Research Article
Hittite Journal of Science and Engineering, 2024, 11 (1) 15-23
ISSN NUMBER: 2148-4171
DOI: 10.17350/HJSE19030000327

COVID-19 Diagnosis from Blood Gas Using Multivariate Linear Regression

Faruk Ayata[1] Ebubekir Seyyarer[2]
[1]Van Yuzuncu Yil University, Computer Technologies, Van, Türkiye
[2]Van Yuzuncu Yil University, Computer Engineering, Van, Türkiye

Article History:
Received: 2023/09/20
Accepted: 2024/02/73
Online: 2024/03/31

***Correspondence to**: Faruk Ayata, Van Yüzüncü Yıl University, Computer Technologies, Van, TÜRKİYE*
E-Mail: farukayata@yyu.edu.tr;
Phone: +90 531 500 33 66.

This article has been checked for similarity.

ABSTRACT

With the impact of the COVID-19 outbreak, almost all scientists and nations began to show great interest in the subject for a long time. Studies in the field of outbreak, diagnosis and prevention are still ongoing. Issues such as methods developed to understand the spread mechanisms of the disease, prevention measures, vaccine and drug research are among the top priorities of the world agenda. The accuracy of the tests applied in the outbreak management has become extremely critical. In this study, it is aimed to obtain a function that finds the positive or negative COVID-19 test from the blood gas values of individuals by using Machine Learning methods to contribute to the outbreak management. Using the Multivariate Linear Regression (MLR) model, a linear function is obtained to represent the COVID-19 dataset taken from the Van province of Turkey. The data set obtained from Van Yüzüncü Yıl University Dursun Odabaş Medical Center consists of blood gas analysis samples (109 positive, 1146 negative) taken from individuals. It is thought that the linear function to be obtained by using these data will be an important method in determining the test results of individuals. Gradient Descent optimization methods are used to find the optimum values of the coefficients in the function to be obtained. In the study, the RMSProp optimization algorithm has a success rate of 58-91.23% in all measurement methods, and it is seen that it is much more successful than other optimization algorithms.

Keywords:
Multivariate linear regression; COVID-19; Blood gases.

INTRODUCTION

The COVID-19 pandemic is an infectious disease that emerged in the city of Wuhan, China, in 2019 and quickly spread worldwide. Diagnosis of COVID-19 is typically performed using various methods such as symptoms, imaging techniques, and laboratory tests [1, 2]. Common symptoms include fever, cough, shortness of breath, muscle aches, fatigue, headache, and loss of taste or smell [3]. Imaging techniques are another auxiliary method in the diagnosis of COVID-19. The most commonly used methods include computerized tomography (CT) and chest X-rays. These imaging methods can help detect damage, inflammation, and other abnormalities in the lungs [4, 5]. The most common laboratory tests include polymerase chain reaction (PCR) tests, antigen tests, and antibody tests. The PCR test helps diagnose COVID-19 by detecting the genetic material of the virus in the patient's respiratory samples. In addition to these, blood gas values are important parameters that needs to be monitored during the disease management process [6]. Hemoglobin, composed of heme and globin, which transport gases in the blood, is present in all living organisms. Hemoglobin's primary function is to transport oxygen (O_2) from the lungs to peripheral tissues and carry carbon dioxide (CO_2) from tissues to the lungs. An increase in carboxyhemoglobin levels has been observed in COVID-19 patients receiving treatment in intensive care units [6]. Diagnosing whether or not someone is infected with COVID-19 remains a current challenge. Machine learning methods, which have provided solutions to many contemporary problems, can also play a significant role in the fight against COVID-19.

Machine learning techniques can be applied in various fields such as tracking the spread of the virus, making diagnoses, optimizing treatment, and discovering potential vaccines. COVID-19 diagnosis is categorized

Cite as:
Ayata F, Seyyarer E. COVID-19 Diagnosis From Blood Gas Using Multivariate Linear Regression. Hittite Journal of Science and Engineering 2024;11 (1):15-23. doi:10.17350/hjse19030000327

as a Non-deterministic Polynomial problem, and analytical methods of Machine Learning that provide exact results are not preferred for solving such problems [7]. This is because as the dimensions of the problem increase, analytical methods become cost-prohibitive. Instead of analytical methods, heuristic or metaheuristic approach algorithms inspired by nature are recommended, which provide approximate results and are optimization-based, such as Artificial Neural Networks [7, 8], Genetic Algorithms [7, 9], Particle Swarm Optimization [7, 10], and others. The reason for this is that these methods can provide approximate results at a reasonable cost even as the size of the problem increases. Such methods do not guarantee exact results [11, 12].

Intuitive methods are techniques aimed at finding a satisfactory solution through trial and error without guaranteeing that the solution is optimal or even close to optimal. These methods are typically used for complex problems that cannot be solved with analytical methods or in situations where an approximate solution is acceptable [7]. Metaheuristic methods, on the other hand, are optimization techniques that combine multiple intuitive methods to obtain a better solution [13]. These methods aim to explore the search space more efficiently than simple heuristic methods by using various strategies such as randomization, neighborhood search, and adaptive memory. Metaheuristic methods are often used in complex optimization problems where the search space is vast and navigation is challenging [8, 13].

Multivariate linear regression is based on the assumption that the dependent variable can be expressed as a linear combination of independent variables. Therefore, it is one of the most suitable metaheuristic methods to use when data exhibits a linear relationship [14, 15]. Multivariate linear regression can be employed in crises such as the COVID-19 pandemic for tasks such as predicting the course of the outbreak, determining the disease burden, assessing healthcare resource requirements, planning pandemic control strategies, and evaluating their effectiveness.

Özen et al. (2021) use machine learning methods to make predictions for the detection of COVID-19 cases. They employed Python and R programming languages and utilized Prophet, Polynomial Regression, ARIMA, Linear Regression, and Random Forest models for their predictions. They reported that the Polynomial regression method provided the best prediction results [15]. Saadatmand et al. (2021) developed a model for the detection of oxygen therapy needs of COVID-19 patients. This model includes five different methods: Logistic Regression, Random Forest, XGBoost, C5.0, and Artificial Neural Networks. They used data obtained from two local hospitals in Iran to create their dataset. Test results showed that the Logistic Regression and Artificial Neural Networks used in the model achieved the highest accuracy rate [16]. Mohan et al. (2021) used their Ensemble Learning, Autoregressive, and Moving Regressive (EAMA) hybrid model to detect COVID-19. The EAMA model, also known as a community learning, autoregressive model, and moving average model, used data from the Ministry of Health and Family Welfare in India and Worldometers. Their analysis allowed for detailed predictions of active cases and deaths at the state level in India [17]. Pinter et al. (2020) proposed a hybridization model consisting of a network-based fuzzy inference system and a multi-layer perceptron-empirical competitive algorithm for the prediction of COVID-19. The performance evaluation of the proposed model used metrics such as Mean Absolute Percentage Error, Root Mean Square Error (RMSE), and coefficient of determination (R-squared). The analysis indicated promising results for the proposed method in disease prediction [18]. Elaziz et al. (2020) used CXR images to distinguish COVID-19 cases. They proposed a new feature selection method and utilized a modified mantis-ray search optimization algorithm based on Modified Reflective and Fitness-Oriented Differential Evolution (MRFODE) to determine relevant subset features. Test results indicated promising accuracy values in classifying COVID-19 patient samples [19].

In this study, a method is aimed to be developed for the detection of the COVID-19 virus from blood gas data frequently used in laboratory tests and for monitoring the disease following transmission. The method includes a multivariate linear regression model, six gradient descent-based optimization algorithms to minimize error, and a dataset obtained from the Van province of Turkey for COVID-19.

MATERIAL AND METHODS

Optimization Algorithms

The primary objective of optimization algorithms is to minimize the error quantity. Methods based on Gradient Descent, such as Stochastic Gradient Descent (SGD), Momentum Stochastic Gradient Descent (MMT), Adadelta Gradient Descent (AGD), RMSProp (RMP), Adagrad (ADD), and Adam (ADM), are among the most commonly used optimization algorithms. These algorithms are employed to reduce the error quantity and improve the model's performance. Table 1 displays these algorithms [20- 24].

Objective Functions

The choice of the objective function is dependent on the goal of the optimization algorithm. This goal is typically determined as enhancing the model's accuracy or minimizing the error quantity. Integral of Absolute Er-

Table 1. Most popular gradient descent methods [20- 24].

Algorithm Name	Formula	Description
SGD	$W_{t+1} = W_t - a\dfrac{\partial L}{\partial W_t}$	The current derivative ($\partial L/\partial w_t$) updates the current weight (w_t) by multiplying it with the learning rate (a).
MMT	$w_{t+1} = w_t - aV_t$ $V_t = \beta V_{t-1} + (1-\beta)\dfrac{\partial L}{\partial W_t}$	The initial value of V_t is 0. β is between 0 and 1, and it is commonly taken as 0.9.
ADG	$w_{t+1} = w_t - \dfrac{a}{\sqrt{S_t + \epsilon}} \cdot \dfrac{\partial L}{\partial w_t}$ $S_t = S_{t-1} + \left[\dfrac{\partial L}{\partial w_t}\right]^2$	S starts as 0 initially. ϵ is typically set to a very small value (10^{-7}).
RMP	$w_{t+1} = w_t - \dfrac{a}{\sqrt{S_t + \epsilon}} \cdot \dfrac{\partial L}{\partial w_t}$ $S_t = \beta S_{t-1} + (1-\beta)\left[\dfrac{\partial L}{\partial w_t}\right]^2$	S starts as 0 initially, a = 0.001, β = 0.9, and ϵ is chosen as 10^{-6}.
ADD	$w_{t+1} = w_t - \dfrac{\sqrt{D_{t-1}}}{\sqrt{S_t + \epsilon}} \cdot \dfrac{\partial L}{\partial w_t}$ $D_t = \beta D_{t-1} + (1-\beta)[\Delta w_t]^2$ $S_t = \beta S_{t-1} + (1-\beta)\left[\dfrac{\partial L}{\partial w_t}\right]^2$ $\Delta w_t = w_t - w_{t-1}$	S and D are initialized to 0, β is set to 0.95, and ϵ is chosen as 10^{-6}.
ADM	$W_{t+1} = W_t - \dfrac{a}{\sqrt{\hat{S}_t + \epsilon}} \cdot \hat{V}_t$ $\hat{V}_t = \dfrac{V_t}{1-\beta_1^t}$ $\hat{S}_t = \dfrac{S_t}{1-\beta_2^t}$ $V_t = \beta_1 V_{t-1} + (1-\beta_1)\dfrac{\partial L}{\partial W_t}$ $S_t = \beta_2 S_{t-1} + (1-\beta_2)\left[\dfrac{\partial L}{\partial W_t}\right]^2$	S and V are initially set to 0, a = 0.001, β1 = 0.9, β2 = 0.999, and ϵ is chosen as 10^{-8}.

ror (IAE), Integral of Time multiplied by Absolute Error (ITAE), and Mean Squared Error (MSE) are commonly used objective functions in control systems and optimization problems. IAE measures the system's error, i.e., the deviation between actual and desired values, and is used for minimization. ITAE evaluates the performance of a control system by considering both the error and response time. MSE measures the error quantity and is used for minimization as well. MSE is also often used as a performance metric in regression problems [25, 26]. Table 2 displays the mathematical formulas of the objective functions.

Table 2. Objective Functions [25, 26].

Method Name	Formula
IAE	$\int_0^t \lvert e(t) \rvert \, dt$
ITAE	$\int_0^t t \lvert e(t) \rvert \, dt$
MSE	$\frac{1}{n}\int_0^t (e(t))^2 \, dt$

Feature Selection

In the field of machine learning, datasets are growing exponentially day by day and their quantitative numbers as well as qualitative features are increasing. The increase in the number of features in datasets can lead to different behaviors of machine learning methods. Even when behaviors do not change, it excessively increases the costs of methods. Dimension reduction techniques are used to reduce these costs. Feature Selection (FS) and Feature Extraction (FE) are the most common dimension reduction techniques. FE creates new and more effective features by using existing ones. FS, on the other hand, selects

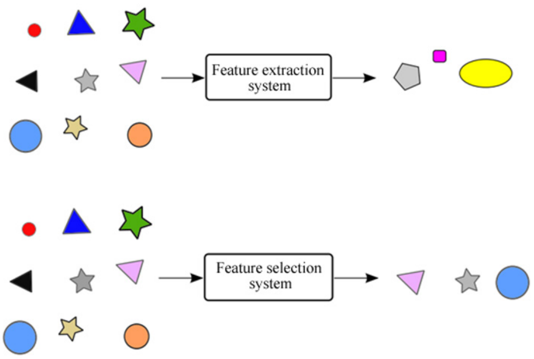

Figure 1. The Difference Between FE and FS [29].

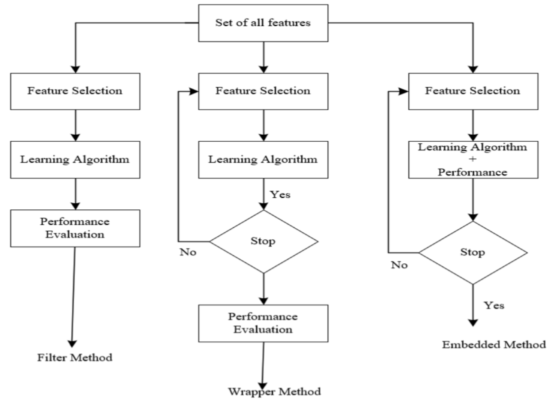

Figure 2. An Overview of FS Main Headings [27].

Table 3. Attributes of blood gas examination records for COVID-19 testing.

	1.	2.	3.	4.	...	...	686.	...	...	1255.
Age	84	64	82	70	...	...	51	32	70	18
Gender	E	K	K	E	...	...	E	E	E	K
Base(B,ox)	-0.15	-1.50	7.90	0.00	...	...	5.10	0.70	-17.60	-0.70
Base(Ecf)	1.25	-2.50	6.89	-3.90	...	...	6.90	2.70	-11.00	1.60
Base(Ecf,ox)	1.15	-2.60	8.57	0.00	...	...	6.40	2.40	-11.80	1.00
Ca++	1.04	0.94	1.08	1.05	...	...	1.45	1.20	1.30	1.23
$cHCO_3$	24.65	23.20	30.09	21.40	...	...	29.10	25.00	12.30	23.90
Cl	104.75	108.00	106.00	106.00	...	...	95.00	106.00	110.00	103.00
ctHb	13.95	14.50	13.43	12.20	...	...	12.90	16.30	18.30	14.30
ctO_2e	16.85	18.40	18.14	14.40	...	...	8.60	13.20	1.80	7.00
FHHb	9.93	7.80	2.66	13.30	...	...	50.30	36.30	92.20	63.20
FMetHb	0.03	-0.60	0.55	0.70	...	...	0.70	0.40	0.50	0.70
FO2Hb	88.48	91.00	95.60	84.40	...	...	47.50	57.70	6.70	35.10
Glukoz	139.75	100.00	129.83	105.00	...	...	124.00	129.00	154.00	110.00
K+	3.27	3.90	3.70	5.80	...	...	4.00	4.10	6.30	3.00
Na+	134.50	136.00	144.14	140.00	...	...	134.00	141.00	148.00	141.00
p50e	28.53	25.42	24.00	27.22	...	...	27.77	25.60	39.92	28.81
pCO2	54.10	28.80	40.31	34.70	...	...	40.10	47.60	103.00	45.40
pCO2(T)	54.10	28.80	40.31	34.70	...	...	40.10	47.60	103.00	45.40
pH	7.35	7.47	7.49	7.39	...	...	7.49	7.38	6.93	7.38
pH(T)	7.35	7.47	7.49	7.39	...	...	7.49	7.38	6.93	7.38
pO2	76.15	61.80	95.12	53.00	...	...	27.20	30.70	13.10	22.90
pO2(T)	76.15	61.80	95.12	53.00	...	...	27.20	30.70	13.10	22.90
sO2	89.90	92.10	97.30	86.40	...	...	48.60	61.40	6.80	35.70
Result		Positive						Negative		

the features among the existing ones that have the most influence on the outcome. Therefore, instead of raw data, a dataset that is more effective on the results is provided as input data to machine learning methods. Using these methods, performance improvement and cost reduction can be achieved. The difference between FE and FS is shown in Fig. 1 [27-29].

FE is applied to data sets such as images and audio, while FS is applied to more sensitive data. Since the dataset used in the study does not involve image data and is related to the field of medicine, FS methods are employed. As seen in Fig. 2, FS is examined under three main headings (Filter, Wrapper, Embedded) [27, 29, 30].

Table 4. Features to be used in the MLR model.

Age	Base(B,ox)	Ca++	Result
84	-0.15	1.04	Positive
64	-1.50	0.94	Positive
82	7.90	1.08	Positive
70	0.00	1.05	Positive
...	...	...	...
...	...	...	...
61	2.00	1.06	Negative
60	7.41	1.05	Negative
51	5.10	1.45	Negative
32	0.70	1.20	Negative
...	...	...	...
...	...	...	...

Spiral methods are preferred because they take into account the dependence between features and have high accuracy performance from filter methods. There are three types of spiral methods used in regression or classification methods [27, 30].

Spiral methods have different methods such as Forward Selection, Backward Selection, and Stepwise (Exhaustive) Selection. In the Forward Selection method, the feature that most affects the performance of the machine learning method is selected from the feature pool, then the second most influential feature is selected, and so on until the feature selection reaches the stopping criterion. A stopping criterion is considered to be selecting all features with a P-value below 0.05. In the Backward Selection method, the opposite path of Forward Selection is followed. Feature selection begins by removing the feature that has the least impact on the performance of the machine learning method from all features. Then, the second least influential feature is removed. This process continues until all features with a P-value above 0.05 are removed to create the best feature subset. Stepwise Selection is a combination of both methods. Each feature is compared to all other features and selected for the best feature subset [27, 30, 31].

Problem Formulation

Table 3 displays the features of the dataset obtained from the hospital. The Blood Gas data set, where the MLR model is applied, is obtained upon an official request. This dataset consists of examination records conducted by Van Yüzüncü Yıl University Dursun Odabaş Medical Center between 01.11.2020 and 31.12.2020. The Blood Gas records contain 22 attributes (Base(B,ox), Base(Ecf), Base(Ecf,ox), Ca++, $cHCO_3$, C_l, ctHb, ctO_2e, FHHb, FMetHb, FO_2Hb, Glucose, K+, Na+, p50e, pCO_2, pCO2(T), pH, pH(T), pO_2, $pO_2(T)$, sO_2). Together with age and gender information as shown in Table 3, a dataset with 24 features is obtained. In the result section, there are labels indicating whether the COVID-19 virus is positive or negative. In other words, the obtained dataset has two classes (positive, negative). The positive class has a total of 109 samples (87 training, 22 testing), and the negative class has a total of 1146 samples (916 training, 229 testing), making a total of 1255 samples. In MLR models, a single point is initially selected for coefficients. These points are generally zero, one, or a randomly selected value. In the conducted study, a population is used for the initialization of coefficients.

Feature selection methods (Filter, Wrapper, and Embedded) are applied to the dataset in Table 3. The features that are common among the most effective ones obtained in all three methods (Age, Base(B,ox), Ca++) are shown in Table 4. The training and testing processes of MLR are conducted using these selected features.

The resulting function is formulated with the most effective features as shown in Eq. 1.

$$Y = \theta_1 X_{Yas} + \theta_2 X_{\text{Base}}(B, ox) + \theta_3 X_{Ca++} \quad (1)$$

Y, X_{Age}, $X_{(Base(B,ox))}$, $X_{(Ca++)}$ values are read from the dataset to obtain the coefficients θ1, θ2, θ3. These coefficients are generally initially set to 0 or 1 or randomly generated between two values. Using the obtained coefficients after the operations, the X input values are substituted in Equation 1 to obtain Y' (Y to the power of Y). The difference between the real value Y and the predicted value Y' constitutes our error amount.

Experimental Study

All modules of the conducted study are shown in Figure 3. After applying feature selection methods to the entire dataset, outputs are generated with six different optimization algorithms and three different objective functions. In other words, 6*3 outputs are compared, and the results are analyzed. Solving MLR problems requires finding a function using many inputs and their corresponding output values. Therefore, after the regression process, the most suitable values found represent a function that includes all input and output values, replacing the θ coefficients in Equation 1. To find appropriate values for these coefficients, an initial value is assigned, and the process is initiated. With this initial value, the Dik Inis optimization algorithms are used to gradually obtain the optimal θ coefficient.

In this study, it starts with stochastic initial coefficients because the aim is to reach a solution as quickly as possible.

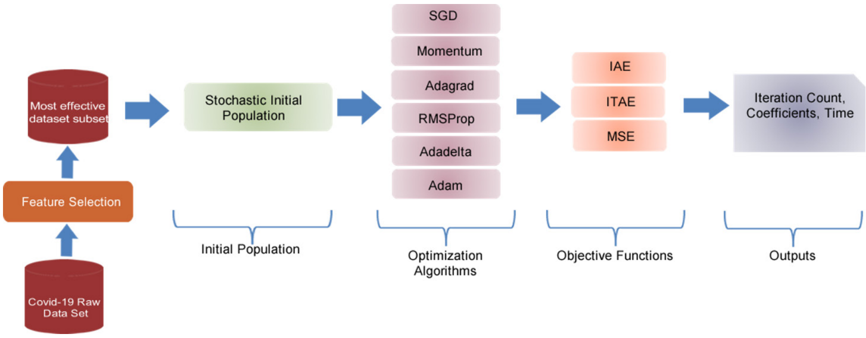

Figure 3. All modules of the study.

From the initial population, one candidate solution that is closest to the result must be selected to initiate the application. After individually processing all candidate solutions through optimization and objective functions once, the candidate solution with the lowest result is chosen to proceed. The obtained ideal candidate solution continues with regression steps for up to 1000 iterations. At the end of the regression process, the coefficients of the function that represents the entire dataset will have been determined.

RESULTS AND DISCUSSION

The outputs obtained from the application with a starting population size of 10 are listed in Table 5. In Table 5, the averages of the outputs obtained from 10 runs are compared. Here, the outputs of the applications using different optimization algorithms and objective functions (θ_1, θ_2, θ_3, and the number of iterations) are listed. Additionally, the "time" column of the table provides the processing times for each objective function for all optimization methods. Adadelta and Adagrad optimization algorithms achieve results with the lowest number of iterations for all objective functions. SGD, Momentum, and Adam optimization algorithms have high numbers of iterations for the IAE objective function. Among the objective functions, MSE reaches results with the lowest number of iterations for all optimization algorithms.

In Fig. 4, the errors and the number of iterations for the application using the IAE objective function with all optimization algorithms are shown graphically. Here, applications for all optimization algorithms are run 10 times, and the

Table 5. Results table for a population size of 10.

Obj. Func.	Opt. Alg.	Iterations Number	θ_1	θ_2	θ_3	Time
IAE	SGD	1000	0.3533127117	0.7155792544	0.9136478838	00:27:52
	MMT	1000	0.3539140773	0.7161656774	0.9141953364	
	ADG	1000	0.0000619687	0.0000620018	0.0000621926	
	RMP	1000	0.9425588422	0.9871959691	0.9895221032	
	ADD	1000	0.0000525305	0.0000527744	0.0000528254	
	ADM	1000	0.5689422754	0.4302199096	0.3704841926	
ITAE	SGD	1000	0.3532622052	0.7154840038	0.9135009433	00:32:59
	MMT	1000	0.3538630228	0.7160694295	0.9140472966	
	ADG	1000	0.0000614089	0.0000618772	0.0000618805	
	RMP	1000	0.9427011017	0.9867973079	0.9885063650	
	ADD	1000	0.0000525305	0.0000527744	0.0000528254	
	ADM	1000	0.5675193230	0.4295664937	0.3697876234	
MSE	SGD	1000	0.3532656701	0.7154717486	0.9135007489	00:28:45
	MMT	1000	0.3538664716	0.7160571368	0.9140470657	
	ADG	1000	0.0000614950	0.0000614423	0.0000614426	
	RMP	1000	0.9447178907	0.9855191028	0.9872083973	
	ADD	1000	0.0000525305	0.0000527744	0.0000528254	
	ADM	1000	0.5678662791	0.4290865789	0.3693919793	

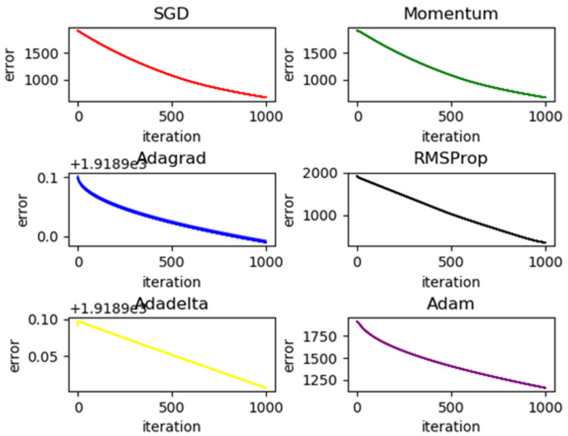

Figure 4. Outputs with the IAE objective function.

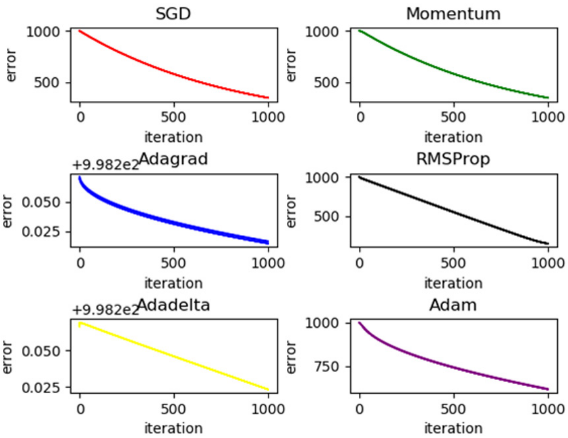

Figure 5. Outputs with the ITAE objective function.

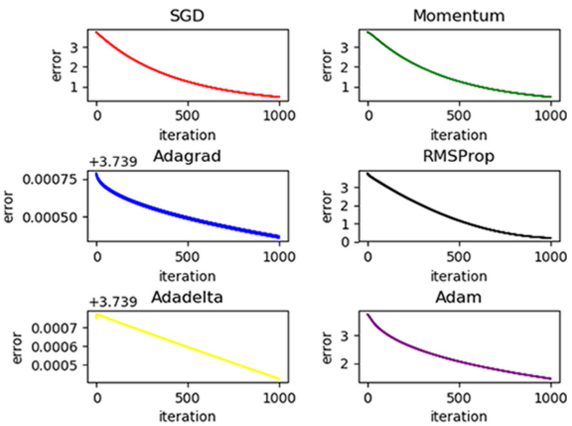

Figure 6. Outputs with the MSE objective function.

results are presented graphically. Fig. 5 displays the results of applications using the ITAE objective function and different optimization methods graphically.

Fig. 6 shows the results of the applications using the MSE objective function and different optimization methods graphically.

Results obtained for different optimization algorithms based on stochastic initialization and different objective functions are shown in Fig.s 4, 5, and 6. As seen in the figures, very similar results are obtained. Table 5 provides a comparative view of the coefficients, iteration counts, and processing times obtained from experiments with different parameters. The obtained coefficients are used for testing, i.e., a validation step is performed. The aim of the validation section is to confirm the accuracy of the obtained coefficients. Here, the success rates of the coefficients obtained with 251 test data (229 negative and 22 positive) are analyzed. Using various optimization algorithms and objective functions, many coefficient vectors (θ_1, θ_2, θ_3) are obtained. With six optimization algorithms (SGD, MMT, ADG, RMP, ADD, ADM) and three objective functions (IAE, ITAE, MSE), 6x3 coefficient vectors are obtained. When the values are put together, a matrix of size 18x3 is formed. By substituting each row of the 18-row matrix into Eq. 1 separately, processing is performed with the 251-test data set. As a result of the processing, 18x251 Y' (predicted results) are obtained. The total error rates of the obtained estimated data are calculated using the MAPE method. In addition, correct prediction counts are analyzed at certain threshold values (20%, 30%, 40%). If the error is below these threshold values, it is considered a correct prediction; if it is above, it is considered an incorrect prediction. Table 6 shows the success rates of all the studies obtained through thresholding and the MAPE method.

When examining the success rates, it is observed that the performance of the SGD, MMT, and RMP algorithms used with the IAE objective function is successful. In particular, RMP exhibits a success rate ranging from 58% to 91.23% across all measurement methods.

CONCLUSION

The aim of this study is to obtain a testing method for the COVID-19 pandemic. Individuals' COVID-19 test results, whether positive or negative, are determined using blood gas values. Multivariate linear regression modeling is carried out with a stochastic initial population method, six different optimization algorithms, and three different objective functions.

The method for generating the initial population is a crucial step in multivariate linear regression and all other machine learning algorithms, yet it remains an area with limited research. The most commonly used method to date is the traditional stochastic initialization method. In studies initiated with the traditional stochastic initialization method, the initial populations and obtained outputs always differ. Therefore, studies initiated with stochastic methods need to be run multiple times (e.g., 10, 50, or 100 times), and

Table 6. Success Rates of the Studies.

Initial Pop	Objective Func.	Opt. Alg.	Threshold			MAPE
			20% Success Rate	30% Success Rate	40% Success Rate	Success Rate (%)
Stokastik	IAE	SGD	3.187	21.513	73.705	64.71
		MMT	3.585	21.912	74.103	64.75
		ADG	0	0	0	0.01
		RMP	58.565	79.282	91.235	77.90
		ADD	0	0	0	0.01
		ADM	5.577	6.772	7.968	42.67
	ITAE	SGD	3.187	21.513	73.705	64.71
		MMT	3.585	21.912	73.705	64.74
		ADG	0	0	0	0.01
		RMP	58.565	79.282	90.836	77.88
		ADD	0	0	0	0.01
		ADM	5.577	6.772	7.9681	42.60
	MSE	SGD	3.187	21.513	73.705	64.71
		MMT	3.585	21.912	73.705	64.74
		ADG	0	0	0	0.01
		RMP	58.565	78.884	90.836	77.85
		ADD	0	0	0	0.01
		ADM	5.577	6.772	7.968	42.57

the average values of the obtained outputs are presented in the literature, which is more acceptable. In terms of processing time, it is observed that the IAE objective has a slight advantage in the study. Additionally, the RMP optimization algorithm is found to have a success rate ranging from 58% to 91.23% across all measurement methods. Thus, a success rate of 91.23% is achieved in the modeling aimed at the COVID-19 pandemic. The desired outcome of this study is to contribute to the field of healthcare.

In future studies, changing the modeling method is aimed at achieving more successful results. Contributions to the field of healthcare are crucial in today's world. Therefore, one of the fundamental duties of every individual should be to serve society and humanity.

ACKNOWLEDGEMENT

The author received financial support from Van Yüzüncü Yıl University Scientific Research Projects Coordination Unit for the research, authorship and/or publication of this article.

CONFLICT OF INTEREST

Authors approve that to the best of their knowledge, there s not any conflict of interest or common interest with an institution/organization or a person that may affect the review process of the paper.

The authors of this study declare that they have received an ethical permission from the Van Yüzüncü Yıl University Dursun Odabaşı Medical Center dated 20.05.2021 and numbered 52545.

AUTHOR CONTRIBUTION

Faruk Ayata: Conceptualization, Methodology, Software, Validation, Writing- original draft. Ebubekir Seyyarer: Data curation, Visualization, Investigation, Supervision.

References

1. Raji P, Lakshmi G. R. D. Covid-19 pandemic analysis using regression. medRxiv. 2020. doi: 10.1101/2020.10.08.20208991.
2. Bandil S, Rathore S. Study of Haematological Parameters in Malaria. Asian J. Med. Res. 2019; volume(8): PT08-PT12. doi: 10.21276/ajmr.2019.8.3.pt3.
3. Demircan S. Öksürük Sesi Kayıtlarından Spektral Özellikler ile Otomatik COVID-19 Tespiti. Avrupa Bilim ve Teknoloji Dergisi. 2022. Ejosat Special Issue 2022 (ICAENS-1): 492-495. DOI: 10.31590/ejosat.1083052
4. Demir F. B. And Yılmaz, E. X-Ray Görüntülerinden COVID-19 Tespiti için Derin Öğrenme Temelli Bir Yaklaşım. Avrupa Bilim ve Teknoloji Dergisi. 2021. Ejosat Special Issue 2021 (RDCONF): 627-632. DOI: 10.31590/ejosat.1039522
5. Taş F, Özüdoğru O, & Bolatlı G. Bilgisayarlı Tomografi Bulguları

Negatif Olan Covid-19 Hastalarının; Epidemiyolojik, Klinik ve Laboratuvar Sonuçları Açısından Değerlendirilmesi. Selçuk Sağlık Dergisi. 2021. Covid-19 Özel:18-32. Retrieved from https://dergipark.org.tr/en/pub/ssd/issue/57170/757740

6. Şahin Gökçe H, Özensoy Güler Ö, Şimşek E, Karagülleoğlu Z, and Çarhan A. COVID-19 Hastalarında Yeni Bir Yaklaşım Olarak Oksihemoglobin Karboksihemoglobin, Kan Gazı Değerlerinin İncelenmesi. Longitudinal Bir Çalışma. Türk Yoğun Bakım Hemşireleri Derneği Yayın Organı. 2022; 26 (3): 92-99.
7. Dokeroglu T, Sevinc E, Kucukyilmaz T, and Cosar A. A survey on new generation metaheuristic algorithms. Comput. Ind. Eng., 2019. doi: 10.1016/j.cie.2019.106040.
8. Villarrubia G, De Paz J. F, Chamoso P, and De la Prieta F. Artificial neural networks used in optimization problems. Neurocomputing. 2018. doi: 10.1016/j.neucom.2017.04.075.
9. Pandey M, Zakwan M, Sharma P. K, and Ahmad Z. Multiple linear regression and genetic algorithm approaches to predict temporal scour depth near circular pier in non-cohesive sediment. ISH J. Hydraul. Eng. 2020. doi: 10.1080/09715010.2018.1457455.
10. Shihabudheen K. V, Mahesh M, and Pillai G. N. Particle swarm optimization based extreme learning neuro-fuzzy system for regression and classification. Expert Syst. Appl. 2018. doi: 10.1016/j.eswa.2017.09.037.
11. Zhijian L, Guangyao F, Dekang S, Di W, Jiacheng G, Shicong Z, Xinyan Y, Xianping L, Lei A. A novel distributed energy system combining hybrid energy storage and a multi-objective optimization method for nearly zero-energy communities and buildings. Energy. 2022; Volume (239). ISSN 0360-5442.
12. On O, AE-S. E, Mohamed T, ve Abualigah L. Ebola Optimizasyon Arama Algoritması: Doğadan Esinlenen Yeni Bir Metasezgisel Optimizasyon Algoritması. IEEE Access.2022; Volume (10): 16150-16177. doi: 10.1109/ACCESS.2022.3147821.
13. Ponce-Ortega J.M, Hernández-Pérez L.G. Optimization of Process Flowsheets through Metaheuristic Techniques. Springer, UK. 2019.
14. Güler E, & Yerel Kandemir S. Lineer ve Kübik Regresyon Analizleri Kullanılarak OECD Ülkelerinin CO2 Emisyonlarının Tahminlemesi. Avrupa Bilim ve Teknoloji Dergisi. Ejosat Special Issue 2022 (ICAENS-1). 2022; 175-180. DOI: 10.31590/ejosat.1079187
15. Özen N. S, Saraç S, & Koyuncu M. COVID-19 Vakalarının Makine Öğrenmesi Algoritmaları ile Tahmini: Amerika Birleşik Devletleri Örneği. Avrupa Bilim ve Teknoloji Dergisi. Ejosat. 2021; 134-139. DOI: 10.31590/ejosat.855113
16. Saadatmand S, Salimifar K, Mohammadi R. Predicting the necessity of oxygen therapy in the early stage of COVID-19 using machine learning. Med Biol Eng Comput. 2022; 60, 957–968. https://doi.org/10.1007/s11517-022-02519-x
17. Mohan S, A J, Abugabah A, M A, Kumar Singh S, Kashif Bashir A, Sanzogni L. An approach to forecast impact of Covid-19 using supervised machine learning model. Softw Pract Exp. 2021 Apr 1:10.1002/spe.2969. doi: 10.1002/spe.2969.
18. Pinter G, Felde I, Mosavi A, Ghamisi P, & Gloaguen R. COVID-19 pandemic prediction for hungary; a hybrid machine learning approach. Mathematics. 2020; Volume 8(6): 890. https://doi.org/10.3390/math8060890
19. Elaziz M. A, Hosny K. M, Salah A, Darwish M. M, Lu S, & Sahlol A. T. New machine learning method for image-based diagnosis of COVID-19. PLoS One. 2020; Volume 15(6). https://doi.org/10.1371/journal.pone.0235187
20. Haji S. H, & Abdulazeez A. M. Comparison of optimization techniques based on gradient descent algorithm: A review. PalArch's Journal of Archaeology of Egypt/Egyptology. 2021; Volume 18(4): 2715-2743.
21. Iqbal I, Odesanmi G. A, Wang J, & Liu L. Comparative Investigation of Learning Algorithms for Image Classification with Small Dataset. Applied Artificial Intelligence. 2021; Volume 35(10): 697-716. https://doi.org/10.1080/08839514.2021.1922841
22. Song C. Y, Pons A, & Yen K. AG-SGD: Angle-Based Stochastic Gradient Descent. Ieee Access. 2021; Volume (9): 23007-23024. https://doi.org/10.1109/Access.2021.3055993.
23. Truong T. T, & Nguyen H. T. Backtracking Gradient Descent Method and Some Applications in Large Scale Optimisation. Part 2: Algorithms and Experiments. Applied Mathematics and Optimization. 2021; Volume 84(3): 2557-2586. https://doi.org/10.1007/s00245-020-09718-8
24. Zaman S. M, Hasan D. M, Sakline R. I, Das D, & Alam M. A. A Comparative Analysis of Optimizers in Recurrent Neural Networks for Text Classification. 2021 Ieee Asia-Pacific Conference on Computer Science and Data Engineering (Csde). 2021. https://doi.org/10.1109/Csde53843.2021.9718394
25. Abdul-Adheem W. R, Ibraheem I. K, Humaidi A. J, Alkhayyat A, Maher R. A, Abdulkareem A. I, & Azar A. T. Design and analysis of a novel generalized continuous tracking differentiator. Ain Shams Engineering Journal. 2021.
26. Seyyarer E, Karci A, & Ates A. Effects of the stochastic and deterministic movements in the optimization processes. Journal of the Faculty of Engineering and Architecture of Gazi University. 2022; Volume37(2):949-965. https://doi.org/10.17341/gazimmfd.887976
27. Venkatesh B, & Anuradha J. A review of feature selection and its methods. Cybernetic and information Technologies. 2019; Volume 19(1): 3-26.
28. Khaire U. M, & Dhanalakshmi R. Stability of feature selection algorithm: A review. Journal of King Saud University-Computer and Information Sciences. 2022; Volume 34(4): 1060-1073.
29. Remeseiro B, & Bolon-Canedo V. A review of feature selection methods in medical applications. Computers in biology and medicine. 2019; Volume (112).
30. Zebari R, Abdulazeez A, Zeebaree D, Zebari D, & Saeed J. A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. Journal of Applied Science and Technology Trends. 2020; Volume 1(2): 56-70.
31. Şener Y. Makine Öğrenmesinde Değişken Seçimi (Feature Selection) Yazı Serisi: Sarmal (Wrapper) Yöntemler ve Python Kodları. 2020. Erişim Adresi: https://yigitsener.medium.com