International Conference on Advances in Materials, Computing and Communication Technologies
In Association with International Journal of Scientific Research in Science and Technology
Volume 9 | Issue 1 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)

Covid-19 Risk Prediction Using Machine Learning Techniques

Ms. M. Angel[1] , Dr. S. Ebenezer Juliet[2]
[1]PG Student, Department of CSE , VV College Of Engineering, Tirunelveli, Tamil Nadu, India
[2]Associate Professor, Department of CSE, VV College Of Engineering, Tirunelveli, Tamil Nadu, India

ABSTRACT

Nowadays Machine Learning (ML) Techniques are used for accurate COVID-19 risk prediction. Several prediction methods are being popularly used to handle forecasting problems. In this paper the capability of ML models to forecast the number of upcoming patients affected by COVID-19 is presented because COVID-19 is considered as a potential threat to mankind. Various risk prediction factors of COVID-19 which performs well in forecasting the new confirmed cases, death rate as well as recovery rate are identified. Two standard forecasting models, such as Support vector Machine,(SVM) and Linear Regression(LR) have been used for prediction to forecast the threatening factors of COVID-19. The results produced by the above techniques proved it a promising prediction mechanism for the current scenario of the COVID-19 pandemic. Also the experimental results proved that the SVM performs better than Linear Regression.

Keywords: COVID-19, Support vector machine, Linear Regression, Machine learning Techniques

I. INTRODUCTION

Machine learning (ML) has proved itself as a prominent field of study over the last decade by solving many very complex and sophisticated real-world problems. The application areas included almost all the real-world domains such as health- care, autonomous vehicle, business applications, natural language processing, intelligent robots, gaming, climate modelling, voice, and image processing. ML algorithms' learning is typically based on trial and error method quite opposite of conventional algorithms.

One of the most significant areas of ML is forecasting, numerous standard ML algorithms have been used in this area to guide the future course of actions needed in many application areas including weather forecasting, disease forecasting, stock market forecasting as well as disease prognosis. Various regression and neural network models have wide applicability in predicting the conditions of patients in the future with a specific disease. There are lots of studies performed for the prediction of different diseases using machine learning techniques such as coronary artery disease, cardiovascular disease prediction, and breast cancer prediction. In particular, the study is focused on live forecasting of COVID-19 confirmed cases and study is also focused on the

forecast of COVID-19 outbreak and early response. These prediction systems can be very helpful in decision making to handle the present scenario to guide early interventions to manage these diseases very effectively.

This study aims to provide an early forecast model for the spread of novel corona virus, also known as SARS-CoV-2, officially named as COVID-19 by the World Health Organization (WHO). COVID-19 is presently a very serious threat to human life all over the world. At the end of 2019, the virus was first identified in a city of China called Wuhan, when a large number of people developed symptoms like pneumonia. It has a diverse effect on the human body, including severe acute respiratory syndrome and multi-organ failure which can ultimately lead to death in a very short duration. Hundreds of thousands of people are affected by this pandemic throughout the world with thousands of deaths every coming day. Thousands of new people are reported to be positive every day from countries across the world. The virus spreads primarily through close person to person physical contacts, by respiratory droplets, or by touching the contaminated surfaces.

The most challenging aspect of its spread is that a person can possess the virus for many days without showing symptoms. The causes of its spread and considering its danger, almost all the countries have declared either partial or strict lockdowns throughout the affected regions and cities. Medical researchers throughout the globe are currently involved to discover an appropriate vaccine and medications for the disease. Since there is no approved medication till now for killing the virus so the governments of all countries are focusing on the precautions which can stop the spread. Out of all precautions, "be informed" about all the aspects of COVID-19 is considered extremely important. To contribute to this aspect to find formation, numerous researchers are studying the different dimension so the pandemic and produce the results to help humanity.

To contribute to the current human crisis our attempt in this study is to develop a forecasting system for COVID-19. This problem of forecasting has been considered as a regression problem in this study, so the study is based on some state-of-art supervised ML regression models such as linear regression (LR), support vector machine (SVM). The data set as been pre-processed and divided into two subsets: training set (70%records) and testing set (30% records). The performance evaluation has been done in terms of important measures including precision, Recall, F1 score, Support and Accuracy values. This study has some key findings which are listed below. Different ML algorithms seem to perform better in different class predictions. Most of the ML algorithms require an ample amount of data to predict the future, as the size of the data set increases the model performances improve.ML model based forecasting can be very useful for decision- makers to contain pandemics like COVID- 19.The csv file of Novel Corona virus 2019 data set available at https://www.kaggle.com/sudalairajkumar/nove l-corona- virus-2019-dataset is downloaded. Aseparat - This paper consists of various sections. Section I presents the introduction, Section II contains the description of the data sets and methods used in this study. Section III presents methodology, Section IV presents the results, and Section V summarizes the results of research work and presents the conclusion.

II. MATERIALS AND METHODS

A. DATA SET

The aim of this study is the future forecasting of COVID-19 spread focusing on the number of new positive cases ,the number of deaths, and the number of recoveries. The dataset used in the study has been obtained from the Machine learning provided by the Center for Systems Science and Engineering, Kaggle

Repository The repository was primarily made available for the open source. A collection of database for machine learning oriented is available with the Kaggle Machine Learning Repository which will be available from Internet and it is open source. The data set are accommodated and preserved in the center for Machine Learning and Intelligent Systems in the University of California, Irvine. Each dataset contains separate webpage that presents the entire facts about its inclusion and any pertinent research that is examining it. COVID 19 data set contain information such as Fever, cough, breath in, breath out, throat pain, joint pain, diarrhea, tholach, old peak, pneumonia. These features are given as input to the LR and SVM based COVID-19 risk prediction model Data preprocessing is a data mining technique that involves transforming raw data into an understandable format. Here the covid-19 data set is transformed into machine readable format.

Table 1. Sample COVID-19 data set from kaggle repository

Age	Sex	Pneumonia	trestbps	chol	thalach	oldpeak	Fever	Cough	breath_difficult	throat_pain	joint_pain	Diarrhea	Target
50	0	3	145	233	150	2.3	1	0	0	0	0	1	1
30	0	2	130	250	187	3.5	0	1	0	0	0	2	1
50	0	1	130	204	172	1.4	0	0	0	2	0	2	1
20	0	1	120	236	178	0.8	0	1	0	2	0	2	1
20	1	0	120	354	163	0.6	0	1	1	2	0	2	1

Table 2.Sample COVID-19 data set with missing data values

age	Sex	Pneumonia	trestbps	chol	Thalach	Oldpeak	fever	Cough	breath_difficult	throat_pain	joint_pain	diarrhea	target
50	0	3	145	233	150	-	1	0	-	0	1	-	1
30	0	2	130	250	187	3.5	0	1	0	0	0	2	1
50	0	1	130	204	172	-	0	0	0	-	1	2	1
20	0	1	120	236	-	-	0	1	0	-	0	2	-
20	1	0	120	354	163	2.5	0	1	-		0	2	1

B. SUPERVISED MACHINE LEARNING MODELS

A supervised learning model is built to make a prediction when it is provided with an unknown input instance. In this learning technique, the learning algorithm takes a data set with input instances along with their corresponding regress or to train the regression model. The trained model then generates a prediction for the given unforeseen input data or test datasets. This learning method may use regression techniques and classification algorithms for predictive models' development. Two regression models have been used in this study of COVID-19 future forecasting

- Linear Regression
- Support Vector Machine

1) LINEAR REGRESSION

In regression modeling, a target class is predicated based on the inndependent features [14]. This method is used to find out the relationship between

independent and dependent variables and also for forecasting. Linear regression a type of regression modelling is the most usable statistical technique for predictive analysis in machine learning. Each observation in linear regression depends on two values, one is the dependent variable and the second is the independent variable. Linear regression determines a linear relationship between these dependent and independent variables. There are two factors (x; y) that are involved in linear regression analysis. The equation below shows how y is related to x known as regression.

$y = \beta 0 + \beta 1x + \varepsilon$ (1)
or equivalently
$E(y) = \beta 0 + \beta 1x$ (2)

Here, ε is the error term of linear regression. The error term here uses to account the variability between both x and y represents y-intercept, represents slope. To put the concept of linear regression in the machine learning context, in order to train the model x is represented as input training data set, y represents the class labels present in the input data set. . The goal of the machine learning algorithm then is to find the best values for β0 (intercept) and β1 (coefficient) to get the best-fit regression line. To get the best fit implies the difference between the actual values and predicted values should be minimum, so this minimization problem can be represented as:

$$minimize \frac{1}{n} \sum_{i=1}^{n} (pred_i - y_i)^2$$

$$g = \frac{1}{n} \sum_{i=1}^{n} (pred_i - y_i)^2 \quad (3)$$

Here, g is called a cost function, which is the root mean square of the predicted value of y (prediction) and actual y (yi), n is the total number of data points.

III. METHODOLOGY

The study is about novel corona virus also known as COVID-19 predictions. The COVID-19 has proved a present potential threat to human life. It causes tens of thousands of deaths and the death rate is increasing day by day throughout the globe. To contribute to this pandemic situation control, this study attempts to perform future forecasting on the death rate, the number of daily conformed infected cases and the number of recovery cases in the upcoming 10 days. The forecasting has been done by using two ML approaches that are appropriate to this context The proposed system uses Linear Regression and Support Vector Machine algorithms for prediction of COVID 19 predictions for different diseases using machine learning techniques such as fever, cough , breathing problem, throat pain, joint pain, diarrhea, tholach, old peak, pneumonia, trestbps is focused on live forecasting of COVID-19 confirmed cases and study is also focused on the forecast of COVID-19 outbreak and early response.

Various regression models have wide applicability in predicting the conditions of patients in the future with a specific disease .To contribute to the current human crisis our attempt in this study is to develop a forecasting system for COVID-19. These prediction systems can be very helpful in decision making to handle the present scenario to guide early interventions to manage these diseases very effectively. In proposed model data is divided in to training and testing values. Machine learning based Linear Regression and Support Vector Machine model is built with training data and then Precision, Recall, F1 sore

After the initial data pre-processing step, the data set has been divided into two subsets: a training set (70 %) to train the models and testing set (30 %). The learning models such as SVM, LR ,have been used in this study. These models have been trained on the 300 data set cases, recovery, and death patterns. The

learning models have then been evaluated based on important metrics such as Precision, Recall, F1 Score, Support values and reported in the results. The proposed approach used in the study has been shown as Figure1. Support values are calculated with testing set. From the symptoms such as Fever, cough, breath in, breath out, throat pain, joint pain, diarrhoea, toolache, old peak, pneumonia this system attempt to forecast the number of people that can be affected in terms of new infected cases and deaths including the number of expected recoveries for the upcoming days. Two machine learning models LR and SVM, have been used to predict the accuracy values.

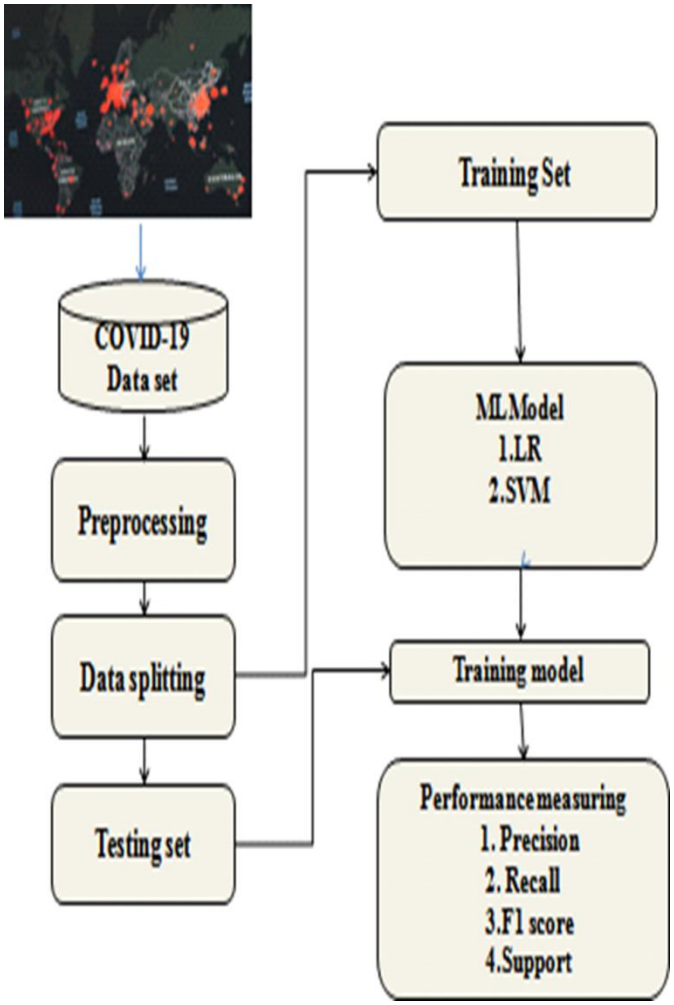


Figure 1: Block Diagram of Proposed System

2) SUPPORT VECTOR MACHINE

A support vector machine (SVM) is a type of supervised ML algorithm used for both regression and classification. SVM regression being a non-parametric technique depends on a set of mathematical functions. The set of functions called kernel transforms the data inputs into the desired form. SVM solves the regression problems using a linear function, so while dealing with problems of non-linear regression, it maps the input vector(x) to n-dimensional space called a feature space

This mapping is done by non-linear mapping techniques after that linear regression is applied to space. Putting the concept in ML context with a multivariate training data set (xn) with N number of observations with yn as a set of observed responses. The linear function can be depicted as

$$f(x) = x^j\beta + b \quad (4)$$

SV classification:

$$\min_{f,\xi} \frac{\| \|_K^2}{2} + C \sum_{i=1}^{l} \xi_i \quad \text{y f(x) } \geq 1 - \xi_i \text{ , for all i } \xi_i \geq 0 \quad (5)$$

SVM classification, Dual formulation:

$$\min_{\alpha_i} \sum_i \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad 0 \leq \alpha_i \leq C, \quad (6)$$

$$\text{for all i;} \quad \sum_{i=1} \alpha_i y_i = 0 \quad (7)$$

Variables □i are called slack variables and they measure the error made at point (xi,yi). Training SVM becomes quite challenging when the number of training points is large. A number of methods for fast SVM training have been proposed.

IV. EVALUATION PARAMETERS

In this study, we evaluate the performance of each of the learning models in terms Performance Metrics such as Precision ,Recall, F1 Score, Support values.

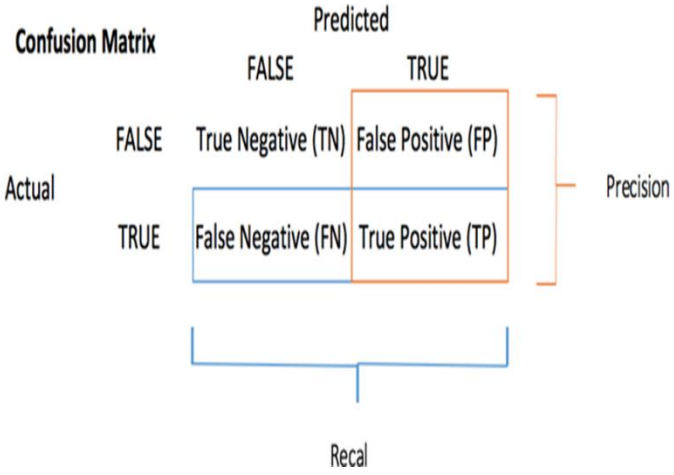
A. PERFORMANCE METRICS

The proposed COVID-19 risk prediction system is implemented using PYTHON under windows environment. The proposed machine learning based prediction system is trained and tested with dataset obtained from the Kaggle machine learning repository site. The COVID-19 data set is given as input to the

algorithm and the data is divided into training data and test data. The training set for the COVID-19 consists of 70% of the total data set and the testing set is 30% of the total data. The performance of risk prediction system with both machine learning approaches (Linear Regression algorithm, Support Vector Machine algorithm) are compared in terms of performance metrics obtained from confusion matrix which is shown in table 3.

- True Positive (TP) - The Extracted data set containing COVID-19 module is classified as COVID-19 cases
- False Positive (FP) - The Extracted data set without COVID-19 module is classified as COVID-19 cases

Table 3 show the performance metric obtained from confusion matrix



- True Negative (TN) - The Extracted data set without COVID -19 module is classified as non-COVID-19 cases
- False Negative (FN) - The Extracted data set containing COVID-19 module is classified as non-COVID 19 cases

Both machine learning algorithms are tested with same data set with varying percentage of training and testing data. All the performance matrices are measured and tabulated. The accuracy of prediction is measured by using the following formula.

Accuracy(A) = (TP+TN) / (TP + TN + FP + FN) (8)

Sensitivity means that how accurately a Covid-19 test identifies people as presence of Covid-19 cases .

Recall(R) or Sensitivity=TP(TP+FN) (9)

Specificity means that how accurately a Covid-19 test identifies people who do not have Covid-19 case.

Precision(P) or (1–Specificity)= TP(TP+FP) (10)

Here several performance metrics are used to measure the accuracy of classification

V. RESULTS AND DISCUSSION

This study attempts to develop a system for the future fore casting of the number of cases affected by COVID- 19 using machine learning methods. The data set used for the study contains information about the daily reports of the number of newly infected cases, the number of recoveries. The proposed system uses Linear Regression and Support Vector Machine algorithms for prediction of COVID 19 predictions for different diseases using machine learning techniques such as fever, cough , breathing problem, throat pain, joint pain, diarrhea, tholach, old peak, pneumonia, trestbps is focused on live forecasting of COVID-19 confirmed cases and study is also focused on the forecast of COVID-19 outbreak and early response.

Various regression models have wide applicability in predicting the conditions of patients in the future with a specific disease .To contribute to the current human crisis our attempt in this study is to develop a forecasting system forCOVID-19.These prediction systems can be very helpful in decision making to handle the present scenario to guide early interventions to manage these diseases very effectively. In proposed model data is divided in to training and testing values. Machine learning based

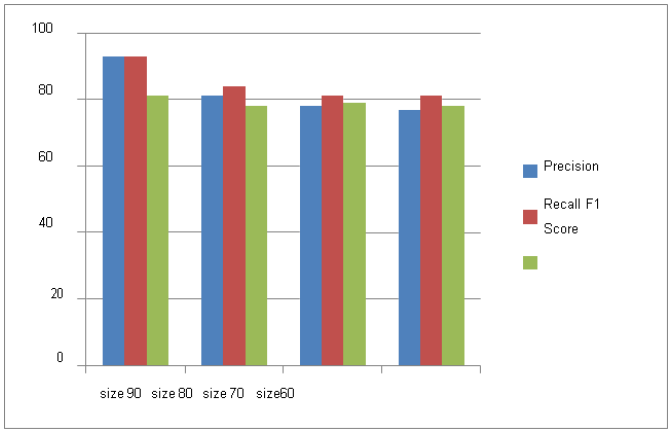
Linear Regression and Support Vector Machine model is built with training data and then Precision, Recall, F1 sore, Support values are calculated with testing set. From the symptoms such as Fever, cough, breath in, breath out, throat pain, joint pain, diarrhea, tholach, old peak, pneumonia this system attempt to forecast the number of people that can be affected in terms of new infected cases and deaths including the number of expected recoveries for the upcoming days. Two machine learning models LR and SVM, have been used to predict the accuracy values.

A. LINEAR REGRESSION MODEL PERFORMANCE METRICS

Table 5.Show the performance metrics measured using Linear Regression model Precision, Recall, Support, F1 score and Accuracy values of Linear Regression based

S.NO	Training size(%)	Precision	Recall	F1 Score
1	90	0.93	0.93	0.81
2	80	0.81	0.84	0.78
3	70	0.78	0.81	0.79
4	60	0.77	0.81	0.78
5	50	0.77	0.83	0.78

Figure 2. Shows the Accuracy obtained from the result of COVID 19 Linear Regression based model for varying size of Training and Testing data Covid-19 Prediction Accuracy of Linear Regression model for varying size of training and testing data



B. SUPPORT VECTOR MACHINE PERFORMANCE METRICS

Table 6. Shows the performance metrics calculation of SVM based model It contains the values of performance metrics such as Precision, Recall, F1score, support and accuracy for varying size of training and testing data set.

S.NO	Training Size(%)	Precision	Recall	F1 Score
1	90	0.93	0.93	0.87
2	80	0.78	0.81	0.77
3	70	0.79	0.84	0.82
4	60	0.77	0.81	0.78
5	50	0.76	0.84	0.80

Figure 4.Shows the Accuracy obtained from the result of COVID 19 Support Vector Machine based model for varying size of Training and Testing data

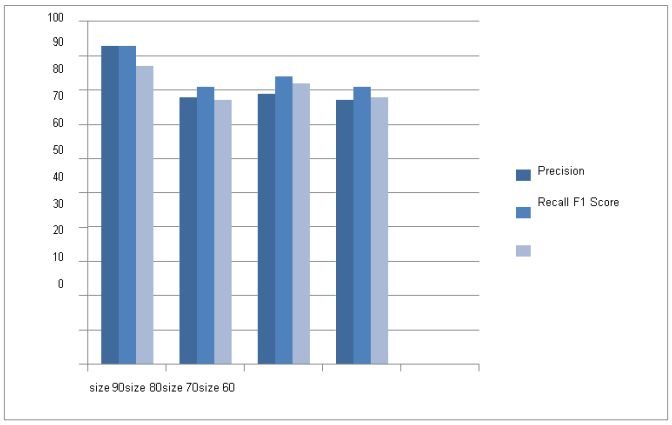


Table 7.Covid-19 Prediction Accuracy of Support Vector Machine based model for varying size of training and testing data Compares the Accuracy of Linear Regression and Support Vector Machine based Prediction model

Test size	LR	SVM
90	87.09	89.09
80	76.4	75.4
70	75.92	79.12
60	75.4	75.4
50	75.4	77.63

Figure 5.Shows the graphical representation of Accuracy obtained through Linear Regression based model

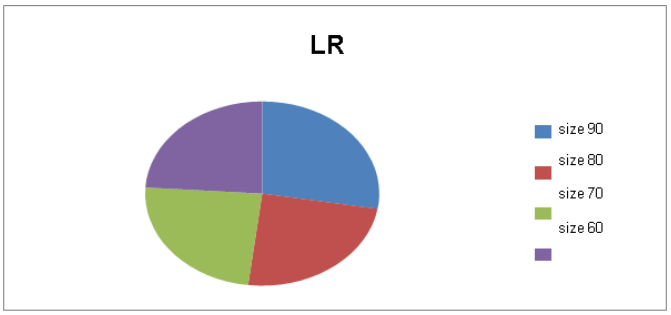
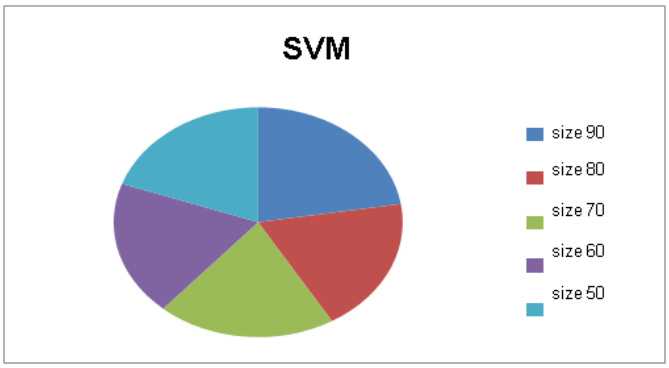


Figure 6.Graphical representation of Accuracy obtained through Support Vector Machine based model



C. PREDICTION INTERVALS OF SVM FOR FORECASTING UNCERTAINTY

A prediction interval is a quantization of the uncertainty on a prediction. It provides a probabilistic upper and lower bounds on the estimate of an outcome variable . To evaluate this uncertainty prediction intervals on LR is performed. Among two machine learning models in general, SVM performs better in LR cases.

VI. CONCLUSION

The precariousness of the COVID-19 pandemic can ignite a massive global crisis. Some researchers and government agencies throughout the world have apprehensions that the pandemic can affect a large proportion of the world population . In this study, an ML-based prediction system has been proposed for predicting the risk of COVID-19 outbreak globally. The system analyses dataset containing the day-wise actual past data and makes predictions for upcoming days using machine learning algorithms. The results of the study prove that LR and SVM perform best in the current forecasting domain given the nature and size of the dataset. According to the results of these two models SVM produces best results in all scenarios because of the ups and downs in the dataset values. The system has analysed Covid-19 data set for risk predictions in upcoming days using machine learning algorithms such as Linear Regression and Support Vector Machine. From the observations it is found that SVM perform forecasting with better accuracy than Linear Regression. From the experimental result it is observed that SVM based model has produced Precision-93%, Recall-93%, F1score-87% and Accuracy 93% .The combination of LR and linear SVM is an effective way to improve forecasting performance. In future efforts can be made to improve the accuracy of proposed system and the system will be improved to perform real time live forecasting.

VII. REFERENCES

[1]. F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.- W. Tao, J.-H. Tian, Y.-Y. Pei et al., “A new corona virus associated with human respiratory disease in,” Nature, vol. 579, no. 7798, pp. 265– 269, 2020.

[2]. Data analytics and visualization using Tableau utilitarian for COVID-19 (Coronavirus) Nihau Akhtar 1,*, Nazia Tabassum 2, Asif Perwej 3 and Yusuf Perwej 4. 2020, 03(02), 028–050

[3]. Y. Grushka-Cockayne and V. R. R. Jose, “Combining prediction intervals in the m4 competition,” Int. J. Forecasting, vol. 36, no. 1, pp. 178–185, Jan. 2020.

[4]. N. C. Mediaite. Harvard Professor Sounds Alarm on ‘Likely’ Coronavirus Pandemic: 40% to 70% of World Could be Infected This Year. Accessed: Feb. 18, 2020. Online]. Available: harvard-professor-sounds-alarm-on- likely-

%coronavirus-pandemic-40- to-70-of-world-could-be-infected-this-year/

[5]. Neural network based country wise riskprediction of COVID-19 Ratnabali pali Arif ahmed sekh ,samrjit kar,dilipk ,current version 2020M. Li, J. Chen, and Y. Deng, "Scaling features in the spreading of covid-19," arXiv preprint arXiv:2002.09199, 2020

[6]. D. S. Kermany et al., "Identifying medical diagnoses and treat- able diseases by image-based deep learning," Cell, vol. 172, no. 5, pp. 1122–1131.e9, Feb. 2018.

[7]. P. Rajpurkar et al., "Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists," PLOS Med., vol. 15, no. 11, 2018, Art. no. e1002686.

[8]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[9]. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 248–255.

[10]. Y. Mori, S.-E. Kudo, T. Berzin, M. Misawa, and K. Takeda, "Computer-aided diagnosis for colonoscopy," Endoscopy, vol. 49, no. 08, pp. 813–819, Aug. 2017.

[11]. R. Zhang et al., "Automatic detection and classification of colorec- tal polyps by transferring low- level CNN features from nonmedical domain," IEEE J. Biomed. Health Informat., vol. 21, no. 1, pp. 41–47, Jan. 2017.

[12]. Johns Hopkins University Data Repository. Cssegisanddata. Accessed: Mar. 27, 2020. Online]. Available: https://github.com/CSSEGISandData

[13]. M. R. M. Talabis, R. McPherson, I. Miyamoto, J. L. Martin, and D. Kaye, "Analyticsdefined,"inInformationSecurityAnalytics,M.R.M.Talabis,

[14]. R. McPherson, I. Miyamoto, J. L. Martin, and D. Kaye, Eds. Boston, MA, USA: Syngress, 2015, pp. 1–12. Online]. Available: http://www.sciencedirect.com/science/article/pii/B978012800 20700 00010

[15]. H.-L.Hwa,W.-H.Kuo,L.-Y.Chang,M.-Y.Wang,T.-H.Tung,K.-J.Chang, and F.-J. Hsieh, "Prediction of breast cancer and lymph node metastatic statuswithtumourmarkersusinglogisticregression models,' 'J.Eval.Clin. Pract., vol. 14, no. 2, pp. 275–280, Apr.2008.

[16]. R. Tibshirani, "Regression shrinkage and selection via the lasso," J. Roy. Stat. Soc., Ser. B, Methodol., vol. 58, no. 1, pp. 267–288, Jan.1996.

[17]. A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," Technometrics, vol. 12, no. 1, pp. 55–67, Feb. 1970.

[18]. E. Cadenas, O. A. Jaramillo, and W. Rivera,``Analysis and forecastingof wind velocity in chetumal, quintana roo, using the single exponential smoothing method," Renew. Energy, vol. 35, no. 5, pp. 925_930, May 2010.

[19]. J. Lupón, H. K. Gaggin, M. de Antonio, M. Domingo, A. Galán, E. Zamora,J. Vila, J. Peña_el, A. Urrutia, E. Ferrer, N. Vallejo, J. L. Januzzi, and A.Bayes-Genis, ``Biomarker-assist score for reverse remodeling predictionin heart failure: The ST2-R2 score," Int. J. Cardiol., vol. 184,pp. 337_343, Apr. 2015.

[20]. J.-H. Han and S.-Y. Chi, ``Consideration of manufacturing data to apply machine learning methods for predictive manufacturing," in Proc. 8th Int. Conf. Ubiquitous Future Netw. (ICUFN), Jul. 2016, pp. 109_113.

[21]. C. Willmott and K. Matsuura, ``Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," Climate Res., vol. 30, no. 1, pp. 79_82, 2005.

[22]. R. Kaundal, A. S. Kapoor, and G. P. Raghava,``Machine learning techniques in

disease forecasting: A case study on rice blast prediction," BMC Bioinf., vol. 7, no. 1, p. 485, 2006.

[23]. S. Baran and D. Nemoda, ``Censored and shifted gamma distribution based EMOS model for probabilistic quantitative precipitation forecasting," Environmetrics, vol. 27, no. 5, pp. 280_292, Aug. 2016.

[24]. Y. Grushka-Cockayne and V. R. R. Jose,``Combining prediction intervals in the m4 competition," Int. J. Forecasting, vol. 36, no. 1, pp. 178_185, Jan. 2020.

[25]. N. C. Mediaite. Harvard Professor Sounds Alarm on `Likely' Coronavirus Pandemic: 40% to 70% of World Could be Infected This Year. Accessed: Feb. 18, 2020. Online]. Available: https://www.mediaite.com/news/ harvard-professor- sounds-alarm-on-likely-%coronavirus-pandemic-40-to-70-of-world-could-be-infected-this-year/

[26]. BBC. Coronavirus: Up to 70% of Germany Could Become Infected_ Merkel. Accessed: Mar. 15, 2020. Online]. Available: https://www.bbc. com/news/world- us-canada-51835856