the following results.

ISSN 2515-8260 Volume 10, Issue 3, Winter 2023

INDIA: Predictions based on the above models using the India Covid dataset.

Forecasting of new confirmed cases using Linear Regression

Forecasting of new confirmed cases using Random forest

Forecasting of new confirmed cases using KNN

Forecasting of new confirmed cases using Decision

Forecasting of new confirmed cases using Elastic net

Forecasting of new confirmed cases using SVR

USA: Predictions based on the above models using USA Covid dataset.

Forecasting of new confirmed cases using Elastic net

Forecasting of new confirmed cases using Decision Tree

Forecasting of new confirmed cases using KNN

Forecasting of new confirmed cases using Random Forest

Forecasting of new confirmed cases using Linear Regression

Forecasting of new confirmed cases using SVR

TEST EVALUATION:

Model	**Mean Squared Error (MSE)**
SVR	6530414588.075543
Linear regression	64688240.0
KNN	32951162.0
DT	10220.761421030496
Elastic net	64688230.0
Random forests	16261209.043276483

The decision Tree gives the best results compared to other models.

CONCLUSION:

The pandemic of a COVID-19 outbreak has been predicted globally using a machine learning (ML)-based prediction algorithm. The software recognizes a collection that comprises day-by-day real prior data using ML techniques and predicts forthcoming moments. The poll's improved version is that DT works efficiently in the development research domain given the type and quantity of the dataset. To a certain extent, KNN and LR also perform well for forecasting to anticipate new situations. These two models' findings indicated that the quantity of original circumstances will decline and that the repossession frequency would be high. Due to the fluctuations in the dataset values, SVR performs poorly in every case. Setting up a precise hyper plane between the dataset's specified values was really challenging

REFERENCES:

1. Sina F. Ardabili , Amir Mosavi , Pedram Ghamisi , Filip Ferdinand , Annamaria R. VarkonyiKoczy , Uwe Reuter , Timon Rabczuk and Peter M. Atkinson ,"COVID-19 Outbreak Prediction with Machine Learning",Algorithms 2020, 13(10), 249; https://doi.org/10.3390/a13100249
2. Dr. R. Gayathri ,Dr. G. R. Damodaran ,"An Enhanced Resource Allocation Algorithm to Allocate
3. Hospital Beds during the COVID19 Pandemic", IEEE 2021.
4. Furqan Rustam, Aijaz Ahmad Reshi, Arif Mehmood, Saleem Ullah, Byung-Won, Waqar Aslam, And Gyu Sang Choi," Covid-19 Future Forecasting Using Supervised Machine Learning Models", IEEE, Volume 8 2020.
5. Rishi Raj Sharma, Mohit Kumar, Shishir Maheshwari, and Kamla Prasan Ray, "EVDHMARIMA-Based Time Series Forecasting Model and Its Application for COVID-19 Cases", IEEE, VOL. 70, 2021.
6. Vartika Bhadana, Anand Singh Jalal, Pooja Pathak, "A Comparating Study of Machine Learning Models for COVID-19 prediction in India", IEEE, 2021.
7. Aayush Jain, Tanay Sukhdeve, Himanshu Gadia, Dr. Satya Prakash Sahu, Satya Verma," COVID-19 Prediction using Time Series Analysis", ICAIS-2021.
8. Arti M.K, Kushagra Bhatnagar, "Modeling and Predictions for COVID 19 Spread in India", ResearchGate, April 2020 DOI: 10.13140/RG.2.2.11427.81444.
9. Fu-Yuan Cheng, Himanshu Joshi, Pranai Tandon, Robert Freeman, David L Reich, Madhu Mazumdar, Roopa Kohli-Seth, Matthew A. Levin, Prem Timsina and Arash Kia, "Using Machine Learning to Predict ICU Transfer in Hospitalized COVID-19 Patients", J.Clin.Med.2020, doi:10.3390/jcm9061668.

10. Eman D. El Desouky, Prediction of the Epidemic Peak of Covid19 in Egypt,medRxiv,2020, DOI: https://doi.org/10.1101/2020.04.30.20086751.

11. Dr. R. Gayathri, Dr. G. R. Damodaran, “An Enhanced Resource Allocation Algorithm to Allocate Hospital Beds during the COVID19 Pandemic”,IEEE,2021,DOI: 10.1109/ICIRCA51532.2021.9544938Reddy, G. Raghupal, and G. Radha Devi. "Security Privacy Content and Impact of Trust in Social Networks." International Journal of Advanced Information Science and Technology (IJAIST), vol. 6, issue 11, pp. 394-398.

12. G. Raghupal Reddy, and A. Shravan Kumar “An Automated Research For Analysing And Visualising The Data Using Machine Learning Analytics” Journal Of Critical Reviews, Vol. 06, Issue 05, pp. 470-475.