How to Cite:
Prasanth, K., Saran, G., Sathish, K. K., & Pradeep, P. (2022). Computational tracking and intelligent forecasting of COVID-19 dynamic propagation. *International Journal of Health Sciences*, *6*(S1), 13478–13491. https://doi.org/10.53730/ijhs.v6nS1.8366

Computational tracking and intelligent forecasting of COVID-19 dynamic propagation

Prasanth K
Department of Information Technology, K.S.Rangasamy College of Technology, Tiruchengode, Namakkal, TamilNadu, India

Saran G
Department of Information Technology, K.S.Rangasamy College of Technology, Tiruchengode, Namakkal, TamilNadu, India

Sathish Kumar K
Department of Information Technology, K.S.Rangasamy College of Technology, Tiruchengode, Namakkal, TamilNadu, India

Pradeep P
Department of Information Technology, K.S.Rangasamy College of Technology, Tiruchengode, Namakkal, TamilNadu, India

***Abstract*---**Coronavirus is a quickly spreading viral sickness that contaminates people and the day to day existence of individuals, their wellbeing, and the economy of a nation are impacted because of this lethal viral illness. Many endeavors have been directed to track down a reasonable and quick method for recognizing contaminated patients in a beginning phase. The spread of COVID-19 in the entire world has seriously jeopardized the humankind. The assets of the absolute biggest economies are worried because of the huge infectivity and contagiousness of this sickness. The ability of Machine Learning models and Deep Learning models can be actually used to estimate the quantity of impending cases impacted by COVID-19 which is by and by viewed as a likely danger to humankind. Specifically, seven standard determining models, in particular LR, LASSO, SVM, NN, XGB Regressor, Random Forest Regressor have been utilized in this review to estimate the compromising elements of COVID-19. Three kinds of assumptions are made by all of the models, similar to the amount of as of late spoiled cases, the amount of passings, and the amount of recoveries. To conquer the issue, proposed strategy utilizing the relapse methods is taken on to foresee the quantity of COVID-19 cases in next 10 days to come. In light of the outcomes, impact of preventive estimates like social seclusion and lockdown can be followed to overrule on the spread of COVID-19. Notwithstanding this

Manuscript submitted: 18 March 2022, Manuscript revised: 9 April 2022, Accepted for publication: 27 May 2022

in view of the ailments, a contaminated individual is investigated and mortality forecast is made. Three models specifically Decision Tree, Random Classifier and Gradient Boosting are utilized for this arrangement. Both the relapse and characterization utilized in the proposed work is viewed as powerful in investigating.

***Keywords*---** COVID-19, Machine Learning, Deep Learning, lockdown.

Introduction

Coronavirus is brought about by Corona infection, which is an exceptional sort of infection and it upgrades the current illness in human's body which makes it an extremely perilous infection. The contraction of COVID means: 'CO' represents crown, 'VI' for the infection, and 'D' for the illness. This infection brings about wheezing, hard to inhale, terrible stomach related framework, and liverwort, impacts severely human sensory system, and furthermore hurts creatures like cows, ponies, and pigs that are kept, raised, and utilized by individuals and different wild creatures. In 2002-2003 the pestilence of Severe Acute Respiratory Syndrome (SARS) and in 2012 the eruption of the Middle East Respiratory Syndrome (MERS) have shown the likelihood of transferrable recently shown up COVID-19 in human to human and creature to human as well as the other way around, however there are exceptionally less instances of this sort, they do exists. In late December 2019, the impact of mystery pneumonia in the entire world is a recognizable subject of study. The course of COVID-19 can go off in unexpected directions where the condition of an evidently consistent open minded self-destructs rapidly to an essential express; this could astonish even the most skilled specialists. To upgrade clinical forecast, Artificial Intelligence (AI) models could be important aides since they can distinguish complex examples in huge datasets; an ability the human cerebrum is bumbling at. Man-made intelligence instruments have been selected to battle COVID-19 on different scales, from epidemiological demonstrating, to individualized conclusion and anticipation expectation. AI (ML) is a sort of AI that permits programming applications to turn out to be more exact at anticipating results without being unequivocally modified to do as such. AI calculations utilize verifiable information as contribution to anticipate new result values. Profound Learning (DL) is essential for a more extensive group of ML techniques in light of Artificial Neural Networks (ANN) with portrayal learning. DL models are prepared by utilizing enormous arrangements of named information and brain network structures that gain includes straightforwardly from the information without the requirement for manual component extraction. Learning can be managed, semi-directed or solo. DL models like Deep Neural Networks (DNN), Deep Belief Networks (DBN), Deep Reinforcement Learning (DRL), Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) have been applied to fields including PC vision, discourse acknowledgment, regular language handling, machine interpretation, bioinformatics, drug plan, clinical picture examination, environment science, material investigation and prepackaged game projects, where they have delivered outcomes similar to and now and again unbelievable human master execution. ANNs were motivated by data handling and appropriated correspondence hubs in natural frameworks. ANNs have different

contrasts from organic cerebrums. In particular, fake brain networks will more often than not be static and representative, while the natural cerebrum of most living life forms is dynamic (plastic) and simple. The principle objective of this proposed work is to apply administered ML and DL answers for produce a learning model equipped for identifying and foreseeing the necessary fields that aides in viable expectation and alleviation of COVID-19 and anticipate the mortality of COVID-19 patients. Thus by joining parts of man-made consciousness with COVID-19, savvy arrangements and forecasts are more fit for adjusting to anticipate the quantity of death, impacted and recuperated patients.

Related Works

To gauge the development of tainted populaces, disease transmission experts and numerical modelers have fostered various measurable and numerical models to reproduce the spread of illness as far as helpless, contaminated, recuperated, and perished populaces. These models incorporate the helpless irresistible recuperation (SIR) model and its inferred models, like the defenseless uncovered irresistible recuperation (SEIR) model [1] [2].

Summed up direct relapse models utilized for the investigation of illness impacting factors, like straight relapse, Poisson relapse, and Logistic relapse, have restricted capacity to manage complex nonlinear connections among trademark factors [3-5]. Simultaneously, because of the restricted avoidance and control assets in the flare-up period, contrasted and the informative investigation of impacting factors, quick distinguishing proof of high-risk bunches is of more commonsense importance for working on the nature of logical navigation. Man-made knowledge has been generally utilized in tainting attestation, game-plan, and evaluation of influencing components [6-8], and has been applied in the field of COVID-19 disease gauge, request, and medicine sufficiency, showing mind boggling benefits.

Considering the forecast and examination of the spread and improvement pattern of the pestilence, the expectation models worked by homegrown and unfamiliar researchers mostly center around the unique model and measurable model. As indicated by the relationship among different components, the powerful differential conditions are built to recreate the improvement pattern of applicable components. Subsequently, COVID-19 has been broadly utilized in infection transmission and investigation. The dynamic models of COVID-19 transmission principally incorporate SIR, SEIR, SEIHR, and SEQIR models [9-11]. Furthermore, a few mixture calculations in light of AI and profound learning are likewise utilized for expectation research with time series qualities [12-14]. Scarcely any new papers could be found on the examination of contamination spreading design in various nations [15]. There have been a ton of past examinations that utilized AI and measurable ways to deal with catch the examples and patterns of various shifting occasions connected with irresistible illnesses [16].

Works can likewise be found on utilization of SIR models [17] for contact following and forecast of positive number of cases. Notwithstanding, a typical suspicion of such works is to treat the thickness of the Indian populace to be homogeneous and not considering the quantity of tests led.

Time series investigation is famously used to figure various infections like SARS, Ebola, pandemic flu and dengue. Tandon et al. [18] have played out a concentrate by utilizing ARIMA (2,2,2) model to foresee the future number of COVID-19 cases in India and have made determination that cases in India will increment dramatically and anticipated that the cases decline from their expected estimate. They had mirrored the significance of social separating and sterilization in work to diminish human openness to the infection. The intention of their review was to help the public authority and clinical labor force to be ready for the present circumstance. The downside of their review was that they had not thought about the impact of number of COVID-19 test that brought about sure cases and couldn't reflect to the place of asymptomatic cases and its impact on the spread of the infection.

One more comparative review was finished by Tyagi et al. [19] where they finished up to have anticipated the quantity of COVID-19 cases till the finish of month June and anticipated the necessity of ICU beds, ventilators and disconnection beds in India utilizing ARIMA(1,3,1) model. The principle objective of their work was to anticipate the clinical necessity (ICU 10% of the dynamic cases, BEDS, ventilators for 5% of the dynamic cases and confinement bed). This work the point is for concentrating on the impact of lockdown and open information utilizing time series expectation models considering the quantity of tests led relating to the quantity of positive cases.

In every one of the current frameworks, different elements are thought of and are restricted distinctly to the forecast of the presence of the infection. The proposed work conquers the limits of the current frameworks by foreseeing the different sorts of component cases on a specific date and furthermore anticipate the demise of the individual.

Proposed Work

The proposed work flow for the Computational Tracking and Forecasting the COVID-19 Dynamic Propagation is as follows:

Data collection

Data arrangement is portrayed as the technique for social event, assessing and taking apart exact pieces of information for research using standard endorsed methods. An analyst can study their speculation considering gathered information. All around, data grouping is the fundamental and most critical stage for research, no matter what the field of investigation. The strategy of data collection is different for different fields of study, dependent upon the essential information. The most basic goal of information assortment is guaranteeing that data rich and dependable information is gathered for measurable investigation so information driven choices can be made for research. The dataset for this work is a seat mark dataset accessible in Kaggle. The dataset has been downloaded from the Kaggle and further utilized for the work.

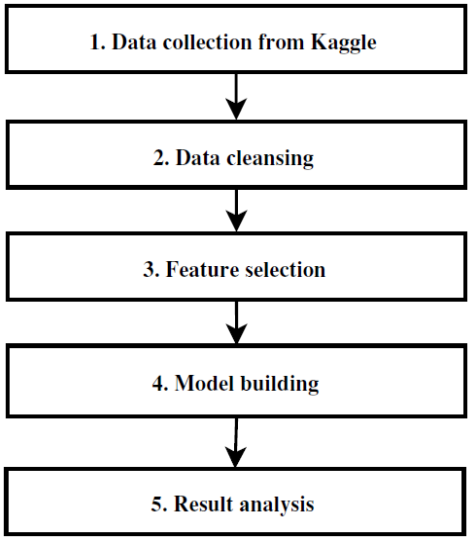


Fig. 1 Workflow for the Proposed Work

Data Cleansing

Information cleaning is the most common way of fixing or eliminating mistaken, undermined, erroneously arranged, copy, or fragmented information inside a dataset. While joining various information sources, there are numerous open doors for information to be copied or mislabeled. Assuming information is erroneous, results and calculations are questionable, despite the fact that they might look right. There is nobody outright method for endorsing the specific strides in the information cleaning process on the grounds that the cycles will differ from dataset to dataset. Yet, it is significant to lay out a layout for your information cleaning process so you realize you are doing it the correct way without fail. While the methods utilized for information cleaning might differ as indicated by the sorts of information your organization stores, you can follow these essential strides to delineate a system for your association.

a) Remove duplicate or irrelevant observations

Eliminate undesirable perceptions from your dataset, including copy perceptions or unimportant perceptions. Copy perceptions will happen most frequently during information assortment. Whenever you consolidate informational indexes from different spots, scratch information, or get information, there are chances to make copy information. De-duplication is probably the biggest region to be considered in this cycle. Superfluous perceptions are the point at which you notice perceptions that don't squeeze into the particular issue you are attempting to examine. This can make investigation more effective and limit interruption from your essential objective — as well as making a more sensible and more performant dataset.

b) Fix structural errors
Underlying mistakes are the point at which you measure or move information and notice weird naming shows, errors, or wrong capitalization. These irregularities can cause mislabeled classifications or classes. For instance, you might find "N/A" and "Not Applicable" both show up, however they ought to be broke down as a similar classification.

c) Filter unwanted outliers
Frequently, there will be one-off perceptions where, initially, they don't seem to fit inside the information you are dissecting. Assuming you have a genuine motivation to eliminate an anomaly, as inappropriate information passage, doing so will help the exhibition of the information you are working with. In any case, in some cases the presence of an exception will demonstrate a hypothesis you are dealing with. Since an exception exists, doesn't mean it is erroneous. This progression is expected to decide the legitimacy of that number. Assuming that an exception ends up being unimportant for examination or is a mix-up, consider eliminating it.

d) Handle missing data
You can't overlook missing information on the grounds that numerous calculations won't acknowledge missing qualities. There are several methods for managing missing information. Nor is ideal, yet both can be thought of.

1. As a first choice, you can drop perceptions that have missing qualities, however doing this will drop or lose data, so be aware of this before you eliminate it.
2. As every subsequent choice, you can enter missing qualities in light of different perceptions; once more, there is a potential chance to lose uprightness of the information since you might be working from suspicions and not genuine perceptions.
3. As a third choice, you could adjust how the information is utilized to explore invalid qualities.

Feature Selection

Highlight extraction alludes to the method involved with changing crude information into mathematical elements that can be handled while protecting the data in the first informational collection. It yields improved outcomes than applying AI straightforwardly to the crude information. During the course of model structure, include determination is utilized to choose most significant highlights out of the relative multitude of elements. It lessens the intricacy of the forecast model.

Model Building

An AI model is worked by gaining and summing up from preparing information, then applying that obtained information to new information it has never seen before to make expectations and satisfy its motivation. Absence of information will keep you from building the model, and admittance to information isn't sufficient. You should prepare the datasets to run as expected and see a steady improvement in the forecast rate. Calculations gain from information. They track

down connections, foster grasping, simply decide, and assess their certainty from the it they're given to prepare information. Furthermore, the more the preparation information is, the more the model performs.

DL and other ML models require an underlying arrangement of information, called preparing information, to go about as a gauge for additional application and usage. This information is the establishment for the program's developing library of data. When a model is prepared on a preparation set, it's generally assessed on a test set. Intermittently, these sets are taken from a similar generally dataset, however the preparation set ought to be named or improved to expand a calculation's certainty and precision.

Result Analysis

We perform result examination with two straightforward yet various models one is generative and the another is discriminative. Then, at that point, the methodology is tried to give improved outcomes.

1. Beginning models are first pictured and exploratory examination is made. This incorporates - plot diagrams, take a gander at connections, shared data, ROC bends and such.
2. Utilizing the bits of knowledge from (1), we see that the elements we have are adequate. In the event that they don't appear to be sufficient, we put a work in further developing them (make new, eliminate the superfluous ones, use include choice/extraction strategies)
3. When the highlights are great, we begin improving the model - change parts, boundaries, and use troupe techniques.
4. At last execution measurements are gotten and they are considered in contrast to different ML and DL models are assessed.

Algorithms Used

1. Random Forest

A Random Forest is a group method fit for performing both relapse and arrangement errands with the utilization of various choice trees and a procedure called Bootstrap and Aggregation, generally known as stowing. The essential thought behind this is to join various choice trees in deciding the last result as opposed to depending on individual choice trees. Irregular Forest has numerous choice trees as base learning models. We arbitrarily perform column testing and component inspecting from the dataset shaping example datasets for each model. This part is called Bootstrap. We want to move toward the Random Forest relapse procedure like some other AI method.

Plan a particular inquiry or information and get the source to decide the necessary information. Ensure the information is in an open configuration else convert it to the expected organization. Indicate every single recognizable oddity and missing information focuses that might be expected to accomplish the necessary information. Set the gauge model that you need to accomplish. Test for forecast with the regressor.

Random Forest Regressor

```
[ ] from sklearn.ensemble import RandomForestRegressor
    regressor = RandomForestRegressor(max_depth=2, random_state=0, n_estimators=100)
    regressor.fit(X_train, y_train)
    y_pred = regressor.predict(X_test)
    from sklearn.metrics import mean_squared_error as mse
    from sklearn.metrics import mean_absolute_error as mae
    from sklearn.metrics import r2_score
    print('MSE =', mse(y_pred, y_test))
    print('MAE =', mae(y_pred, y_test))
    print('R2 Score =', r2_score(y_pred, y_test)*100)
    print('Variance =',explained_variance_score(y_test,y_pred)*100)

    MSE = 703144107089.5659
    MAE = 637523.7515920112
    R2 Score = 96.5556726467539
    Variance = 97.16805702594824
```

2. Support Vector Machine

Support Vector Regression (SVR) is a managed learning calculation that is utilized to anticipate discrete qualities. Support Vector Regression involves a similar guideline as the Support Vector Machines (SVMs). The essential thought behind SVR is to track down the best fit line. In SVR, the best fit line is the hyperplane that has the most extreme number of focuses.

Not at all like other relapse models that attempt to limit the blunder between the genuine and anticipated esteem, the SVR attempts to fit the best line inside an edge esteem. The limit esteem is the distance between the hyperplane and limit line.The fit time flightiness of SVR is more than quadratic with how much tests which makes it trying to scale to datasets with in excess of two or three 10000 models. Straight SVR gives a quicker execution than SVR yet just thinks about the direct piece. The model created by SVR relies just upon a subset of the preparation information, on the grounds that the expense work disregards tests whose forecast is near their objective.

Support Vector Machine

```
[ ] from sklearn.svm import SVR
    regressorpoly=SVR(kernel='poly',epsilon=1.0)
    regressorpoly.fit(X_train,y_train)
    pred=regressorpoly.predict(X_test)
    print('MSE =', mse(pred, y_test))
    print('MAE =', mae(pred, y_test))
    print('R2 Score =',r2_score(y_test,pred)*100)
    print('Variance =',explained_variance_score(y_test,pred)*100)

    MSE = 27693226703047.31
    MAE = 4016099.294737766
    R2 Score = -11.95529693626014
    Variance = 0.03987112523440306
```

3. XGBoost Regressor

Outrageous Gradient Boosting (XGBoost) is an open-source library that gives a proficient and viable execution of the inclination helping calculation. Soon after its turn of events and starting delivery, XGBoost turned into the go-to strategy and frequently the vital part in winning answers for a scope of issues in AI contests. Relapse prescient displaying issues include foreseeing a mathematical worth, for example, a dollar sum or a level. XGBoost can be utilized straightforwardly for relapse prescient displaying.

The legitimacy of this assertion can be induced by being familiar with its (XGBoost) objective capacity and base students. The goal work contains misfortune work and a regularization term. It tells about the differentiation between certifiable characteristics and expected values, i.e how far the model results are from the veritable characteristics. The most widely recognized misfortune capacities in XGBoost for relapse issues is reg:linear, and that for double grouping is reg:logistics.

Outfit learning includes preparing and consolidating individual models (known as base students) to get a solitary forecast, and XGBoost is one of the gathering learning strategies. XGBoost hopes to have the base students which are consistently awful at the rest of that when every one of the expectations are consolidated, terrible forecasts counteracts and better one summarizes to frame last great expectations.

XGB REGRESSOR

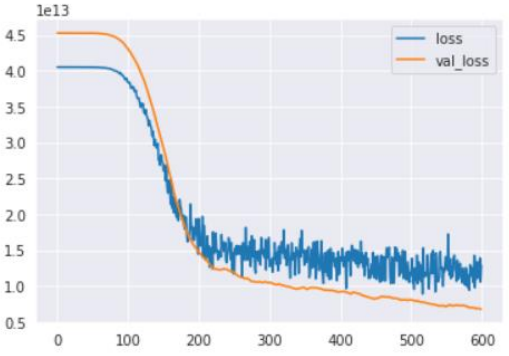
```
[ ] from xgboost import XGBRegressor
    from sklearn.metrics import mean_absolute_error
    XGBModel = XGBRegressor()
    XGBModel.fit(X_train,y_train , verbose=False)

    # Get the mean absolute error on the validation data :
    XGBpredictions = XGBModel.predict(X_test)
    MAE = mean_absolute_error(y_test , XGBpredictions)
    print('MSE =', mse(XGBpredictions, y_test))
    print('MAE =', mae(XGBpredictions, y_test))
    print('R2 Score =',r2_score(y_test,XGBpredictions)*100)
    print('Variance =',explained_variance_score(y_test,XGBpredictions)*100)

    [06:37:45] WARNING: /workspace/src/objective/regression_obj.cu:152: reg:
    MSE = 13956688860.358358
    MAE = 84584.86835252192
    R2 Score = 99.94357734971214
    Variance = 99.94463825830104
```

4. Multilayer Perceptron

Multi-facet perceptron (MLP) is an enhancement of feed forward brain organization. It comprises of three sorts of layers — the info layer, yield layer and secret layer. The info layer gets the information sign to be handled. The expected errand, for example, expectation and grouping is performed by the result layer. An erratic number of stowed away layers that are set in the middle of the info and result layer are the genuine computational motor of the MLP. Like a feed forward network in a MLP the information streams in the forward bearing from contribution to yield layer. The neurons in the MLP are prepared with the back spread learning calculation. MLPs are intended to inexact any persistent capacity and can take care of issues which are not directly divisible.



5. Logistic Regression

Logistic Regression is a course of demonstrating the likelihood of a discrete result given an information variable. The most well-known calculated relapse models a paired result; something that can take two qualities like valid/bogus, yes/no, etc. Multinomial calculated relapse can display situations where there are multiple conceivable discrete results. Calculated relapse is a helpful investigation technique for grouping issues, where you are attempting to decide whether another example squeezes best into a class.

Logistic Regression

```
[ ]  # import the class
     from sklearn.linear_model import LogisticRegression

     # instantiate the model (using the default parameters)
     logreg = LogisticRegression()

     # fit the model with data
     logreg.fit(X_train,y_train)

     #
     y_pred=logreg.predict(X_test)
     print('MSE =', mse(y_pred, y_test))
     print('MAE =', mae(y_pred, y_test))
     print('R2 Score =', r2_score(y_pred, y_test)*100)
     print('Variance =',explained_variance_score(y_test,prediction)*100)
```

```
MSE = 900899895134.2982
MAE = 613275.8421052631
R2 Score = 96.6466429003085
Variance = 97.88775325176039
```

6. Linear Regression

Linear Regression (LR) is a straight model, for example a model that expects a direct connection between the information factors (x) and the single result variable (y). Even more expressly, that y still up in the air from an immediate blend of the data factors (x). At the point when there is a solitary info variable (x), the strategy is alluded to as basic straight relapse. Whenever there are different information factors, writing from measurements frequently alludes to the strategy as numerous direct relapse.

Various procedures can be utilized to plan or prepare the straight relapse condition from information, the most widely recognized of which is called Ordinary Least Squares. It is normal to subsequently allude to a model arranged

this way as Ordinary Least Squares Linear Regression or simply Least Squares Regression. At times the information should be changed to meet the prerequisites of the investigation, or stipend must be made for exorbitant vulnerability in the X variable. On the off chance that the prerequisites for direct relapse investigation are not met, alterative powerful nonparametric techniques can be utilized. In certain informational indexes, the straight line goes through the beginning at 0,0, and afterward improved on conditions can be utilized. Straight relapse is normally used to anticipate the worth of the Y variate at any worth of the X variate, however now and again the backwards forecast is required, in view of an alternate methodology.

Linear Regression

```
[ ]  from sklearn.linear_model import LinearRegression
     linear_regression = LinearRegression()
     linear_regression.fit(X_train,y_train)
     y_predd=logreg.predict(X_test)
     print('MSE =', mse(y_predd, y_test))
     print('MAE =', mae(y_predd, y_test))
     print('R2 Score =', r2_score(y_predd, y_test)*100)
     print('Variance =',explained_variance_score(y_test,y_predd)*100)
```

```
MSE = 900899895134.2982
MAE = 613275.8421052631
R2 Score = 96.6466429003085
Variance = 96.35806213687509
```

7. Lasso Regression

The Word "LASSO" represents Least Absolute Shrinkage and Selection Operator. It is a factual recipe for the regularization of information models and component choice. Rope relapse is a regularization procedure. It is utilized over relapse strategies for a more exact forecast. This model purposes shrinkage. Shrinkage is where information values are contracted towards a main issue as the mean. The rope method empowers basic, inadequate models (for example models with less boundaries). This specific sort of relapse is appropriate for models showing elevated degrees of multi-collinearity or when you need to mechanize specific pieces of model choice, similar to variable determination/boundary end.
Tether Regression utilizes L1 regularization method. It is utilized when we have more number of highlights since it naturally performs include determination. L1 regularization adds a punishment that is equivalent to the outright worth of the size of the coefficient. This regularization type can bring about inadequate models with not many coefficients. A few coefficients could become zero and get wiped out from the model. Bigger punishments bring about coefficient esteems that are more like zero (ideal for delivering easier models).

Lasso Regression

```
from sklearn.linear_model import Lasso
lasso = Lasso(alpha=0.3,max_iter=100000)
lasso.fit(X_train,y_train)
y_predd=lasso.predict(X_test)
print('MSE =', mse(y_predd, y_test))
print('MAE =', mae(y_predd, y_test))
print('R2 Score =', r2_score(y_predd, y_test)*100)
print('Variance =',explained_variance_score(y_test,y_predd)*100)
```

```
MSE = 140430602494.56732
MAE = 265367.4480254334
R2 Score = 99.3484445121385
Variance = 99.55631071105611
```

8. Decision Tree

The decision tree classifier makes the arrangement model by building a choice tree. Every hub in the tree indicates a test on a trait, each branch diving from that hub compares to one of the potential qualities for that property. Each leaf addresses class marks related with the example. Examples in the preparation set are arranged by exploring them from the base of the tree down to a leaf, as per the result of the tests along the way. Beginning from the root hub of the tree, every hub parts the occasion space into at least two sub-spaces as per a property test condition. Then dropping down the tree limb comparing to the worth of the trait, another hub is made. This interaction is then rehashed for the sub-tree established at the new hub, until all records in the preparation set have been ordered. The choice tree development process as a rule works in a hierarchical way, by picking a trait test condition at each progression that best divides the records. There are many estimates that can be utilized to decide the most effective way to divide the records.

Classification with Gradient Boosting

```
[ ] gradientBoostingClassifier = GradientBoostingClassifier(learning_rate=1)
    gradientBoostingClassifier.fit(X_train_oversampled, y_train_oversampled)

    y_pred = gradientBoostingClassifier.predict(X_test)

    plot_confusion_matrix(gradientBoostingClassifier, X_test, y_test)
    print(classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

         0.0       0.86      0.84      0.85      4032
         1.0       0.48      0.53      0.51      1146

    accuracy                           0.77      5178
   macro avg       0.67      0.68      0.68      5178
weighted avg       0.78      0.77      0.77      5178
```

9. Gradient Boosting

Gradient Boosting is an iterative useful inclination calculation, i.e a calculation which limits a misfortune work by iteratively picking a capacity that focuses towards the negative slope; a frail theory. In Gradient Boosting, every indicator attempts to develop its ancestor by decreasing the mistakes. Yet, the intriguing thought behind Gradient Boosting is that as opposed to fitting an indicator on the information at every emphasis, it really fits another indicator to the lingering mistakes made by the past indicator.

Classification with Gradient Boosting

```
[ ] gradientBoostingClassifier = GradientBoostingClassifier(learning_rate=1)
    gradientBoostingClassifier.fit(X_train_oversampled, y_train_oversampled)

    y_pred = gradientBoostingClassifier.predict(X_test)

    plot_confusion_matrix(gradientBoostingClassifier, X_test, y_test)
    print(classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

         0.0       0.86      0.84      0.85      4032
         1.0       0.48      0.53      0.51      1146

    accuracy                           0.77      5178
   macro avg       0.67      0.68      0.68      5178
weighted avg       0.78      0.77      0.77      5178
```

Conclusion

Coronavirus has been a significant general wellbeing challenge around the world. Until this point in time, in excess of 235 nations have been impacted by the pandemic. Brazil is the second country with the biggest number of affirmed instances of the sickness, behind just the United States. Considering that, it is essential to figure the spatial dispersion of the aggregate instances of Covid-19, so general wellbeing administrators can assess better techniques to slow the advancement of the infection and forestall the presence of a subsequent wave. A ML based information driven anticipating strategy has been utilized to gauge the conceivable number of positive instances of COVID-19 in India for the following 10 days. The quantity of recuperated cases, everyday positive cases, and expired cases has likewise been assessed by utilizing and bend fitting. The expectation aftereffects of various numerical models are different for various boundaries and in various districts. As a general rule, the fitting impact of Logistic model might be awesome among the three models.

References

[1] C Connell McCluskey. Complete global stability for an SIR epidemic model with delay Distributed or discrete. Nonlinear Analysis: Real World Applications, 11(1):55–59, 2010.

[2] Chao Fan, Xiangqi Jiang, and Ali Mostafavi. “A network percolation-based contagion model of flood propagation and recession in urban road networks”, 2020

[3] J. Ren et al., "A Novel Intelligent Computational Approach to Model Epidemiological Trends and Assess the Impact of Non-Pharmacological Interventions for COVID-19," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 12, pp. 3551-3563, Dec. 2020.

[4] Guo Changman, Guo Min, Liu Yuanyuan, et al. Application of Machine learning Algorithms in Predicting HIV Infection a- mong Men. Chinese Health Statistics, 2019, 36(1): 28-31, 35.

[5] A. U. Mandayam, R. A.C, S. Siddesha and S. K. Niranjan, "Prediction of Covid-19 pandemic based on Regression," 2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Bangalore, India, 2020, pp. 1-5.

[6] O.Tutsoy, Ş.Çolak, A.Polat and K. Balikci, "A Novel Parametric Model for the Prediction and Analysis of the COVID-19 Casualties," in IEEE Access, vol. 8, pp. 193898-193906, 2020.

[7] N. Zheng et al.,"Predicting COVID-19 in China Using Hybrid AI Model," in IEEE Transactions on Cybernetics, vol.50, no.7, pp.2891- 2904, July 2020

[8] R. G. Babukarthik, V. A. K. Adiga, G. Sambasivam, D. Chandramohan and J. Amudhavel, "Prediction of COVID-19 Using Genetic Deep Learning Convolutional Neural Network (GDCNN)," in IEEE Access, vol. 8, pp. 177647-177666, 2020.

[9] Zareie B, Mohammad A R, Mansournia A,et al. A model for COVID-19 prediction in Iran based on China parameters. Archivesof Iranian medicine,2020,23(4):244-248.

[10] Y. Zhao, Y. He."COVID-19 Outbreak Prediction Based on SEIQR Model," 2020 39th Chinese Control Conference, 2020, pp.1133-1137.

[11] Y. -C. Chen, P. -E. Lu, C. -S. Chang and T. -H. Liu, "A Time-Dependent SIR Model for COVID-19 With Undetectable Infected Persons," in IEEE Transactions on Network Science and Engineering, vol. 7, no. 4, pp. 3279-3294, 1 Oct.-Dec. 2020.

[12] Z. Liu, J. Zuo, R. lv, Y. Sun and H. Kang, "Research on Time Series Problem Model Based on Dynamic Network NAR and Multiple Regression," 2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE), Beijing, China, 2020, pp. 416-419, doi: 10.1109/ICAICE51518.2020.00088.

[13] Y. -H. Wu et al., "JCS: An Explainable COVID-19 Diagnosis System by Joint Classification and Segmentation," in IEEE Transactions on Image Processing, vol. 30, pp. 3113-3126, 2021.

[14] F. Rustam."COVID-19 Future Forecasting Using Supervised Machine Learning Models," in IEEE Access, vol. 8, pp. 101489-101499, 2020.

[15] Hillmer SC, Tiao GC. An arima-model-based approach to seasonal adjustment. J. Am. Stat. Assoc. 1982;77(377):63–70. doi: 10.1080/01621459.1982.10477767.

[16] K. Kalpakis, D. Gada, V. Puttagunta, Distance measures for effective clustering of arima time-series. In: Proceedings 2001 IEEE international conference on data mining, pp. 273–280. IEEE (2001)

[17] LaViola JJ. Double exponential smoothing: an alternative to Kalman filter-based predictive tracking. Proc. Worksh. Virt. Environ. 2003;2003:199–206.

[18] Le TT, Andreadakis Z, Kumar A, Roman RG, Tollefsen S, Saville M, Mayhew S. The covid-19 vaccine development landscape. Nat. Rev. Drug Discov. 2020;19(5):305–306. doi: 10.1038/d41573-020-00073-5.

[19] Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ, Collaboration HAS, et al. Covid-19: consider cytokine storm syndromes and immunosuppression. Lancet (London, England) 2020;395(10229):1033. doi: 10.1016/S0140-6736(20)30628-0.