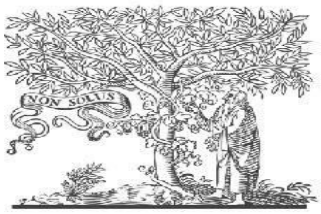
International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT



ELSEVIER
SSRN

IJIEMR Transactions, online available on 19th Nov 2022. Link :http://www.ijiemr.org/downloads.php?vol=Volume-11&issue=Issue 11

DOI: 10.48047/IJIEMR/V11/ISSUE 11/09

Title COVID-19 FUTURE FORECASTING USING SUPERVISED MACHINE LEARNING MODELS

Volume 11, ISSUE 10, Pages: 59-66

Paper Authors
Aravapalli V Pushpavathi , G Venkata Rami Reddy

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COVID-19 FUTURE FORECASTING USING SUPERVISED MACHINE LEARNING MODELS

Aravapalli V Pushpavathi , G Venkata Rami Reddy
Student , Professor
arvapallivani@gmail.com , gvr_reddi@jntuh.ac.in
contact no: 9849268292 , 9849466982
M.Tech Software Engineering,
School of Information Technology JNTUH, Hyderabad , India

ABSTRACT: COVID-19, which is currently thought to pose a threat to humanity. This study illustrates the capacity of ML models to predict the impending COVID-19 patient population. In order to improve decision-making regarding the future course of action, machine learning (ML) based forecasting methods have demonstrated their usefulness to foresee in perioperative outcomes. The detection prioritisation of negative aspects considering a threat was a long-standing requirement in many application domains where ML models were deployed. Many prediction techniques are currently in use to address forecasting issues. In this study, the threatening factors of COVID-19 were specifically predicted using standard forecasting models like linear regression (LR), least absolute shrinkage selection operator (LASSO), support vector machine (SVM), exponential smoothing (ES), random forest (RF), long short term memory (LSTM). Each of the models makes three different projections, including the number of newly infected cases, the number of fatalities, the number of recoveries in the coming days. The study's findings demonstrate that using these techniques in the present COVID-19 pandemic scenario is a promising methodology.

Keywords *– linear regression, least absolute shrinkage selection operator, support vector machine, exponential smoothing, random forest, long short term memory*

1. INTRODUCTION

Over the past ten years, machine learning (ML) has established itself as a significant subject of research by finding solutions to numerous extremely challenging complicated practical issues. As opposed to traditional calculations, which comply with programming guidelines in light of choice proclamations like if-else, ML calculations frequently advance by preliminary mistake. One of the main uses of ML is guaging; a few normal ML calculations have been applied in this field to coordinate the future game-plan expected in an assortment of utilization spaces, for example, weather conditions determining, illness estimating, securities exchange guaging, sickness expectation. Various relapse brain network models are broadly relevant in determining future patient circumstances thinking about a given illness. Various exploration have been led to gauge different infections utilizing AI draws near, including coronary corridor

illness, cardiovascular sickness, bosom malignant growth. The review is especially centered around foreseeing Coronavirus affirmed cases as well as the flare-up early response to the sickness. While pursuing choices to deal with the flow circumstance direct early activities to deal with specific infections effectively, these forecast calculations can be very useful. This work looks to make an early figure model considering the novel Covid, otherwise called SARS-CoV-2, which the World Wellbeing Association has given the authority name Coronavirus (WHO). As of now, Coronavirus represents an intense danger to human existence on a worldwide scale.

The infection was initially found toward the finish of 2019 in Wuhan, China, where numerous patients experienced pneumonia-like side effects. It has various consequences for the human body, including multi-organ disappointment serious intense respiratory disorder, the two of which can rapidly bring about death. This pandemic is influencing a huge number of people around the world, a great many individuals are passing on each day. Consistently, a huge number of new people from different countries all through the world are accounted for to be positive. The infection predominantly spreads by close private contact between people, respiratory beads, or contacting sullied objects. The way that an individual could have the infection considering a few days without showing side effects makes its proliferation the most troublesome. Practically each of the nations have forced either fractional or severe lockdowns across the affected locales urban communities because of the reasons for its spread the gamble it presents. Right now, clinical scientists from everywhere the world are attempting to find a reasonable immunization treatment choices thinking about the disease. Since there is currently no perceived treatment to kill the infection, state run administrations overall are focusing on safeguard estimates that can slow its spread. The most essential security measure is "be educated" about any Coronavirus related data. Various scientists are researching the different features of the pandemic to add to this data give the discoveries that will help mankind. among this work, we desire to add to the current helpful emergency by making a Coronavirus estimating framework. taking into account the impending days, anticipating is finished considering the accompanying three vital infection related factors: 1) the amount of spic and span, affirmed cases. The amount of lethal occurrences, the amount of recuperations.

2. LITERATURE REVIEW

Machine learning strategies considering forecasting:

The meaning of solid compelling procedures that can construe from perceptions the stochastic reliance between past future is essential because of the developing accessibility of a lot of verifiable information the prerequisite to perform precise determining of future conduct in a few logical applied spaces. Since the 1960s, straight factual methods like ARIMA models an affect the guaging field. As of late, AI models have acquired fame have

shown to areas of strength for be. The conversation of nearby learning methods as a helpful device considering working among fleeting information, the formalization of one-step determining issues as managed learning undertakings, the job of the estimating methodology while moving from one-move toward various step guaging are the three primary subjects of this part's outline of AI procedures in time series guaging.

Regression models considering prognostic prediction:

The utilization of numerous relapse models in determining the course of patients among a scope of illnesses is extremely normal. Nonetheless, a ton of specialists don't approve the vital assumptions prior to using such models. Analysts habitually "overfit" the information by making models among deficient example measures such a large number of indicator factors. It is unrealistic that models made thusly will breeze through the approval assessment on an alternate patient example. The specialist is ignorant that over fitting has occurred without performing such an approval. There are information decrease methods that can essentially improve the presentation of relapse models when the proportion of patients encountering endpoints to potential indicators is low (suppose less than 10). At the point when model presumptions are entirely analyzed, moves are made, (for example, picking another model or changing the information) when suspicions are abused, the model definition strategy doesn't bring about over fitting the information, relapse models can make expectations that are more precise than different techniques like delineation recursive parceling.

Using machine learning algorithms considering breast cancer risk prediction diagnosis:

Applications considering AI in medication incorporate distinguishing the sort of harmful cells. One of the diseases that outcome in a critical number of fatalities every year is bosom malignant growth. It is the most predominant type of malignant growth the main worldwide reason for mortality thinking about ladies. There are two sorts of carcinogenic cells: harmless (B) dangerous (M) (M). Support Vector Machine (SVM), Decision tree (CART), Innocent Bayes (NB), k Closest Neighbors are only a couple of the calculations accessible thinking about the order expectation of K-nearest neighbor (kNN). Support Vector Machine (SVM) on the Wisconsin Bosom Malignant growth dataset is used in this venture. Moreover, KNN, Naives Bayes, Truck are prepared on the dataset, the exactness of expectation considering every calculation is thought about.

Ridge regression: biased estimation considering non orthogonal problems:

It has been exhibited in different relapses that on the off chance that the expectation vectors are not symmetrical, boundary assessments in view of least remaining amount of squares are almost certain to be poor, while perhaps not off-base. An assessing technique in view of X′X's askew little sure augmentations is proposed. The edge follow, a procedure considering

showing the outcomes of nonorthogonality in two aspects, is presented. Then, at that point, it is shown how to improve X′X to get one-sided gauges among lower mean square mistake.

3. METHODOLOGY

The way that an individual could have the infection considering a few days without showing side effects makes its spread the most troublesome. Virtually each of the nations have forced either fractional or severe lockdowns across the influenced locales urban communities because of the reasons for its spread the gamble it presents. At present, clinical scientists from everywhere the world are attempting to find a reasonable inoculation treatment choices thinking about the disease. Since there is currently no perceived treatment to kill the infection, state run administrations overall are focusing on safeguard estimates that can slow its spread. The most urgent security measure is "be educated" about any Coronavirus related data. Various specialists are researching the different aspects of the pandemic to add to this data give the discoveries that will help humankind. A few calculations are utilized to gauge Coronavirus cases.

Disadvantages:

Less accuracy

This work outlines the capacity of machine learning (ML) models to foresee the quantity of impending Coronavirus patients impacted, which is presently remembered to represent a danger to mankind. In this review, the compromising elements of Coronavirus have been anticipated utilizing an assortment of ordinary guaging models, including linear regression (LR), least absolute shrinkage selection operator (LASSO), support vector machine (SVM), exponential smoothing (ES), Random forest (RF), long short term memory (LSTM). Every one of the models makes three unique projections, including the quantity of recently tainted cases, the quantity of fatalities, the quantity of recuperations before long. The review's discoveries exhibit that involving these procedures in the ongoing climate of the Coronavirus pandemic.

Advantages:

Important metrics such as mean square error (MSE), mean absolute error (MAE), root mean square error have been used to evaluate performance (RMSE).

These are a couple of the review's significant decisions:

• Certain AI techniques seem to improve in different class expectations.

• To conjecture the future, most of ML calculations need a sizable measure of information; as the dataset size develops, so do the model exhibitions.

• Anticipating utilizing ML models can assist chiefs with managing among pandemics like Coronavirus.

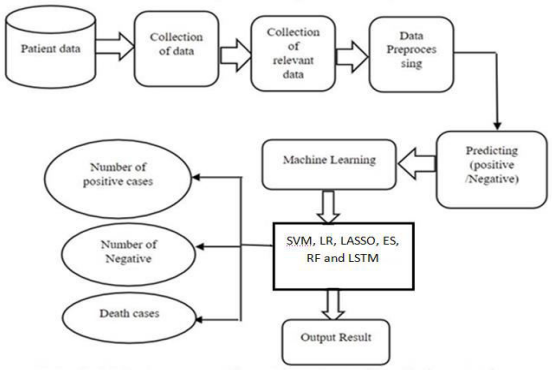


Fig.2: System architecture

MODULES:

To implement aforementioned project we have designed following modules:

- **Upload Covid19 Dataset**
  The user will upload the Covid-19 dataset from Kaggle in this module. csvformat.
- **Preprocess Dataset**
  This module will read the dataset extract the relevant information. The dataset has over 9999 records, the application uses 9969 records to train machine learning algorithms 30 records to test anticipate future cases. The dataset is now ready among train test data.
- **Run SVM Algorithm**
  In this module train SVM then perform forecasting calculate error rate.
- **Run Linear Regression Algorithm**
  In this module run linear regression algorithm.
- **Run Lasso Algorithm**
  In this module run Lasso algorithm.
- **Run ES Algorithm**
  In this module run ES Algorithm
- **Run Random forest Algorithm**
  In this module run Random forest algorithm
- **Run LSTM Algorithm**
  In this module run LSTM algorithm
- **All Algorithms Error Rate Graph**
  In this module graph x-axis represents algorithm name y-axis represents Error rate .Among all algorithms LSTM is giving less error rate.

4. IMPLEMENTATION

ALGORITHMS:

SVM:

A profound learning framework known as a Support vector machine (SVM) utilizes regulated figuring out how to characterize or foresee the way of behaving of groupings of information. Directed learning frameworks in simulated intelligence AI give input planned yield information that are marked thinking about arrangement.

LR:

An ML calculation in view of regulated learning is direct relapse. It executes a relapse activity. Relapse utilizes free factors to show an objective expectation esteem. It generally decides how elements estimating collaborate.

LASSO:

A shrinkage variable determination procedure considering straight relapse models is called rope relapse investigation. Finding the subset of indicators among the most minimal forecast mistake considering a quantitative reaction variable is the point of rope relapse.

ES:

The outstanding window capability is an overall strategy considering smoothing time series information known as remarkable smoothing. Remarkable capabilities are utilized to apply loads that decline dramatically over the long haul, instead of the essential moving normal, which loads past perceptions similarly.

RF:

An amazingly popular administered AI approach called the Irregular Backwoods Calculation is used to tackle arrangement

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

relapse issues. We know that a backwoods is comprised of many trees, the more trees there are, the more strong the timberland will be.

LSTM:

Long transient memory organizations, or LSTMs, are utilized in profound learning. Different recurrent neural networks (RNNs) can learn long haul conditions, especially in issues including succession expectation.

5. EXPERIMENTAL RESULTS

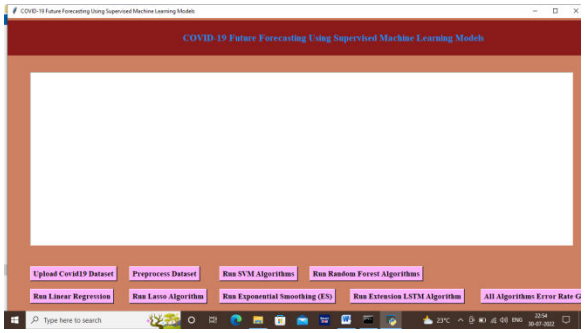


Fig.3: Home screen

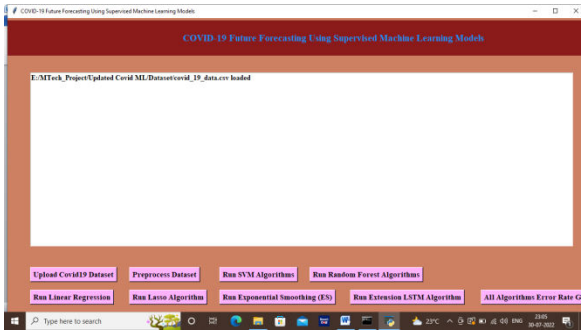


Fig.4: Upload dataset

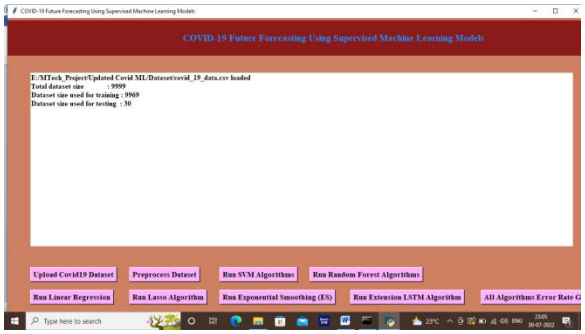


Fig.5: Preprocess dataset

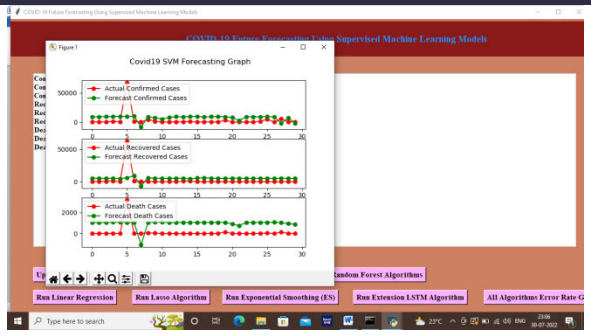


Fig.6: SVM algorithm

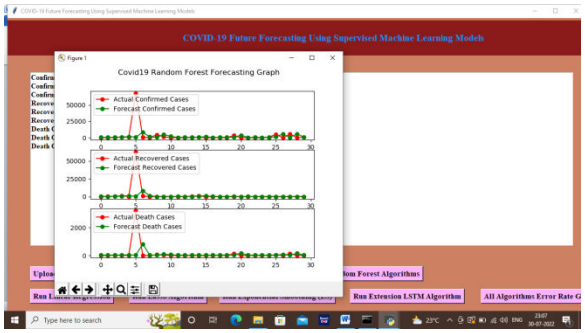


Fig.7: Random forest algorithm

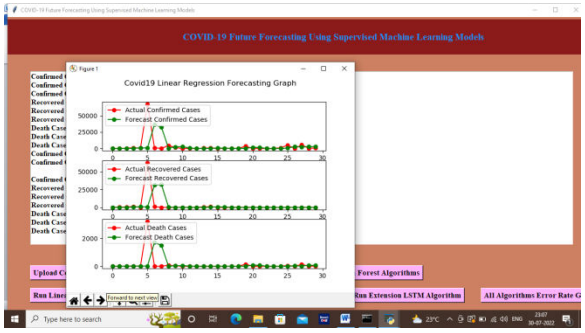


Fig.8: Linear regression algorithm

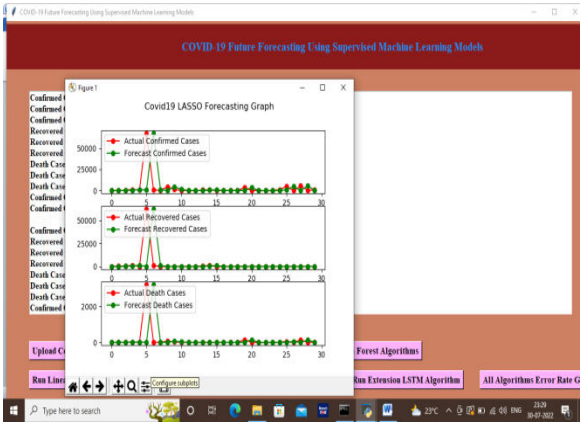


Fig.9: LASSO regression algorithm

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

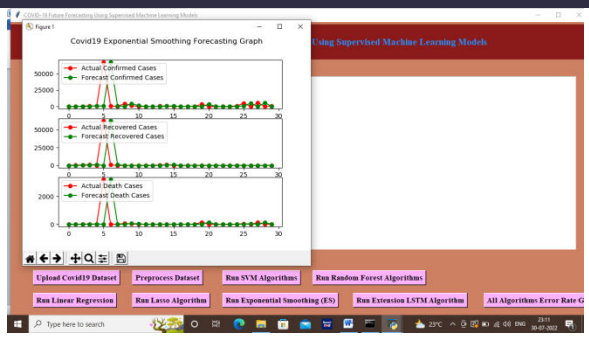


Fig.10: ES algorithm

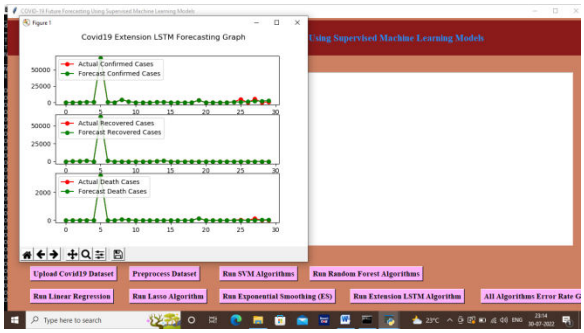


Fig.11: Extension LSTM algorithm

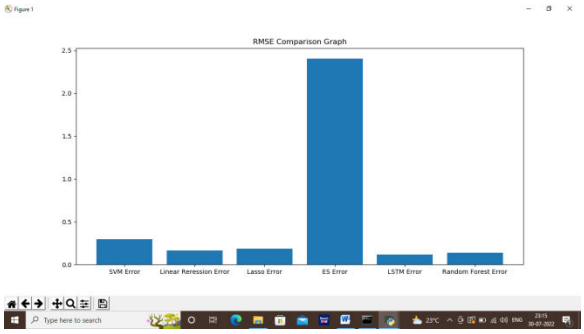


Fig.12: Accuracy comparison graph

6. CONCLUSION

The Coronavirus pandemic's weakness can possibly ignite an extreme overall disaster. There are worries among specific researchers government associations all over the planet that the pandemic might influence a critical segment of the total populace. In this review, an machine learning (ML)-based expectation approach considering overall Coronavirus pandemic gamble has been proposed. Every one of the models makes three unique projections, including the quantity of recently contaminated cases, the quantity of fatalities, the quantity of recuperations before very long. The review's discoveries show that applying these procedures to the ongoing Coronavirus pandemic situation is a promising strategy. The results will show the viability of different calculations likewise demonstrate that LSTM is the best calculation to foresee Coronavirus cases precisely. By and large, we reach the determination that model expectations made in light of the current conditions are exact might be valuable in grasping future occasions. Subsequently, the review's forecasts can likewise be a tremendous assistance to the experts in assisting them with choosing how to deal with the Coronavirus circumstance make a brief move. This study will be dynamically worked on in the following illustrations. We then plan to explore the expectation technique using the refreshed dataset utilize the most exact appropriate ML strategies thinking about estimating.

REFERENCES

1.S. Makridakis, E. Spiliotis, V. Assimakopoulos, “Statistical machine learning forecasting methods: Concerns ways forward,” PloS one, vol. 13, no. 3, 2018.

2. G. Bontempi, S. B. Taieb, Y.-A. Le Borgne, “Machine learning strategies considering time series forecasting,” in European business intelligence summer school. Springer, 2012, pp. 62–77.

3. F. E. Harrell Jr, K. L. Lee, D. B. Matchar, T. A. Reichert, “Regression models considering prognostic prediction: advantages, problems, suggested solutions.”

Cancer treatment reports, vol. 69, no. 10, pp. 1071–1077, 1985.

4.P. Lapuerta, S. P. Azen, L. LaBree, "Use of neural networks in predicting the risk of coronary artery disease," Computers Biomedical Research, vol. 28, no. 1, pp. 38–52, 1995.

5.K. M. Anderson, P. M. Odell, P. W. Wilson, W. B. Kannel, "Cardiovascular disease risk profiles," American heart journal, vol. 121, no. 1, pp. 293–298, 1991.

6. H. Asri, H. Mousannif, H. Al Moatassime, T. Noel, "Using machine learning algorithms considering breast cancer risk prediction diagnosis," Procedia Computer Science, vol. 83, pp. 1064–1069, 2016.

7. F. Petropoulos S. Makridakis, "Forecasting the novel coronavirus covid-19," Plos one, vol. 15, no. 3, p. e0231236, 2020.

8.G. Grasselli, A. Pesenti, M. Cecconi, "Critical care utilization considering the covid-19 outbreak in lombardy, italy: early experience forecast during an emergency response," Jama, 2020.

9. WHO. Naming the coronavirus disease (covid-19) the virus that causes it. [Online]. Available: https://www.who.int/emergencies/diseases/novelcoronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it

10.C. P. E. R. E. Novel et al., "The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (covid-19) in china," ZhonghualiuxingbingxuezazhiZhonghualiuxingbingxuezazhi, vol. 41, no. 2, p. 145, 2020.