

COVID-19 mortality rate prediction for India using statistical neural networks and gaussian process regression model

Dhamodharavadhani S, Rathipriya R

Department of Computer Science, Periyar University, Salem-India.

Abstract

The primary purpose of this research is to identify the best COVID-19 mortality model for India using regression models and is to estimate the future COVID-19 mortality rate for India. Specifically, Statistical Neural Networks (Radial Basis Function Neural Network (RBFNN), Generalized Regression Neural Network (GRNN)), and Gaussian Process Regression (GPR) are applied to develop the COVID-19 Mortality Rate Prediction (MRP) model for India. For that purpose, there are two types of dataset used in this study: One is COVID-19 Death cases, a Time Series Data and the other is COVID-19 Confirmed Case and Death Cases where Death case is dependent variable and the Confirmed case is an independent variable. Hyperparameter optimization or tuning is used in these regression models, which is the process of identifying a set of optimal hyperparameters for any learning process with minimal error. Here, sigma (σ) is a hyperparameter whose value is used to constrain the learning process of the above models with minimum Root Mean Squared Error (RMSE). The performance of the models is evaluated using the RMSE and 'R2 values, which shows that the GRP model performs better than the GRNN and RBFNN.

Keywords: Covid-19; India; mortality rate; mortality prediction; regression model; hyperparameter tuning; GPR; GRNN; RBFNN.

DOI: <https://dx.doi.org/10.4314/ahs.v21i1.26>

Cite as: Dhamodharavadhani S, Rathipriya R. COVID-19 mortality rate prediction for India using statistical neural networks and gaussian process regression model. *Afri Health Sci.* 2021;21(1):194-206. <https://dx.doi.org/10.4314/ahs.v21i1.26>

Introduction

At the end of December 2019 in Wuhan, China, it was first reported that a human infection was caused by a novel coronavirus (nCov) or Wuhan virus or 2019-nCov¹. One of the biggest challenges of this epidemic is a human-to-human transition of nCov. The coronavirus (COVID-19) infected cases increase at an exponential rate worldwide. On 30 January 2020, the World Health Organization (WHO) issued a worldwide health emergency warning notice², describing that 2019-nCoV is of critical global concern. The morbidity and mortality rates for the COVID -19 are uncertain at the early stage³ especially for young ones and aged people. WHO has estimated the reproduction factor (R0) of nCov is 2.7. To monitor the massive and rapid spread of the nCov, public health sectors took reliable preventative measures. They imposed curfew or lockdown infested cities in China, the United States, India, and other coun-

tries also. This is to limit the social distance between people and to avoid the spread of this novel virus via humans to humans.

For the last decade, machine learning techniques have gain momentum and play a vital role in many domains of research fields. Notably, it has a tremendous impact on data analytics and data science. It enables a better understanding of the data and its methods, allows for future assumptions based on past data / empirical data, and automatically classifies the type of data (known as classification). Machine Learning (ML) techniques also can be used to develop standard mortality models.

Generally, ML tools used for prediction of the mortality rate of epidemic diseases in advance which helps the public healthcare authorities to develop and design an effective and efficient plan to reduce deaths. Authors Deprez et al. ⁴, used machine learning algorithms to fit and assess the mortality model by detecting the weaknesses of different mortality models. Artificial Neural Networks (ANNs)⁵ used to identify and forecast latent mortality variables with higher predictive accuracy. In ⁶, the authors used neural networks to expand the Lee-Carter model to several predictions of populations.

Corresponding author:

Rathipriya R,
Department of Computer Science, Periyar
University, Salem-India.
Email: rathi_priyar@periyaruniversity.ac.in

Gaussian process models have typically been used extensively in engineering-based optimization applications (Razavi et al. 7). A combination of GPR and adaptive neuro-fuzzy inference system (ANFIS) 8 used in groundwater level forecasting. In 9, an extensive comparative study was carried out between several surrogate models, comprising GPR, using simulation-optimization methodology with uncertainty parameters. In the end, they had concluded that the GPR models and their ensemble were efficient methods concerning prediction accuracy. GRNN model was built in 10 as a new computational method for the field of incidence prediction of infectious diseases. Han, et al. 11 developed a GRNN network with a one-dimensional input and output layer to predict blood, and sexually transmitted infections are occurring. In 12, —authors implemented a comparison analysis on Back Propagation Neural Network (BPNN), Generalized Regression Neural Network (GRNN), and Radial Basis Function Neural Network (RBFNN) network for prediction of the evaporation. The results showed that the GPR is a successful technique compared with artificial neural network approaches. In 13, —a large-scale comparison analysis was presented for major machine learning models such as multilayer perceptron, Bayesian neural networks, radial base functions, generalized neural regression networks (also kernel regression), K-nearest neighbor regression, CART regression trees, support vector regression, and Gaussian time series prediction processes. The authors observed that the performance of these models was solely dependent on the data set, having different impacts.

The RBF and GRNN 14 were applied to data of patients with heart disease for the medication outcome. The results showed that RBF performed well for prescribing medicine for the patient. In 15, the authors claimed that the Gaussian process approach performed better than the standard generalized linear model (GLM) for the Phenomenological forecasting of dengue disease incidence. Gholam Ali Montazer et. al 16 reviewed various learning methods for defining network parameters such as widths, centers, and synaptic weights of the RBF neural network. In 17, extensive neural regression networks were proposed as an automated technique for forecasting time series. This technique is intended to achieve an efficient and fast tool for automatically predicting a vast amount of time series. From these works, one could clearly understand the applications of GPR, GRNN, and RBFNN in various research domains.

In this work, the GPR model with optimized hyperparameter (Co-variance, mean), GRNN with optimized hyperparameter (spread), and RBFNN with optimized hyperparameter (spread) are applied to develop COVID -19 mortality models for two types of dataset. Moreover, to evaluate the performance of these models, RMSE, a quantifiable measure will be used, and the MRP model with low RMSE will be selected as the best model for predicting the COVID-19 mortality rate for India. The purpose of this study is to predict mortality rates against multiple COVID-19 confirmed cases using machine learning techniques that capture patterns that cannot be identified by a standard statistical mortality model.

The rest of the paper is organized as follows. Section 2 details the methods and materials for Covid-19 mortality rate forecasting for India. The outcomes and analysis of this study are given in Section 3. This work is summarized in Section 4 with potential future research.

Methods and materials

A. Hyperparameter Optimization

The goal of hyperparameter optimization in regression models is to find the parameters of a given regression technique, which returns the best output on a validation set while training and testing the model 15. This is as shown in equation (1)

$$H^* = \arg \min_{o \in O} f(o) \quad (1)$$

Where $f(O)$ is an objective score to minimize RMSE calculated on the validation dataset; H^* is the set of hyperparameters that gives the lowest RMSE score, and o is any value in the problem domain O .

Even though hyperparameter optimization is costly in terms of computational time, it yields good prediction accuracy than traditional regression models.

B. Gaussian Process Regression

Gaussian process is a machine learning technique used to make uncertain predictions. It is defined as a finite set of random variables distributed jointly by the Gaussians 15. These random variables represent the value for a function $f(x)$ at input x in regression problems. It is represented as $\{f(x) : x \in X\}$ defined by the mean function $\mu(x)$ and the covariance function $k(x, x')$ so it can be represented as

$$f(\cdot) \sim GP(\mu(\cdot), k(\cdot, \cdot)) \quad (2)$$

Usually, a zero-mean Gaussian process before equation (2) is the prior distribution over functions $f(\cdot)$ Same

detailed explanation found in 16. This model is named, in literature, a substitute for the objective function. The substitute can be configured more easily than the intent feature. A GP method determines the next set of hyperparameters to test by choosing the best hyperparameters acting on this surrogate function on the actual objective function.

A covariance function as defined in equation (3) is used to represent the covariance between pairs of random variables in GPR.

$$K_{ij} = k(x_i, x_j) = \alpha \exp\left\{-\frac{\|x_i - x_j\|^2}{2\sigma_1^2}\right\} \quad (3)$$

Here, hyperparameters are
 $\sigma_1 = \text{Characteristic lengthscale}$
 $\alpha = \text{Signal variance}$

GPR model-based hyperparameter optimization

1. Initialize hyperparameters for GPR model
2. Define an objective function of GPR model
3. Specify the selection criteria (ie. Minimal RMSE) for evaluating hyperparameters which have to choose next from the surrogate model
4. Update the surrogate model by (score, hyperparameter) pair
5. Repeat steps 2–4 until maximum iterations or time is reached.

C. GRNN

A special case of Radial Basis Networks (RBN) is the Generalized Regression Neural Network (GRNN) 28. The structure of a GRNN with two layers is comparatively simple and fixed. The first is the sequence, and the second is a summation. When the input is passed through each unit in the pattern layer, the input-response relationship will be "memorized" and stored within the unit. As a result, the number of units in the pattern layer equals the number of individual values in the training set. In each pattern unit, a Gaussian PDF is applied to the network input, so that it is defined as equation (4)

$$\text{Theta} = \text{EXP}[-0.5 * (A - t) \cdot (A - t) / (\text{Sigma}^2)] \quad (4)$$

where Theta is the output of the Pattern Unit, A is the origin, t is the vector of training stored in the unit, and Sigma is a positive variable known as the "distance" or "smooth parameter" or —smoothing factor—. If Theta is determined, the calculation is transferred to the summation layer $P = \text{SUM}(P \cdot \text{Theta}) / \text{SUM}(\text{Theta})$ where P is the conditional prediction of P and Q is the solution in the sample of training.

D. RBFNN

RBFNN 30 is an artificial neural network that uses radial functions as the activation functions shown in equation (5). RBFNN is a three-layer neural network of

feed-forwards. The first layer is linear, transmitting only the input signal, while the next layer is nonlinear, using Gaussian functions. The third layer integrates a linear representation of the Gaussian outputs. Only the tap weights between the hidden layer and the output layer shift during preparation.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (5)$$

E. Nonlinear Autoregressive Neural Network

The Nonlinear autoregressive neural network is a type of ANN that is suitable for estimating future input variable values. The NAR Network helps to forecast future values of the time. It supported the use of a re-feeding mechanism through its historical precedent, in which a predicted value would serve as feedback for new predictions at more advanced points in time. In equation (6) represents as predict series y(t) given d past values of y(t).

$$y(t) = f(y(t-1), \dots, y(t-d)) \quad (6)$$

F. Root Mean Squared Error (RMSE)

RMSE is the square root of the square differences measured between predicted and actual COVID-19 Death cases 8. It represents as in equation (7)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{Predicted Death Case}_i - \text{Actual Death Case}_i)^2} \quad (7)$$

Where, n= number of samples, $P_i = i$ th Predicted value, $A_i = i$ th Actual value

G. Correlation coefficient (R2)

It measures a linear relationship between the predicted and actual COVID-19 death cases. It represents as in equation (8)

$$R = \frac{\sum_{i=1}^n (a_i - \bar{a}_i)(b_i - \bar{b}_i)}{\sqrt{\sum_{i=1}^n (a_i - \bar{a}_i)^2} \sqrt{\sum_{i=1}^n (b_i - \bar{b}_i)^2}} \quad (8)$$

Where a_i is the actual COVID-19 death case value, b_i is the predicted COVID-19 death case value; \bar{a} is the mean of real COVID-19 death case value, \bar{b} is the predicted mean COVID-19 death cases value, and n is the total number of data points.

Proposed methodology

Dataset Description

A dataset has been downloaded from the Kaggle website (www.kaggle.com) for predicting the COVID-19 death cases for India. This dataset contains India's COVID-19 Confirmed cases and Death cases from January 20, 2020, to April 30, 2020, which is used for training and testing models.

First, these data are pre-processed to eliminate missing values and inappropriate values. These data can be used to create two types of datasets. They are:

- Dataset 1 contains two attributes, such as COVID-19 confirmed cases and death cases. Here, 'death case' is a predictive attribute and 'confirmed case' is a response attribute.
- Dataset 2 contains a time series of COVID-19 death cases.

In this paper, three models (such as GPR, GRNN, and RBFNN) are constructed with the appropriate model

parameter values and used in these two datasets to validate the predicted results concerning given the available datasets.

Figure 1 illustrates the proposed methodology. In general, residues or errors are an inevitable part of any predictive or regression models. Similarly, there are errors in the GPR, GRNN, and RBFNN models. To provide a predictive model with high accuracy, this study explores a hybrid approach, including regression methods and the non-linear auto-regression (NAR) neural network (NAR-NN) time series forecasting model. Therefore, trends in residues or errors are detected and predicted by the NAR-NN model. Combining the predicted residual values of each model with the predicted value of each model will provide greater predictive accuracy. The following steps are used to develop the hybrid model:

Step 1: Download data from the website and pre-process the dataset. Create Dataset 1(D1) and Dataset2(D2).

Step 2: Initialize Model Parameters and define hyperparameters for GPR, GRNN, and RBFNN.

Step 3: Input datasets D1 and D2 into the GPR model, GRNN model, and RBFNN model respectively, and predict COVID-19 death cases (Prednew) for 'n' period ahead or for given set of confirmed cases.

Step 4: The residuals produced by these models are extracted and converted into time-series data.

Step 5: Input these residuals into the NAR-NN time series forecasting model and predict the residual values (Ferr) for three models separately. It is shown graphically in figure 3.

Step 6: Ferr is added with PredNew to generate an optimized prediction value.

Step 7: Return optimized predicted values as output

Table 1: Model Parameters Setup

Model Parameters	GPR Model	GRNN Model	RBFNN Model	NAR-NN Model
Hidden Layer (HL)	N/A	Fixed Architecture	Fixed Architecture	Fixed Architecture
Number of Neurons in HL	N/A	10	10	15
Training Algorithm	N/A	Bayesian Regularization	Bayesian Regularization	Bayesian Regularization
Kernel Functions	Exponential	N/A	N/A	N/A
Hyperparameter	Covariance (σ) Mean (μ)	Spread/ Smoothing factor (σ)	Spread/ Smoothing factor (σ)	N/A
Performance	RMSE	RMSE	RMSE	MSE

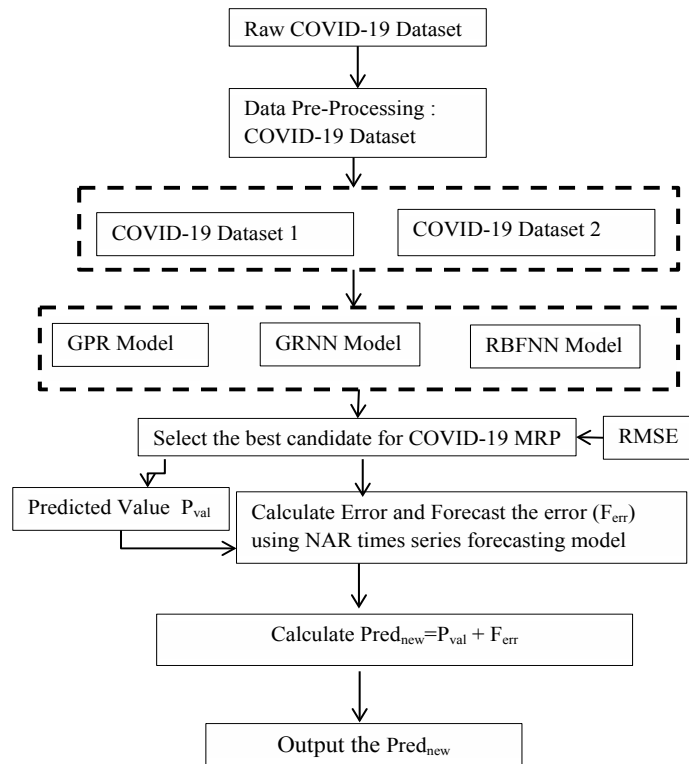


Figure 1: Proposed Methodology for COVID-19 MRP Model

Table 1 shows the parameter setup for all three models. Figure 2 illustrates the operating principles of the hyperparameter tuning used with the GPR model, the

GRNN model, and the RBFNN model to achieve the minimum RMSE value for these models, respectively. Similarly, figure 3 describes the working principle of the NAR-NN model for error forecasting for these models.

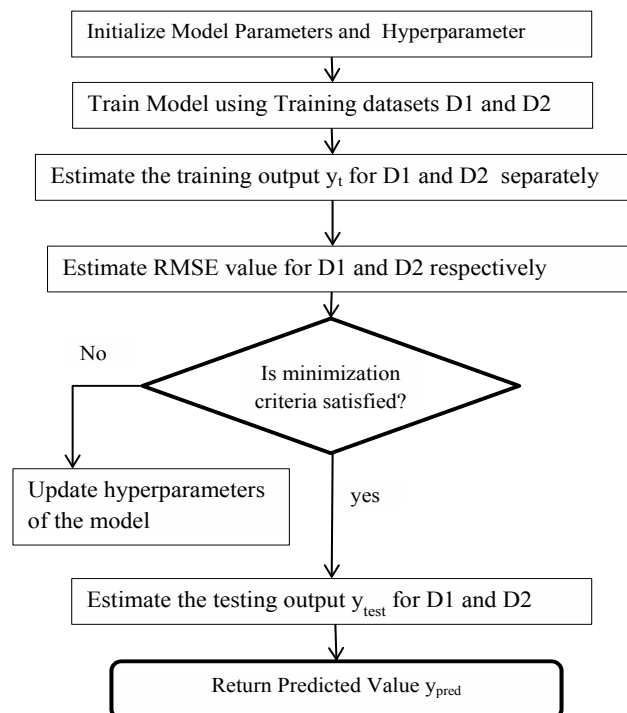


Figure 2: Workflow of Model with Hyperparameter Tuning

Outcomes and Discussions

This section contains results of Gaussian Process Regression, and two different statistical neural networks (SNNs): GRNN and RBFNN models for dataset1 and

dataset² are presented and discussed. The performance of these models was compared.

The benchmark performance metrics such as Root Mean Squared Error (RMSE), an error measure, and

Correlation coefficient accuracy measure (R) are used to estimate the COVID-19 death cases prediction accuracy. By using Hyperparameter tuning, the predictive efficiency of these three models can be improved. The goal of hyperparameter tuning is to optimize the value of hyperparameters of each model to minimize the RMSE value of these models. Based on RMSE value, the best model is selected for COVID-19 death cases prediction and mortality rate prediction.

Table 2 shows the values of the performance metrics such as RMSE and R2 for three models. While comparing these models, it is seen that the GPR model has low RMSE value and high R2 value for both datasets. It signifies that the GPR model performs better than two statistical neural networks for the COVID-19 dataset. The spread value (σ) of GRNN and RBFNN is 4 and 1.76, respectively. Table 3 and Table 4 display the values of hyperparameters for statistical neural networks and Gaussian Process Regression models.

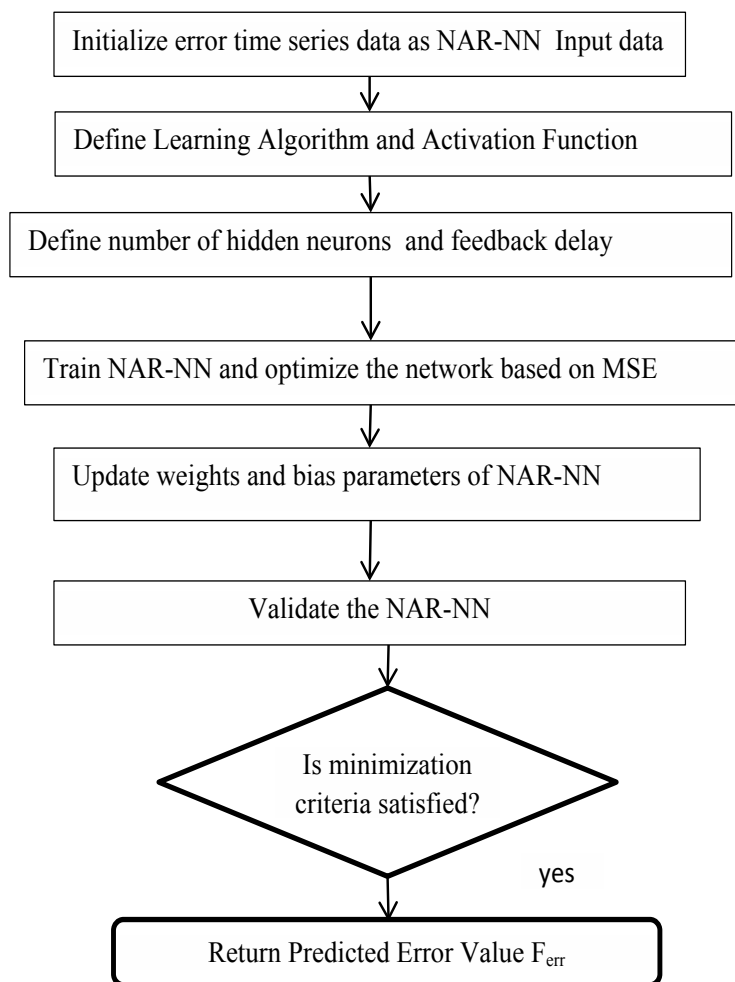


Figure 3: Workflow of NAR-NN Time Series Forecasting

Table 2: Performance Metrics for Datasets

Model	RMSE_D1	RMSE_D2	R ² _D1	R ² _D2
GRNN	2.44605585	3.249880129	0.999938	0.9999646
RBFNN	2.959403436	3.264847469	0.999937	0.999954
GPR	0.177218356	0.16828063	1	0.9999998

Table 3: Hyperparameter Value for GRNN and RBFNN (Spread)

Dataset	GRNN	RBFNN
D1		2
D2		4

Table 4: Hyperparameter Value for GPR

Dataset	Sigma M	Sigma F	Sigma
D1	6863.8	255	6.8
D2	1008.5	79.1	2.6

Table 5: Predicted Value Y_{pred} for Dataset (D2) using Standard Models

Date	GPR	GRNN	RBFNN
01-May-20	1152	1234	1234
02-May-20	1152	1317	1317
03-May-20	1151	1403	1402
04-May-20	1151	1490	1489
05-May-20	1150	1579	1576
06-May-20	1150	1667	1664
07-May-20	1149	1754	1750
08-May-20	1149	1840	1835
09-May-20	1148	1922	1918
10-May-20	1148	2002	1997
11-May-20	1147	2078	2074
12-May-20	1147	2150	2147
13-May-20	1146	2218	2216
14-May-20	1145	2281	2281
15-May-20	1145	2340	2341
16-May-20	1144	2395	2398
17-May-20	1144	2445	2450
18-May-20	1143	2491	2498
19-May-20	1143	2533	2543
20-May-20	1142	2571	2583

Tables 5 and 6 show the predicted number of COVID-19 death cases for all three standard and hybrid models using the time series dataset (i.e. D2). A hybrid model is the combination of standard models and NAR-NN based error forecasting models. There is no difference in the predicted values for the standard and

hybrid GPR model since their RMSE value is about 0.2 approximately. Figure 4 shows the number of predicted death cases versus days from May 1, 2020, to May 20, 2020. From this point of view, the two SNNs have nearly equal efficiency, but their RMSE values are higher than that of the GPR models.

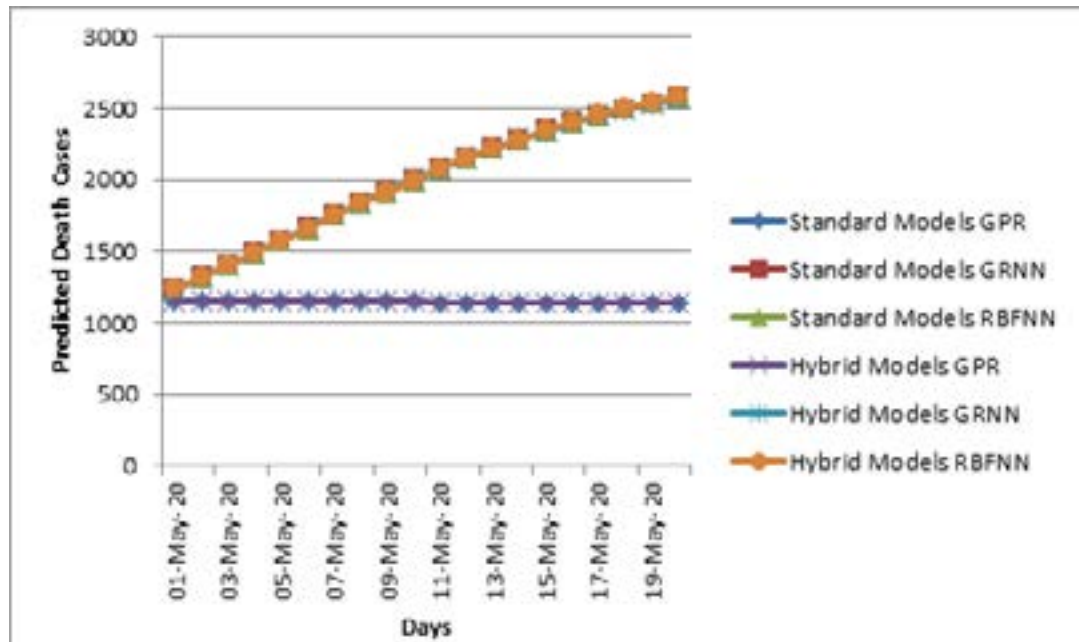


Figure 4: Comparison of Standard and Hybrid Models for D1

Table 6: Predicted Value Y_{pred} for Dataset (D2) using Hybrid Models

Date	GPR	GRNN	RBFNN
01-May-20	1152	1233	1233
02-May-20	1152	1317	1317
03-May-20	1151	1404	1404
04-May-20	1151	1490	1488
05-May-20	1150	1577	1574
06-May-20	1150	1667	1664
07-May-20	1149	1757	1751
08-May-20	1149	1839	1834
09-May-20	1148	1920	1916
10-May-20	1148	2003	1999
11-May-20	1147	2084	2076
12-May-20	1147	2148	2145
13-May-20	1146	2220	2213
14-May-20	1146	2286	2282
15-May-20	1145	2338	2340
16-May-20	1144	2395	2396
17-May-20	1144	2451	2452
18-May-20	1143	2492	2500
19-May-20	1143	2534	2540
20-May-20	1142	2570	2580

Table 7 shows the predicted number of COVID-19 death cases using the dataset (D1) for three standard and hybrid models. The spread value (σ) of GRNN and RBFNN is 2 and 1.28, respectively.

From the presented table 7, it is evident that there is no difference in the predicted values for standard and hybrid GPR models using the dataset (D1) since their RMSE value is about 0.18 approximately. Figure 5 shows the number of death cases predicted against the number of COVID-19 confirmed cases. From this

perspective, there is no significant difference between the standard and hybrid models for the dataset (D1), as their RMSE values are negligible. However, in this case, the GPR models perform better than the SNNs. Table 8 shows the calculated Mortality Rate Prediction (MRP) for COVID-19 predicted death cases using the dataset (D1). MRP is defined as equation (9)

$$MRP = \frac{\text{Number of Predicted Death Cases}}{\text{Number of Confirmed Cases}} \times 100 \quad (9)$$

Table 7: Predicted Death Cases for Dataset (D1)

Confirmed cases	Standard Models			Hybrid Models		
	GPR	GRNN	RBFNN	GPR	GRNN	RBFNN
40000	1149	1334	1449	1149	1333	1449
45000	1145	1458	1659	1148	1458	1665
50000	1142	1536	1787	1141	1534	1787
55000	1138	1582	1860	1133	1583	1860
60000	1134	1610	1898	1133	1609	1898
65000	1130	1628	1919	1130	1628	1918
70000	1126	1639	1929	1125	1638	1929
75000	1122	1647	1935	1122	1647	1935
80000	1118	1654	1938	1116	1653	1937
85000	1115	1660	1939	1114	1659	1939
90000	1111	1665	1940	1109	1663	1940
95000	1107	1670	1940	1106	1669	1940
100000	1103	1674	1940	1101	1673	1940

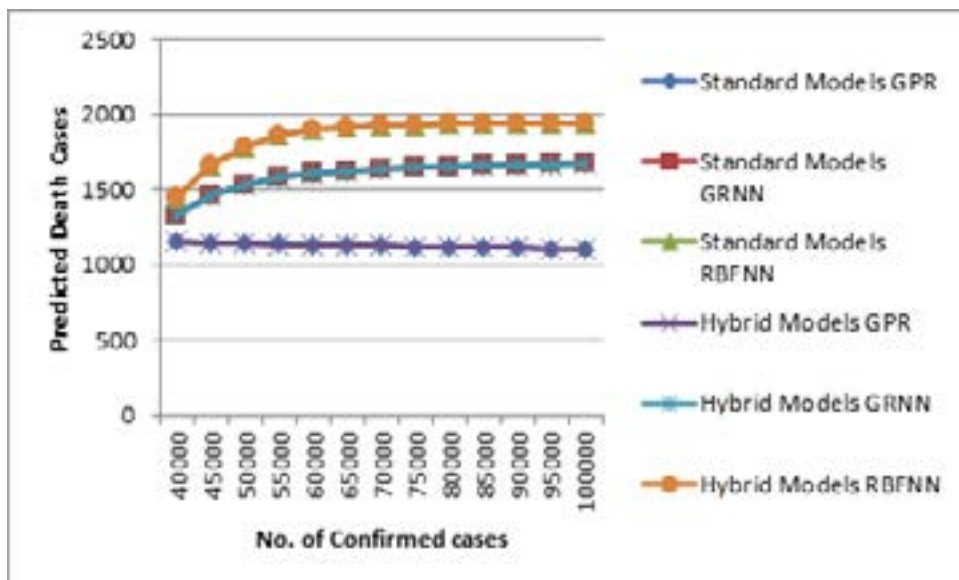


Figure 5: Comparison of Standard and Hybrid Models for D1

Table 8: Mortality Rate Prediction for Dataset (D1)

Confirmed cases	Standard Models			Hybrid Models		
	GPR	GRNN	RBFNN	GPR	GRNN	RBFNN
40000	2.87	3.34	3.62	2.87	3.33	3.62
45000	2.54	3.24	3.69	2.55	3.24	3.70
50000	2.28	3.07	3.57	2.28	3.07	3.57
55000	2.07	2.88	3.38	2.06	2.88	3.38
60000	1.89	2.68	3.16	1.89	2.68	3.16
65000	1.74	2.50	2.95	1.74	2.50	2.95
70000	1.61	2.34	2.76	1.61	2.34	2.76
75000	1.50	2.20	2.58	1.50	2.20	2.58
80000	1.40	2.07	2.42	1.40	2.07	2.42
85000	1.31	1.95	2.28	1.31	1.95	2.28
90000	1.23	1.85	2.16	1.23	1.85	2.16
95000	1.17	1.76	2.04	1.16	1.76	2.04
100000	1.10	1.67	1.94	1.10	1.67	1.94

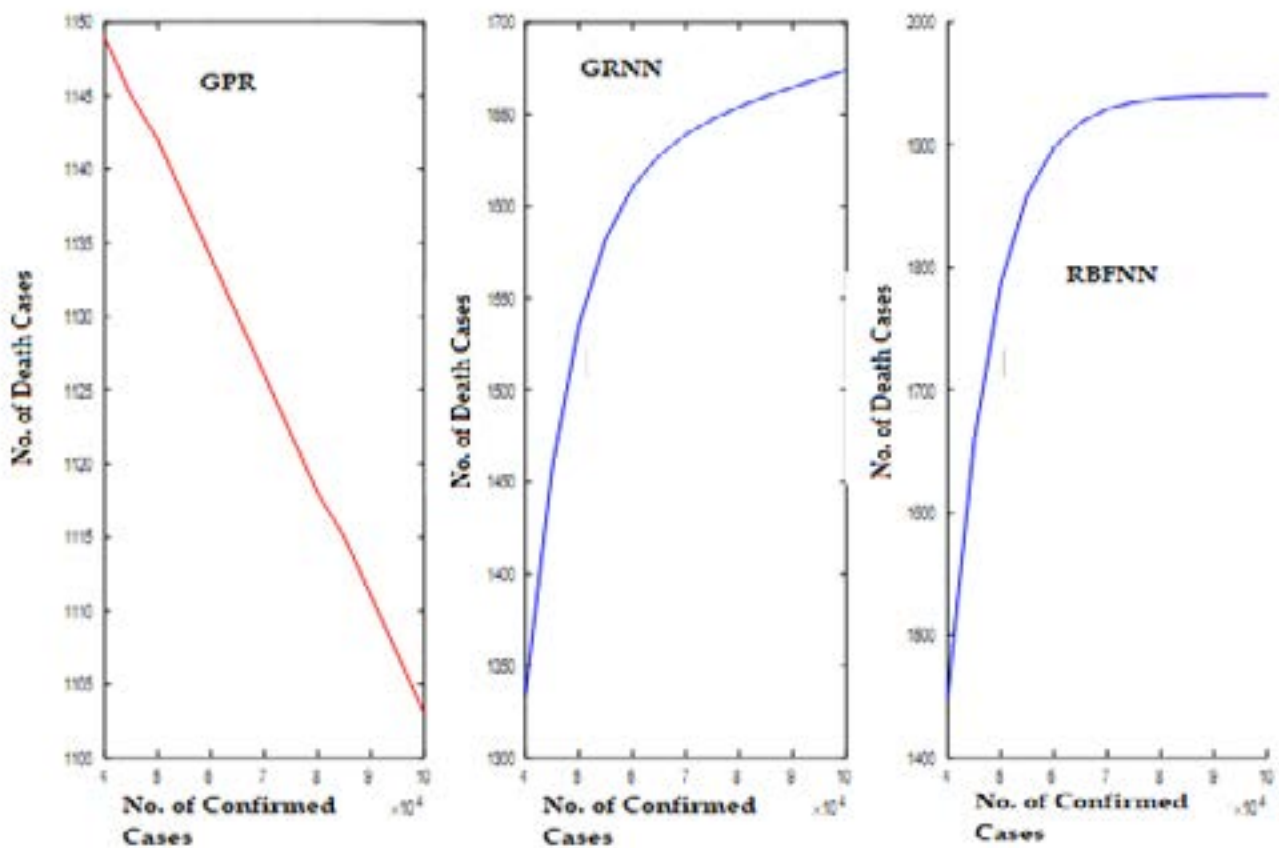


Figure 6: Predicted Curve for Dataset (D1) using GPR, GRNN, and RBFNN

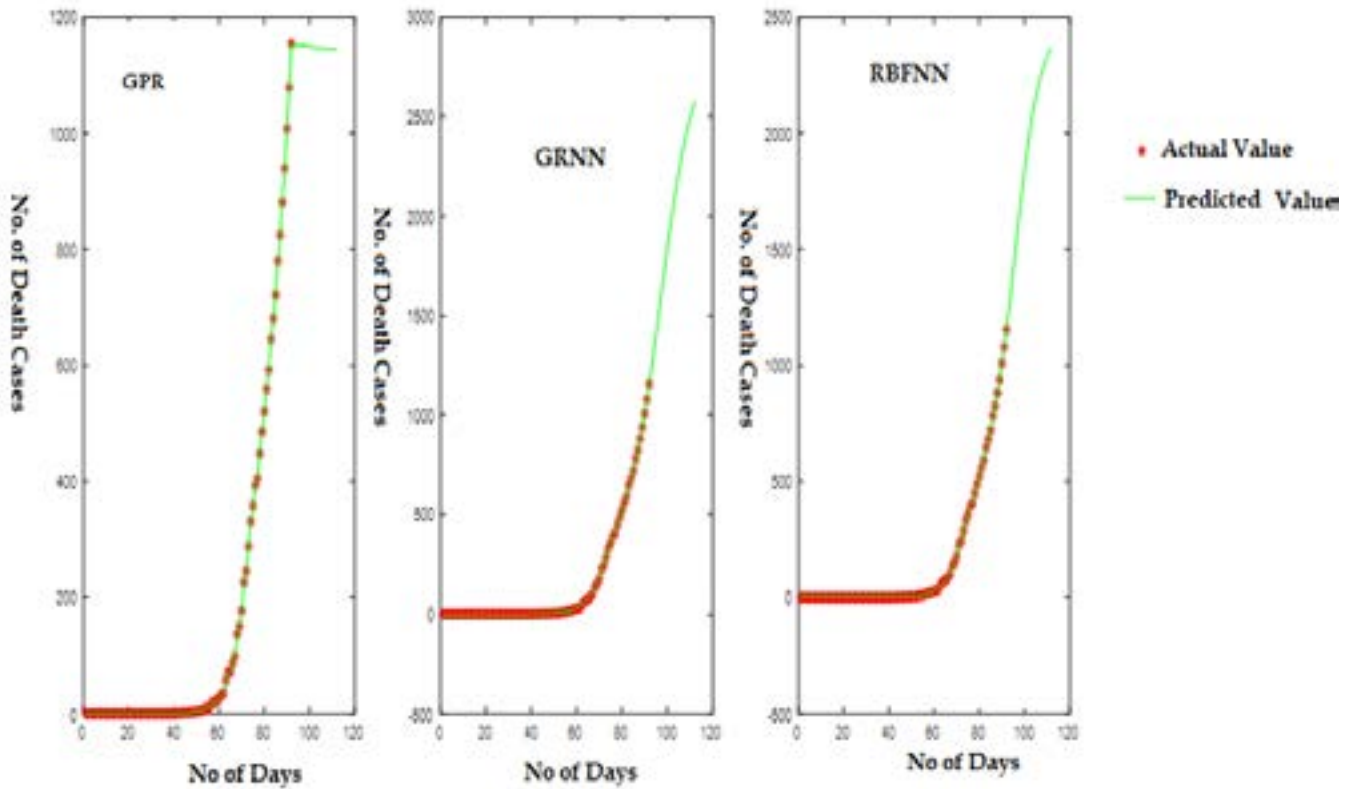


Figure 7: Predicted Curve for Dataset (D2) using GPR, GRNN, and RBFNN

Figure 6 illustrates the predicted curve for COVID-19 death cases versus the number of confirmed cases for India. Here, the X-axis indicates the number of confirmed cases, and Y-axis shows the number of death cases predicted. The GPR model shows a gradual decrease in the number of death cases, while the SNN models show an increasing pattern.

Figure 7 displays the predicted curve for COVID-19 death cases versus the number of days since the first COVID-19 case for India.

Here, the X-axis indicates the number of days, and Y-axis shows the number of death cases predicted. For the dataset (D2) also, the GPR model shows a gradual decrease in the number of death cases, while the SNN models show an increasing pattern.

The performance of these models is compared based on RMSE value, as shown in figure 8. From the results, it can be found that GPR performs better than statistical neural networks for two types of COVID-19 dataset.

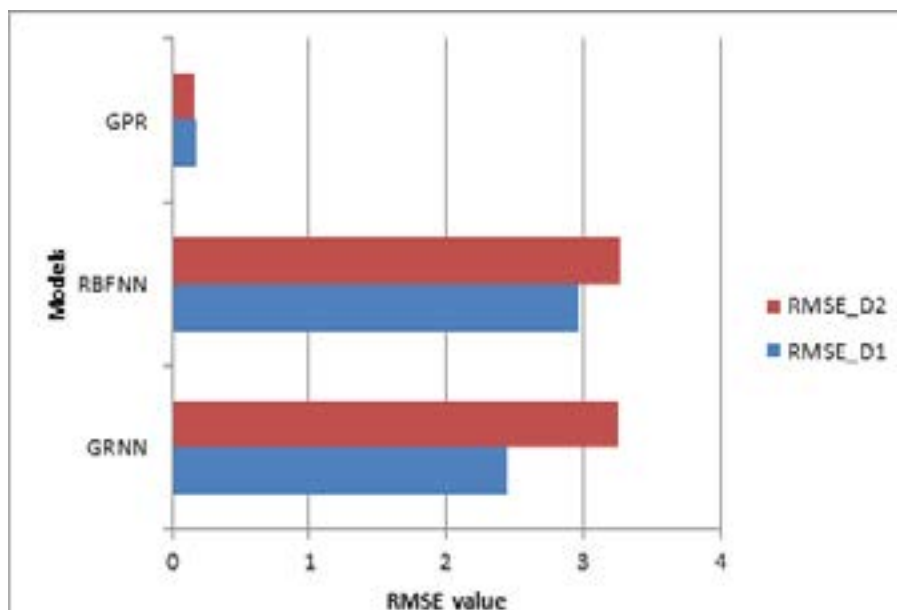


Figure 8: Comparison of RMSE values with a different model

Conclusion

This article proposed a methodology that hybridized the regression model (GPR) and the SSN (GRNN and RBFNN) models with the NAR-NN time series forecasting model to achieve higher predictive accuracy in the prediction of COVID-19 death cases. The NAR-NN time series forecasting model was used to predict errors that should be included in the expected value. The Gaussian Process Regression (GPR) model for the two datasets has yielded a relatively good result in terms of optimized predicted values for death cases in the COVID-19 epidemiological data.

The proposed method is capable of providing a predictive tool for assessing its current state of infection, severity, and help government and health care workers for better decision making to reduce the mortality rate in India.

The first author acknowledges the —UGC- Special Assistance Programme (SAP) for the financial support to her research under the —UGC-SAP at the level of DRS-II (Ref.No.F.5-6/2018/DRS-II (SAP-II)), 26 July 2018 in the Department of Computer Science, Periyar University.

Conflict of interest

The authors have no conflict of interest.

References

1. WHO | Novel Coronavirus – China. WHO. Archived from the original on 23 January 2020. Retrieved 1 February 2020.
2. Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). World Health Organization (WHO). 30 January 2020. Archived from the original on 31 January 2020. Retrieved 30 January 2020.
3. Sparrow Annie. How China's Coronavirus Is Spreading and How to Stop It, Foreign Policy. Archived from the original on 31 January 2020. Retrieved 2 February 2020.
4. S. Dhamodharavadhani and R. Rathipriya, "Novel COVID-19 Mortality Rate Prediction (MRP) Model for India Using Regression Model With Optimized Hyperparameter," *Journal of Cases on Information Technology*, vol. 23, no. 4, pp. 1–12, 2021.
5. R. Devipriya, S. Dhamodharavadhani, and S. Selvi., "SEIR model FOR COVID-19 Epidemic using DELAY differential equation," *Journal of Physics: Conference Series*, vol. 1767, no. 1, p. 012005, 2021.
6. Dhamodharavadhani S, Rathipriya R., Chatterjee JM. "COVID-19 Mortality Rate Prediction for India Using

Statistical Neural Network Models," *Frontiers in Public Health*, vol. 8, 2020.

7. Deprez Philippe, Pavel V Shevchenko, Mario V Wüthrich. Machine learning techniques for mortality modeling, *European Actuarial Journal* 2017, 7: 337–52.
8. Hainaut Donatien. A neural-network analyzer for mortality forecast, *Astin Bulletin*, 2018, 48: 481–508.
9. Richman Ronald, Mario V Wüthrich. A Neural Network Extension of the Lee-Carter Model to Multiple Populations", Rochester: SSRN, 2018.
10. Razavi S, Tolson BA, Burn DH. *Review of surrogate modeling in water resources*, *Water Resources Research*, 2012;48(7) doi: 10.1029/2011wr011527.
11. Raghavendra NS, Deka PC. Multistep ahead groundwater level time-series forecasting using Gaussian Process Regression and ANFIS, *Advances in Intelligent Systems and Computing*. Vol. 396. New Delhi, India: Springer; 2016.
12. S. Sivabalan, S. Dhamodharavadhani, and R. Rathipriya, "Arbitrary walk with minimum length based route identification scheme in graph structure for opportunistic wireless sensor network," *Swarm Intelligence for Resource Management in Internet of Things*, pp. 47–63, 2020.
13. Roy DK., Datta B. Trained meta-models and evolutionary algorithm based multi-objective management of coastal aquifers under parameter uncertainty, *Journal of Hydroinformatics*, 2018.
14. Ebden M. Gaussian Processes for Regression: A Quick Introduction, August 2008
15. Hong S, Zhou Z. Application of Gaussian process regression for bearing degradation assessment, *Information Science and Service Science and Data Mining (ISSDM)*, 2012 Sixth International Conference on New Trends in, 2012.
16. S. Sivabalan, S. Dhamodharavadhani, and R. Rathipriya, "Opportunistic Forward Routing Using Bee Colony Optimization," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 5, pp. 1820–1827, 2019.
17. Rasmussen CE. Gaussian processes in machine learning", *Adv. Lect. Mach. Learn.* 14 (2), 2004
18. Ažman K, Kocijan J. Dynamical systems identification using Gaussian process models with incorporated local models. *Eng. Appl. Artif. Intell.* 24 (2), 2011.
19. Huber M. Recursive Gaussian process: on-line regression and learning. *Pattern Recognit. Lett*, 2014
20. Rasmussen CE, Williams CKI. *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, 2006.
21. Saha S, Saha B, SaxenAA, Goebel K. *Distributed prognostic health management with gaussian process regression*, 2010 IEEE Aerosp. Conf., pp. 1–8, Mar. 2010.

22. Ažman K, Kocijan J. Dynamical systems identification using Gaussian process models with incorporated local models. *Eng. Appl. Artif. Intell.* 24 (2), 2011.
23. V. Kaleeswaran, S. Dhamodharavadhani, and R. Rathipriya, "A Comparative Study of Activation Functions and Training Algorithm of NAR Neural Network for Crop Prediction," 2020 4th *International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2020.
24. V. Kaleeswaran, S. Dhamodharavadhani, and R. Rathipriya, "Multi-crop Selection Model Using Binary Particle Swarm Optimization," *Innovative Data Communication Technologies and Application*, pp. 57–68, 2021.
25. Huber M. Recursive Gaussian process: on-line regression and learning", *Pattern Recognit. Lett*, 2014
26. Dhamodharavadhani S, Rathipriya R. "A pilot study on climate data analysis tools and software," 2016 *Online International Conference on Green Engineering and Technologies (IC-GET)*, 2016.
27. Rasmussen CE, Williams CKI. Gaussian Processes for Machine Learning, MIT Press, Cambridge, 2006.
28. Huber MF. Recursive Gaussian process regression, IEEE Int. Conf. Acoust. Speech Signal Process, 2013
29. Wang JM, Fleet DJ, Hertzmann A. Gaussian process dynamical models for the human motion, IEEE Trans. *Pattern Anal. Mach. Intell.* 30 (2), 2008
30. Seyedzadeh S, Rahman F, Glesk I. *Machine learning for estimation of building energy consumption and performance: a review, Vis. in Eng.*, 2018.
31. Ahmad Jobran AL-Mahasneh, Sreenatha G Anavatti, Matthew A. Garratt. Review of Applications of Generalized Regression Neural Networks in Identification and Control of Dynamic Systems. *Neural and Evolutionary Computing*, 2018
32. Dhamodharavadhani S, Rathipriya R. "Region-Wise Rainfall Prediction Using MapReduce-Based Exponential Smoothing Techniques," *Advances in Intelligent Systems and Computing*, pp. 229–239, 2018.
33. Sharma N, OM H. Usage of Probabilistic and General Regression Neural Network for Early Detection and Prevention of Oral Cancer. *The Scientific World Journal*, 2015.
34. Han Q, Su H, Wang CC, Shan XW, Chang WW, Xu ZW. Prediction on the incidence of blood and sexually transmitted diseases with models of ARIMA and GRNN. *Modern Preventive Medicine*, 2012(6):1337–40.
35. Solomon Asante-Okyere, Chuanbo Shen, Yao Yevenyo Ziggah, Mercy Moses Rulegeya, Xiangfeng Zhu. Investigating the Predictive Performance of Gaussian Process Regression in Evaluating Reservoir Porosity and Permeability. *Energies*. 2018, 11(12), 3261; <https://doi.org/10.3390/en11123261>
36. Ahmed Nesreen K, Atiya Amir F, Gayar Neamat El, El-Shishiny Hisham. An Empirical Comparison of Machine Learning Models for Time Series Forecasting", *Econometric Reviews*, 29: 5, 2010, 594 — 621
37. Abdul Shaikh, Abdul Hannan, Ramesh Manza, Rakesh Ramteke. Generalized Regression Neural Network and Radial Basis Function for Heart Disease Diagnosis. *International Journal of Computer Applications* 13(13), 2010
38. S. Dhamodharavadhani and R. Rathipriya, "Variable Selection Method for Regression Models Using Computational Intelligence Techniques," *Research Anthology on Multi-Industry Uses of Genetic Programming and Algorithms*, pp. 742–761, 2021.
39. Johnson LR. *Phenomenological forecasting of disease incidence using heteroskedastic Gaussian processes: A dengue case study*, arXiv:1702.00261, 2017.
40. Gholam Ali Montazer, Davar Giveki, Maryam Karami, Homayon Rastegar. Radial Basis Function Neural Networks: A Review", *Computer Reviews Journal*. Vol. 1 No. 1 (2018): Advanced Computing.
41. S. Dhamodharavadhani and R. Rathipriya, "Enhanced Logistic Regression (ELR) Model for Big Data," *Handbook of Research on Big Data Clustering and Machine Learning*, pp. 152–176, 2020.
42. Francisco Martínez, Charle F, Antonio Jesús Rivera Rivas, María P Frías. Automatic Time Series Forecasting with GRNN: A Comparison with Other Models. *Advances in Computational Intelligence*, 2019, 10.1007/978-3-030-20521-8_17