Journal of Applied Intelligent Systems & Information Sciences
Vol. 3, pp. 17-35, October 2022.
Available at: **www.journal.research.fanap.com**
DOI: https://doi.org/10.22034/j.jaisis.2022.327647.1042

Prediction of Daily New Death Cases Due to Covid-19 Using Artificial Neural Networks and Genetic Algorithm

Milad Shahvaroughi Farahani[1,*], Amirhossein Esfahani[2]

[1]*Department of Finance, Faculty of Management, Khatam University, Tehran, Iran*

[2]*Department of Accounting, Eslamshahr University, Tehran, Iran*

ABSTRACT

The main goal of this article is to predict the mortality rate of Covid-19 using an artificial neural network (ANN) and genetic algorithm (GA). GA and ANN have been used as feature selection and prediction models respectively. We have predicted new death cases in the Netherlands using an artificial neural network and genetic algorithm. Different technical indicators such as short moving average (SMA), moving average convergence divergence (MACD), relative strength index (RSI), etc. have been used as input variables along with daily historical data. The result shows that machine learning (ML) can be a powerful tool in different subjects such as forecasting the Covid-19 mortality rate. On the other hand, GA can be a powerful method as a feature selection method and facilitating calculations. Using this technique for Covid-19 prediction is the novelty of this paper. We tried to use a different statistical population.

***KEYWORDS*:** Artificial Intelligence, Artificial Neural Network, Covid-19, Genetic Algorithm, Machine Learning.

1. Introduction

Artificial Intelligence (AI) has been around us for about 70 years, but its possibilities have significantly risen in today's world and each time, we can see new capabilities and applications of AI which have made our work easier than before (Manju, 2019). Undoubtedly, the major challenge faced by many researchers is uncertainty. This kind of uncertainty introduces an inescapable risk factor that has been in theories consistently (Asgharnezhad et al, 2022). When you try to predict something, for example, financial, economic, social, or other events, there are multiple factors such as internal or external factors which impact your model, however, you cannot consider all of them. At the same time, you need to consider or create some hypotheses or assumptions for facilitating the situation and solution. Therefore, uncertainty is the nature of each predictive model. In some problems and issues, there is too much data (which often is called *Big Data*) for decision-making about issues such as prediction of Covid-19 new cases and new death cases, prediction of stock price, customer credit rating, etc. So, it could be hard and complicated for decision-making and it should be time-consuming because of the huge number of variables and parameters which exist (Smith & Nobanee 2020). AI and machine learning can improve and overcome these challenges. On the other hand, it can optimize and boost your solutions and save your time because it can consider different parameters and conduct complicated tasks. Machine learning is an application of AI that enables systems to learn and improve by experience without being explicitly programmed. Machine learning focuses on developing computer programs that can access data and use it to

[*] *Corresponding Author, Email: m.shahvaroughi@khatam.ac.ir*

Received: 12 February 2022; **Accepted**: 19 July 2022; **Published Online**: 25 July, 2022

learn. These processes include learning (the acquisition of information and rules for using the information), reasoning (using the rules to reach approximate or definite conclusions), and self-correction. In addition, AI can be the science and engineering of making machines intelligent, especially intelligent computer programs. Therefore, AI can be characterized as a series of systems, methods, and technologies that display intelligent behavior by analyzing their environments and taking actions with some degree of autonomy toward achieving pre-specified outcomes (Biallas & O'Neill 2020).

AI includes a lot of subfields and subsets. So, people might confuse or make a mistake and misuse the word interchangeably or incorrectly. So, it is better to know the concepts behind some words such as machine learning, artificial neural network, expert system, deep learning (DL), etc. (Shahvaroughi Farahani 2019). We have covered these words which are about AI architecture in the literature review section.

The general objective of AI is to make machines such as computers think and act like a human or do things that require intelligence when done by humans (Lake et al., 2017).

Machine learning is one of the significant and popular applications of AI. Like AI, machine learning acts based on learning and recognition but ML, for example, can make predictions using statistical algorithms and performs tasks beyond what it was explicitly programmed for. The core idea of ML is to generalize from experience which can happen with repetition and duplicating tasks. ML has many applications such as speech recognition, fraud detection, natural language processing, dynamic pricing, and others. Data is the most vital factor when using ML algorithms because the learning process depends on input data.

Today, there is a huge amount of data around the world about every addressable problem. So, we need to extract significant features and find the related parameters to increase the accuracy of the model. This process is called data mining. As a result, different models and algorithms have emerged. Therefore, in this article, we used an artificial neural network as a prediction method. On the other hand, GA has been used as a feature selection. As we mentioned earlier, some technical indicators have been used as input variables. As it is clear, extra explanations about considered algorithms and variables are presented in the methodology section.

One of the main issues in the 21st century is the emergence of Covid-19. Covid-19 is a kind of virus that has many casualties. So, if it can be well predicted, the devastating effects and costs can be reduced (Sathyamala, 2022). Covid-19 has had many devastating effects such as death, business closures, recession, rising depression, and so on. Now, by using AI and machine learning methods along with global vaccination, the rate of new death cases is decreased but it still exists. By applying AI-based and ML models, it is possible to predict the next outbreak waves, as well as the required vaccination, and announce the end time of the Corona, and so on.

As we mentioned earlier, the main problem is the prediction of Covid-19 new death cases. As a solution, AI-based solutions are used for Covid-19 prediction which can be considered the main contribution of the paper. Then, GA is used for feature selection and finding the most related and appropriate indicators. Therefore, ANN is calculated and run before and after GA. This can help to recognize GA impact and its role in calculations and network performance. Finally, different loss functions such as MSE, and AE. etc., are used.

The structure of the paper is as the following:

The 1st part is the introduction, some information about AI and ML and their importance. The 2nd section is about ML background and overview, or literature review, AI architecture, and its sub-branches. The 3rd section is about using ANN and GA as feature selection and prediction models and using these two models to predict daily new death cases in the Netherlands. After that, there are a conclusion and remarks, and further recommendations. Finally, references and usable resources.

2. Overview and Background

In this part, we have tried to look at the history and evolution of ML along with AI as complementary material. On the other hand, we have reviewed some research and papers about the application of AI and ML in the prediction of Covid-19.

The history of ML in figure 1 with related and expert phrases is presented in the following picture (Dramsch, 2020):

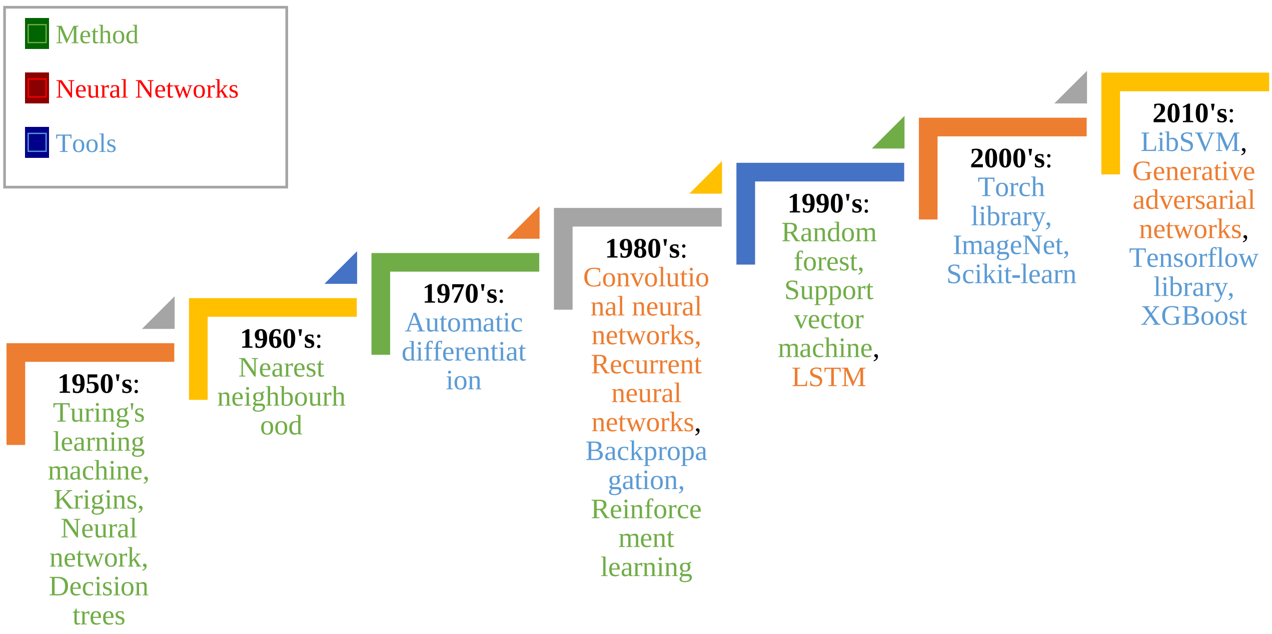

Fig 1. ML history

Green colors, red colors, and blue colors are a method, neural network, and tools respectively. As it is observable, in each phase, ML progressed and it is now more thorough than before while it could do different and various tasks such as graphical tasks, reasoning base tasks, decision making, etc.

During the early stages of machine learning (ML), experiments had been carried out involving theories of computers recognizing patterns in data and learning from them. Today, after building upon those foundational experiments, machine learning is becoming more complex. While machine learning algorithms have been around us for a long time, the ability to apply complex algorithms for big data applications more rapidly and effectively is a more recent development.

ML can be used in automating manual data entry to more complex tasks such as fraud detection, insurance risk assessment, etc. A major part of what makes machine learning so valuable is its ability to detect what the human eye misses. Machine learning models can catch complex patterns that would have been overlooked during human analysis.

Almost any task that can be completed with a data-defined pattern or a set of rules, can be automated with machine learning. The following table shows the ML techniques[2].

Table 1. ML components

Name	Definition
Regression	Regression methods help to predict or explain a particular numerical value based on a set of prior data.
Classification	Classification methods predict or explain a class value
Clustering	Group or cluster observations that have similar characteristics
Dimensionality reduction	Remove the least important information (sometimes redundant columns) from a data set

[2] https://towardsdatascience.com/

Ensemble methods	Combining several predictive models (supervised ML) to get higher quality predictions more than each of the models could provide on its own.
Neural network (NN) and deep learning (DL)	The objective of neural networks is to capture non-linear patterns in data by adding layers of parameters to the model
Transfer learning	It refers to re-using part of a previously trained neural net and adapting it to a new but similar task.
Supervised learning	Allows you to collect data or produce a data output from a previous ML deployment
Unsupervised learning	Helps you find all kinds of unknown patterns in data.
Reinforcement learning	Taking suitable action to maximize reward in a particular situation.
Natural language processing	A widely used technique to prepare the text for machine learning

For a better understanding of the relationship between categories and AI components, figure 2 is depicted (Mohammadrezaei et al., 2019).

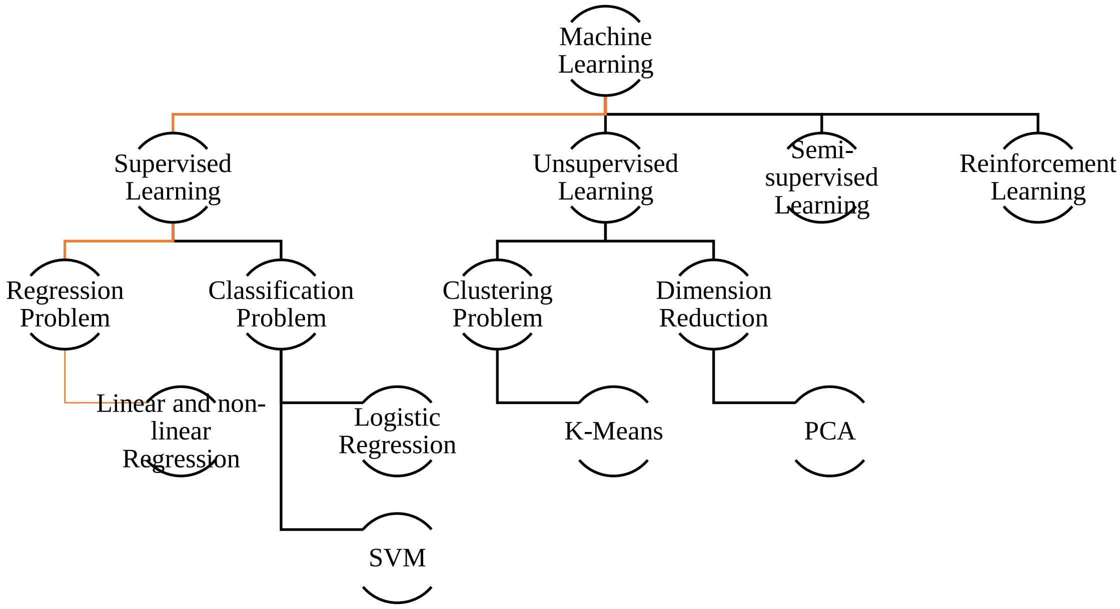

Fig 2. ML subsets

One of the main phenomena in the last years was the emergence of Covid-19. At first, it was almost an unknown disease. Gradually and step by step, we could get some knowledge from different dimensions such as economic, social, medical, etc. (Berenbaum, 2021). Because Covid-19 is related to human health and can be so dangerous, we need to pay attention to different guidelines such as washing hands, wearing a mask, vaccination, social distancing, etc. (Trogen & Caplan, 2021).

Now, there are different kinds of vaccines such as Pfizer, Moderna, Sputnik V, Barekat, Sinopharm, etc. from different countries around the world. The mortality rate due to Covid-19 has increased but these vaccines are not sufficient, and they create immunity in a short period. So, we need to inject booster dose seasonally.

Occasionally, the prevalence and death rate of Covid-19 disease increases for a variety of reasons, including poor hygiene and reduced vaccine safety. These factors increase new cases and new death cases. So, it can cause many social, behavioral, and economic problems such as inequality and exclusion, psychological or behavioral disorders, or recession. As a result, AI-based, and ML models can boost the possibility of the prediction of Covid-19 new waves and new cases.

Because Covid-19 is a significant issue along with different dimensions and causes and effects, to get more information about Covid-19 please see Barouki et al., 2021.

The next table presents different research about the applications of AI in the prediction of Covid-19. In this table, the latest and recent papers about the applications of ML algorithms in the prediction of Covid-19 mortality rate, new cases, and the latest achievements have been shown.

Table 2. Literature review

No	Author(s) (year)	Journal Name	Objectives	Findings
1	Dos Santos Santana et al., (2021)	Journal of Medical Internet Research	Effectively prioritize patients who are symptomatic for testing to assist in early Covid-19 detection in Brazil, addressing problems related to inefficient testing and control strategies	Gender, fever, and dyspnea are the main sign which has been classified by models. KNN and LR outperformed other models. The DT was considered the most suitable model.
2	Zawbaa et al., (2021)	International Journal of Clinical Practice	Prediction and forecasting of different countries' daily confirmed cases and daily death-cases	The results proved useful in modeling and forecasting the end status of the virus spreading based on specific regional and health support variables.
3	Gray et al., (2021)	BMJ Health & Care Informatics	Training machine learning models to predict Covid-19 cases growth and understanding the social, physical, and environmental risk factors associated with higher rates of SARS-CoV-2 infection in Tennessee and Georgia counties	African American and Asian racial demographics present comparable and contrasting, patterns of risk depending on locality
4	Rios et al., (2021)	Scientific reports	Presented a temporal analysis of the number of new cases and deaths among countries using artificial intelligence	1. They showed the historical infection path taken by specific countries and emphasize changing points that occur when countries move between clusters with small, medium, or large numbers of cases. 2. They estimated new waves for specific countries using the transition index.
5	Malki et al., (2021)	Environmental science and pollution research	Applying machine learning approaches to predict the spread of Covid-19 in many countries.	Covid-19 infections will greatly decline during the first week of September 2021 when it will be going to an end shortly afterward.
6	Anđelić et al., (2021)	Health informatics journal	Using GA to obtain mathematical models for estimation of epidemiology curve	The estimated epidemiology curve for each country obtained from considered equations was almost identical to the real data contained within the data set.
7	Muhammad et al., (2021)	SN computer science	Prediction of Covid-19 infection (positive and negative cases in Mexico) using ML algorithms such as logistic regression,	The decision tree model has the highest accuracy of 94.99% while the support vector machine model has the highest sensitivity of 93.34% and the

			decision tree, SVM, naïve Bayes, and ANN.	Naïve Bayes model has the highest specificity of 94.30%.
8	Kwekha-Rashid et al., (2021)	Applied Nanoscience	Detect the role of machine-learning applications and algorithms in investigating various purposes that deal with Covid-19.	The results show that machine learning can produce an important role in Covid-19 investigations, prediction, and discrimination. machine learning can be involved in health provider programs and plans to assess and triage Covid-19 cases. Supervised learning showed better results than other unsupervised learning algorithms by having 92.9% testing accuracy. In the future recurrent supervised learning can be utilized for superior accuracy.
9	Pinter et al., (2020)	Mathematics	Using machine learning approaches such as ANFIS and MLP-ICA to predict Covid-19 in Hungary.	The models predict that by late May, the outbreak and the total mortality will drop substantially.
10	Kumar et al., (2020)	MedRxiv	Prediction of Covid-19 using chest X-Ray images through deep feature learning model with SMOTE and machine learning classifiers	The model achieved an accuracy of 0.973 on random forest and 0.977 using XGBoost predictive classifiers.
11	Tuli et al., (2020)	Internet of Things	Predicting the growth and trend of the Covid-19 pandemic using machine learning and cloud computing	Using the proposed Robust Weibull, the results show that the considered model can make statistically better predictions than the baseline. The baseline Gaussian model shows an over-optimistic picture of the Covid-19 scenario.
12	Xiong et al., (2022)	Infectious diseases of poverty	Comparing different machine learning techniques for predicting Covid-19 severity	Among the three models, RF yielded better overall performance with the highest AUC of 0.970 than SVM of 0.948 and LR of 0.928, RF also achieved a favorable sensitivity of 96.7%, specificity of 69.5%, and accuracy of 84.5%. SVM had a sensitivity of 93.9%, specificity of 79.0%, and accuracy of 88.5%. LR also achieved a favorable sensitivity of 92.3%, specificity of 72.3%, and accuracy of 85.2%. The results indicated that RF could be a useful predictive tool to identify patients with severe Covid-19, which may facilitate effective care and further optimize resources.
13	Kuo et al., (2022)	International Journal of Medical Informatics	The accuracy of machine learning approaches using non-image data for the prediction of Covid-19: A meta-analysis	The results show that non-image data can be used to predict Covid-19 with acceptable performance. Further, class imbalance and feature selection are suggested to be incorporated whenever building models for the prediction of Covid-19, thus improving further diagnostic performance.
14	Mohan et al., (2022)	Computers in Biology and Medicine	predicting the impact of the third wave of Covid-19 in India using hybrid statistical machine learning models: A time series forecasting and sentiment analysis approach	A spike in daily confirmed and cumulative confirmed cases was predicted in India in the next 180 days based on the past time series data. The results were validated using various analytical tools and evaluation metrics, producing a root mean square error (RMSE) of 0.14 and a mean absolute percentage error (MAPE) of 0.06. The NLP processing results revealed negative sentiments in most articles and blogs, with few exceptions.

3. METHODOLOGY

As a case study, we are going to predict new death cases in the Netherlands using ML tools such as artificial neural networks and genetic algorithms. So, we used daily data from Feb 2020 to April 2021. Different indicators such as short moving average (SMA), exponential moving average (EMA), and moving average convergence divergence (MACD) have been used as input variables. These indicators are derived from the technical analysis that is used to predict stock prices (Gocken et al., 2016). There is the main difference between our paper and Gocken et al., 2016 article. On one hand, these indicators are used as input variables **for the prediction of Covid-19 new death cases** which can be a kind of novelty. On the other hand, we have tried to train ANN before and after using GA to compare the results and get a better sense means the impact of GA on ANN to obtain better results. Because, in different studies and technical analyses these indicators are important and the target variable had a significant correlation to the input variable; so, we tried to consider different types of input variables whose fundamental is based on the output. We used ANN as a prediction method. On the other hand, GA has been used for feature selection and choosing the most important input variables.

There are considered input variables in table 3 which are used to improve and increase the predictability.

Table 3. Indices as input variables with formulas

No	Indicator	Formula
1	Short Moving Average (SMA-5)	SMA (5) $=\frac{(DDC1+DDC2+\cdots+DDC5)}{5}$
2	Short Moving Average (SMA-6)	SMA (6) $=\frac{(DDC1+DDC2+\cdots+DDC6)}{6}$
3	Short Moving Average (SMA-10)	SMA (5) $=\frac{(DDC1+DDC2+\cdots+DDC10)}{10}$
4	Short Moving Average (SMA-20)	SMA (5) $=\frac{(DDC1+DDC2+\cdots+DDC20)}{20}$
5	Exponential Moving Average (EMA-5)	EMA (5) $_{Today}=\frac{\text{DDC Today}*\text{k}+\text{EMA}(5)Yestarday*(1-k)}{5}$ K$=\frac{2}{5+1}.EMA(5)0 = SMA(5)$
6	Exponential Moving Average (EMA-6)	EMA (6) $_{Today}=\frac{\text{DDC Today}*\text{k}+\text{EMA}(6)Yestarday*(1-k)}{6}$ K$=\frac{2}{6+1}.EMA(5)0 = SMA(6)$
7	Exponential Moving Average (EMA-10)	EMA (10) $_{Today}=\frac{\text{DDC Today}*\text{k}+\text{EMA}(10)Yestarday*(1-k)}{10}$ K$=\frac{2}{10+1}.EMA(10)0 = SMA(10)$
8	Exponential Moving Average (EMA-20)	EMA (20) $_{Today}=\frac{\text{DDC Today}*\text{k}+\text{EMA}(20)Yestarday*(1-k)}{20}$ K$=\frac{2}{20+1}.EMA(20)0 = SMA(20)$
9	Triangular Moving Average (TMA-5)	TMA (5) $=\frac{(SMA(1)+SMA(2)+\cdots+SMA(5))}{5}$
10	Triangular Moving Average (TMA-6)	TMA (6) $=\frac{(SMA(1)+SMA(2)+\cdots+SMA(6))}{6}$
11	Triangular Moving Average (TMA-10)	TMA (10) $=\frac{(SMA(1)+SMA(2)+\cdots+SMA(10))}{10}$
12	Triangular Moving Average (TMA-20)	TMA (20) $=\frac{(SMA(1)+SMA(2)+\cdots+SMA(20))}{20}$
13	M-death cases	M_{DDC}=DDC $_{Today}$-DDC $_{Yesterday}$
14	ACC-death cases	Acc_{DDC}=M DDC $_{Today}$-M DDC $_{Yesterday}$
15	MACD	MACD=EMA (12)-EMA (26)
16	RSI	RSI=100-$\frac{100}{1+RS}.RS = \frac{Average\ positive}{Average\ negative}$
17	Daily Death Cases (DDC)	-

As you can see, SMA (5) means the average of five last death cases. DDC stands for Daily Death Cases. Because these indicators are derived from technical analysis, the idea and results can be interesting.

3.1. *Artificial Neural Network (ANN)*

ANN has different applications such as function approximation, clustering, etc. (Abiodun et al., 2018). In this paper, we used ANN as a prediction method. The efficiency of the neural network is not hidden from anyone, and in recent years we have seen the increasing use of neural network-based methods in various fields.

Our team and others have checked the applications of ANN in different fields or industries such as economic, social, mathematics, medical, etc. along with brilliant results. Since ANN has some characteristics and qualifications such as speeding up the calculation, being user-friendly, compatibility with complex data structures, and so on, ANN is used to predict daily new death cases in Netherland.

As it is clear, ANN contains three layers: 1. Input layer 2. Hidden layer 3. Output layer. As we mentioned earlier (Table 3), 16 indicators have been used as input variables. The number of hidden layers is obtained through trial and error. We used a multilayer perceptron (MLP) neural network along with backpropagation (BP) training. Levenberg-Marquardt (ML) is used as an optimization algorithm to find the minimum error point. 1000 iterations are used. Probably, there is no clear formula for setting the optimum iteration number, but you can figure out this issue by an iterative process by initializing the iteration number by a small number like 100 and then increasing it linearly. This process will be repeated until the MSE of the test does not decrease and may even increase. In the case of no improvement observation, we will increase the iterations and retrain the network again. You should iterate until the error does not significantly decrease. To avoid overfitting, you should split the data set in training and test, and the error should be similar for both subsets. We used 70% and 30% of data for training and testing datasets.

Figure 3 represents the research methodology.

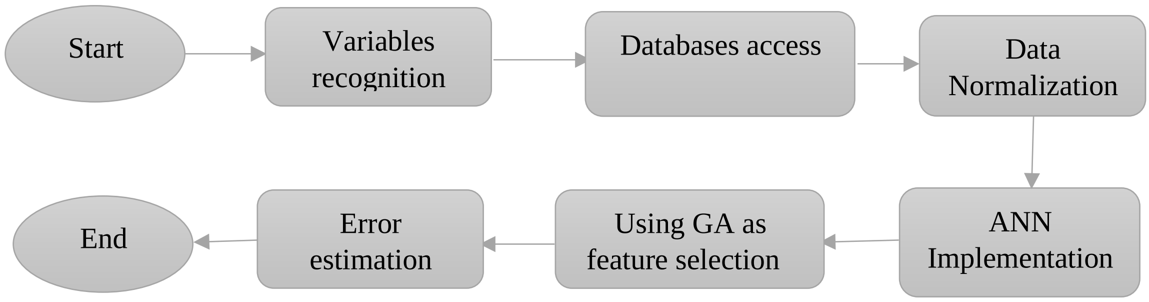

Fig 3. Research Methodology

Firstly, we need to select the most appropriate variables as inputs. The second step is finding a validated dataset for accessing data. To prepare data as input variables, it needs to be preprocessed and normalized. So, they are scaled. As a result, very large and very small-scale data is removed. Now, it is possible to implement ANN as a predictive model. Then, GA is used for feature selection and finding the most related and important variables. After selecting and determining the considered inputs, model fitting and error estimation are presented.

Figure 4 presents the architecture of the proposed neural network (Gocken et al., 2016).

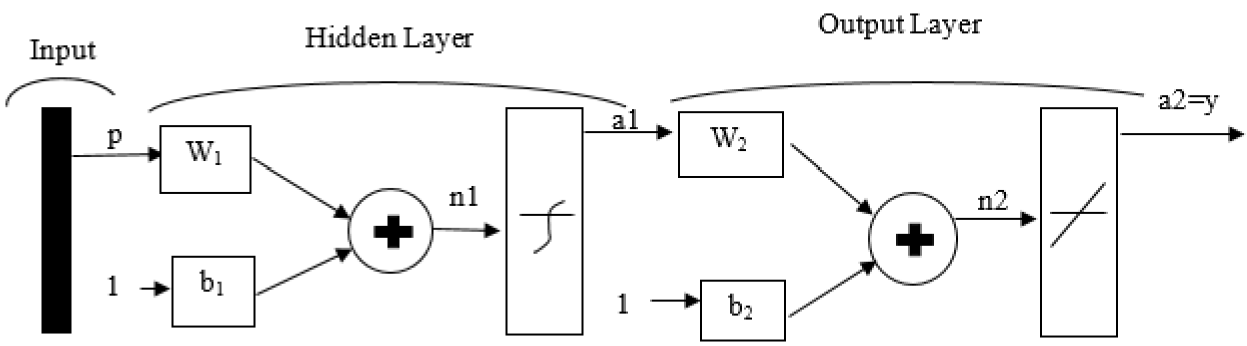

Fig 4. The architecture of the proposed Neural Network

In figure 4, P is the input pattern, b_1 is the vector of bias weights on the hidden neurons, and W_1 is the weight matrix between the 0^{th} (i.e., input) layer and 1^{st} (i.e., hidden) layer. a_1 is the vector containing the outputs from the hidden neurons, n_1 is the vector containing net inputs going into the hidden neurons, a_2 is the column vector coming from the second output layer, and n_2 is the column vector containing the net inputs going into the output layer. W_2 is the synaptic weight matrix between the 1^{st} (i.e., hidden) layer and the 2^{nd} (i.e., output) layer and b_2

is the column-vector containing the bias inputs of the output neurons. Each row of the W_2 matrix contains the synaptic weights for the corresponding output neuron.

First, we need to normalize data using the following equation:

$$\widetilde{S_i} = \frac{(S_i - S_{min})}{S_{max} - S_{min}} . i = 1 \dots . N \quad (1)$$

In equation 1, numerator i is the amount of data.

3.2. Genetic Algorithm (GA)

GA is an evolutionary algorithm that imitates biological evolution behavior (Katoch et al., 2021). Before any computations, we need to be confident about the amount and structure of data means a structure, missing value, unrelated data, and other cases. One of the main methods which are very common along with brilliant results is GA. Because GA creates multiple optimal solutions, requires less information, a large set of solution space, etc. it is very ideal and suitable for feature selection. Binary coding has been used along with 22 bits which 17bits show the existence or nonexistence of a variable. 5 other bits are used to find the number of the hidden layer (2^5=32). We start the algorithm by using a 20 random number population as the initialization process. The best population (with the lowest error) remains in each epoch until the termination condition. 100 epochs have been used and it can increase due to conditions. Like ANN, 70% and 30% of data are used for training and testing respectively. To get better results, crossover and mutation have been used. This process is as the following:

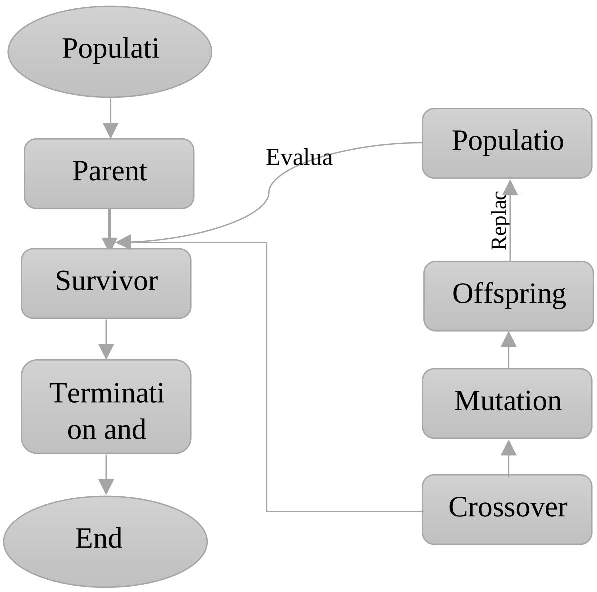

Fig 5. Genetic algorithm process

Among 20 parents and 20 children, 20 best individuals will be selected as new generations. This will repeat until the termination condition. For example, reaching the 100 best generations. The maximum number of generations is 2000. The crossover of two-parent strings which is called offspring leads to new solutions. The crossover rate is usually between 0.8-0.9.

Table 4 represents the considered parameters and some details about tuning:

Table 4. GA parameters

Output Error	**Output Activation Function**	**Input Activation Function**	**Mutation Rate**	**Crossover Rate**	**Number of Generation**	**Population size**	**Max Itr**
MSE	Logistic	Logistic	0.1	0.9	50	20	2000
Selection parents			**Mutation**		**Crossover**		
Roulette wheel method			Binary Method		One-point method		

4. Findings and Results

4.1. Artificial Neural Network (ANN)

As the first step, we need to normalize data. As we mentioned earlier, 17 variables are used: 16 variables as input and a variable as the target.

256 instances are used for training and 108 instances for validation and testing.

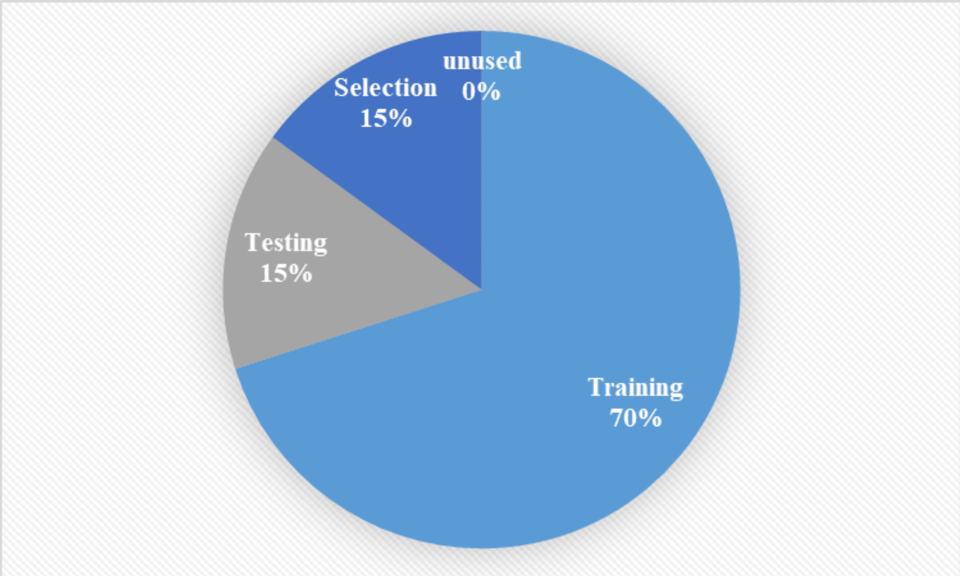

Fig 6. Data pie chart

The correlation coefficient is important between variables and can give some beneficial information about the relationship between inputs and target variables. Fig7 shows the correlation coefficient between variables:

	SMA-20	SMA5	SMA6	SMA10	TMA5	TMA6	TMA10	TMA20	MCLOSE	ACCCLOSE	MACD	RSI	EMA5	EMA6	EMA10	EMA20
SMA-20	1	0.92	0.94	0.97	0.96	0.97	0.99	1	0.011	0.031	-0.019	0.19	0.75	0.78	0.88	0.99
SMA5		1	0.99	0.98	0.98	0.98	0.96	0.89	0.052	0.026	-0.049	-0.027	0.89	0.91	0.96	0.93
SMA6			1	0.99	0.99	0.99	0.97	0.9	0.012	0.024	-0.074	-0.026	0.89	0.91	0.96	0.94
SMA10				1	0.99	0.99	0.99	0.94	0.027	0.016	-0.026	0.059	0.85	0.87	0.95	0.96
TMA5					1	1	0.99	0.94	-0.026	0.045	-0.055	0.018	0.86	0.88	0.95	0.96
TMA6						1	0.99	0.95	-0.021	0.038	-0.047	0.039	0.85	0.87	0.95	0.97
TMA10							1	0.97	-0.011	0.041	-0.037	0.094	0.81	0.83	0.92	0.98
TMA20								1	0.003	0.033	-0.0058	0.24	0.7	0.73	0.84	0.98
MCLOSE									1	-0.025	-0.22	0.21	-0.073	-0.083	-0.045	-0.024
ACCCLOSE										1	0.11	-0.052	0.041	0.039	0.029	0.037
MACD											1	0.21	-0.11	-0.12	-0.075	-0.043
RSI												1	-0.23	-0.21	-0.1	0.14
EMA5													1	1	0.97	0.81
EMA6														1	0.98	0.83
EMA10															1	0.91
EMA20																1

Fig 7. Input correlations

The minimum correlation is -0.230896 between the variable's RSI and EMA5. The maximum correlation is 0.998943 between the variables TMA5 and TMA6.

It is possible to analyze the correlation coefficient between inputs and target variables too because we can recognize the most important variables which can increase predictability.

Table 5. Input-target correlation

SMA5	EMA10	SMA6	SMA10	RTMA6	TMA5	EMA6	EMA5
0.8825	0.8769	0.8755	0.8743	0.8563	0.8559	0.8375	0.8289
TMA10	**EMA20**	**SMA20**	**TMA20**	**MCLOSE**	**ACCCLOSE**	**MACD**	**RSI**
0.8222	0.7994	0.7921	0.7561	0.2472	0.1015	0.0654	0.0607

The network structure including three layers and the number of hidden layers is obtained through trial and error. In the beginning, we supposed 10 neurons as the following figure:

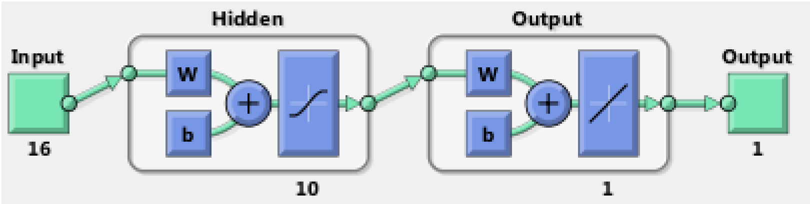

Fig 8. Network structure

After searching and trying different layers, the considered and suitable structure is as figure 9:

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared	Stop Reason
1	[16-2-1]	37	0.121901	6.729333	6.832392	8.203398	-820.524272	0.973997	0.948095	All iterations done
2	[16-40-1]	721	0.120449	3.472768	6.384306	8.302291	383.417742	0.993432	0.986556	All iterations done
3	[16-25-1]	451	0.104142	6.323607	7.411357	9.602284	-7.946447	0.976228	0.952612	All iterations done
4	**[16-16-1]**	**289**	**0.12485**	**4.432392**	**6.236379**	**8.009626**	**-420.073347**	**0.989137**	**0.978296**	**All iterations don**
5	[16-10-1]	181	0.108827	4.993647	6.055251	9.188855	-606.505036	0.985669	0.970937	All iterations done
6	[16-21-1]	379	0.115586	5.647566	7.124435	8.651552	-179.986609	0.981632	0.963144	All iterations done
7	[16-13-1]	235	0.096895	6.207115	6.910349	10.320496	-444.557665	0.976434	0.953194	All iterations done
8	[16-19-1]	343	0.111451	6.008173	6.499199	8.972532	-236.636408	0.978989	0.958349	All iterations done
9	[16-17-1]	307	0.103305	6.329805	6.846508	9.680088	-295.7035	0.976845	0.953739	All iterations done
10	[16-14-1]	253	0.110183	4.854208	7.955166	9.075774	-469.528549	0.987348	0.973934	All iterations done
11	[16-15-1]	271	0.096441	6.383535	7.424617	10.369051	-365.607259	0.975729	0.951638	All iterations done

Fig 9. Considered network structure

Figure 10 shows the regression and fitting data for training, validation, and testing. As it is clear, for all three parts R^2 is 0.9867 and it can be a good sign of well-training.

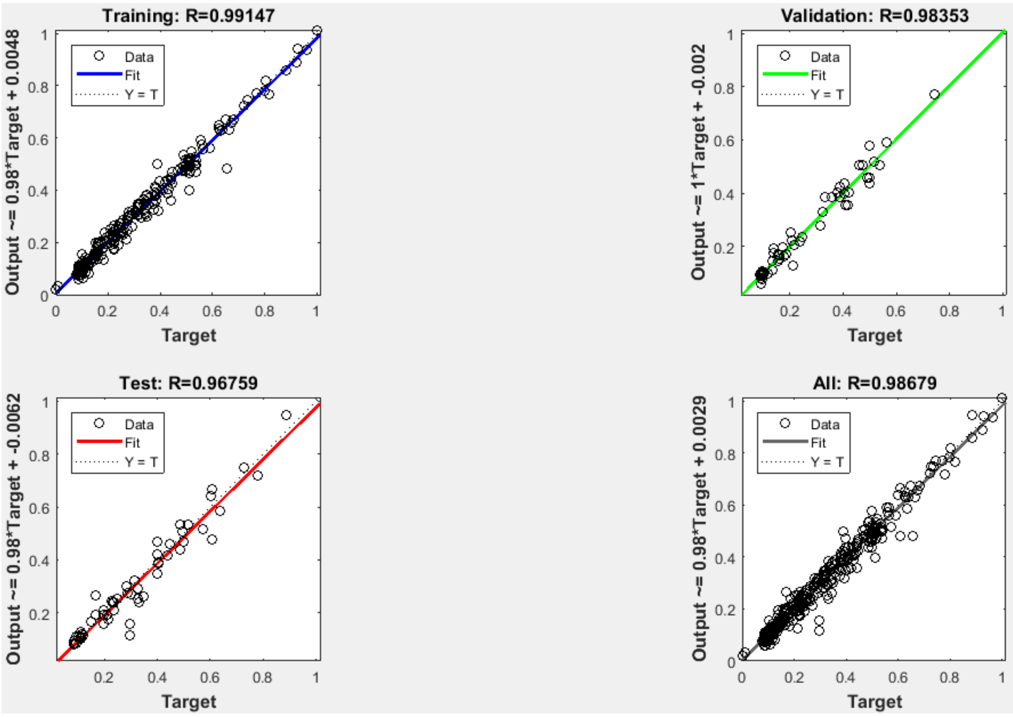

Fig 10. Regression and fitting data

Error estimation during each epoch has been depicted in figure 11.

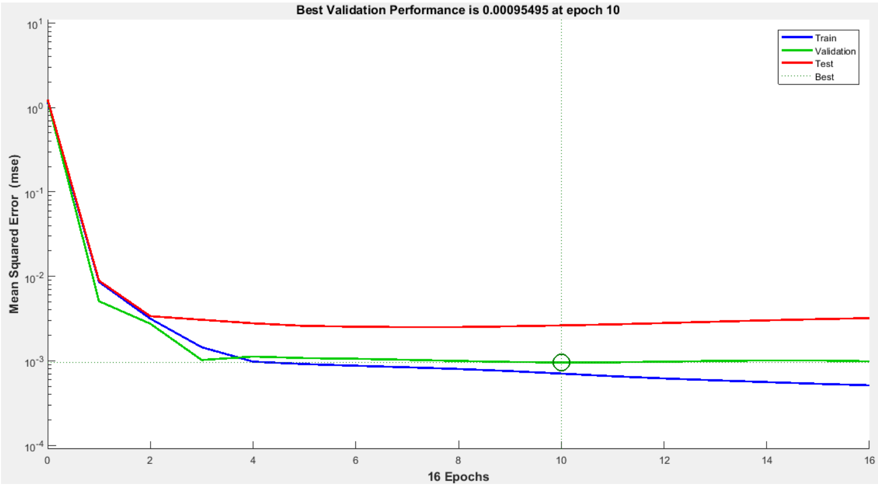

Fig 11. Training and testing error

The network considered the optimal solution and good training at epoch 10 with the best validation performance equal to 0.00095. For accessing more information about error distribution during each iteration and training error, please see the appendix (figure I to figure III). In figure I, the error is divided through 20 binaries. This figure shows the difference between the predicted and the actual value. Because these error values show how the predicted values are different from the target values, therefore, these values can be negative. Y-axis represents the number of samples from the database that are within a certain boundary. For example, there is a boundary with an error of 0.001012, which is approximately 80 in height for training and between 80 and 100 for verification and testing. This means that there are many instances and observations in the database that have errors in this range.

Figure II shows the training and learning during each epoch. The gradient is a measure of change in all weights concerning the change in error rate. The gradient is considered the slope of the function. A higher gradient means a steeper slope and the ability to learn faster, and a zero gradient means the model stops learning. μ is a control parameter used to train the neural network. μ selection directly affects the convergence error. In the Levenberg-Marquardt algorithm, μ must be a number between 0.8 and 1. In the neural network, it is used to estimate the inverse Hessian matrix, which is a complex function. Figure III shows the training NN parameters. Due to the lack of improvement in the results and network learning after 16 repetitions, network training and learning in the tenth repetition have stopped. Of course, one of the criticisms of the neural network is the rapid convergence and finding the relative minimum and getting caught in the trap of the relative optimal answer, which makes the absolute optimal answer not be achieved.

Finally, you can see the actual versus output regression in the following figure (Fig. 12).

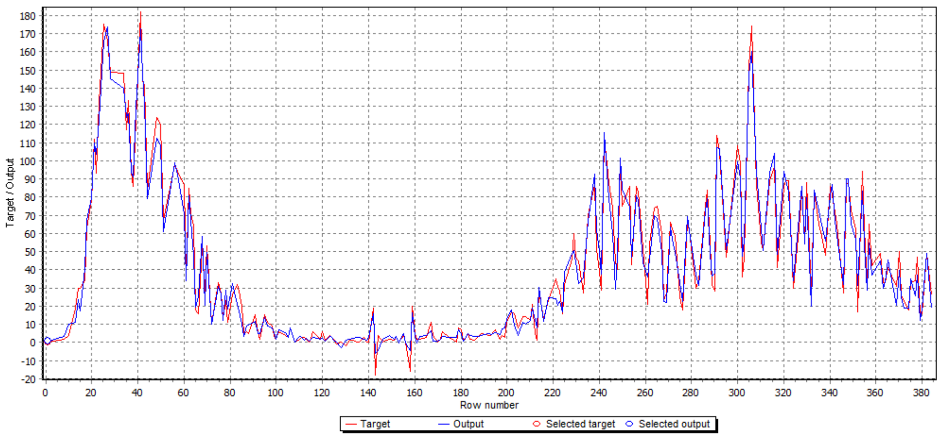

Fig 12. Actual vs. output plot

The red line shows the actual data while the blue line is the prediction or output. In most cases, there are some coverages and overlapping between the red line and the blue line.

4.2. Genetic Algorithm (GA)

In this paper, as we mentioned earlier, GA has been used as a feature selection. So, after running GA, considered variables have been recognized.

A graphical representation of the resultant deep architecture is depicted. It contains a scaling layer, a neural network, and an un-scaling layer. The yellow circles represent scaling neurons, the blue circles' perceptron neurons, and the red circles' un-scaling neurons. The number of inputs is 13, and the number of outputs is 1. The complexity, represented by the number of hidden neurons, is 3.

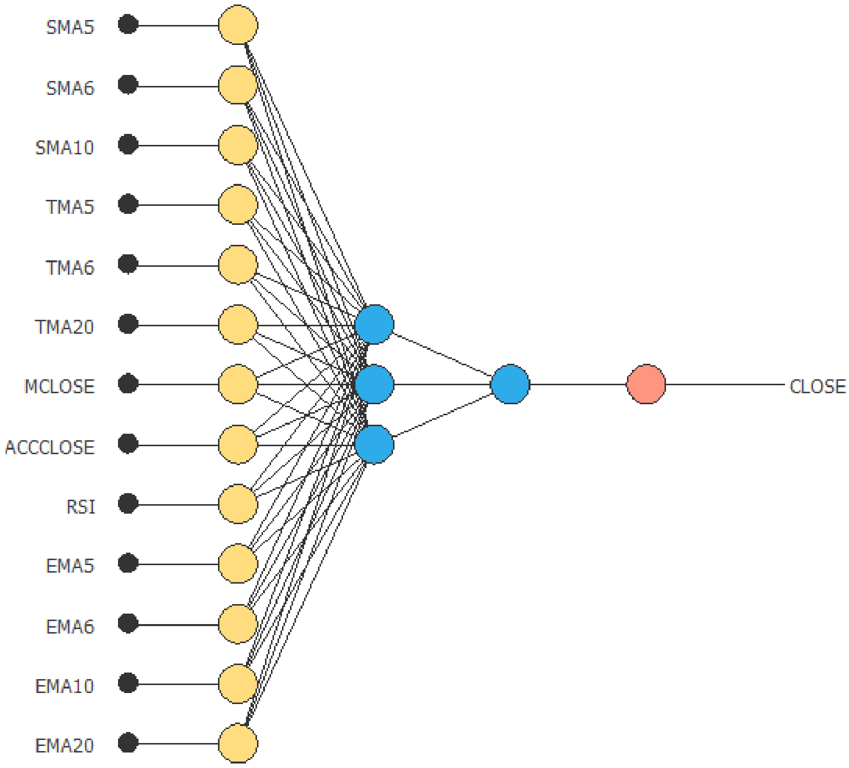

Fig 13. GA final architecture

More details about optimum training error, generation numbers, etc. can be observed in table 6.

Table 6. GA error results

Parameters	Value
The optimal number of inputs	13
Optimum training error	0.0198567
Optimum selection error	0.0303468
Generations number	100
Elapsed time	00:31

The following plot shows the training and selection error in each iteration. The blue line represents the training error, and the orange line represents the selection error. The initial value of the training error is 35.9619 and the final value after 527 epochs is 0.0152546. The initial value of the selection error is 45.836 and the final value after 527 epochs is 0.0370253.

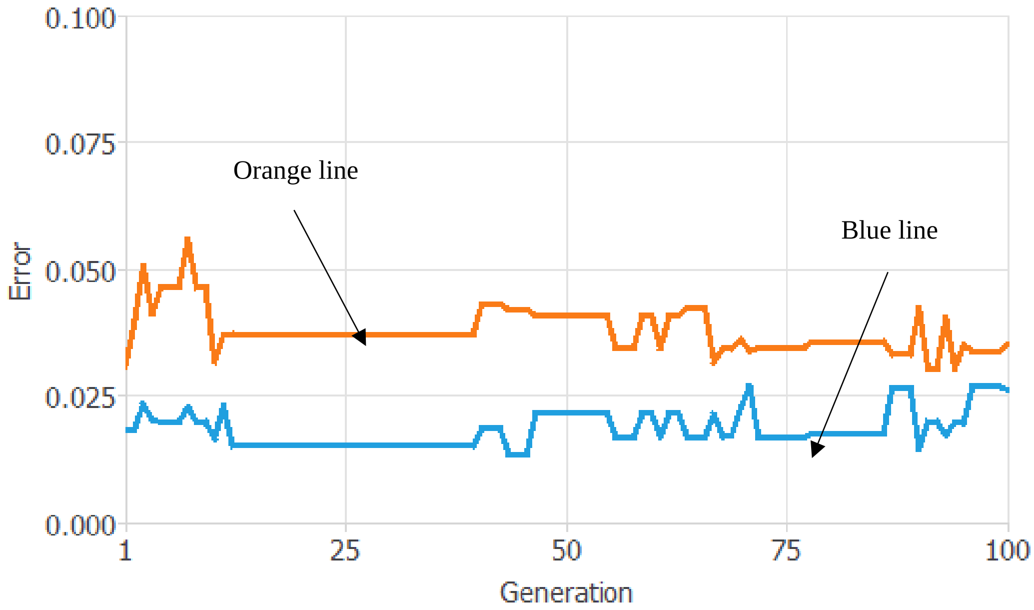

Fig 14. GA error plot

To get more information about the history of the mean of selection error in each generation during the GA input selection process and training network by using the quasi-Newton method, please see the appendix (figure IV & figure V). Figure IV shows the mean selection error per generation during the input selection process using the genetic algorithm. The initial value is equal to 0.12 and the final value after performing the desired repetitions is equal to 0.01. In figure 5, the blue line indicates the learning error, and its initial value after multiple repetitions is equal to 45.0124 and its final value is equal to 0.005592. The orange line indicates the selection error so that its initial value is 45.0124 and its final value after the desired repetitions is 0.005592.

Finally, different error estimations such as SSE, MSE, etc. have been presented in the following table:

Table 7. Loss functions

Criteria	Training	Selection	Testing
Sum Square Error	176.825	40.3461	30.9336
Mean Square Error	0.690723	0.747151	0.572844
Root Mean Square Error	0.831098	0.864379	0.756865
Normalized Square Error	17.1538	18.7584	18.502
Minkowski Error	186.292	41.8213	34.6514

4.3. Artificial Neural Network (after feature selection)

The main goal of this section is to examine the efficiency of GA as feature selection based on the experimental results. In summary, brief results are presented.

The last and final variables which were selected by GA in the previous section are 13.

The first step is training the network and finding the best structure. So, we trained the network, and the following results are achieved:

ID	Architecture	# of Weights	Fitness	Train Error	Validation Error	Test Error	AIC	Correlation	R-Squared	Stop Reason
3	[13-21-1]	316	0.194462	3.575968	4.893844	5.142395	-440.308677	0.991646	0.983236	All iterations done
4	[13-13-1]	196	0.192772	4.037988	4.55478	5.187477	-649.687974	0.988732	0.977559	All iterations done
5	[13-28-1]	421	0.228573	3.365977	5.173112	4.374976	-245.559157	0.992945	0.985887	All iterations done
6	[13-25-1]	376	0.181526	3.970714	4.778418	5.508865	-293.921784	0.990018	0.980041	All iterations done
7	[13-31-1]	466	0.194767	3.689375	4.271657	5.13435	-132.440926	0.990063	0.980182	All iterations done
8	[13-29-1]	436	0.214629	3.52565	4.695405	4.659208	-203.879836	0.991811	0.983624	All iterations done
9	[13-26-1]	391	0.188624	4.93394	4.949902	5.301552	-209.1894	0.984848	0.969182	All iterations done
10	**[13-27-1]**	**406**	**0.228956**	**3.775817**	**4.450932**	**4.367649**	**-246.604705**	**0.99079**	**0.981598**	**All iterations done**

Fig 15. Top 5 best network structures

The best network structure is [13-27-1] means 13 input variables, 27 nodes in a hidden layer, and output or target. This selection is based on the best fitness.

Training parameters are shown in the following table:

Table 8. Training parameters

Criteria	Training	Validation
Absolute error	1.807555	2.624603
Network error	0.000119	0
Error improvement	0.000011	
Iteration	37	
Training speed, iter/sec	46.249999	
Architecture	[13-6-1]	
Training algorithm	Levenberg-Marquardt	
Training stop reason	No error improvement	

LM algorithm is used as an optimization algorithm. This algorithm can minimize the network training error.

Then, in the following figure, there is the output of regression analysis based on training, validation, and testing.

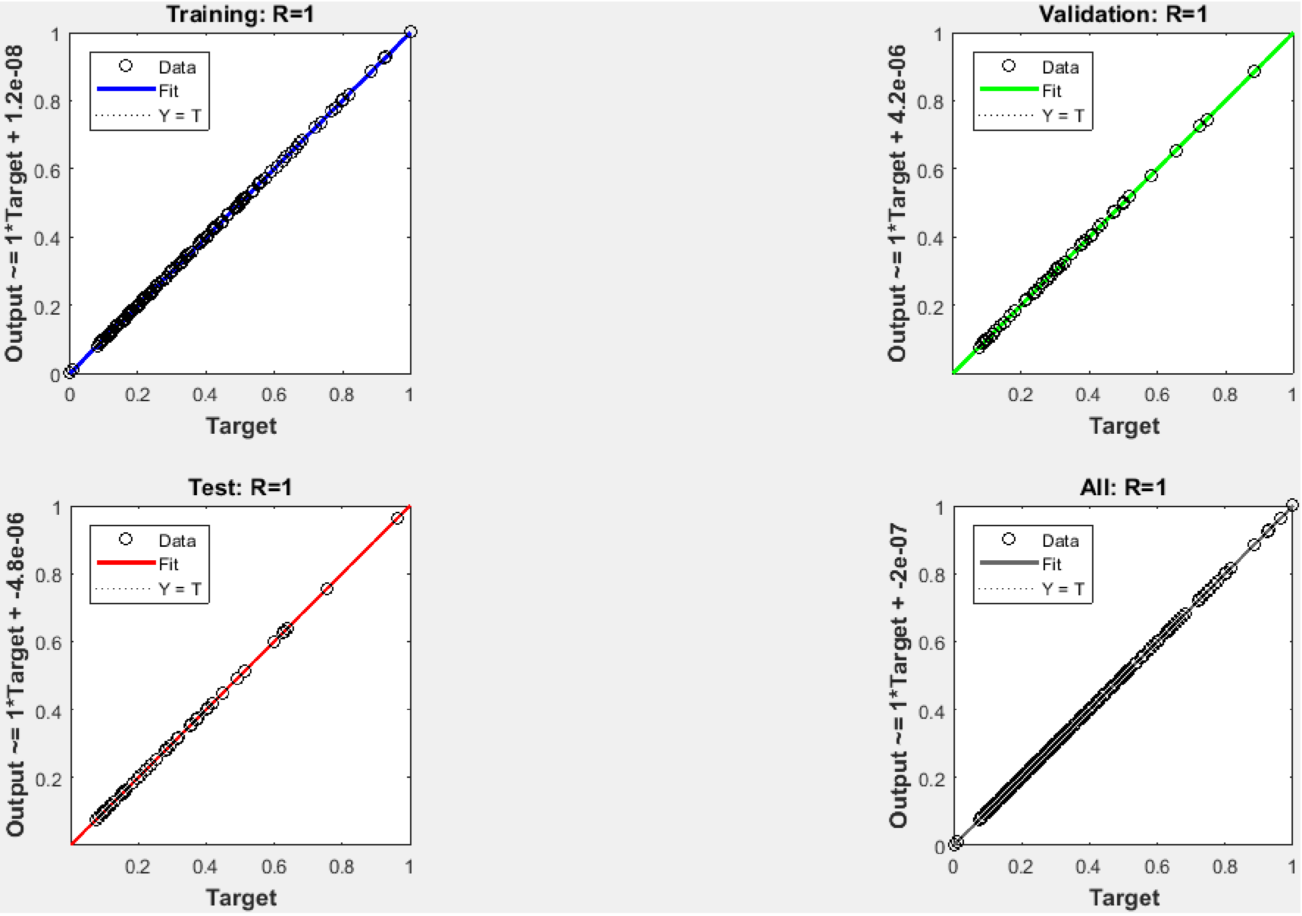

Fig 16. Regression and fitting data

As it is clear, for all three cases (i.e., training, validation, and testing), the rate of R-squared is maximum. So, the GA could recognize and remove unrelated and insignificant variables. As a result, it can increase the model predictability.

Then, training and testing errors are observable in figure 17.

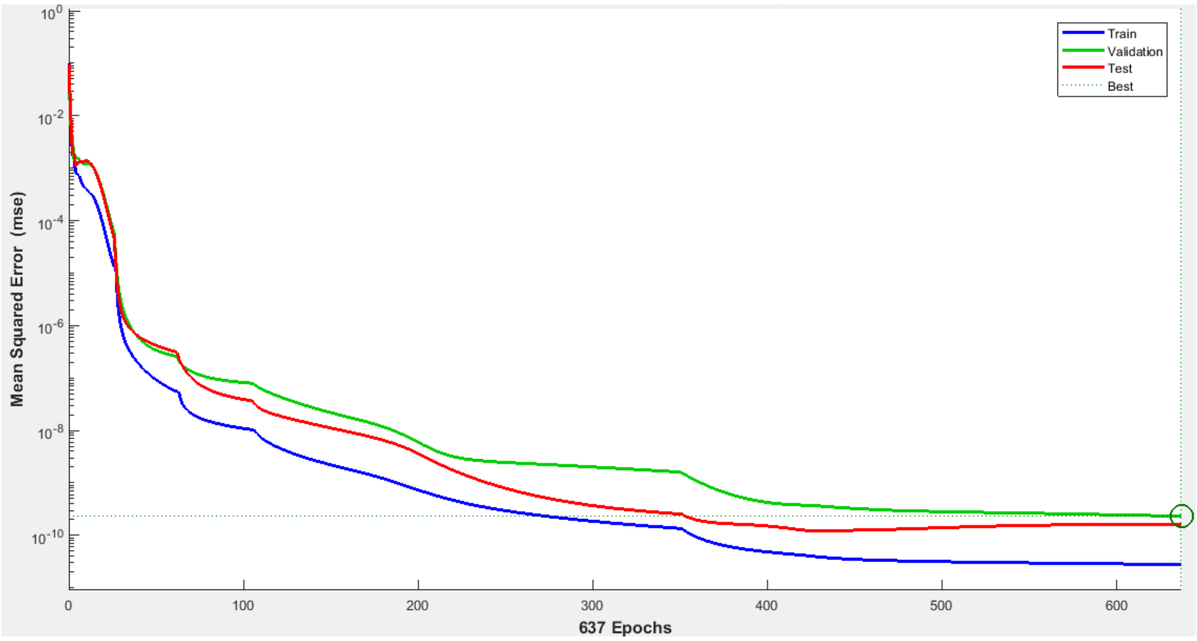

Fig 17. Training, validation, and testing errors

Fig. 17 can induce that the number of iterations has increased. The best validation performance is 2.3227e-10 at epoch 637. After that, the actual versus output graph can be seen in Fig. 18:

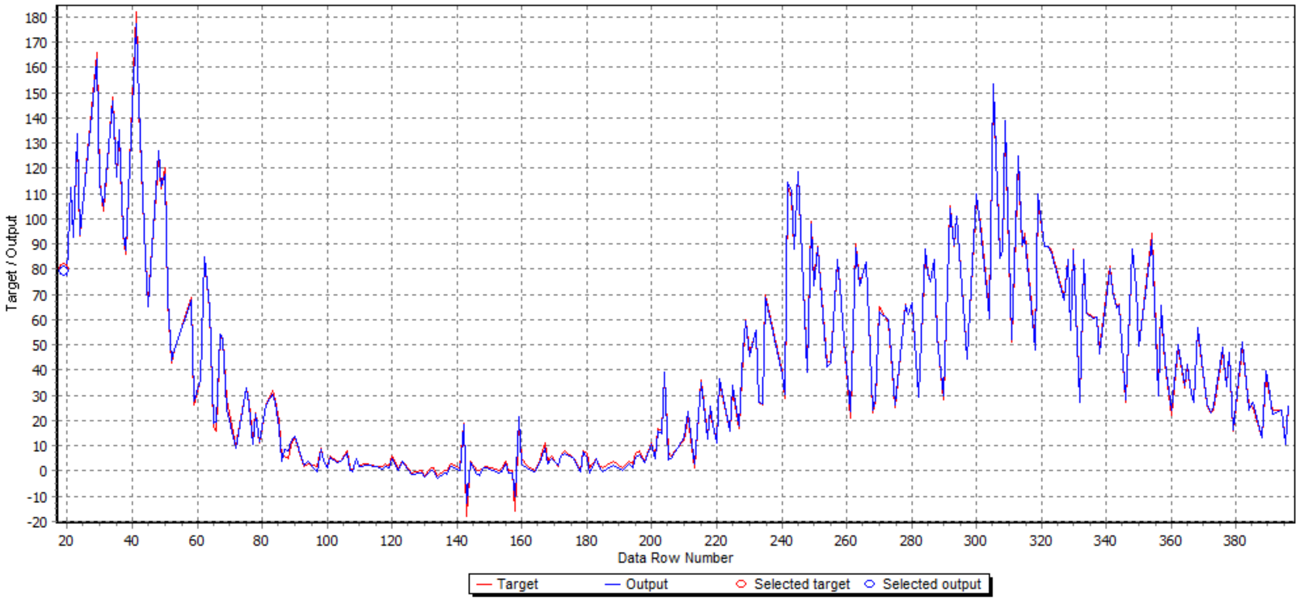

Fig 18. Actual vs. output

These two lines (red and blue lines) are fully coincidental. This is a sign of good prediction and fitting. Finally, summary details of testing results are presented in Table 9.

Table 9. Testing summary

Criteria	Target	Output	AE	RE
Mean	41.6785	41.5304	1.1318	2.70E+13
Std. Dev	39.9677	40.1248	0.8921	1.37E+14
Min	-18	-12.6337	0.0109	0.000213
Max	182	177.4320	5.7036	1.23E+15
Correlation	0.999367			
R-squared	0.99871			

The approximation error in a data value is the discrepancy between an exact value and some approximation to it. This error can be expressed as an absolute error (the numerical amount of the discrepancy) or as a relative error (the absolute error divided by the data value). The following formula shows the relative error formula:

$$\delta = \left|\frac{\upsilon_A - \upsilon_E}{\upsilon_E}\right| \qquad (2)$$

Where

δ: Relative error

υ_A: Actual value observed

υ_E: Expected value

The second formula is related to absolute relative error:

$$\sigma = \upsilon_A - \upsilon_E \qquad (3)$$

Where

σ: Absolute error

υ_A: Actual value observed

υ_E: Expected value

CONCLUSIONS

In this paper, we used machine learning-based methods such as artificial neural networks (ANN) and genetic algorithms (GA) to predict daily new death cases in the Netherlands. Some indicators are considered input variables based on technical indicators such as MACD, SMA, EMA, TMA, etc. ANN is used as a prediction model. On the other hand, GA is used for feature selection and finding the most important input variables. To make a better understanding of the results, we used diagrams and figures. Among 16 input variables, 13 variables were selected by GA. The results showed that ANN has a goodness of fit with high predictability power and high R-squared means of more than 98%. On the other hand, GA is very effective in finding the most important indicators as input variables.

There are some notes and points when applying ANN or AI-based models:

- These models are very sensitive to setting up their parameters such as initial population, training rate, input variables, etc. it is called tuning. It is important because it can impact the performance of the network and its accuracy or speed of calculation. As a result, tuning can increase or decrease the predictability of the network or model.
- Sometimes you may face local minima or maxima trap. It means that you may find the optimal solution which is not the real answer (solution). You may experience early convergence or divergence. Therefore, under-training and over-training are important.
- Meta-heuristic algorithms are fantastic and interesting because of their exploration and exploitation power. They can help you robust your models and decrease the chance of local minima or maxima traps. Thus, novel meta-heuristic algorithms such as the Aquila optimizer (AO) algorithm, Bald Eagle (BA) optimization algorithm, and Chimp optimization algorithm (ChOA) are suggested.

References

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. Heliyon, 4(11), e00938.

Anđelić, N., Baressi Šegota, S., Lorencin, I., Mrzljak, V., & Car, Z. (2021). Estimation of COVID-19 epidemic curves using genetic programming algorithm.

Asgharnezhad, H., Shamsi, A., Alizadehsani, R., Khosravi, A., Nahavandi, S., Sani, Z. A., ... & Islam, S. M. S. (2022). Objective evaluation of deep uncertainty predictions for covid-19 detection. Scientific Reports, 12(1), 1-11.

Barouki, R., Kogevinas, M., Audouze, K., Belesova, K., Bergman, A., Birnbaum, L., ... & HERA-COVID-19 working group. (2021). The COVID-19 pandemic and global environmental change: Emerging research needs. Environment international, 146, 106272.

Berenbaum, M. R. (2021). On COVID-19, cognitive bias, and open access. Proceedings of the National Academy of Sciences, 118(2).

Dos Santos Santana, Í. V., da Silveira, A. C., Sobrinho, Á., e Silva, L. C., da Silva, L. D., Santos, D. F., ... & Perkusich, A. (2021). Classification models for COVID-19 test prioritization in Brazil: machine learning approach. Journal of medical Internet research, 23(4), e27293.

Dramsch, J. S. (2020). 70 years of machine learning in geoscience in review. Advances in geophysics, 61, 1-55.

Göçken, M., Özçalıcı, M., Boru, A., & Dosdoğru, A. T. (2016). Integrating metaheuristics and artificial neural networks for improved stock price prediction. Expert Systems with Applications, 44, 320-331.

Gray, J. D., Harris, C. R., Wylezinski, L. S., & Spurlock Iii, C. F. (2021). Predictive modeling of COVID-19 case growth highlights evolving racial and ethnic risk factors in Tennessee and Georgia. BMJ Health & Care Informatics, 28(1).

Gupta, N. (2017). A Literature Survey on Artificial Intelligence. International Journal of Engineering Research & Technology (IJERT) ICPCN, 5(19).

Katoch, S., Chauhan, S. S., & Kumar, V. (2021). A review on the genetic algorithm: past, present, and future. Multimedia Tools and Applications, 80(5), 8091-8126.

Kumar, R., Arora, R., Bansal, V., Sahayasheela, V. J., Buckchash, H., Imran, J., ... & Raman, B. (2020). Accurate prediction of COVID-19 using chest X-ray images through deep feature learning model with SMOTE and machine learning classifiers. MedRxiv.

Kunwar, M. (2019). Artificial intelligence in finance: Understanding how automation and machine learning is transforming the financial industry.

Kuo, K. M., Talley, P. C., & Chang, C. S. (2022). The Accuracy of Machine Learning Approaches Using Non-Image Data for the Prediction of COVID-19: A Meta-Analysis. International Journal of Medical Informatics, 104791.

Kwekha-Rashid, A. S., Abduljabbar, H. N., & Alhayani, B. (2021). Coronavirus disease (COVID-19) cases analysis using machine-learning applications. Applied Nanoscience, 1-13.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. Behavioral and brain sciences, 40.

Lin, T. C. (2019). Artificial Intelligence, Finance, and the Law. Fordham L. Rev., 88, 531.

Malki, Z., Atlam, E. S., Ewis, A., Dagnew, G., Ghoneim, O. A., Mohamed, A. A., ... & Gad, I. (2021). The COVID-19 pandemic: prediction study based on machine learning models. Environmental science and pollution research, 28(30), 40496-40506.

Mohammadrezaei, M. R., Shiri, M. E., & Rahmani, A. M. (2019). Detection of fake accounts in social networks based on One-Class Classification. The ISC International Journal of Information Security, 11(2), 173-183.

Mohan, S., Solanki, A. K., Taluja, H. K., & Singh, A. (2022). Predicting the impact of the third wave of COVID-19 in India using hybrid statistical machine learning models: A time series forecasting and sentiment analysis approach. Computers in Biology and Medicine, 144, 105354.

Muhammad, L. J., Algehyne, E. A., Usman, S. S., Ahmad, A., Chakraborty, C., & Mohammed, I. A. (2021). Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset. SN computer science, 2(1), 1-13.

Pinter, G., Felde, I., Mosavi, A., Ghamisi, P., & Gloaguen, R. (2020). COVID-19 pandemic prediction for Hungary; a hybrid machine learning approach. Mathematics, 8(6), 890.

Rios, R. A., Nogueira, T., Coimbra, D. B., Lopes, T. J., Abraham, A., & Mello, R. F. D. (2021). Country transition index based on hierarchical clustering to predict the next COVID-19 waves. Scientific reports, 11(1), 1-13.

Sathyamala, C. (2022). COVID-19: The Political Economy of a Global Pandemic. Development and Change.

Shi, Z. (2019). Advanced artificial intelligence (Vol. 4). World Scientific.

Trogen, B., & Caplan, A. (2021). Risk compensation and COVID-19 vaccines. Annals of internal medicine, 174(6), 858-859.

Tuli, S., Tuli, S., Tuli, R., & Gill, S. S. (2020). Predicting the growth and trend of the COVID-19 pandemic using machine learning and cloud computing. Internet of Things, 11, 100222.

Xiong, Y., Ma, Y., Ruan, L., Li, D., Lu, C., & Huang, L. (2022). Comparing different machine learning techniques for predicting COVID-19 severity. Infectious diseases of poverty, 11(1), 1-9.

Zawbaa, H. M., El-Gendy, A., Saeed, H., Osama, H., Ali, A. M., Gomaa, D., ... & Abdelrahim, M. E. (2021). A study of the possible factors affecting COVID-19 spread, severity, and mortality and the effect of social distancing on these factors: Machine learning forecasting model. International Journal of Clinical Practice, 75(6), e14116.