Appraisal of Predictive Techniques using Computational Methods

Ozoh Patrick[1*], Olayiwola Moruf[2], Adegoke Adebisi[2]

[1]Department of Information and Communication Technology, Osun State University, Osogbo, Nigeria
[2]Department of Mathematical Sciences, Osun State University, Osogbo, Nigeria
[*]corresponding author: patrick.ozoh@uniosun.edu.ng
Received: 15 November 2021; Accepted: 29 December 2021

Abstract
This research focuses on theorems governing computational methods. Computational methods are a group of techniques that acts on big content with data to get outcomes with the prospect of obtaining a decision-making process. The computational techniques that are considered in this study include the neural network, Kalman filter adaptation algorithm, Box-Jenkins technique, and regression analysis. The theorems considered in this research are mathematical methods. The predictive techniques covered consist of mathematical analysis of solving problems, which are computational methods. These theorems encompass predictive values, which come about being considered. These methods are based on these computational methods. The major objective of this research is to utilize quantitative techniques in solving complex challenges. The techniques are enumerated in this research. The evaluation of these methods is also investigated. The Kalman filter adaptation technique is the most accurate when compared with the neural network, Box-Jenkins technique, and the regression analysis when tested on data. This will enable a reliable method to be applied for modeling data and for making accurate and reliable decision making.

Keywords: Predictive models, Computational methods, Behavioral pattern, Machine learning, Evaluation.

Introduction
Prediction models can be applied in various ways. Kang *et al.* (2021) utilized different ways to predict and identify the reliability of techniques applied to modeling. These include data processing, data identification, feature processing, creation of models, preventing model oversizing, and identification of reliability models, together with artificial neural network techniques. According to Chancellor & De Choudhury (2020), machine learning techniques are applied to predict mental disorders, depression, suicidal rates, and levels of anxiety. As a result of limitations inherent in quantitative techniques, modeling of various disorders is a big issue (Chen, 2020). The preparedness with analytic examination concerning the procedure of the system is the base of the approach. Successes achieved in techniques, structure, and findings from the research are evaluated in the research. The evaluation of the research come up with requisitions and issues concerning the research (Chen *et al*., 2020).

For modeling data, Ozoh *et al*. (2014) exhibited a previous technique for modeling was by applying various computational methods. A major challenge to using these techniques is results are less than accurate. The minimum variation error is lowered as presented by Wu *et al*. (2021). A study of predictions of bioinformatics and production of drugs was introduced, and the study was divided into groups. This includes 3D structure-based system, similarity-based system, machine learning-based method, and deep learning-based method. (Zhao *et al*., 2020). Up-to-date procedures found within computational methods, together with enlarged computational activities in recent times have led to more procedures within computational methods. (Hewing *et al*., 2020). A certain objective of the study is to summarize previous works on computational methods. Several computational methods go through lower time, together with cost for systems. Altogether, different blueprints containing different techniques are enumerated. A mixture comprising various techniques are used to resolve challenges with problems (Lin *et al*., 2020)

In Agumah *et al.* (2020), different prediction techniques were developed for various applications. The study proposes different methods utilized in predictions and identifying accuracy in models. It gives a thorough study of their dominance and compares these techniques to select an accurate technique that will result in inefficient results. Achievements obtained in developing machine learning models, and the abundance of computational techniques have resulted in more interest in computational methods. Machine learning provides a big chance to discover plenty of data in error-free ways (Hewing *et al*., 2020). The main function of the research categorizing previous studies with computational methods. By introducing computational techniques, error values for financial transactions can be reduced (Kumar *et al*., 2020). The research dwells on utilizing computational techniques. These include artificial neural networks, fuzzy logic, etc for predicting stock market fluctuations.

State-of-the-art techniques provide various methods for computational analysis. Together with analytical techniques, these data unlock new techniques for reliable data. The paper encapsulates computational methods for analysis and discusses their applications, problems, and future control of the process (Efremova & Teichmann, 2020). Computer simulation techniques allow the description of important characteristics and changes between them, excluding sampling without significant computational features. In the past, a group of techniques was applied to sampling, defined as collective variables (CVs), with which sampling can be utilized. This research gives a presentation of simulation and the characteristics of CVs with their techniques, together with current achievements with techniques in computational methods (Allison, 2020).

The contributions of this research are summarized as follows:

1. This review carries out research and chooses a reliable technique to model and predict data.
2. This study recommends an efficient method to predict data known as the Kalman filter adaptation algorithm, which offers reliable modeling performance in relation to artificial neural networks, Box-Jenkins, and regression analysis.
3. This study confirms that the Kalman filter algorithm has the most reliable statistical

estimates than an artificial neural network, Box-Jenkins, and regression analysis.
This study is composed of the following:
Section 1 proposes modeling and predictions, and an investigation of the accomplishment of artificial neural network, Box-Jenkins, regression analysis and the Kalman filter adaptation algorithm. The research contributions to this study are also discussed.
Section 2 introduces techniques applied in this study. It also gives the concepts of mathematical models used.
Section 3 concerns evaluations of methods used and a description of results from the study.

Methodology

This research explores the concepts, theories, models, and data collection for machine learning techniques. The rules and guidelines for the methods are introduced. The models for covering the methodologies are also represented in the following sections. The computational techniques are thoroughly explained in detail.

Neural Network

With giant achievements in computing, the neural network has been utilized to solve complex problems (Yang *et al*., 2021). The research proposes a critical study of current processes in the neural network. This research investigates recent papers, for the most part, research administered over the last three decades. The neural network systems used in the research include multi-layer perceptron neural network (MLPNN), convolutional neural network (CNN), and recurrent neural network (RNN). Chaabouni & Baroni (2021) explore this computational technique trained with deep-learning techniques, which develop communication systems. The recognized changes is described by various requirements. The uncertainties involved in the models are detected in the obtained models. The study introduces this technique to obtain the variables (Khoo *et al*. (2021). Baek *et al*. (2021) proposed reliable predictions at a conference. The research investigated network systems containing inter-connected theorems, and evaluated the system to result in a transformed and well-integrated system. The system enabled the fast infusion of modeling problems, and on condition that they get meaning to the consequence of the system.

A neural network as discussed by Khoo *et al.* (2021) is given as follows:

$$\begin{aligned} y_1 &= f_1(x_t, x_{t-1}, \ldots, x_{t-n}) \\ y_2 &= f_2(x_t, x_{t-1}, \ldots, x_{t-n}) \\ &. \\ &. \\ &. \\ y_t &= f_t(x_t, x_{t-1}, \ldots, x_{t-n}) \end{aligned} \tag{1}$$

$f_1, f_2, \ldots$ are models. The relationship pattern $f_1, f_2, \ldots$ is given as follows:

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j (\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} y_{t-1}) = \varepsilon_t \qquad (2)$$

$\alpha_j (j = 0,1,...)$, $\beta_{ij} (i = 0,1,..., p; j = 0,1,...)$ are variables

Kalman Filter Adaptation Algorithm

According to Alfian *et al.* (2021), the Kalman filter has the capacity to capture errors. The study utilizes this technique to lower errors. If the error is big, then accuracy becomes reduced (Alfian *et al.*, 2021). Although when there are outliers and error deviations, the reliability obtained models cannot be confirmed. (Lai *et al.*, 2921) propose a narrative approach by incorporating model-based and data-driven techniques depending on the sequential extended Kalman filter (SEKF), to enhance the reliability of the system. The consequence of the investigation expresses that the proposed SEKF is more appropriate with outstanding reliability.

An iterative procedure can be applied to eliminate the noise produced by error functions (Lin *et al*., 2021). Madhavan (2021) presents a stochastic model for a systemic problem. Kalman filter algorithm is used to evaluate similar variables for a model. The method is described as follows (Arroyo-Marioli *et al*., 2021) :

$$\begin{aligned} y_{1i} &= y_{10} + y_{10} x_{0i} + \beta_{11} y_{1i} + \varepsilon_{oi} \\ y_{2i} &= y_{20} + y_{21} x_{1i} + y_{22} x_{2i} + \beta_{21} y_{1i} + \varepsilon_{1i} \\ y_{3i} &= y_{30} + y_{32} x_{2i} + \beta_{31} y_{1i} + \beta_{32} y_{2i} + \varepsilon_{2i} \\ &. \\ &. \\ &. \\ y_{ti} &= y_{t0} + y_{(t,t-1)} x_{(t-1,i)} + \beta_{t1} y_{1i} + \beta_{t,2} y_{2i} + ... + \beta_{(t,t-1)} y_{(t-1,i)} + \varepsilon_{(t-1,i)} \end{aligned} \qquad (3)$$

$y_{(t-1,i)}$ and $\beta_{(t,t-1)}$ $(t = 0,1,2,...)$ are unpredictable, y_{ti} is the product, and $\varepsilon_{(t-1,i)}$ are sustained variables. x_t is denoted by price; temperature and humidity are known by y_t . The dependent coefficient is denoted as y_{ti} . Equation (3) is denoted by

$$y_{ti} = b_{t0} + b_{(t,t-1)} x_{(t-1,i)} + \alpha_{t1} y_{1i} + ... + \alpha_{(t,t-1)} y_{t-1} \qquad (4)$$

$$\varepsilon_{ti} = y_{ti} - y_{ti} = bt_0 - b_{(t,t-1)} x_{(t-1,i)} - \alpha_{ti} \hat{y}_{1i} - \alpha_{t2} y_{2i} - ... - \alpha_{(t,t-1)} y_{(t-1,i)} \qquad (5)$$

for $t = 1, 2, ,...n$

Equation (5) is the difference between observed and predicted variables.

$$SSE = \sum_{i=1}^{n} (y_{ti} - b_{t0} - b_{(t,t-1)} x_{(t-1,i)} - \alpha_{t1} y_{1i} - \alpha_{t2} y_{2i} - ... - \alpha_{(t,t-1)} y_{(t-1,i)})^2 \qquad (6)$$

The standard deviation (SD) is denoted by:

$$SD = \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{n} \qquad (7)$$

x_i are real values, μ is average, and n is numbers. The numerical calculation for standard error for values utilized in this research is denoted by:

$$SE = \frac{SD}{\sqrt{n}} \tag{8}$$

According to Zhong, (2021), the traditional data envelopment analysis (DEA) technique was utilized for reliability tests. It consists of issues. Examples of such issues are being over-elaborated with errors. In addition, the evaluation of original units us re-assessed when a new evaluation is required. This research utilizes machine learning algorithms to allow the shortcomings assessment technique. The created model is, more usable, provides functional instruments for enhanced cost-effectiveness (Zhong, 2021). Accuracy within obtained results are investigated by utilizing the statistical estimates. These statistical parameters are calculated by utilizing the following connection:

$$RMSE = \sqrt{\frac{(e_t)^2}{N}} \tag{9}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{n}\left|\frac{e_t}{y_t}\right| * 100 \tag{10}$$

where, e_t is error, y_t is real, and N is number

Box-Jenkins Technique

The Box-Jenkins technique used in this study applies autoregressive integrated-moving average (ARIMA) process as a time series method. The idea of using a mathematical model to describe the behavior of a physical phenomenon is well established. In particular, it is possible to derive a model based on physical laws, which it is possible to calculate the value of some time-dependent quantity nearly exactly at any instant of time. If exact calculation were possible, such a model would be entirely deterministic. The adequacy of the Box Jenkins approach to time series modeling was discussed by (Asamoah et al., 2012). The Box Jenkins method used in this study was able to estimate models in order to meet the demand requirements of customers. Results obtained from the study indicated that the method is adequate for time-series data modeling.

It may be possible to derive a probability or stochastic model, which models the probability of the future behavior of a value lying between two specified limits. The models for time series are stochastic models. A sequence of numbers is produced in the stochastic process. The *backward shift operator* E is applied to the computation of the Box-Jenkins technique and is defined by $Bz_t = z_{t-1}$.

Hence, $B^n z_t = z_{t-n}$

Also,

$\nabla z_t = z_t - z_{t-1} = (1 - B)\, z_t$

The stochastic models employed are based on the idea that an observable time series z_t is

transformed to the process a_t, with $\psi_1, \psi_2, \ldots$ as weights. i.e.

$$z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots = \mu + \psi(B) a_t \tag{11}$$

In general,

$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$

Let $t, t-1, t-2, \ldots$ be $z_{t-1}, z_{t-2}, \ldots$

Also let $\tilde{z}_t = z_t - \mu$ be the series of deviations from μ The autoregressive model given by Equation (1) may be written economically as $\phi(B)\tilde{z}_t = a_t$

The model contains *p+2* unknown parameters $\mu, \phi_1, \phi_2, \ldots, \phi_p, \sigma_a^2$.

The autoregressive (AR) model is a special case of the linear filter model of Equation (11), for example,

$\tilde{z}_{t-1}$ can be eliminated by substituting

$$\tilde{z}_{t-1} = \phi_1 \tilde{z}_{t-2} + \phi_2 \tilde{z}_{t-3} + \ldots + \phi_p \tilde{z}_{t-p-1} + a_{t-1}$$

Similarly, $\tilde{z}_{t-2}$ can be substituted. Hence,

$$\tilde{z}_t = \phi^{m+1} \tilde{z}_{t-m-1} + a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \ldots + \phi^m a_{t-m}$$

Generally,

$\phi(B)\tilde{z}_t = a_t$ is written as

$$\tilde{z}_t = \phi^{-1}(B) a_t = \psi(B) a_t$$

with $\psi(B) = \phi^{-1}(B) = \sum_{j=0}^{\infty} \psi_j B^j$

This is given as

$$\varphi(B) z_t = \phi(B)\nabla^d z_t = \theta_o + \theta(B) a_t \qquad (12)$$

where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$$

$\phi(B)$ and $\theta(B)$ are polynomial operators in *B* of degrees *p* and *q*.

This process is referred to as an ARMA (*p*, *q*) process. The ARIMA model can be expressed explicitly in terms of current and previous shocks.

A linear model can be written as the output z_t from the linear filter

$$z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + ... = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j} = \psi(B) a_t \quad (13)$$

whose input is a white noise, or a sequence of uncorrelated shocks a_t with mean 0 and Common variance σ_a^2. Operating on both sides of Equation (13) with the generalized autoregressive operator

$\varphi(B)$, then,

$\varphi(B) z_t = \varphi(B)\psi(B) a_t$,

However, since $\varphi(B) z_t = \theta(B) a_t$, it follows that

$\varphi(B)\psi(B) = \theta(B)$

Therefore, the ψ weights can be determined in the expansion

$$\left(1-\varphi_1 B - ... - \varphi_{p+d} B^{p+d}\right)\left(1+\psi_1 B + \psi_2 B^2 + ...\right)\left(1-\theta_1 B - ... - \theta_q B^q\right) \quad (14)$$

Thus, the ψ_j weights of the ARIMA process can be determined recursively through the equations $\psi_j = \varphi_1 \psi_{j-1} + \varphi_2 \psi_{j-2} + ... + \varphi_{p+d} \psi_{j-p-d} - \theta_j,\ j > 0$ with $\psi_0 = 1, \psi_j = 0$ for $j < 0$, and $\theta_j = 0$ for $j > p$. It is noted that for *j* greater than the larger of $p+d-1$ and *q*, the ψ weights satisfy the homogenous difference equation defined by the generalized autoregressive operator, that is,

$\varphi(B)\psi_j = \phi(B)(1-B)^d \psi_j = 0$

where *B* now operates on the subscript *j.*

Regression Analysis

Regression techniques belong to the class of causal models. The regression technique, As described by Wu (2018), is discussed in this section. A principal purpose of regression is to forecast known or assumed variables. To make predictions or estimates, we must identify the effective predictors of the variable of interest. A regression using only one predictor is called simple regression. Where variables are multiple or more predictors, multiple regressions analysis is employed.

To develop a regression model, begin with a hypothesis about how several variables might be related to another variable and the relationship, i.e.
Given a straight line, and letting

$$\varepsilon^2 = \Sigma (y_i - y_p)^2 \quad (15)$$

y_p = predicted values, and y_i = actual values. Also, let

$$y_p = m\, x_i + b \qquad (16)$$

Combine Equation (15) and Equation (16),

$$\varepsilon^2 = \Sigma\, (y_i - m\, x_i - b)^2 =$$

$$\Sigma\, (m^2 x_i^2 + 2mbx_i - 2mx_i y_i + b^2 - 2by_i + y_i^2) \qquad (17)$$

Equation (17) is at a minimum, i.e.

$$\frac{ds^z}{dm} = 0 \text{ , and } \frac{ds^z}{db} = 0 \qquad (18)$$

Solving for Equation (18), gives

$$\frac{ds^z}{dm} = 2m\Sigma x_i^2 + 2b\Sigma x_i - 2\Sigma(x_i y_i) = 0 \qquad (19)$$

Solving for Equation (18), gives

$$\frac{ds^z}{db} = 2m\Sigma x_i + 2\Sigma b - 2\Sigma y_i = 0 \qquad (20)$$

To simplify the notations, let

$Sx = \Sigma\, x_i$

$Sy = \Sigma\, y_i$

$Sxy = \Sigma\, (x_i y_i)$

$Sxx = \Sigma\, (x_i^2)$

$\Sigma\, b = nb$

From Equation (19),

$$Sxy = mSxx + bSx \qquad (21)$$

From Equation (20),

$$Sy = mSx + nb \qquad (22)$$

The optimized values for *m* and *b* are respectively given as:

$$m = \frac{n\sum x_i y_i - \sum x_i y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$b = \frac{\sum y_i}{n} - \frac{m\sum x_i}{n}$$

Data Collection

Data collection is covered in this section. The quantitative techniques are tested on data collected to identify an accurate and reliable technique. When checking for the reliability, methods are investigated on data procured. This data consists of different structures of toxicity threats by internet users on their websites (Open Data Source, 2021).

The chosen dataset taken from the website is below. The methods are investigated based on this dataset.

Table 1: Dataset

id	comment_text	toxic	severe_toxic	obscene	threat	insult	identity_hate
0000997932d777bf	Explanation	0	0	0	0	0	0
000103f0d9cfb60f	D'aww! He matches	0	0	0	0	0	0
000113f07ec002fd	Hey man, I'm really r	0	0	0	0	0	0
0001b41b1c6bb37e	"	0	0	0	0	0	0
0001d958c54c6e35	You, sir, are my hero	0	0	0	0	0	0
00025465d4725e87	"	0	0	0	0	0	0
0002bcb3da6cb337	COCKSUCKER BEFOR	1	1	1	0	1	0
00031b1e95af7921	Your vandalism to th	0	0	0	0	0	0
00037261f536c51d	Sorry if the word 'no	0	0	0	0	0	0
00040093b2687caa	alignment on this su	0	0	0	0	0	0
0005300084f90edc	"	0	0	0	0	0	0
00054a5e18b50dd4	bbq	0	0	0	0	0	0

Results and Discussion

The outcome of this research is shown in this part of the study. The output is implemented using the MATLAB programming language. Table 2 focuses on the reliability of the predicted values for the machine learning techniques by calculating and differentiating their independent *root mean square error* (*RMSE*) and the *mean average percentage error* (*MAPE*).

Table 2: Performance of Machine Learning Techniques

Statistical Variables	ANN	Kalman	Box-Jenkins	Regression
Root mean square error	5.5	7.3E-15	7.1	8.3
Mean average percentage error	4.1	3.9E-15	6.5	7.8

The *RMSE* and *MAPE* of the artificial neural network, Box-Jenkins technique, and the regression analysis are higher than the *RMSE* and *MAPE* of the Kalman filter adaptation algorithm. Table 2 indicate the *RMSE* for the Kalman filter adaptation algorithm is 7.3E-15, and 3.9E-15 for *MAPE*, which are very much better than an artificial neural network (5.5), Box-Jenkins (7.1), and regression analysis (8.3) for *RMSE* and for *MAPE,* the artificial neural network is 4.1*,* Box-Jenkins is *6.5* and regression analysis is 7.8. As a result, the Kalman filter adaptation algorithm is more reliable than the artificial neural network.

The statistical variables for computational methods are computed to study the statistical inference of results obtained from the research. The statistical variables are weighed up indicating reliability with forecast values to actual values for studying the correctness with the output. Table 3 shows values of the Kalman filter adaptation algorithm (0.1%) is more reduced than that of the ANN technique (11.3%), Box-Jenkins technique (13.2) and the regression analysis (14.1). As a result, the Kalman filter adaptation algorithm has better reliability when weighed up with the ANN technique, Box-Jenkins technique and the regression technique.

Table 3: Statistical Variables

Method	**Mean**	**Std. Dev.**	**Minimum**	**Maximum**
Kalman	98.5	0.1%	78.6	113.9
ANN	98.7	11.3%	81.3	119.6
Box-Jenkins	89.9	13.2	88.2	121.7
Regression	89.4	14.1	89.2	122.5

Conclusion

The use of implementing the accurate and reliable methods is important in modeling and predictions, otherwise poor results might occur. This research focuses on the reliability of evolving predictive models utilizing machine learning techniques. These models are evaluated on archival data utilizing these methods. This research suggests that the Kalman filter adaptation algorithm is more constructive in contrast to the artificial neural network, Box-Jenkins technique and the regression analysis.This validates the Kalman filter adaptation algorithm to be more functional in predicting data.

For further work, research will be taken up to evolve predictive models of variables influencing internet usage. This area of research is also known as IoT. This could be a part of future technology. Frameworks will be constructed based on these variables.

References

Allison, J. R. (2020). Computational methods for exploring protein conformations. *Biochemical Society Transactions*, 48(4), 1707-1724.

Agamah, F. E., Mazandu, G. K., Hassan, R., Bope, C. D., Thomford, N. E., Ghansah,A. and Chimusa, E. R. (2020). Computational/in silico methods in drug target and lead prediction. *Briefings in bioinformatics*, 21(5), 1663-1675.

Arroyo-Marioli, F., Bullano, F., Kucinskas, S. and Rondón-Moreno, C. (2021). Tracking R of COVID-19: A new real-time estimation using the Kalman filter. *PloS one,* 16(1), e0244474.

Alfian, R. I., Ma'arif A. and Sunardi, S. (2021). Noise Reduction in the Accelerometer and Gyroscope Sensor with the Kalman Filter Algorithm. *Journal of Robotics and Control (JRC),* 2(3): 180-189.

Asamoah, D., Annan, J., & Arthur, Y. (2012). Time Series Analysis of Electricity Meter Supply in Ghana. International Journal of Business ans Social Science, 3(19), 16– 22.

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., and Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557), 871-876.

Chaabouni, R., Kharitonov, E., Dupoux, E. and Baroni, M. (2021). Communicating artificial neural networks develop efficient color-naming systems. *Proceedings of the National Academy of Sciences,* 118(12).

Chancellor, S. and De Choudhury, M. (2020). Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine*, 3(1): 1-11.

Chen, B. W., Xu, L. and Mavrikakis, M. (2020). Computational methods in heterogeneous catalysis. *Chemical Reviews*, 121(2), 1007-1048.

Chen, Z., Zhao, P., Li, F., Wang, Y., Smith, A. I., Webb, G. I. and Song, J. (2020). Comprehensive review and assessment of computational methods for predicting RNA post-transcriptional modification sites from RNA sequences. *Briefings in bioinformatics,* 21(5): 1676-1696.

Efremova, M. and Teichmann, S. A. (2020). Computational methods for single-cell omics across modalities. *Nature methods,* 17(1), 14-17.

Hewing, L., Wabersich, K. P., Menner, M. and Zeilinger, M. N. (2020). Learning-based model predictive control: Toward safe learning in control. *Annual Review of Control, Robotics, and Autonomous Systems*, 3, 269-296.

Kang, J., Chen, T., Luo, H., Luo, Y., Du, G. and Jiming-Yang, M. (2021). Machine learning predictive model for severe COVID-19. *Infection, Genetics and Evolution 90*: 104737.

Khoo, Y., Lu, J. and Ying, L. (2021). Solving parametric PDE problems with artificial neural networks. *European Journal of Applied Mathematics,* 32(3): 421-435.

Kumar, G., Jain, S. and Singh, U. P. (2021). Stock market forecasting using computational intelligence: A survey. *Archives of Computational Methods in Engineering,* 28(3): 1069-1101.

Lai, X., Yi, W., Cui, Y., Qin, C., Han, X., Sun, T. and Zheng, Y. (2021). Capacity estimation of lithium-ion cells by combining model-based and data-driven methods based on a sequential extended Kalman filter. *Energy*, 216, 119233.

Lin, X., Li, X. and Lin, X. (2020). A review on applications of computational methods in drug screening and design. *Molecules*, 25(6), 1375.

Lin, X., Li, W., Li S., Ye J., Yao C. and He, Z. (2021). Combined adaptive robust Kalman filter algorithm. *Measurement Science and Technology*, 32(7): 075015.

Madhavan, P. G. (2021). Stochastic Formulation of Causal Digital Twin: Kalman Filter Algorithm arXiv preprint arXiv:2105.05236.

Open Data Kit (2021). A post by ODK available at https://getodk.org/

Ozoh, P., Abd-Rahman, S; Labadin, J; Apperley, M. (2014). Modelling Electricity Consumption Using Modified Newton's Method, *International Journal of Computer Applications*, 86(13): 27-31.

Wu, C. (2018). Regression Technique. Retrieved from http://www.historyofinformation.com/expanded.php?id=2706.

Wu, Z., Lv, H., Meng, Y., Guan, X. and Zang, Y. (2021). The determination of flood damage curve in areas lacking disaster data based on the optimization principle of variation coefficient and beta distribution. *Science of The Total Environment,* 750: 142277.

Yang, X., Guan, J., Ding, L., You, Z., Lee, V. C., Hasan, M. R. M. and Cheng, X. (2021). Research and applications of artificial neural network in pavement engineering: a state-of-the-art review. *Journal of Traffic and Transportation Engineering (English Edition).*

Zhao J., Cao Y. and Zhang, L. (2020). Exploring the computational methods for protein-ligand binding site prediction. *Computational and structural biotechnology journal*, 18: 417-426.

Zhong, K., Wang, Y., Pei, J., Tang, S. and Han, Z. (2021). Super efficiency SBM-DEA and neural network for performance evaluation. *Information Processing & Management,* 58(6), 102728.