***TERM PAPER***

***AFE BABALOLA UNIVERSITY ADO-EKITI, EKITI***

***COLLEGE OF ENGINEERING***

***COURSE CODE: ENG 384***

***COURSE TITLE: ENGINEERING LAW AND MANAGERIAL ECONOMICS***

***TITLE:- ENGINEERING STRATEGIES FOR HANDLING COVID-19 FOR ENVIRONMENT HEALTH AND ECONOMICAL SUSTAINABILITY***

***BY***

***COLLINS OLUWASEUN ADEBAYO***

**Abstract**

Nowadays, sustainable development is considered a key concept and solution in creating a promising and prosperous future for human societies. Nevertheless, there are some predicted and unpredicted problems that epidemic diseases are real and complex problems. Hence, in this research work, a serious challenge in the sustainable development process was investigated using the classification of confirmed cases of COVID-19 (new version of Coronavirus) as one of the epidemic diseases. Hence, binary classification modeling was used by the group method of data handling (GMDH) type of neural network as one of the artificial intelligence methods. For this purpose, the Hubei province in China was selected as a case study to construct the proposed model, and some important factors, namely maximum, minimum, and average daily temperature, the density of a city, relative humidity, and wind speed, were considered as the input dataset, and the number of confirmed cases was selected as the output dataset for 30 days. The proposed binary classification model provides higher performance capacity in predicting the confirmed cases. In addition, regression analysis has been done and the trend of confirmed cases compared with the fluctuations of daily weather parameters (wind, humidity, and average temperature). The results demonstrated that the relative humidity and maximum daily temperature had the highest impact on the confirmed cases. The relative humidity in the main case study, with an average of 77.9%, affected positively, and maximum daily temperature, with an average of 15.4 °C, affected negatively, the confirmed cases.

TABLE OF CONTENT

TITLE PAGE……………………………………………………………………………………………………….1

ABSTRACT …………………………………………………………………………………………………………2

CHAPTER 1: INTRODUCTION…………………………………………………………………...........4-5

CHAPTER 2: MATERIALS AND METHODS…………………………………………………………6-15

CHAPTER 3 CONCLUSIONS:……………….…………………………………………………………..15

REFERENCE ………………………………………………………………………………………………………..16

**CHAPTER 1**

**Introduction**

The concept of sustainable development as a new concept, process, and undeniable fact has emerged in the policies of major governments, and plays a key role in the development of human societies. Sustainable development is, generally, a combination of the three social, economic, and environmental goals in which political goals are involved. In fact, sustainable development is the advancement of the quality of all aspects of life of today’s generation without creating negative impacts on the lives of future generations. While it may seem easy to create sustainability in theory, the process of sustainable development faces many unforeseen problems and obstacles that slow down the process. One of the anticipated problems is the emergence of epidemic diseases that not only have a negative impact on the economy but also cause social problems, both of which are fundamental to sustainable development. Although this is a temporary and transient problem, it has the potential to disrupt the process, which can have years of adverse effects. The COVID-19 (a new version of Coronavirus) is one of the newest and most serious challenges facing governments, and there has not been much research into this problem . Although the understanding of COVID-19 is limited, interim guidance on laboratory biosafety was introduced by the World Health Organization (WHO). Kampf et al. investigated the persistence of coronaviruses on inanimate surfaces and ways to deal with it. They found that the period of persistence is nine days for coronaviruses, and some disinfectants, such as 62%–71% ethanol, 0.5% hydrogen peroxide, or 0.1% sodium hypochlorite can be very efficient in dealing with this virus. Lai et al. have evaluated the outbreak of COVID-19 and its challenges. Based on their results, they made some recommendations for the prevention of more outbreaks of the virus.

In another study, the role of inanimate surfaces in the outbreak of coronaviruses is investigated by Kampf. Based on the obtained results, he provided some recommendations about the impact of surface disinfection to prevent further viral spread. Telles has conducted an overview of the behaviors of viruses. Some datasets were investigated, and the obtained results show that dynamic mathematical modeling was essential to predict behaviors of viruses. The role of media coverage on the public was evaluated by Wen et al. Their obtained results indicated that misleading and biased media coverage could have a negative impact on individuals’ mental health. Chen et al. carried out an investigation for predicting the number of confirmed cases of COVID-19. They evaluated the trend of transmission and recovery rates based upon time, and a mathematical model was proposed. The obtained results show that the proposed model had a suitable performance to predict the confirmed cases. In another study, the trend of the COVID-19 outbreak was estimated in China by Li and Feng. Their results show that rapid and dynamic strategies can be useful to diminish and constrain the current crisis. Chinazzi et al. evaluated the effect of travel constraints to reduce the COVID-19 outbreak. The obtained results indicated that travel restrictions are highly effective in reducing the spread of this new coronavirus .  
It is worth mentioning that the impact of temperature on virus spread and survival show different results. There have been some investigations on the effects of environmental parameters on epidemic diseases, too. The flu virus spreads quickly in cold and dry conditions, while it is completely inactive at temperatures above 30 °C. However, the epidemic of one type of coronavirus, MERS-CoV, was between April and August, which meant that the virus spread quickly in warm temperature, low wind speed, low relative humidity, and high ultraviolet index .  
A review of previous studies shows that very few researchers have addressed the challenges of the COVID-19 (new version of Coronavirus) in the sustainable development process and, also, unpredictable problems and complexity of the issues require the use of highly capable approaches like artificial intelligence methods to understanding these types of issues. Hence, due to the importance of the subject, the present study investigated the feasibility of artificial intelligence in the classification of confirmed cases of COVID-19, which is a severe challenge in the sustainable development process, and it is an imperative task. In addition, statistical analysis was carried out, and the obtained results were discussed. It is worth mentioning that this type of analysis has not been used in previous research.

**CHAPTER 2**

**Materials and Methods**

Two approaches were used in the current study, as follows:

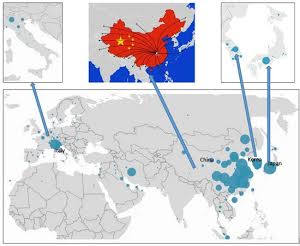
* The possible correlations among the trends of confirmed cases in different case studies were investigated, and then a binary classification model was constructed to predict and classify using the group method of data handling (GMDH) algorithm based upon some critical factors; maximum, minimum, and average temperature, the density of a city, relative humidity, and wind speed were considered as the input dataset and the number of confirmed cases was selected as the output dataset for 30 days.
* Regression analysis was used, and a trend of the confirmed cases of COVID-19 analyzed in the five provinces with the highest confirmed cases, including Hubei, Guangdong, Henan, Zhejiang, and Hunan, and the daily fluctuations of confirmed cases were compared with fluctuations of weather parameters.

Conditions of analysis

* The environmental and urban parameters in the analysis included density, sex ratio, average age, elevation, maximum, minimum, and average temperature, relative humidity, and wind.
* For daily analysis of the possible trend between confirmed cases of COVID-19 and environmental factors, the data of Hubei province was used.
* The climate data is based on the stations situated in the capital of the provinces or regions because the population is generally higher in these areas.
* The analysis period was from 28 January 2020 to 26 February 2020 (30 days).
* The analysis of the possible correlations about trends of confirmed cases in different case studies was based on the average values in one month.

*2.1. Case Study*

To carry out the analysis of correlation among environmental factors and confirmed cases of COVID-19, a set of data, including 42 provinces in China, Japan, South Korea, and Italy, were used. The case studies are based on the most confirmed cases of COVID-19 and available data, It is worth mentioning that the quarantine on travel in and out of Wuhan and the suspension of flights, trains, public buses, and the metro system began on 23 January 2020 and also on 24 January 2020 in 15 cities in Hubei. The estimated incubation period of COVID-19 is about 2–14 days.



*2.2. Group Method of Data Handling (GMDH)*

Artificial Intelligence includes a wide range of methods and algorithms that work based on machine intelligence and has many applications in various fields of science], including fuzzy logic theory and application artificial intelligence techniques and sociology, risk assessment and hazard identification , machine learning, and meta-heuristic algorithms and clustering techniques. The group method of data handling (GMDH) type of neural network is one of these algorithms that was proposed by Ivakhnenko. GMDH is a self-organization algorithm that has been used successfully for pattern recognition, optimization of complex systems modeling, and prediction problems, and it is also called the polynomial of the Ivakhnenko equation. This algorithm can predict the value of

y  
ˆ

i

from an approximate function (

f  
∧

), for each input vector (*X*), which is shown in Equation (1). The basic form of a relation between input and output data can be declared as a discrete type of the Volterra functional series, referred to as the Kolmogorov–Gabor polynomial. Equation (2) shows the underlying neural network map, which is also called the polynomial of Ivakhnenko.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| y ∧ | = | f ∧ | ( | x    i1 | , | x    i2 | , | x    i3 | , ……… | x    im | ) |

i=(1,2,3,……, m)

(1)

y = a +   
∑  
i=1  
m

|  |  |
| --- | --- |
| b | i |

x

i

 +   
∑  
i=1  
m

|  |  |  |  |
| --- | --- | --- | --- |
| ∑ j=1 m  c    ij | x    i | x    j | + |

∑  
i=1  
m

∑  
j=1  
m

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | ∑ k=1 m | d    ijk | | x    i | x    j | x    k | + |

∑  
i=1  
m

∑  
j=1  
m

∑  
k=1  
m

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ∑ l=1 m  e    ijkl | x    i | x    j | x    k | x    l | ,………… |

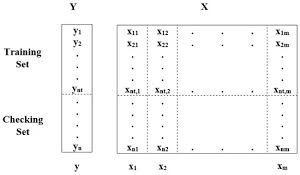
(2)

where *Y* is the output, *m* is the number of data, and *X*1, *X*2, *X*3… *Xm* is the input variables vector. In many cases, the quadratic and bivariate form of this polynomial is used as Equation (3).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| y ˆ | = G ( | x    i | , | x    j | ) = | a    0 | + | |  |  | | --- | --- | | a | 1 | | x    i | + | |  |  | | --- | --- | | a | 2 | | x    j | + | |  |  | | --- | --- | | a | 3 | | x    i | x    j | + | |  |  | | --- | --- | | a | 4 | | |  |  | | --- | --- | | x | 2 | | i | + | |  |  | | --- | --- | | a | 5 | | |  |  | | --- | --- | | x | 2 | | j |

(3)

Input and output variables of the GMDH algorithm, where *X*= (*X*1, *X*2, *X*3, …. *Xm*) is the input dataset and *Y*= (*y*1, *y*2, *y*3,…. *yn*) is the output dataset. In this algorithm, the input dataset is imported to the initial layer and then, after evaluation and optimization, the output is considered as a new input for the next layer. This process is continued and stopped if the algorithm reaches a better answer from layer (n+1) in comparison with layer (n). As mentioned before, to deal with unpredicted and uncertain problems, the GMDH algorithm can be applied as a powerful tool. Hence, a binary classification analysis was done by the GMDH algorithm in the present study.



input and output variables of the group method of data handling (GMDH) algorithm.

**3. Results**

*3.1. Binary Classification Modelling Using GMDH*

Before binary classification modeling using the Gmdh algorithm, a regression analysis was conducted among the total data set for 42 case studies in four countries, including China, Japan, South Korea, and Italy. This analysis shows that there is a low correlation coefficient (R2); hence, it can be concluded that we should evaluate case by case for binary classification modeling, and Hubei province in China was selected as a case study for this part of the analysis.  
Initially, evaluating the parametric correlation of each independent input dataset is necessary for carrying out reliable modeling. Hence, before modeling, a correlation analysis was conducted using Pearson’s correlation coefficient among the dataset for five independent input data, including maximum, minimum, and average temperature, relative humidity, and wind speed. Furthermore, it should be noted that independent input data, including the density of the city, is not considered for evaluating the parametric correlation because the value of this parameter is constant. The mathematical relations of Pearson’s method can be expressed in Equations (4)–(7).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ρ=r=   |  |  | | --- | --- | | S | P  Dxy |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | S | S    X | . S | |  |  | | --- | --- | | S | Y | |   −    −    −    −    −    −    −    −    −  √ | (4) |
| S  P  Dxy    = ∑ xy −  (∑x) (∑y)  n | (5) |
| |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | S | S    X | =  ∑ i n  x  2 i   − | |  |  | | --- | --- | | (∑  x    i  ) | 2 |   n | | (6) |
| S   |  |  | | --- | --- | | S | Y |    =  ∑ i n  y  2 i   −   |  |  | | --- | --- | | (∑  y    i  ) | 2 |   n | (7) |

where

ρ

is the Pearson’s correlation coefficient for two independent parameters that are shown with *X* and *Y*, and

S

P

Dxy

is the covariance between them. The standard deviation of *X* and *Y* are indicated with

|  |  |  |  |
| --- | --- | --- | --- |
| S | |  |  | | --- | --- | | S | X | |

and

|  |  |  |  |
| --- | --- | --- | --- |
| S | |  |  | | --- | --- | | S | Y | |

According to the previous studies, when

|ρ| ⟩ 0.8

, the correlation coefficient is strong, hence the obtained results show that independent input data were correctly selected. Although there is a correlation coefficient of 0.83 between the maximum and average temperatures, it can be accepted because this value is very insignificant.

Secondly, determining the control parameters of the algorithm is an important task because it plays a key role in the fast convergence of the algorithm. There are no specific relations about most of these parameters, and they are determined based upon previous studies, experts, and trial and error. Hence, the selection pressure is dimensionless and has an impact on the sensitivity of the modeling error. It is selected as 0.6 based upon the most recent studies. However, the maximum number of layers and the maximum number of neurons in a layer are considered based upon expert opinions and trial and error. For this purpose, a range of values was determined for the maximum number of layers equal to 5, 10, and 15, and also another range of values was considered including 5, 10, 15, 20, and 30. It is worth mentioning that the confusion matrix is used as the measure of accuracy to evaluate the performances of the binary classification model by the GMDH algorithm. The basic form of a confusion matrix is shown in, and Equations (8) and (9) indicate the mathematical relations of accuracy and error.

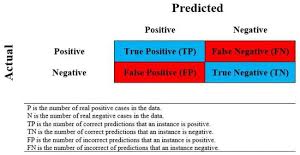
Acc =

TP + TN  
TP+FP+TN+FN

(8)

|  |  |  |
| --- | --- | --- |
| Error = | TP + TN TP+FP+TN+FN | = 1 − Acc |

(9)



[](https://googleweblight.com/i?u=https%3A%2F%2Fwww.mdpi.com%2F2071-1050%2F12%2F6%2F2427%2Fhtm&geid=NSTNR" \l "a.fig_body_display_sustainability-12-02427-f003)

**Figure 3.** Schematic of the confusion matrix.

According to the maximum number of layers and the maximum number of neurons in a layer, in total, 20 models were constructed. The six notable factors, namely the maximum, minimum, and average temperature, the density of cities, relative humidity, and wind speed were considered as the input dataset and the number of confirmed cases was chosen as the output dataset for 30 days. It should be noted that the two classes (label) are assigned and considered for the number of confirmed cases, and this means that, for the number of confirmed cases under 850, people were given label “0” and, for the number of confirmed cases above 850, people were given label “1”. In addition, there is the information of Wuhan city for 30 days as the number of cases studies; for modeling, 75% of cases were considered as training cases and the rest were considered as testing cases [

**CHAPTER 3**

**Conclusions**

In this research study, a serious challenge of sustainable development was investigated using the GMDH algorithm and regression analysis. According to the results, the GMDH algorithm has an appropriate performance to predict and classify the parameters of a case study affected by COVID-19, and the accuracies based on Wuhan datasets were equal to 95.7% and 85.7% for training and testing, respectively. No correlation was found among the different case study datasets in four countries, which might be due to different policies and types of restrictions in each country and means that the prediction of the trend could be made case by case. The results of collinearity diagnostics of the regression analysis demonstrated that the relative humidity and maximum daily temperature and average temperature had the highest share in the expression of changes of output (confirmed cases), respectively. The relative humidity (in the case study with an average of 77.9%) affected positively, and maximum daily temperature (in the case study with an average of 15.4 °C) affected negatively, the confirmed cases. The study shows the positive impact of quarantine in decreasing the number of confirmed cases, which was effective after about 14 days, alongside the impact of environmental factors in confirmed cases of COVID-19 and the role of regression analysis and binary classification by using artificial intelligence in the investigations.  
Finally, since the analysis shows the impact of the weather parameters on confirmed cases of COVID-19, the development of a prediction model with more datasets is suggested for future studies.

**REFERENCE**

* World Health Organization (WHO). Laboratory Biosafety Guidance Related to Coronavirus Disease 2019 (COVID-19). Available online: [**https://apps.who.int/iris/bitstream/handle/10665/331138/WHO-WPE-GIH-2020.1-eng.pdf**](https://apps.who.int/iris/bitstream/handle/10665/331138/WHO-WPE-GIH-2020.1-eng.pdf)
* Altamimi, A.; Ahmed, A.E. Climate factors and incidence of Middle East respiratory syndrome coronavirus. *J. Infect. Public Health* **2019**. [[**Google Scholar**](https://scholar.google.com/scholar_lookup?title=Climate+factors+and+incidence+of+Middle+East+respiratory+syndrome+coronavirus&author=Altamimi,+A.&author=Ahmed,+A.E.&publication_year=2019&journal=J.+Infect.+Public+Health&doi=10.1016/j.jiph.2019.11.011)] [[**CrossRef**](https://googleweblight.com/i?u=https://dx.doi.org/10.1016/j.jiph.2019.11.011&hl=)]
* 2019–20 Coronavirus Outbreak. Available online: [**https://en.wikipedia.org/wiki/2019–20\_coronavirus\_outbreak**](https://googleweblight.com/i?u=https://en.wikipedia.org/wiki/2019%25E2%2580%259320_coronavirus_outbreak&hl=)
* New Coronavirus Cases in China. Available online: [**https://www.timesleader.com/interactives/774260/new-coronavirus-cases-in-china**](https://googleweblight.com/i?u=https://www.timesleader.com/interactives/774260/new-coronavirus-cases-in-china&hl=)
* Coronavirus COVID-19 Outbreak: Latest News, Information and Updates. Available online: [**https://www.pharmaceutical-technology.com/special-focus/covid-19/coronavirus-covid-19-outbreak-latest-information-news-and-updates/**](https://googleweblight.com/i?u=https://www.pharmaceutical-technology.com/special-focus/covid-19/coronavirus-covid-19-outbreak-latest-information-news-and-updates/&hl=)