*I. J. Mathematical Sciences and Computing*, **2021, 2, 14-23**
Published Online June 2021 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijmsc.2021.02.02

# Machine Learning Based on Kernel Function Controlled Gaussian Process Regression Method for In-depth Extrapolative Analysis of Covid-19 Daily Cases Drift Rates

**Joseph Isabona**
Department of Physics, Faculty of Science, Federal University Lokoja, PMB, 1154, Lokoja, Kogi State
Emails: josabone@yahoo.com, joseph.isabona@fulkoja.edu.ng

**Divine O. Ojuh**
Department of Physical Sciences, Faculty of Sciences, Benson Idahosa University, Benin City, Edo State
Emails: dojuh@biu.edu.ng

Received: 18 March 2021; Revised: 22 April 2021; Accepted: 03 May 2021; Published: 08 June 2021

**Abstract:** Precise extrapolative mining and analysis of relevant dataset during or after any disease outbreak can assist the government, stake holders and relevant agencies in the health sector to make important decisions with respect to the disease outbreak control and management. While prior works has concentrated on non-stationary long term data, this work focuses on a short term non-stationary and relatively noisy data. Particularly, a distinctive nonparametric machine learning method based kernel-controlled probabilistic Gaussian process regression model has been proposed and employed to model and analyze Covid-19 pandemic data acquired over a period of approximately six weeks. To accomplish the aim, the MATLAB 2018a computational and machine learning environment was engaged to develop and perform the Gaussian process extrapolative analysis. The results displayed high scalability and optimal performance over the commonly used machine learning methods such as the Neural networks, Neural-Fuzzy networks, Random forest, Regression tree, Support Vector machines, K-nearest neighbor and Discriminant linear regression models. These results offer a solid foundation for conducting research on reliable prognostic estimations and analysis of contagious disease emergence intensity and spread.

**Index Terms:** Covid-19 pandemic, short term, Non-stationary data, kernel, Machine learning, Gaussian process regression

## 1. Introduction

In spite of unrelenting efforts to improve on the health care systems and processes globally, the outbreak and spreading of corona virus disease 19 epidemics, also popularly shortened as ‘Covid-19’ that is currently ravaging the whole human wellbeing remain a primary public health challenge.

In Nigeria, the first confirmed COVID-19 pandemic case was made known publicly on 27 February 2020. This happened when an Italian national who visited Lagos around the said date tested positive for the virus [1,2]. Up till date, information of such confirmed COVID-19 pandemic cases are being reported daily globally, including the length and breathe of Africa, and well as our dear Country-Nigeria.

Nevertheless, realistic response to such quantum epidemics counts on propitious intervention, superlatively informed by available up-to-date data sources. Also, the visualization and predictive analysis of such epidemic data can play a vital role in exclusive decision making. Predictive analysis plays a decisive head the cast role in driving fruitful outcomes and scrutinizing pain points [3,4]. Specifically, the provision of detail predictive analysis of such Covid-19 data can assist the government, stake holders and relevant agencies in the health sector and relevant agencies to make important decisions with respect future disease outbreak control and management. The results of the quantitative predictive analysis can as well help to reveal whether various efforts currently being put in place to manageably control or curtail the disease outbreak is effective or not.

 

## 2. Literature Review

There exist many linear or nonlinear disease extrapolative models and methods in literature and among they key ones are autoregressive integrated moving average (ARIMA) [5,6], Neural Networks (NN) [7-20] and nonlinear autoregressive neural networks (NAR) [21]. However, ARIMA, NN and NAR models, including other similar ones usually performed poorly on non-stationary long-term data and noisy data sets [22-28].

In [14, 15, 23, 24], the authors explored the long short-term memory model (LSMM) to tackle the non-stationary long-term data with time series analysis. But the LSMM only predicts well when the data is large [25-28].

In this paper, a machine learning based on Gaussian process regression model that is capable of handling both simple and complex input-output data relationship, including the non-stationary data and noisy Covid-19 data sets is presented. The impact of different kernels (covariance functions), which usually impose some level of characteristics weights on the Gaussian process modelling are also investigated and reported.

## 3. Methodology

In this research work, nonparametric machine learning method based kernel-controlled probabilistic Gaussian process regression model has been proposed and employed to model and analyze Covid-19 pandemic data acquired over a period of approximately six weeks. To accomplish the aim, the MATLAB 2018a computational and machine learning environment was engaged to develop and perform the Gaussian process extrapolative analysis.

*3.1. Gaussian Regression Process*

Gaussian processes (GP) can be described as a finite assemblage of conjointly random variables with (stable) Gaussian distribution. In terms of regression problems, these random variables define and characterize the values of a function f(x) over an x input and y output space.

Given a k training set of data, $\aleph_1 = [(x_i, y_i)]_{i=1}^{t}$, where $x_i$ expresses the input vector and $y_i$ the target, the desire is to efficiently learn a function $f(x_i)$ such that the input vector $x_i$ is transform into the target $y_i$ as defined by:

$$y_i = ((x_i) + \varepsilon_i; i = 1,2,...,t) \tag{1}$$

where

$\varepsilon$ designate the Gaussian zero mean noise and $\sigma_t^2$ the variance.

The observed target can equally be defined using the expression equation (2):

$$y_i \approx N(0, K(X, X) + \sigma_i^2); i = 1,2,...,t) \tag{2}$$

where

$K(X, X)$ expresses the covariance matrix.

Now, let $\aleph_2 = [(x_i^{\ t}, y_i^{\ t})]_{i=1}^{k*}$ is another set of unknown i.i.d. testing data samples like the $\aleph_1$.

The combined distribution of the focused target data values and the expected desired predicted output over the input vector can be described by:

$$\begin{vmatrix} f \\ \vec{f}* \end{vmatrix} X, X* \approx N\left(\vec{0}, \begin{bmatrix} \mathrm{K(X,X)} + \sigma_i^2 \mathrm{I} & \mathrm{K(X,X*)} \\ \mathrm{K(X*,X)} & \mathrm{K(X*,X*)} \end{bmatrix}\right) \tag{3}$$

where

$$X = \begin{bmatrix} -(x_1)T \\ -(x_2)T \\ \vdots \\ -(x_k)T \end{bmatrix} \in \mathrm{R}^{nxm}, \qquad \vec{y} = \begin{bmatrix} (y_1)T \\ (y_2)T \\ \vdots \\ (y_k)T \end{bmatrix} \in \mathrm{R}^{m} \tag{4}$$

$$\vec{f} = \begin{bmatrix} f(x_1)T \\ f(x_2)T \\ \vdots \\ f(x_k)T \end{bmatrix} \in \mathrm{R}^{mxn}, \in = \begin{bmatrix} (\epsilon_1) \\ (\epsilon_2) \\ \vdots \\ (\epsilon_k) \end{bmatrix} \in \mathrm{R}^{m} \tag{5}$$

$$X *= \begin{bmatrix} -(x_1)T \\ -(x_2)T \\ \vdots \\ -(x_k)T \end{bmatrix} \in R^{nxm}, \qquad \vec{y} * = \begin{bmatrix} (y_1 *)T \\ (y_2 *)T \\ \vdots \\ (y_k *)T \end{bmatrix} \in R^{m*} \tag{6}$$

$$\vec{f} * = \begin{bmatrix} f(x_1 *)T \\ f(x_2 *)T \\ \vdots \\ f(x_k *)T \end{bmatrix} \in R^{m*x\,n}, \in= \begin{bmatrix} (\epsilon_1 *) \\ (\epsilon_2 *) \\ \vdots \\ (\epsilon_k *) \end{bmatrix} \in R^{m*} \tag{7}$$

and

$K(X,X) \in R^{mx\,n}$, $K(X,X^*) \in R^{mx\,n^*}$, $K(X^*,X) \in R^{m^*x\,n}$ and $K(X^*,X^*) \in R^{m^*x\,n^*}$

*3.2. Kernels*

The kernel (or covariance function) plus the mean function completely delineates the Gaussian process modeling. There exist a number of kernels which can impose different characteristics weights on the Gaussian process modelling. The ones investigated in this work includes: ardexponential, squaredexponential, ardsquaredexponential, exponential, ardrationalquadratic, rationalquadratic, matern32, ardmatern32, matern52, and ardmatern52.

*3.3 Six Alternative Regression Models for Comparison.*

**(a) A multilayer perceptron (MLP) Neural Network.** MLP is a special popular feedforward artificial neural network (ANN) model with at least three layered nodes. Its input-output prediction function can be expressed as [18, 20]:

$$\mathrm{y(x)} = f_{log}\left(w_o + \vec{w}, \vec{x}\right) + b \tag{8}$$

where $f_{log}$ denote the logistic function, $w_o$ is the weight matrix, $\vec{w}$ and $\vec{x}$ are the kth neurons number and the input variable, b is the bias.

**(b) Decision Regression Trees (DRT):** Regression trees are known distinctive nonlinear regression machine learning model for predictive analysis. Its input-output prediction function can be defined as:

$$\mathrm{y(x)} = \frac{1}{\mathrm{z}} \sum_{i=1}^{z} y_i \tag{9}$$

where z denote the available observation number at a specified cell.

**(c) Support Vector Machine (SVM):** SVM regression is another nonlinear regression machine learning model used for predictive analysis. Its input-output prediction function for non-linear regression can defined as:

$$\mathrm{y(x)} = \sum_{i=1}^{c} (\alpha_i - \alpha_i^{\bullet}).k(x_i..x) + b \tag{10}$$

where $\alpha_i, \alpha_i^{\bullet}$ designate the langrage multipliers, $c$ is the bias term and $k(.)$ is the kernel function.

**(d) Neural-Fuzzy.** SVM is a special hybrid nonlinear supervised learning model which combines neural networks with that of fuzzy logic for robust data predictions. Its input-output ANFIS (adaptive neuro-fuzzy inference system) based prediction function can defined as:

$$\mathrm{y(o)} = \sum_{i=1} \vec{w} f_i{}_i = \frac{\sum_{i=1} w_i f_i}{\sum_{i=1} w_i} \tag{11}$$

where $\vec{w}$ is the output layer. $f_1, f_2$ express the hybridized least square function and gradient descent function

**(e) Random Forest:** Random forest is a predictor made up of randomized base regression trees collection of randomized base regression trees $\{r_n(x, \theta_m, D_n) m \geq 1\}$). Its input-output function can be defined as:

$$r_n(x,, D_n) = E_o\{r_n(x, \theta_m, D_n)\} \tag{12}$$

where $\theta_m$ indicate the i.i.d. randomizing output variable. $E_o$ denotes expectation in correspondence with the random parameter, and the data set Dn.

**(f) K-nearest Neighbors:**K-nearest Neighbors (KNN) is a nonparametric regression method which employs 'feature similarity' to estimate and predict new data points' values. Its input-output function can be defined as [29]:

$$\mathrm{f(x)} = \frac{1}{\mathrm{K}} \sum_{x_i \varepsilon N_o}^{K} y_i \tag{13}$$

where
K= nearest neighbor number
*xo*. = the prediction point
*No*, = training observation number

**(g) Linear Discriminant analysis Regression:** Linear discriminant analysis employs number of prediction equations to estimate cluster membership from predictors set by means of independent variables. Its input-output function can be defined as:

$$\mathrm{D} = a + b_1 x_1 + b_2 x_2 .... + b_m x_m \tag{14}$$

where:
a, b= modelling parameters
*x*. = input variable
*m*, = training observation number

## 4. Data Collection and Transformation Method

The six weeks Covid-19 datasets used in this work were obtained from Nigeria Center for Disease Control (NCDC) website database (https://ncdc.gov.ng/) [30]. The NDCDC is the Nigeria public health institute, established by the Government with the sole responsibility to control, prevent, detect, respond and manage public health emergencies and infectious disease outbreaks. Among other key responsibilities, the institute is also empowered to organize surveillance systems for data collection, synthnesization and provision on different diseases of public health importance.

The data collected from the institute are of two main parts. The first part consist of only confirmed cases of Covid-19 disease infected persons across thirty states in Nigeria since 13th March, 2020. The second part consist confirmed cases of daily infected persons in admission, discharged person who recovered from Covid-19 disease after treatment, persons who died of Covid-19 disease during treatment.

Each portion of the obtained data where processed and standardized before passing it through it the different regression training processes. Their scripts coding and execution was achieved using Matlab 2018a software/user interface. Regression and some of their Key training parameters and algorithms are provided in table 1:

Table 1. Regressors and some of their Key training parameters/algorithms

| Regressor Name | Some Key training parameters/algorithms |
|---|---|
| MLP Neural Networks | Training algorithm: trainlm, No. of layers/neurons: 2:20, NumEpochs = 50 |
| Neural-Fuzzy Networks | NumMfs = 20; MfType = gbellmf, NumEpochs = 100 |
| Support Vector | Kernel function: Gaussian, KFold:5, KernelScale:auto, OptimizeHyperparameters; auto |
| Linear Discriminant Analysis | DiscrimType: linear, Holdout:0.3, Optimizer: bayesopt |
| Random Forest | Method:Bag,NumLearningCycles:200 |
| Binary Decision tree | MaxNumSplits:7, Optimizer: bayesopt |
| K-nearest neighbor | NumNeighbors:2,distfunctn:Euclidean |
| Gaussian Process | Kernel function: Squared exponential, FitMethod: sr, PredictMethod: fic Optimizer: bayesopt |

## 5. Results and Analysis

Here, all the graphics and computations were achieved using Matlab 2018a software. First, we start by looking at the impact of different kernels (covariance functions), which usually impose high degree of characteristics weights on the Gaussian process modelling. Table 2 display the attained prediction performance impact of the ten different kernels on confirmed cases Covid-19 data using four evaluation metrics such as mean absolute error (MAE), standard deviation error (STD), coefficient of determination ($R^2$) and root mean absolute error (RMSE). A lower MAE, STD AND RMSE values indicate superior prediction accuracy. $R^2$ value close or equal to 1 means the predicted out values correlate well with the focused target data. The results reveal that all kernels almost have the same predictive performance results. However, squaredexponential recorded the lowest MAE of 0.0010 compared to others. Therefore, Gaussian process regression with squaredexponential kernel is chosen and employed in the remaining parts of this work.

Table 2. Prediction Errors and R2 Values for 10 different Kernels investigated

| Kernel | MAE | STD | RMSE | $R^2$ |
|---|---|---|---|---|
| ardexponential | 0.0018 | 0.0438 | 0.0442 | 1.0000 |
| squaredexponential | 0.0010 | 0.0020 | 0.0020 | 1.0000 |
| ardsquaredexponential | 0.0060 | 0.0076 | 0.0075 | 1.0000 |
| exponential | 0.0409 | 0.0609 | 0.0409 | 1.0000 |
| ardrationalquadratic | 0.0058 | 0.0075 | 0.0075 | 1.0000 |
| rationalquadratic | 0.0020 | 0.0021 | 0.0021 | 1.0000 |
| matern32 | 0.0034 | 0.0047 | 0.0046 | 1.0000 |
| ardmatern32 | 0.0027 | 0.0038 | 0.0038 | 1.0000 |
| matern52 | 0.0011 | 0.0020 | 0.0020 | 1.0000 |
| ardmatern52 | 0.0062 | 0.0079 | 0.0079 | 1.0000 |

Figs. 1 to 8 display predictive analysis graphs of confirmed cases of Covid-19 data plotted using are the Gaussian process regression model with squaredexponential kernel and seven other models for the purpose of comparison. From the graphs and Table 3, the focus Gaussian process regression model display the best prediction performance with approximately 0.008 RMSE value when compared to Neural networks, Neural-Fuzzy networks, Neural-fuzzy model, Random forest, Regression tree, Support Vector Machines, K-nearest model and Discriminant regression model models that attained 28.87, 6.75, 38.84, 41.58, 40.19, 9.23 and 17.23 RMSE values.

The summarized results of table 4 shows the predictive performance of different regression model on confirmed cases data of daily infected persons in admission, data on discharged persons that recovered from Covid-19 disease after treatment, and lastly data on persons who died of Covid-19 disease since 13th March to 17th May, 2020. Though the Gaussian process regression model display a best prediction results on all Covid-19 data as usual, but we can also observed from the table that neural-fuzzy prediction accuracies are also very high, probably due to its hybrid nature. This results also implies that neural-fuzzy model can be used an alternative model. On the other hand, support vector machine and Random forest regression models performed the worst prediction result.

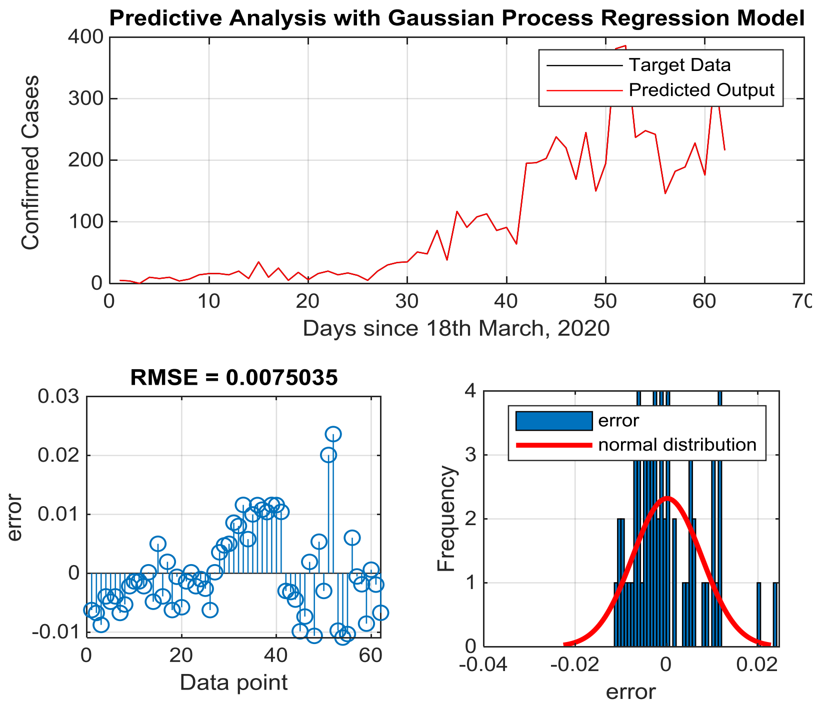


Fig. 1. Predictive analysis plot with Gaussian process regression

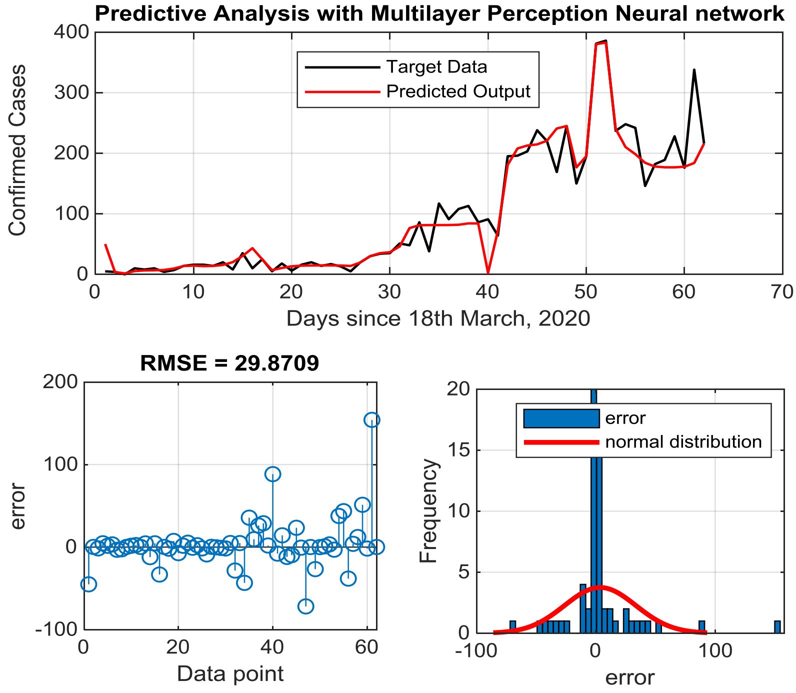


Fig. 2. Predictive analysis plot with Neural network Model

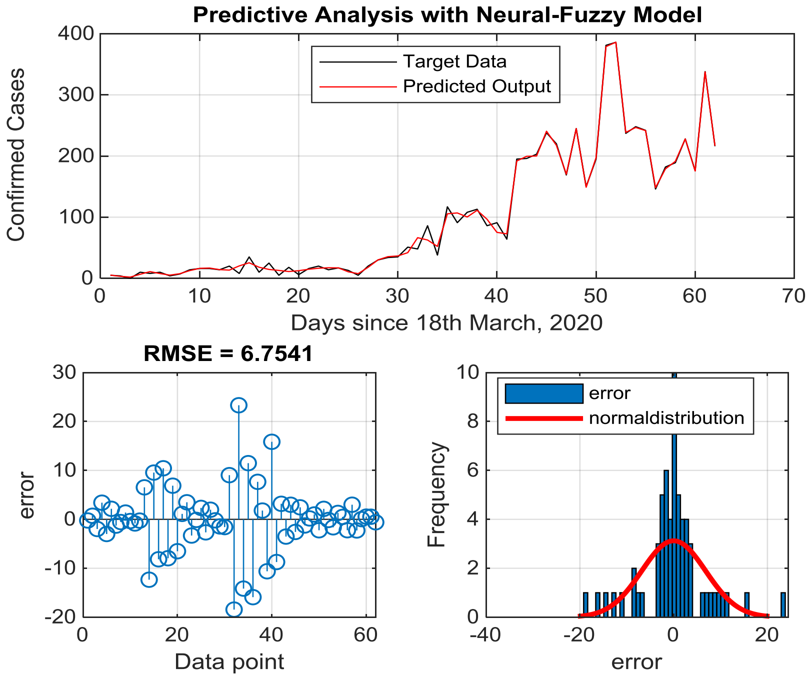


Fig.3. Predictive analysis plot with Neural-Fuzzy model

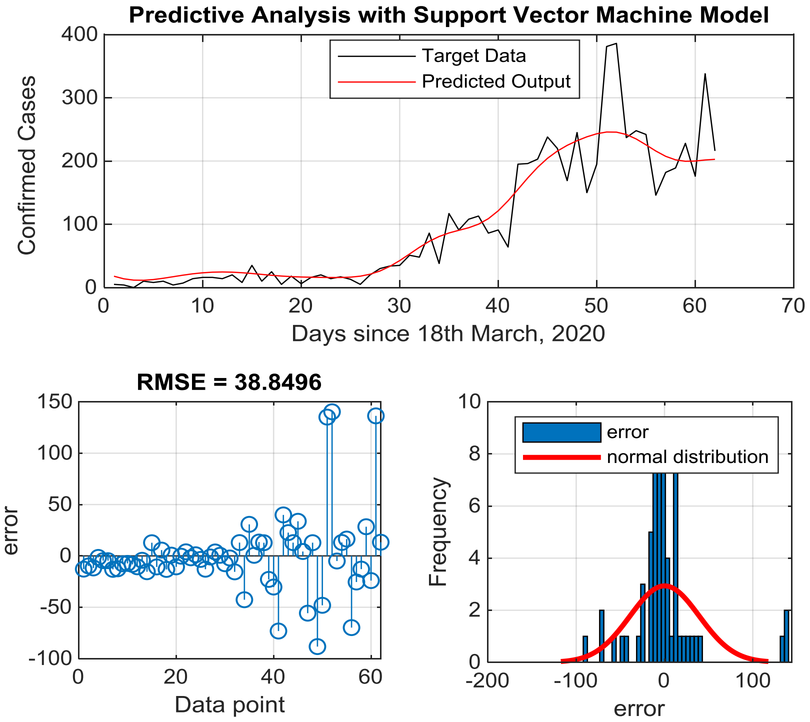


Fig. 4. Predictive analysis plot with Support Vector Machine model

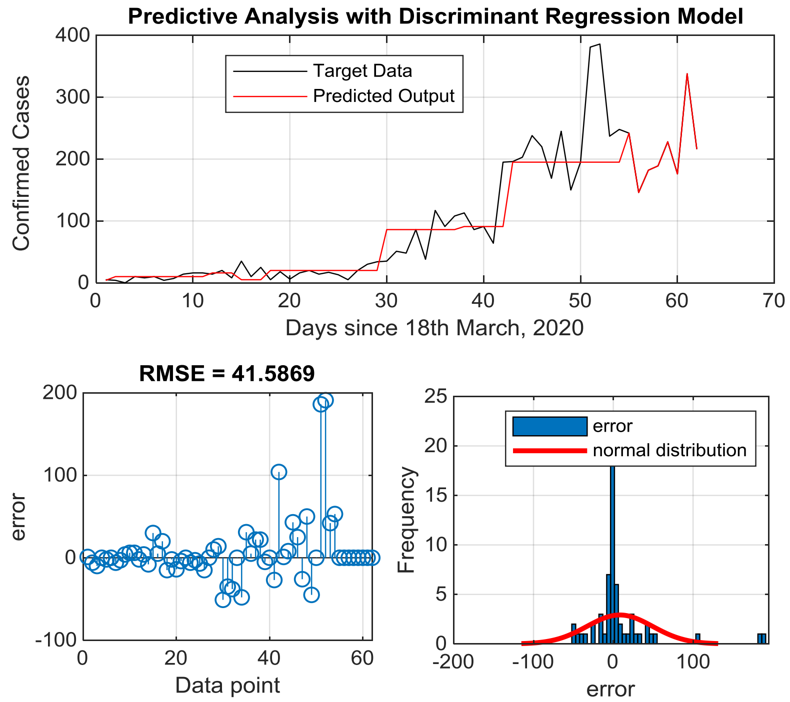


Fig. 5. Predictive analysis plot with Discriminant regression model

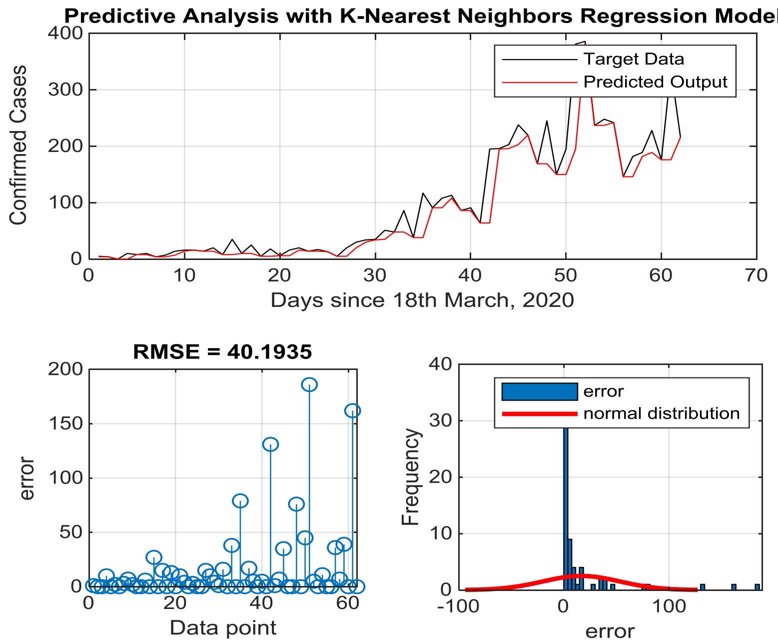


Fig. 6. Predictive analysis plot with K-Nearest Neighbor model

Table 3. Prediction Errors and $R^2$ Values for 8 different Regression Models investigated

| Regression Model | MAE | STD | RMSE | $R^2$ |
|---|---|---|---|---|
| MLP Neural Network | 15.32 | 29.90 | 29.87 | 0.9563 |
| Neural-Fuzzy Network | 4.420 | 274 | 6.81 | 0.9978 |
| Support Vector Machine | 22.75 | 1411 | 37.17 | 0.9260 |
| Discriminant | 20.33 | 1261 | 41.11 | 0.9153 |
| K-nearest Neighbor | 1668 | 1034 | 38.87 | 0.9208 |
| Decision Tree | 6.59 | 407 | 9.81 | 0.9954 |
| Random Forest | 25.25 | 1565 | 42.48 | 0.9126 |
| Gaussian Process | 0.006 | 0.369 | 0.008 | 1.0000 |

Table 4. Prediction Errors and $R^2$ Values for 8 different Regression models investigated

| | Regression Model | MAE | STD | RMSE | $R^2$ |
|---|---|---|---|---|---|
| Prediction results using Data on Confirmed cases of persons on admission | MLP Neural Network | 11.75 | 23.53 | 25.72 | 0.9949 |
| | Neural-Fuzzy Network | 0.02 | 0.02 | 0.02 | 1.0000 |
| | Support Vector Machine | 73.94 | 305.01 | 307.57 | 0.2654 |
| | Discriminant Linear | 8.91 | 19.71 | 20.97 | 0.9966 |
| | K-nearest Neighbor | 41.08 | 215.82 | 216.54 | 0.6959 |
| | Decision Tree | 11.88 | 27.47 | 27.04 | 0.9943 |
| | Random Forest | 88.11 | 300.24 | 276.59 | 0.3109 |
| | Gaussian Process | 0.008 | 0.01 | 0.01 | 1.0000 |
| | | | | | |
| Prediction results using Data on discharge cases of persons that recovered from the disease | MLP Neural Network | 14.93 | 43.44 | 42.44 | 0.8285 |
| | Neural-Fuzzy Network | 0.05 | 0.020 | 0.010 | 1.000 |
| | Support Vector Machine | 20.41 | 76.70 | 76.66 | 0.4547 |
| | Discriminant | 2.68 | 4.97 | 5.02 | 0.9977 |
| | K-nearest Neighbor | 17.62 | 75.16 | 75.96 | 0.4646 |
| | Decision Tree | 3.37 | 5.75 | 5.67 | 0.9970 |
| | Random Forest | 28.42 | 86.98 | 85.90 | 0.3153 |
| | Gaussian Process | 0.01 | 0.01 | 0.01 | 1.0000 |
| | | | | | |
| Prediction results using daily Data of those that died of the disease within the period | MLP Neural Network | 3.51 | 6.52 | 6.59 | 0.5859 |
| | Neural-Fuzzy Network | 0.002 | 0.002 | 0.002 | 1.0000 |
| | Support Vector Machine | 3.26 | 7.28 | 7.58 | 0.4534 |
| | Discriminant | 2.57 | 6.50 | 6.44 | 0.6055 |
| | K-nearest Neighbor | 1.45 | 3.73 | 3.96 | 0.8590 |
| | Decision Tree | 0.88 | 2.00 | 1.97 | 0.9629 |
| | Random Forest | 3.81 | 7.00 | 6.93 | 0.5427 |
| | Gaussian Process | 0.0002 | 0.0004 | 0.0005 | 1.0000 |

## 6. Conclusion

In recent years, prognostic-based data mining by means of machine learning models have been gaining cute recognition and relevance in different research areas of science, medicine engineering and humanity.

In this work, a kernel-based probabilistic Gaussian process regression model, which is a distinctive nonparametric machine learning method has been proposed and employed to model and analysis the acquired data over a period of approximately six weeks. To accomplish the task, the MATLAB 2018a computational and machine learning environment was engaged to develop and perform the Gaussian process extrapolative analysis. By worth of some key evaluation metrics such as mean absolute error, standard deviation error, coefficient of determination and root mean absolute error, the extrapolative results of the adopted model reveal robust performance over the commonly used machine leaning types such as the Neural networks, Neural-Fuzzy networks, Neural-fuzzy model, Random forest, Regression tree, Support Vector Machines, K-nearest model and Discriminant regression model.

In future work, prognostic and prescriptive analysis of global datasets on Covid-19 pandemic with larger geographical areas would be the focus.

## References

[1] "First Case of CORONA Virus Disease Confirmed in Nigeria". Nigeria Centre for Disease Control. 28 February 2020. https://ncdc.gov.ng/news/227/first-case-of-corona-virus-disease-confirmed-in-nigeria, Retrieved 22 May, 2020.

[2] Maclean, Ruth; Dahir, Abdi Latif (28 February 2020). "Nigeria Responds to First Coronavirus Case in Sub-Saharan Africa". The New York Times. Retrieved 22 May, 2020.

[3] Isabona, J and Konyeha. C.C (2013) ‘‘Urban Area Path loss Propagation Prediction and Optimisation Using Hata Model at 800MHz’’, IOSR Journal of Applied Physics (IOSR-JAP), Vol. 3, Issue 4, pp.8-18.

[4] Ebhota, V.C, Isabona, J, and Srivastava, V.M. Environment‑Adaptation Based Hybrid Neural Network Predictor for Signal Propagation Loss Prediction in Cluttered and Open Urban Microcells, Wireless Personal Communications, vol.104:935-948, 2019. DOI 10.1007/s11277-018-6061-2 pp. 1-16, 2019

[5] Liu, L., Luan, R. S., Yin, F., Zhu, X. P. & Lu, Q. Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model. Epidemiology and Infection 144, 144–151, https://doi.org/10.1017/s0950268815001144 (2016).

[6] Benvenuto, D; Giovanetti, M; Vassallo, L; Angeletti, S and Ciccozzi, M. Application of the ARIMA model on the COVID-2019 epidemic dataset, Data in Brief, Vol. 29, pp.1-4, 2020, https://doi.org/10.1016/j.dib.2020.105340.

[7] Akhtar, M; Kraemer, M U. G. and Gardner, L. M. A dynamic neural network model for predicting risk of Zika in real time, BMC Medicine, pp. 1-16; 2017. https://doi.org/10.1186/s12916-019-1389-3

[8] Kuo P-J, Wu S-C, Chien P-C, Chang S-S, Rau C-S, Tai H-L, et al. Artificial neural network approach to predict surgical site infection after free-flap reconstruction in patients receiving surgery for head and neck cancer. Oncotarget 2018;9:13768e82. https://doi.org/10.18632/oncotarget.24468.

[9] Habibi Z, Ertiaei A, Nikdad MS, Mirmohseni AS, Afarideh M, Heidari V, et al. Predicting ventriculoperitoneal shunt infection in children with hydrocephalus using artificial neural network. Childs Nerv Syst 2016;32:2143e51.https://doi.org/10.1007/s00381-016-3248-2.

[10] Nidhi Arora, Jatinderkumar R. Saini,"Estimation and Approximation Using Neuro-Fuzzy Systems", International Journal of Intelligent Systems and Applications (IJISA), Vol.8, No.6, pp.9-18, 2016. DOI: 10.5815/ijisa.2016.06.02.

[11] Koushal Kumar, Gour Sundar Mitra Thakur, "Advanced Applications of Neural Networks and Artificial Intelligence: A Review", International Journal of Information Technology and Computer Science (IJITCS), vol.4, no.6, pp.57-68, 2012. DOI: 10.5815/ijitcs.2012.06.08

[12] Ryusuke Hata, M. A. H. Akhand, Md. Monirul Islam, Kazuyuki Murase, "Simplified Real-, Complex-, and Quaternion-Valued Neuro-Fuzzy Learning Algorithms", International Journal of Intelligent Systems and Applications(IJISA), Vol.10, No.5, pp.1-13, 2018. DOI: 10.5815/ijisa.2018.05.01

[13] Obanijesu Opeyemi, Emuoyibofarhe O. Justice, "Development of Neuro-fuzzy System for Early Prediction of Heart Attack", International Journal of Information Technology and Computer Science (IJITCS), vol.4, no.9, pp.22-28, 2012. DOI: 10.5815/ijitcs.2012.09.03

[14] K Srinivasa Rao, G. Lavanya Devi, N. Ramesh, "Air Quality Prediction in Visakhapatnam with LSTM based Recurrent Neural Networks", International Journal of Intelligent Systems and Applications(IJISA), Vol.11, No.2, pp.18-24, 2019. DOI: 10.5815/ijisa.2019.02.03.

[15] Meghna Sharma, Manjeet Singh Tomer, " Predictive Analysis of RFID Supply Chain Path Using Long Short Term Memory (LSTM): Recurrent Neural Networks", International Journal of Wireless and Microwave Technologies (IJWMT), Vol.8, No.4, pp. 66-77, 2018.DOI: 10.5815/ijwmt.2018.04.05

[16] Isabona, J. and Osaigbovo, I. A. ‘‘Investigating Predictive Capabilities of RBFNN, MLPNN and GRNN Models for LTE Cellular Network Radio Signal Power Datasets’’, FUOYE Journal of Engineering and Technology, Vol. 4(1), pp. 155-159, 2019

[17] Ebhota, V.C, Isabona, J, and Srivastava, V.M. ‘‘Investigation and Comparison of Generalization Ability of Multi-Layer Perceptron and Radial Basis Function Artificial Neural Networks for Signal Power Loss Prediction’’**,** *International Journal on Communications Antenna and Propagation***,** Vol. 9 (1), pp. 43-54, 2019.

[18] Ebhota, V.C, Isabona, J, and Srivastava, V.M. ‘‘Effect of Learning Rate on GRNN and MLP for the Prediction of Signal Power Loss in Microcell Sub-Urban Environment’’, *International Journal on Communications Antenna and Propagation***,** Vol. 9 (1), pp. 36-45, 2019.

[19] Obahiagbon, K and Isabona, J. ‘‘Generalized Regression Neural Network: an Alternative Approach for Reliable Prognostic Analysis of Spatial Signal Power Loss in Cellular Broadband Networks’’, International *Journal of Advanced Research in Physical Science,* vol. 5(10): 35-42, 2018.
[20] Isabona, J, and Srivastava, V.M. ‘‘A Neural Network based Model for Signal Coverage Propagation Loss Prediction in Urban Radio Signal Propagation Loss in Urban Microcell**s,** International Journal of Applied Engineering Research Vol. 11, No 22, pp. 11002-11008, 2016.
[21] Zhou, L. et al. A hybrid model for predicting the prevalence of schistosomiasis in humans of Qianjiang City, China. PloS One 9, e104875, https://doi.org/10.1371/journal.pone.0104875 (2014).
[22] Wang,Y; Xu,C; Zhang, S; Yang, L; Wang,Z; Zhu, Y and Yuan, J. Development and evaluation of a deep learning approach for modeling seasonality and trends in hand-foot-mouth disease incidence in mainland China. *Sci Rep* 9**,** 8046 (2019). https://doi.org/10.1038/s41598-019-44469-9
[23] Saqib M, Sha Y, Wang MD. Early prediction of sepsis in EMR records using traditional ML techniques and deep learning LSTM networks. Conf Proc IEEE Eng Med Biol Soc 2018;2018:4038e41. https://doi.org/10.1109/ EMBC.2018.8513254.
[24] Kam HJ, Kim HY. Learning representations for the early detection of sepsis with deep neural networks. Computing in Biological Medicine 2017;89:248e55.https:// doi.org/10.1016/j.compbiomed.2017.08.015.
[25] Ebhota, V.C, Isabona, J and Srivastava V.M., ‘‘Base line knowledge on propagation modelling and prediction techniques in wireless communication networks’’**,** Journal of Engineering and Applied Sciences (JEAS), Vol. 13 (4), pp. 235-240, 2018. https://doi.org/10.15866/irecap.v9i1.15329
[26] Isabona, J. (2020), Wavelet Generalized Regression Neural Network Approach for Robust Field Strength Prediction in Open and Shadow urban Microcells, Wireless Personal Communications**,** Vol. 114 (3), pp.3635–3653
[27] Isabona, J, and Srivastava, V. M. "Hybrid neural network approach for Predicting Signal Propagation Loss in Urban Microcells," 2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Agra, India, 2016, pp. 1-5.
[28] Ebhota, V.C., Isabona, J, and Srivastava, V.M. ‘‘Improved Adaptive Signal Power loss Prediction using Combined Vector Statistics based Smoothing and Neural Network approach’’, Progress in Electromagnetic Research C, Vol. 82, 155–169, 2018.
[29] Goyala, R; Chandraa, P and Singh, Y, Suitability of KNN Regression in the Development of Interaction Based Software Fault Prediction Models, IERI Procedia 6 (2014) 15 – 21.
[30] Nigeria Centre for Disease Control (NDDC), https://ncdc.gov.ng/ Accessed 17th May, 2020

## Authors' Profiles

**Joseph Isabona**, Ph.D, received Ph.D. and M.Sc. degrees in Communication Electronics, 2013 and 2007 respectively, and a B.Sc in Applied Physics in 2003. He is the author of more than 100 scientific contributions including articles in international refereed Journals and Conferences in the area of Wireless Mobile communications. The Author is a Postdoctoral Research Fellow of the Department of Electronic Engineering, Howard College, University of KwaZulu-Natal, Durban, South Africa. His area of interest includes Signal Processing, Radio Resource Management and Physics of radio signal propagation engineering. She can be reached with josabone@yahoo.com.

**Dr. (Mrs.) O. D. Ojuh** is an academic staff of Benson Idahosa University, Benin City, where she lectures Physics. She obtained her Ph.D. and M.Sc in Theoretical/ Mathematical Physics at the University of Benin City, Nigeria in 2012 and 2007 respectively and a B.Sc in Physics in 1996, from then Edo State University now known as Ambrose Alli University, Ekpoma, Edo State. Her research interests are Computational condensed matter physics/materials Science for renewable energy applications and Physics of radio signal propagation engineering. She can be reached through Email dojuh@biu.edu.ng.

**How to cite this paper:** Joseph Isabona, Divine O. Ojuh," Machine Learning Based on Kernel Function Controlled Gaussian Process Regression Method for In-depth Extrapolative Analysis of Covid-19 Daily Cases Drift Rates ", International Journal of Mathematical Sciences and Computing(IJMSC), Vol.7, No.2, pp. 14-23, 2021. DOI: 10.5815/ijmsc.2021.02.02