www.mecsj.com/
Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185

The impact of COVID-19 on Arabian Gulf countries using the Classical Machine Learning Methods

Reem A KH AlMeshal
The Public Authority for Applied Education and Training (PAAET)
Email: reem_almeshal@hotmail.com

Abstract

The coronavirus COVID-19 pandemic is causing a global health crisis, and a public menace outbreak and emergency for public health, where it is being transmitted via the respiratory tract when a normal person comes in contact with the infected person. The coronavirus COVID-19 has affected countries worldwide. It caused scientists around the world to conduct research to combat this continuous pandemic. In this paper, we proposed a method based on Machine Learning (SVM, RProp, and Decision tree) methods that aim to detect the abnormal data and show the essential differences in normal data with high accuracy rates for the COVID-19 in Kuwait and Arabic Gulf countries. The experiment results show that the accuracy for SVM, RProp, and Decision tree reached (94.75%), (91.99%), and (96.35%) respectively.

Keywords

Machine learning, deep learning, COVID-19, Genetic algorithm, Support Vector Machine, RProp, Decision Tree.

الملخص

تتسبب جائحة فيروس كورونا (COVID-19) في أزمة صحية عالمية وتفشي عام للخطر وحالة طوارئ للصحة العامة ، حيث ينتقل الفايروس للجسد من خلال الجهاز التنفسي عندما يتلامس شخص طبيعي مع الشخص المصاب. أثر فيروس كورونا (COVID-19) على أغلب الدول في جميع أنحاء العالم، حيث دفع العلماء في جميع أنحاء العالم إلى إجراء أبحاث لمكافحة هذا الوباء المستمر. في هذه الورقة ، اقترحنا طريقة تعتمد على أساليب التعلم الآلي (SVM) ، RProp، وشجرة القرار، حيث يهدف النموذج المقترح إلى اكتشاف البيانات غير الطبيعية وإظهار الاختلافات الأساسية في البيانات العادية ذات معدلات الدقة العالية لـ (COVID-19) في دولة الكويت بشكل خاص ودول الخليج العربي بشكل عام. أظهرت النتائج العملية للتجارب أن نسبة الدقة في (SVM) و (RProp) وشجرة القرار بلغت (94.75٪) و (91.99٪) و (96.35٪) على التوالي.

الكلمات المفتاحية: لتعلم الآلي ، التعلم العميق ، فايروس كورونا ، الخوارزمية الجينية ، دعم آلة المتجهات ، RProp، شجرة القرار

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

1. Introduction

Coronavirus appeared for the first time in the Wuhan city of China in December 2019, the government of china reported to the World Health Organization (W.H.O) on 31st December 2019 (Zhang et al., 2020). The virus created a global menace as a public outbreak and emergency for public health and was named COVID-19 by the W.H.O on 11th February 2020. The COVID-19 is classified as a virus of the family of viruses like SARS, and ARDS, where it is being transmitted via the respiratory tract when a normal person comes in contact with the infected person or through other ways which are currently unclear (Khanday et al., 2020).

COVID-19 is a pandemic that has emerged in the present year and has affected countries worldwide. It caused scientists around the world to conduct research to combat this continuous pandemic. COVID-19 research mostly focused on observational research and basic microbiological research is less prevalent (Doanvo et. al, 2020).

In our study, we reviewed some research on COVID-19, where the authors applied both deep learning model and ML methods to predict cases of COVID-19 and mortality rate, detect illness through reviewing X-ray images, CT scan images, etc., as well analyzed available data on COVID-19 through ML. They discovered that ML showed good performance (Khalifa et. al, 2020; Elaziz et. al, 2020; Loey et. al, 2020) for applying as predicting model (Duttaa & Bandyopadhyay (2020; de Souza et. al, 2020; Pinter et. al, 2020; Ünlü & Ersin, 2020). Still, it was noted that the availability of a bigger amount of data on COVID-19 would lead to better and more accurate results (Khanday et. al, 2020; Pinter et. al, 2020). As well, Zame et. al (2020) stated that fields like epidemiology, natural operations research, language processing, statistics, and systems biology as well as advertising and finance could give important and necessary contributions in terms of COVID-19/SARS-CoV-2.

To solve the previous problems, several studies have focused on developing COVID-19 that benefit from different techniques known as classical and deep learning methods. In this paper, the proposed method will focus on Machine Learning (ML) as one of these classical methods. ML methods can automatically detect abnormal data and show the essential differences in normal data with high accuracy rates.

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

2. Related works

To date, many works and researches have suggested several approaches for COVID-19 prediction using different methods such as machine learning and deep learning. In this literature review part of the thesis, we will review the most significant approaches that have been put forward in this field.

In the Khalifa et. al (2020) study, they aimed to classify the potential treatments for coronavirus (COVID-19) on a single human cell according to the type and concentration level of treatment. A deep learning model and machine learning (ML) methods were applied in the present study. The experimental results showed that the proposed DCNN model for treatment classification reached 98.05% in testing accuracy in comparison to a classical machine, while in treatment concentration level prediction, it showed 98.2% accuracy against 98.5% of classical machine learning (ensemble). It was concluded that deep learning and computer algorithms can assist in testing approved treatments on human cells, which can reduce the gap between treatments and revealing of an actual cure.

While in Khanday et. al (2020) study, they conducted their study to detect COVID-19 using clinical text data by applying ML-based approaches. The textual clinical reports were divided into four classes by applying classical and ensemble ML methods. The experimental outcomes showed that logistic regression and multinomial Naive Bayesian classifier gave great results, performing 96.2% of accuracy. The authors noted that more feature engineering was required to obtain better results and the deep learning approach could be used in the future.

Because the COVID-19 pandemic is presenting great challenges to medical research, and to clinical trials particularly, Zame et. l (2020) conducted their study. They aimed to bridge the gap between quantitative research scientists involved in clinical trials influenced by or connected to COVID-19 and the ML community and to assist in uniting those communities together. The authors applied ML for suggesting solutions that it could provide. The main focus of the study was on ML and clinical trials because they were the areas of authors' competence. The authors stated that fields like epidemiology, natural operations research, language processing, statistics, and systems biology as well as advertising and finance could give important and necessary contributions.

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

To predict COVID-19 Pandemic for Hungary, Pinter et. al (2020) conducted their study. They suggested hybrid ML techniques of adaptive network-based fuzzy inference system (ANFIS) as well as multi-layered perceptron-imperialist competitive algorithm (MLP-ICA) as a new tendency in promoting outbreak models to forecast time series of infected people and mortality cases. The experimental results showed that both models provided good results in a viewpoint of forecasting the time series without the assumptions (required by the epidemiological models) and in forecasting COVID-19 outbreak and in estimating general mortality. Still, MLP-ICA performed better than ANFIS with providing accurate results on samples of validation. Based on the obtained results the authors suggested ML as a potential technology to model the outbreak.

Jha et. al (2020) conducted their study to examine a ranked list of features that could point to a predisposition to a mental disorder in the COVID pandemic period. For the study authors applied Bayesian networks and the classical ML method. The experimental results showed that people were stressed according to the gender and age differences and also people with chronic medical state of mental illness were more inclined to mental disorders within a period of the COVID. While in the de Souza et. al (2020) study, they aimed to forecast COVID-19 confirmed cases with 17 days forward in the Amazon region. For that purpose, they compared classical and ML models: autoregressive integrated moving average (ARIMA), Holt-Winters, support vector regression (SVR), k-nearest neighbors regressor (KNN), random trees regressor (RTR), seasonal linear regression with change-points (Prophet), and simple logistic regression (SLR). The test results showed that all models outperform SLG. Two classical approaches performed better: ARIMA and Holt-Winters. It could be because of sudden variations and seasonality of the Amapaense data. As well, both of them were easier to code and tune than ML models.

The study of Youha et. al (2020) aims to reveal clinical and demographic characteristics of COVID-19 of both symptomatic and asymptomatic patients. In the study, they used an extensive multi-algorithm ML pipeline and 68 clinical and demographic and clinical characteristics. ML pipeline included five methods, such as 1) logistic regression (LR), 2) support vector machine (SVM), 3) random forest (RF), 4) gradient boosting (GBM), 5) extreme gradient boosting (XGM). A test result showed remarkable predictive performance of ML in comparison to traditional statistical techniques in identifying important clinical and demographic characteristics of symptomatic and asymptomatic.

The authors concluded that ML could assist in improving case definition and in observation of public health.

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

While Duttaa & Bandyopadhyaya (2020) conducted their study to evaluate how much predicted results are close to original data related to Confirmed-Negative-Released-Death cases of COVID-19. For the study, the authors used the ML method. The result of the experiment showed that the mixed approach of Deep-Learning models showed good results in forecasting given cases of COVID-19. Thus, it was concluded that due to the fast-expanding of COVID-19 it was needed to build an automated model with the basis on the ML approach to take corrective measures after doctors took the decision. While in the Elaziz et. al (2020) study, they suggested a method for the visual diagnosis of COVID-19 cases on chest X-ray. In the study, it was applied to an ML method with two datasets. The experimental results showed that the suggested method reached comparable performance on the following: accuracy, recall, and accuracy assessment metric with the least quantity of characteristics. The proposed approach attained both high performances and resource consumption by choosing the most significant characteristics.

Loey et. al (2020) fulfilled their study to detect face masks in the period of the COVID-19 pandemic. For the study, they proposed a hybrid model with the use of a deep and classical ML. The results of the experiment showed that the applied model super passed the related works in testing accuracy. The major drawback was not trayed most of the classical ML techniques to get the highest accuracy and the lowest consumption time. But in Ünlü & Namlı (2020) study, they applied various ML techniques to predict potential confirmed cases of COVID-19 and associated with this mortality rate for the future. The next models were used in the study: Support Vector Machines (SVM), Long-Short Term Memory (LSTM), Prophet, and Holt-Winters. The experimental results showed that the Prophet model provides the lowest Root Mean Squared Error (RMSE) score in comparison to the other three models.

Doanvo et. al (2020) made a research review and analysis of available information on COVID-19. In the study ML methods were used to reveal key trends in coronavirus literature. The experimental results showed that ML methods can analyze coronavirus research at a massive scale as well as can be used to analyze novel disease literature. It was revealed that COVID-19 research was mostly non-lab-based (e.g., observational), thus COVID-19 lab-based/basic microbiological studies were less prevalent than expected.

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

Due to the small number of benchmark datasets especially on chest CT on COVID-19 in Loey et. al (2020) study. The authors aimed to provide a method of detecting COVID-19 at an early-stage for faster recovery of people around the world. They applied the ML technique with 5 various deep convolutional neural network-based models (GoogleNet, AlexNet, VGGNet16, VGGNet19, ResNet50) were chosen to reveal the patient with coronavirus by using chest CT radiographs digital images. The experimental outcomes that ResNet50 with classical data augmentation in combination with CGAN were the best classifier to reveal the COVID-19 from the chest CT dataset.

Finally, in the Liang et. al (2020) study, they aimed to apply a model of identifying critically ill COVID-19 patients. In the study, the authors proposed a deep learning-based model of survival to prove that it could predict the risk of COVID-19 patients of developing critical sickness based on clinical features at admission. A CPH with LASSO penalty both were applied to define baseline clinical characteristics associated with the later critical illness status. The experimental results showed high performance of the applied model despite missing data (30%) in some cases.

Table 1: Research Summary

Author	**Method**	**Accuracy**	**Weakness**
Khalifa et. al (2020)	- Deep learning model - ML methods	- DCNN model classification 98.05%. - classical **ML** classification 98.5%.	The dataset chosen in the present study is a subset of the in a public online dataset.
Khanday et. al (2020)	- ML algorithms	Naive Bayesian classifier gave excellent results by obtaining 94% precision, 95% f1 score, accuracy 96.2%, and 96% recall.	- Small dataset
Zame et. al (2020)	ML	Authors reviewed opportunities for using ML for clinical trials in the period of COVID-19 to stimulate further research and have highlighted a few cases in which special benefits could be seen.	Authors have provided references for deeper reading but in most cases haven't gone into detail about ML techniques and outcomes.

Pinter et. al (2020)	Hybrid ML approach	Both ML models presented potential in forecasting COVID-19 outbreak and evaluating total mortality.	Because of the small amount of training data, it was essential to fulfill further investigation to examine the true capability of the suggested hybrid model.
Jha et. al (2020)	classical ML	Accuracy reach 80%	Data used for training the model is cross-sectional and thus it was impossible to comment upon the persistence and temporality of the detected effects.
de Souza et. al (2020)	classical and ML models	All models outperform SLG, special Holt-Winters in all scenarios (exponential increase: RMSE = 575, R-Squared = 0.96, sMAPE = 3.09; sudden decrease: RMSE - 262, R-Squared = 0.95, sMAPE = 0.74, stability period: RMSE = 162, R-Squared = 0.98, sMAPE = 0.34). SVR and ARIMA have better performances in isolated scenarios.	The data might differ a little from the website of the Brazilian government, because the counting protocol may diverge from those used by the Amap´a state. As well, case sub-notifications were not treated.
Al Youha et. al	ML algorithm	SVM outperformed all algorithms with an	- A general limitation of the study is the size of the population and potential selection bias

(2020)		accuracy of 70.01%	toward the study population. - Also, the proposed method is not able to characterize the uncertainties in the model forecast well. -
Elaziz et. al (2020)	ML	- 96.09% for the first dataset - 98.09% for the second dataset.	Limitation in the age group (patient with an age range from 40 to 84)
Loey et. al (2020)	Hybrid model using deep and classical ML	SVM in 3 datasets: 1st - 99.64% 2nd - 99.49%, 3rd - 100%	The major drawback is not trayed most of the classical ML techniques to get the highest accuracy and the lowest consumption time.
Ünlü & Ersin (2020)	classical ML methods	Prophet model gives the lowest Root Mean Squared Error (RMSE)	Future predictions are based on prognosticated values after some certain point. Thus, the error rate might be higher than the real values.
Loey et. al (2020)	ML	82.91%	Limited chest CT scan images
Liang et. al (2020)	- Deep learning model - ML	The model showed a certain tolerance to missing data as it still had achieved high efficiency on the external validation set for samples with missing 30% of the data.	More than 50% of the considered patients did not have the required values gathered. Missing data can happen in particular with small or poorly equipped hospitals.

Table (1) showed a summary of the previous studies that interested in COVID-19, the table presents their methods, and the accuracy reached for each study with the limitations. In this paper, we will show the impact of COVID-19 on Arabian Gulf countries using classical machine learning methods.

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

3. Proposed Method

In this section, the methodology for COVID-19 impact in the proposed approach is described. The first step in the proposed methodology is to extract the whole features selection where we select a representative set of attributes from the set of original attributes (in the raw dataset) using different algorithms.

In the next phase, the proposed method will build the classifier models using supervised machine learning to categorize unseen patterns in suitable classes using the KNIME Analytics Platform (Chen et al. 2009). The classifier's models aim to (1) test the accuracy for the refined dataset files and (2) compute the ability for the proposed approach for computing the impact with the appropriate features.

The proposed approach consists of different phases, started with dataset acquisition (raw dataset) that contains the original files for the "Our World in Data" dataset. In the dataset preprocessing phase, the proposed method applies four levels (correlation-based feature selection method, principal component analysis, information gain ratio based feature selection, and minimum redundancy maximum relevance) used with feature selection algorithms for building the refined dataset. Finally, the KNIME tool used for building a model based on a different classifier (Supervised Machine Learning like SVM, RProp, and Decision Tree). Figure (1) below represents the architecture and general framework for the proposed approach (COVID-19 impact):

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

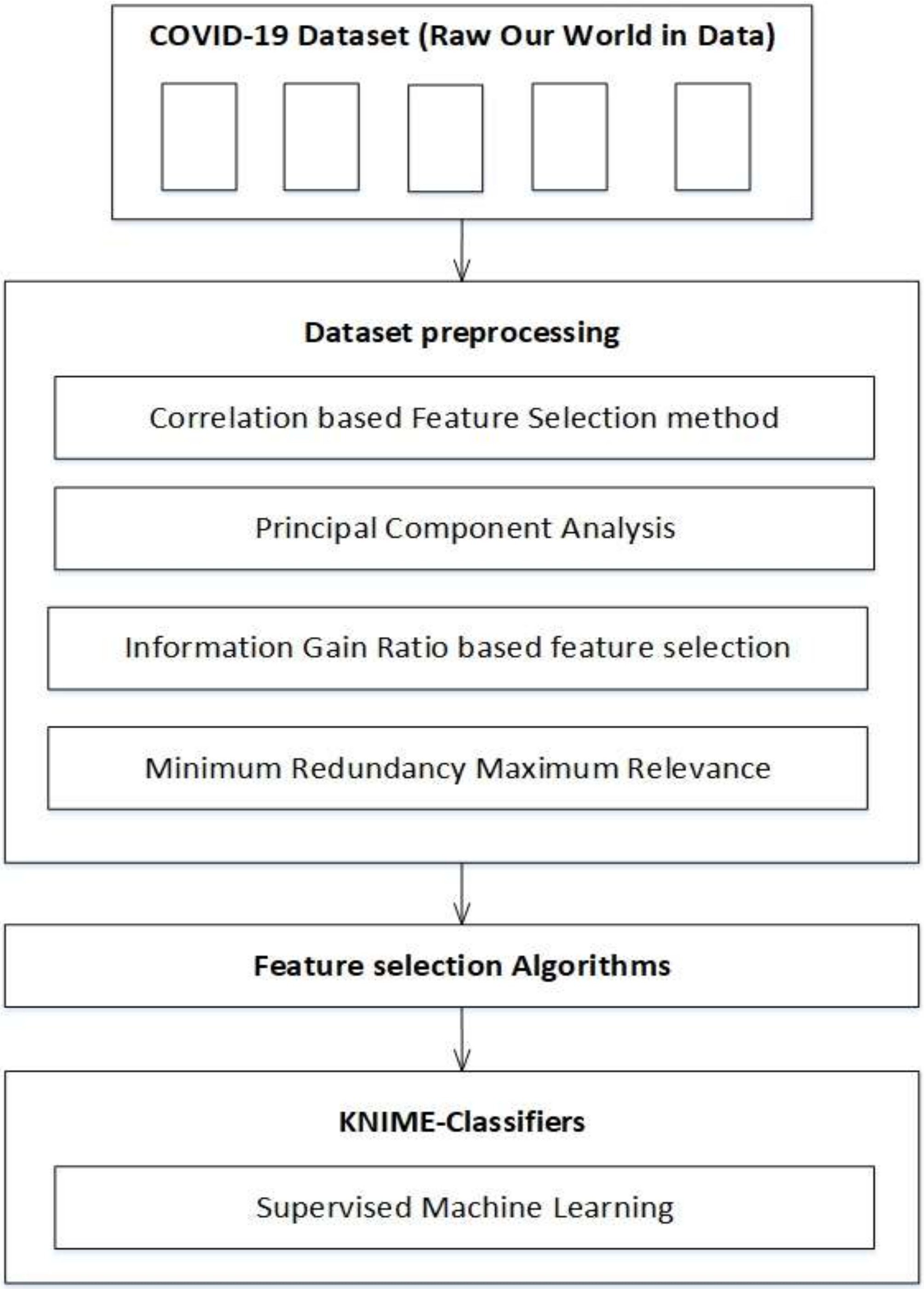

Figure 1: General framework

3.1 COVID-19 Dataset

The proposed COVID-19 dataset is a collection of the COVID-19 data maintained by "Our World in Data". It is updated daily and includes data on confirmed cases, deaths, and testing, where the "Our World in Data" relies on data from Johns Hopkins University. The Johns Hopkins University dashboard and the dataset are maintained by a team at its Center for Systems Science and Engineering (CSSE). It has been publishing updates on confirmed cases and deaths for all countries since January 22, 2020. A feature on the JHU dashboard and dataset was published in The Lancet in early May 2020.1 This has allowed millions of people across the world to track the course and evolution of the pandemic.
Testing data is collected by Our World in Data by browsing public information from official sources. They rely on figures published on official websites, in press releases, and by social media accounts of national authorities usually governments, ministries of health, or centers for disease control. JHU updates its data multiple times each day. This data is sourced from governments, national and subnational agencies across the world a full list of data sources for each country is published on Johns Hopkins GitHub site. It also makes its data publicly available there.

3.2 Dataset Preprocessing

Feature selection (or dataset preprocessing) is a set of techniques aimed at reducing the complexity of the dataset by eliminating some of the non-descriptive, messed values, and non-necessary attributes from the original dataset (Hamid et al. 2016). The feature selection phase is geared to select all representative and appropriate set of attributes from the set of raw attributes (raw dataset). The representative set of the dataset is functions to keep only relevant and important attributes and cross any non-necessary attributes.

In this study, the proposed approach will use several techniques and several algorithms for selecting relevant features from the raw dataset, thus gaining more facilitates for data visualization and data understanding. Table (2) describes the used techniques in the feature selection phase:

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

Table 2: The extraction techniques in dataset preprocessing

Technique	Description
Correlation-based feature selection method (CFS)	CFS aims in the proposed method to have new subsets of features highly correlated with a specific class (classes), and uncorrelated to each other (attributes).
Principal Component Analysis (PCA)	PCA technique aims to identify all uncorrelated features.
Information Gain Ratio based feature selection (IGR)	IGR is used for splitting the attribute pattern distribution into classes, where a gain ratio of attribute decreases as the value of split information increases.
Minimum redundancy maximum relevance	Minimum redundancy maximum relevance is used in the proposed method to punish a feature's relevance based on its redundancy.

3.3 Feature Selection Algorithm

In the next step of the proposed approach, the previous techniques (Correlation-based feature selection method (CFS), Principal Component Analysis (PCA), Information Gain Ratio based feature selection (IGR), and Minimum redundancy maximum relevance) will be working in conjunction with the Genetic (GA) and algorithm as a fitness function, where the GA algorithm will generate a separate Excel file as a refined dataset to be ready for the next phase in the KNIME-Classifiers.

Genetic algorithm is a heuristic research method used in artificial intelligence and computing. It is used to find improved solutions to search for problems based on natural selection theory and evolutionary biology. Genetic algorithms are excellent for searching through large and complex data sets. They can find reasonable solutions to complex problems as they are very able to solve unrestricted and restricted optimization problems. On the other hand, this concludes the final step in the proposed approach and how it started to generate new generations and apply crossover and mutation to meet the fitness function conditions.

3.4 KNIME-Classifiers

The proposed approach uses three kinds of classifier models using the KNIME Analytics Platform. Each classifier works separately for testing the accuracy of the generated files from the previous phase. Figure (2) shows the architecture of the proposed model in KNIME Analytics Platform:

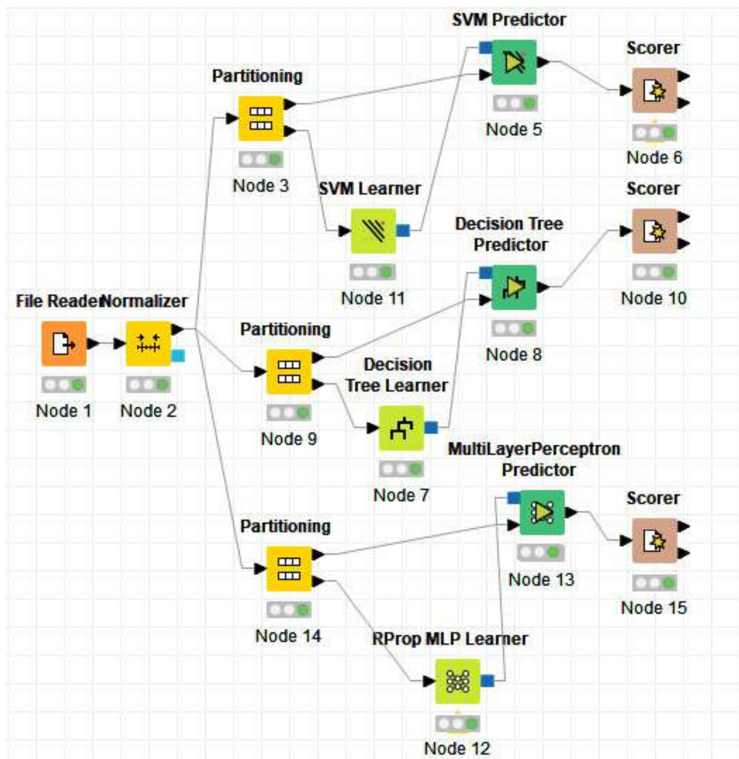

Figure 2: The classifier of KNIME Model

Table (3) describes the classifier model of supervised machine learning using the KNIME Analytics Platform:

Table 3: Supervised Machine Learning Model

Algorithm	KNIME Node	Description
Support Vector Machine (SVM)	- SVM Learner - SVM Predictor	This node trains a support vector machine on the input data (Refined Dataset). It supports several different kernels (HyperTangent, Polynomial, and RBF), and uses an SVM model generated by the SVM learner node to predict the output for given values.
RProp	- RProp MLP Learner - Multi-LayerPerceptron Predictor	Implementation of the RProp algorithm for multilayer feedforward networks.RPROP performs a local adaptation of the weight-updates according to the behavior of the error function.
Decision Tree	- Decision Tree Learner - Decision Tree Predictor	This node induces a classification decision tree in the main memory. The target attribute must be nominal (classes of attack).

4. Results and Analysis

This section presents the experimental findings for the proposed methodology, where the final results depend on three different classifiers (SVM, RProp, Decision Tree). In the first step in this section, the proposed method generates 10 excel files from MATLAB according to the algorithm Genetic algorithm (GA) to get the average of results and to reach the highest accuracy. In the next step, we convert the excel files from XLSX to CSV extension to be ready for the KNIME Analytics Platform. In the last step, we take the results from KNIME separately and classify them as algorithms and classifiers.

www.mecsj.com/
Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185

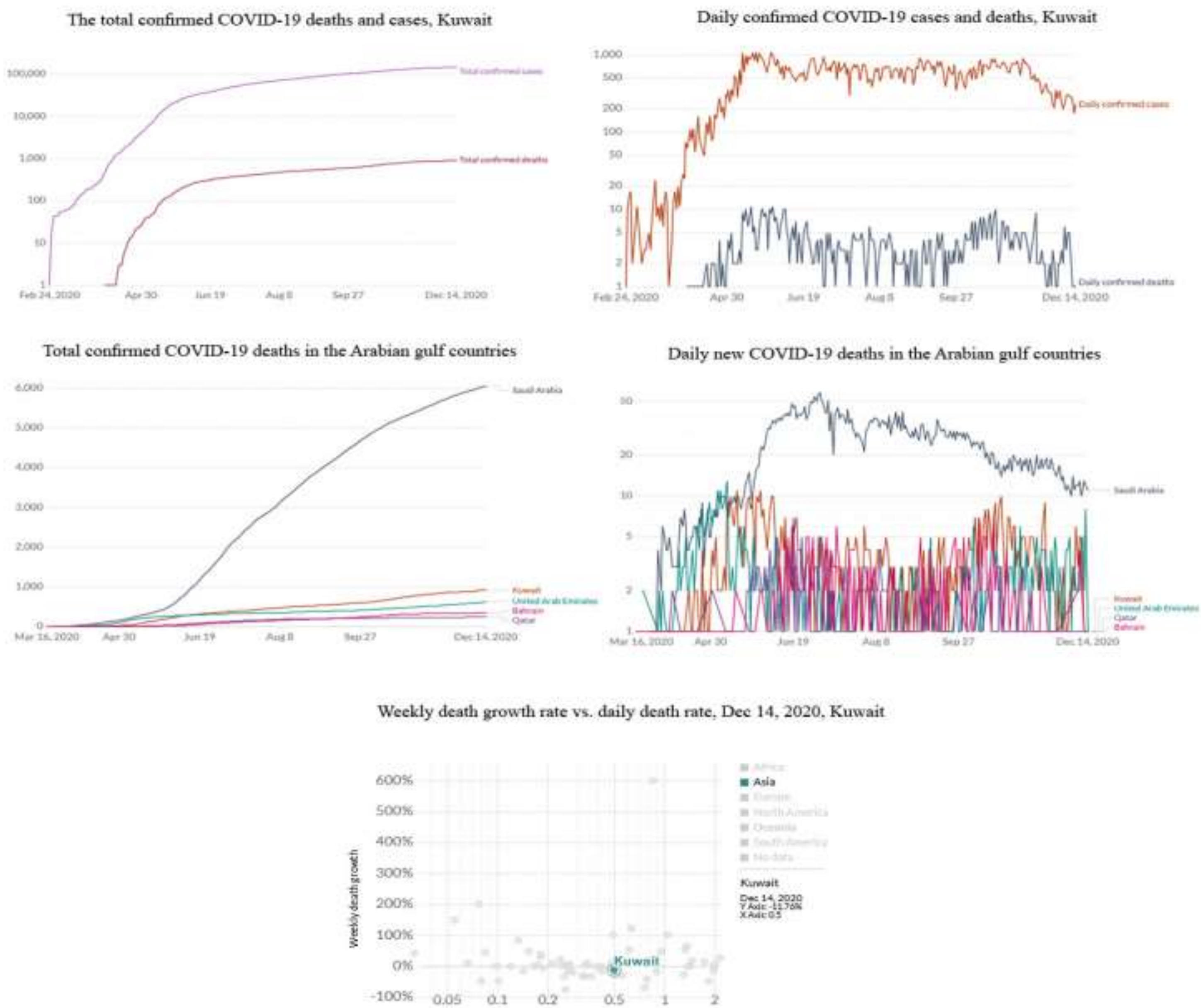

Figure 3: Our World in Data Analysis

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

In the next step, we compared the experimental results for the proposed approach with the "Our World in Data" results and analysis using the following factors figure (3), where the confirmed counts are shown in the "Our World in Data" is lower than the total counts. The main reason for this is limited testing and challenges in the attribution of the cause of death:

1. The total confirmed COVID-19 deaths and cases, Kuwait.
2. Daily confirmed COVID-19 cases and deaths, Kuwait.
3. Total confirmed COVID-19 deaths in the Arabian Gulf countries.
4. Daily new COVID-19 deaths in the Arabian Gulf countries.
5. Weekly death growth rate vs. daily death rate, Dec 14, 2020, Kuwait.

In the experiment results, we tested each of previous cases (total cases & death: Kuwait, daily cases & death: Kuwait, total death: Arabian gulf countries, daily new deaths: Arabian gulf countries, and weekly death growth vs. daily death: Kuwait) with GA results using the three classifiers (SVM, RProp, Decision Tree) to compute Confusion Matrix (the cases number of matching for the attribute rows with their classification match), the figure (4) show the confusion matrix for playing the Decision Tree classifier with the Total death, Arabian Gulf countries:

location \P...	Kuwait	Saudi Arabia	United Ara...	Oman	Qatar	Bahrain
Kuwait	59	0	0	0	0	0
Saudi Arabia	0	59	0	0	0	2
United Arab ...	0	0	54	0	0	7
Oman	0	0	0	61	0	0
Qatar	0	0	0	0	61	0
Bahrain	1	0	0	0	1	58

Correct classified: 352 Wrong classified: 11
Accuracy: 96.97 % Error: 3.03 %
Cohen's kappa (κ) 0.964

Figure 4: Confusion matrix for Decision Tree classifier with total death, Arabian Gulf countries

Table (4) shows the accuracy and error rates for the proposed method using the different classifiers (SVM, RProp, Decision Tree) with all cases (total cases & death: Kuwait, daily cases & death: Kuwait, total death: Arabian gulf countries, daily new deaths: Arabian gulf countries, and weekly death growth vs. daily death: Kuwait):

www.mecsj.com/
Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185

Table 4: The classifiers results with the five cases

	Accuracy (SVM)	Error (SVM)	Accuracy (RProp)	Error (RProp)	Accuracy (D.T.)	Error (D.T.)
Total cases & death, Kuwait	95.53%	4.47%	92.36%	7.64%	96.69%	3.31%
Daily cases & death, Kuwait	96.84%	3.16%	93.53%	6.47%	98.62%	1.38%
Total death, Arabian gulf countries	92.27%	7.74%	91.74%	8.26%	96.97%	3.03%
Daily new deaths, Arabian gulf countries	91.68%	8.33%	89.53%	10.47%	90.81%	9.19%
Weekly death growth vs. daily death, Kuwait	97.46%	2.54%	92.78%	7.22%	98.66%	1.34%

According to the table (4), the accuracy for the total cases & death, in Kuwait with SVM reached (95.53%), with RProp reached (92.36%), and with Decision Tree reached (96.69%). The accuracy for the daily cases & death, in Kuwait with SVM reached (96.84%), with RProp reached (93.53%), and with Decision Tree reached (98.62%). While the accuracy for the total death, in the Arabian Gulf countries with SVM, reached (92.27%), with RProp reached (91.74%), and with Decision Tree reached (96.97%). But for the daily new deaths, in the Arabian Gulf countries with SVM reached (91.68%), with RProp reached (89.53%), and Decision Tree reached (90.81%). Finally, the accuracy for the weekly death growth vs. daily death, in Kuwait with SVM reached (97.46%), with RProp reached (92.78%), and with Decision Tree reached (98.66%).

According to previous results, figure (5) shows the accuracy results for the proposed method using the different classifiers (SVM, RProp, Decision Tree) with all cases, and figure (6) shows the average rates for the proposed method using the different classifiers (SVM, RProp, Decision Tree) with all cases:

www.mecsj.com/
Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185

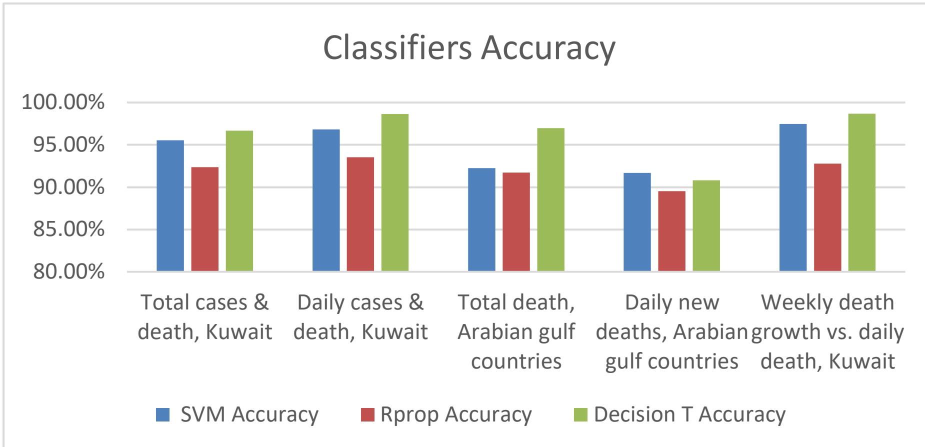

Figure 5: The accuracy results for the different classifiers

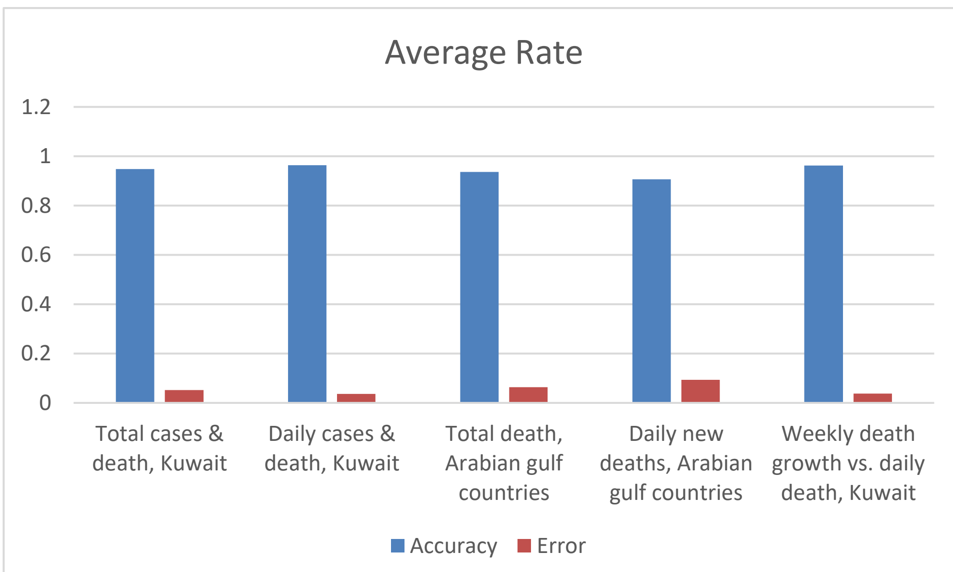

Figure 6: The average rates for the different classifiers

5. Discussion

According to previous results and analysis, the final results for the proposed method using the different classifiers (SVM, RProp, Decision Tree) give accuracy rate reached to (95%) and error rate reached to (5%) for all cases (total cases & death: Kuwait, daily cases & death: Kuwait, total death: Arabian gulf countries, daily new deaths: Arabian gulf countries, and weekly death growth vs. daily death: Kuwait), as it is shown in figure (7):

www.mecsj.com/
Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185

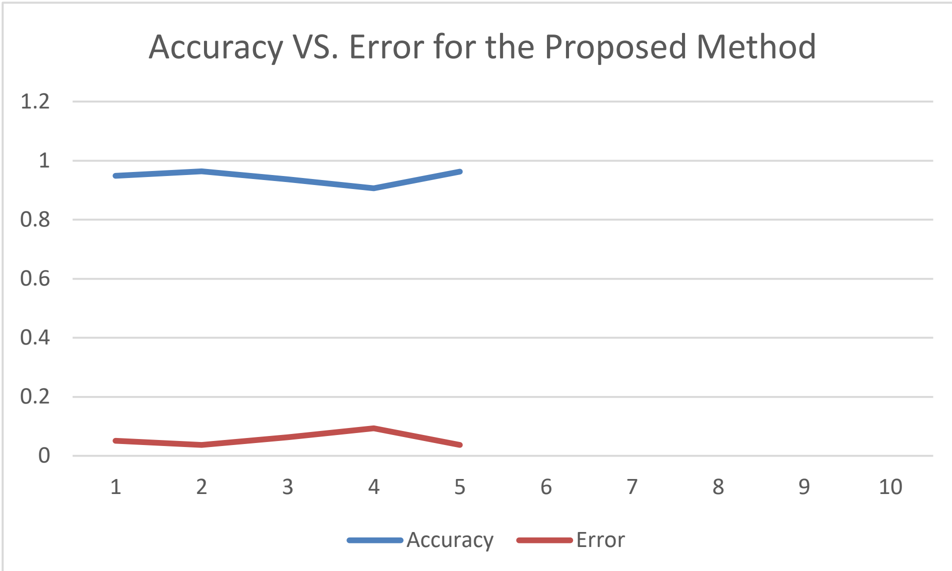

Figure 7: The accuracy and error rates for the proposed method

We can notice also that the Decision Tree classifier gives a high accuracy reached to (96.35%) with a low error rate reached to (3.65%), where the accuracy for SVM classifier reached to (94.75%) with error rate reached to (5.25%). While the RProp classifier gives a low accuracy reached (91.99%) with a high error rate reached to (8.01%), as shown in figure (8):

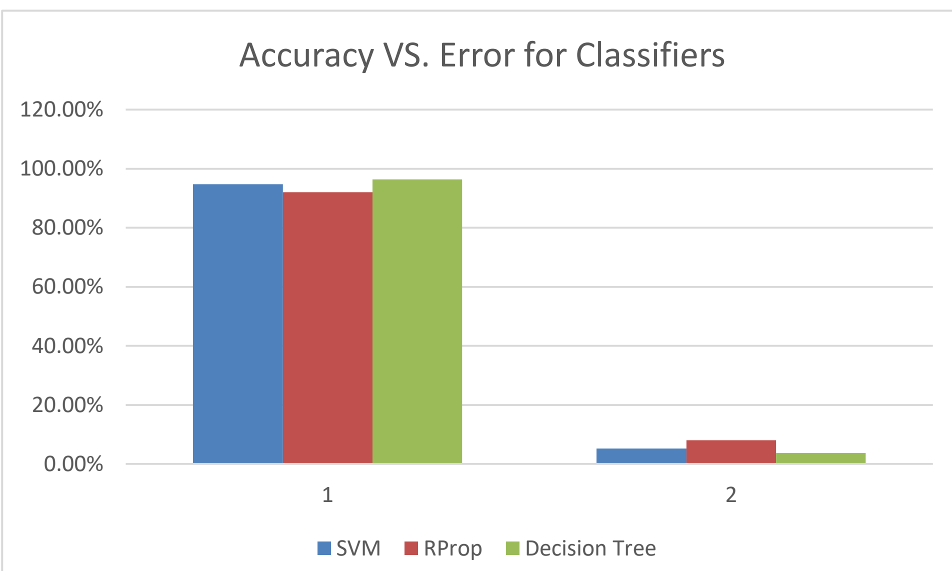

Figure 8: The accuracy and error rates for the classifiers

Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)
Issue (40),2021
ISSN: 2616-9185
www.mecsj.com/

6. Conclusion

Several types of research have suggested several approaches for COVID-19 prediction using different methods such as machine learning and deep learning. In the first section of this study, we reviewed a part of the literature for most studies and reviewed the most significant approaches that have been put forward in this field. In the second section, we presented the proposed method for the impact of COVID-19 on Arabian Gulf countries using the Genetic algorithm, and then we used the KNIME Analytics model for testing the GA results based on three machine learning methods (SVM, RProp, and Decision tree). The experiment results show that the accuracy for SVM, RProp, and Decision tree reached (94.75%), (91.99%), and (96.35%) respectively.

According to the experiments results based on the machine learning methods, we can notice the convergence of results between the proposed method (using the Genetic algorithm) and the Analysis of our world in data in all cases (total cases & death: Kuwait, daily cases & death: Kuwait, total death: Arabian gulf countries, daily new deaths: Arabian gulf countries, and weekly death growth vs. daily death: Kuwait). The KNIME Analytics model for testing the accuracy for the proposed method using the machine learning methods (SVM, RProp, Decision Tree) showed the advantage of the proposed method, where the accuracy reached (94.75%) for the SVM classifier, (91.99%) for the RProp classifier, and (96.35%) for the Decision tree classifier.

References

1. Khalifa, N. E. M., Taha, M. H. N., Manogaran, G., & Loey, M. (2020). A deep learning model and machine learning methods for the classification of potential coronavirus treatments on a single human cell. Journal of Nanoparticle Research, 22(11), 1-13. doi: https://doi.org/10.1007/s11051-020-05041-z.
2. Khanday, A. M. U. D., Rabani, S. T., Khan, Q. R., Rouf, N., & Din, M. M. U. (2020). Machine learning based approaches for detecting COVID-19 using clinical text data. International Journal of Information Technology, 12(3), 731-739. doi: https://doi.org/10.1007/s41870-020-00495-9.
3. Zame, W. R., Bica, I., Shen, C., Curth, A., Lee, H. S., Bailey, S., ... & van der Schaar, M. (2020). Machine learning for clinical trials in the era of COVID-19. Statistics in Biopharmaceutical Research, 1-12. doi: https://doi.org/10.1080/19466315.2020.1797867
4. 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. doi: https://doi.org/10.3390/math8060890
5. Jha, I. P., Awasthi, R., Kumar, A., Kumar, V., & Sethi, T. (2020). Learning the Mental Health Impact of COVID-19 in the United States with Explainable Artificial Intelligence. medRxiv. doi: https://doi.org/10.1101/2020.07.19.20157164
6. de Souza, D. G. B., Júnior, F. T. A., & Soma, N. Y. (2020). Forecasting COVID-19 cases at the Amazon region: a comparison of classical and machine learning models. bioRxiv. doi: https://doi.org/10.1101/2020.10.09.332908
7. Al Youha, S., Alkhamis, M., Al Mazeedi, S., Al-Haddad, M., Jamal, M. H., & Al-Sabah, S. K. (2020). Using Machine Learning to Unveil Demographic and Clinical Features of COVID-19 Symptomatic and Asymptomatic Patients. DOI: https://doi.org/10.21203/rs.3.rs-52330/v1
8. Duttaa, S., & Bandyopadhyaya, S. K. (2020). Machine learning approach for confirmation of COVID-19 cases: positive, negative, death and release. Iberoamerican journal of medicine, 03, 172-177 http://doi.org/10.5281/zenodo.3822623
9. Elaziz, MA, Hosny, KM, Salah, A, Darwish, MM, Lu, S, Sahlol, AT (2020) New machine learning method for image-based diagnosis of COVID-19. PLoS ONE 15(6): e0235187. doi: https://doi.org/10.1371/journal.pone.0235187

10. Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2020). A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Measurement, 167, 108288. doi: https://doi.org/10.1016/j.measurement.2020.108288
11. Ünlü, R., & Namlı, E. (2020). Machine Learning and Classical Forecasting Methods Based Decision Support Systems for COVID-19. CMC-COMPUTERS MATERIALS & CONTINUA, 64(3), 1383-1399.
12. Doanvo, A. L., Qian, X., Ramjee, D., Piontkivska, H., Desai, A. N., & Majumder, M. S. (2020). Machine Learning Maps Research Needs in COVID-19 Literature. bioRxiv. https://doi.org/10.1016/j.patter.2020.100123.
13. Loey, M., Smarandache, F. & Khalifa, N.E.M. (2020) A Deep Transfer Learning Model with Classical Data Augmentation and CGAN to Detect COVID-19 from Chest CT Radiography Digital Images. https://doi.org/10.20944/preprints202004.0252.v1
14. Liang, W., Yao, J., Chen, A., Lv, Q., Zanin, M., Liu, J., ... & Chen, G. (2020). Early triage of critically ill COVID-19 patients using deep learning. Nature communications, 11, 3543, 1-7. doi: https://doi.org/10.1038/s41467-020-17280-8.

Wu, F., Zhao, S., Yu, B., Chen. YM., Wang, W., Song, ZG., Hu, Y., Tao, ZW., Tian, JH., Pei, YY., Yuan, ML., Zhang, YL., Dai, FH., Liu, Y., Wang, QM., Zheng, JJ., Xu, L., Holmes, EC., & Zhang, YZ. (2020). A new coronavirus associated with human respiratory disease in china. Nature 44(59):265–269.

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei. Y., Xia, J., Yu, T., Zhang, X., & Zhang, L. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, *China: a descriptive study*. Lancet 395(10223):507–513.

Chen, J., Huang, H., Tian, S., and Qu, Y. (2009). Feature selection for text classification with Nave Bayes. *Expert Systems with Applications*. vol. 36, issue 3, pp. 54325435.

Mohi Ud Din, M., Rouf, N., Khan, Q.R., Rabani, S.T., Khanday, A.M. (2020). Machine learning based approaches for detecting COVID-19 using clinical text data, *International Journal of Information Technology,* 12(3), 731-739. doi: https://doi.org/10.1007/s41870-020-00495-9.

Hamid, H., Sugumaran, M., and Journaux, L. (2016). Machine Learning Techniques for Intrusion Detection: A Comparative Analysis. *Conference: The International Conference*. DOI: 10.1145/2980258.2980378.