Nat. Volatiles & Essent. Oils, 2021; 8(5): 4583 – 4587

Drugs For Corona Virus Prediction With Artificial Intelligence And Machine Learning

Dr.G.Revathy[1],Dr.E. Sophiya[2],Dr.M.P.Thiruvenkatasuresh[3], Mrs C.P Thamil Selvi[4],Mrs S Irin Sherly[5]

[1]Assistant Professor III, School of Computing, SASTRA Deemed to be University, Thanjavur. Email: revathyjayabaskar@gmail.com

[2]Assistant Professor,Department of Computer Science and Systems Engineering, Sree vidyanikethan Engineering College, Tirupati. Email: sophiya.e@vidyanikethan.edu

[3]Professor, Department of Computer Science and Engineering, Erode Sengunthar Engineering College(Autonomous), Perundurai. Erode, India. Email sureshmpt@gmail.com

[4]ASP -HoD(CSE),PPG Institute of Technology,Coimbatore. Email cpthamil.selvi72@gmail.com

[5]Associate Professor, Department of Information Technology,Panimalar Institute of Technology, Chennai. Email irinkutty@gmail.com

Abstract

The worldwide threat posed by the Severe Acute Respiratory Syndrome - Coronavirus (SARS-CoV-2) outbreak, also known as COVID-2019, has been realised. In terms of infrastructure, data sources, protective clothing, life-risk therapies, and a variety of other resources, the entire world is putting up tremendous effort to combat the spread of this fatal disease. Artificial intelligence researchers are concentrating their knowledge to create mathematical models for assessing and validating anticorona drugs/chemicals. To contribute to the well-being of living society, this article suggests the use of machine learning models with the goal of comprehending its daily exponential behaviour as well as predicting the COVID-2019's future reachability. The KNN algorithm shows the good result when compared with all other machine learning algorithms.

Keywords: Corona Virus, Artificial Intelligence, Machine Learning.

1. Introduction

Coronaviruses are a broad family of viruses that can infect humans and cause serious sickness. The first known severe epidemic, Severe Acute Respiratory Syndrome (SARS), occurred in 2003, and the second severe sickness outbreak, Middle East Respiratory Syndrome (MERS), began in 2012 in Saudi Arabia (MERS). The current coronavirus outbreak was first detected in late December 2019. This new virus is very contagious and has rapidly spread over the world.Because the outbreak has expanded to 18 nations, the World Health Organization (WHO) designated it a Public Health Emergency of International Concern (PHEIC) on January 30, 2020. WHO designated this as COVID-19 on February 11, 2020. WHO declared a pandemic on March 11 after the number of COVID-19 cases grew thirteen times outside of China, with over 118,000 cases in 114 countries and over 4,000 deaths.

Figure 1 Electronic Microscopic Image of Corona Virus

Currently, the global SARS-CoV-2 pandemic need a response. Antiviral medicines are urgently needed to stop the spread of the virus. a scarcity of To date, effective medicines necessitate the development of prediction models. computational technologies that can help and speed up the process medication repurposing options that are now being tested. Dynamic simulations based on virtual reality and molecular docking Antiviral drugs against SARS-CoV-2 are being testedhave already been investigated in this area. Repurposed Machine learning techniques (MLTs) are used to identify drugs. In CoVs' drug discovery enterprise, techniques are little explored.

Previously, MLT-based prediction algorithms were used to construct antiviral predictors However, our group just released 'CoronaVR,' a comprehensive platform for analysing and identifying epitopes for CoVs . The anti-CoVs data used in this investigation came from a comprehensive database of experimentally validated repurposed drugs called 'DrugRepV', which we just released.In this research, we've looked into repurposed drug candidates (against SARS-CoV-2, SARS, and other viruses) and MERS) employing a variety of MLTs such as the Support Vector Machine Artificial Intelligence (AI), SVM, Random Forest (RF), k-Nearest Neighbour (KNN) after exploring the DrugBank repository for anti-Corona chemicals as a result of the predictive models that have been constructed.

Drug repurposing achieved a level of recognition in 2020 that was unfathomable before the COVID-19 epidemic. Due to the isolation of researchers from their laboratory, several used computational methods to propose chemicals for testing against the virus. The idea is that a medicine will be developed as a result of thousands of researchers focusing on the COVID-19 problem and the deluge of articles that followed. While there have been a few tiny breakthroughs and several clinical trials underway in the case of COVID-19 therapies, a therapeutic candidate has yet to emerge from all of this computational repurposing.

2. Related Work

Since the last decade, digital technologies have played a vital role in important health-care issues such as illness prevention. The current global health emergency, COVID-2019, is also seeking technical assistance.In a paper, the authors discussed how trending digital technologies like the internet of things (IoT), big data analytics, artificial intelligence (AI), deep learning, and blockchain technology could be used to develop strategies for epidemic monitoring, detection, and prevention, as well as to determine the impact of the epidemic on the healthcare sector [9].

Author. [10] proposed a research project.To anticipate the spread of COVID-2019, the authors suggested an autoregressive integrated moving average (ARIMA) model. Based on a study on the prevalence and incidence of the COVID-2019, the author anticipated several metrics for the next two days in this paper. The correlogram and ARIMA forecast graph for epidemic incidence and prevalence are also shown in this study. Author [11] provided a time series technique for analysing the COVID-19 outbreak's incidence pattern and projected reproduction number. They used statistical analysis to look into the outbreak's trends in order to highlight the region's current epidemiological stage so that different strategies might be implemented.to be identified in order to combat the COVID-19 epidemic in many countries. In the current situation, it is critical to understand the infection's early transmission patterns in order to plan and regulate effective safety measures.

Author proposed a scientific model of crucial SARS-CoV-2 transmission by analysing the COVID-19 outbreak inside and outside Wuhan using various datasets. They used this to look into the possibility of a

disease outbreak spreading outside of Wuhan [8].Several studies employing exploratory data analysis (EDA) based on various public datasets have recently been undertaken on the epidemiological outbreak of COVID-19.The research focuses on the occurrence of confirmed, deceased, and recovered cases in Wuhan and around the world in order to better understand the hazards and plan containment activities [3].Author emphasised the problem of the criticality of the incubation period for COVID-19 in their study effort, allowing for improved surveillance and control operations to be planned [4].

Singer recently evaluated data from 25 afflicted counties in order to make short-term forecasts about the COVID 2019 outbreak. The study found that disease dissemination rates vary by region and follow either steady or explosive power-law growth with different scaling exponents.The authors used this knowledge to examine the impact of lockdown in various parts of the world [15].Based on the above literature, it is clear that sufficient work has been done on exploratory data analysis to understand the epidemic's current trend, but there is still a lot of room to develop and test effective machine learning-based prediction models so that proactive strategies can be identified to meet immediate needs.

3.Dataset

COVID-2019 day-to-day prevalence data was retrieved from the repository from April 1, 2021 to October 30, 2021. Daily case reports and daily time series summary tables make up the dataset. In this analysis, we used time-series summary tables in CSV format, with three tables for confirmed, death, and recovered cases of COVID-2019, each with six attributes (province/state, country/region, last update, confirmed, death, and recovered cases), and the dataset was updated once each day [2]. MLTS was used to construct multiple prediction models for SARS, including SVM, RF, KNN, and ANN.Using 10-fold cross-validation, the performance of the training/testing dataset containing 198 datasets was calculated . The PCC of the prediction model constructed using the training/testing dataset was 0.92, 0.76, 0.76, and 0.73, respectively, from SVM, RF, KNN, and ANN.

4.**Training and Testing**

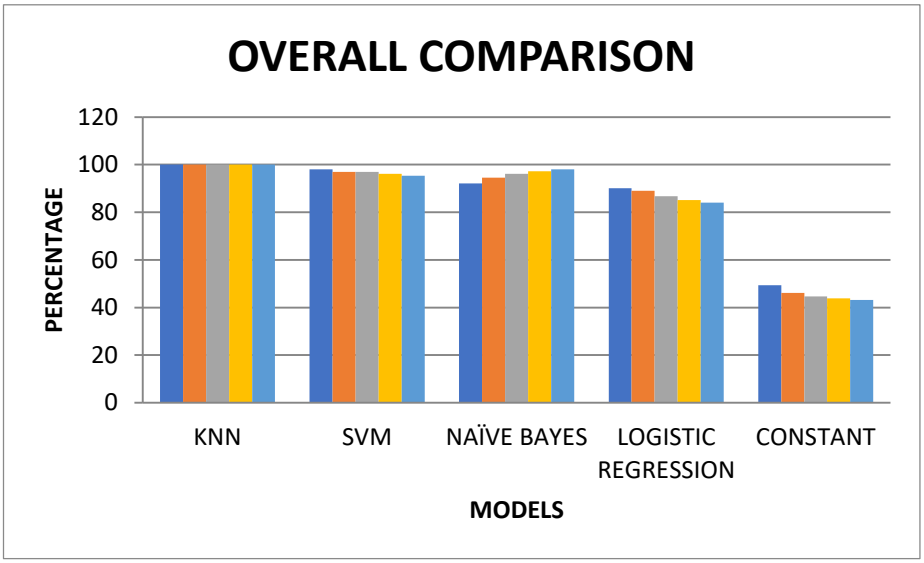
Since the regression approaches for epidemic analysis are trained and tested on real-time data [2], the number of confirmed, recovered, and death cases are utilised as the label for the subsequent day. Following lengthy trials, machine learning techniques are implemented with a polynomial kernel of degree 6 and various coecient values such as gamma=0.01, epsilon=1, and C=0.1. Mean squared error (MSE) is the most commonly used objective function, with root mean square error (RMSE) as a metric function for evaluating regression models. The MSE loss can be calculated using Equation 1. is the number of samples anticipated.The trained models are tested against the training data and then used to anticipate the number of confirmed recovered and death cases over the next 10 days due to the restricted availability of data.

5. Results and Discussions

We used the PaDEL-Descriptor software to create CoVs-specific prediction models from anti-corona chemicals [2].We generated 17,968 features in total, including 1D, 2D, and 3D chemical descriptors and fingerprints. The molecular descriptors are bits of information recorded in a chemical's molecular structure.They are divided into three categories based on their dimensionality: 1D, 2D, and 3D. The molecular weight and other fundamental information calculated from the molecular formula are shown in the 1D descriptors.The signatures calculated from two-dimensional molecular representations, intramolecular hydrogen bonding, and so on are described by 2D descriptors like the number of bonds, connectivity indices, and so on. As the name implies, 3D descriptors describe molecular features related to

the molecule's three-dimensional conformations, such as solvent accessible surface areas, intramolecular hydrogen bonding, and so on.

The fingerprints are another another technique of portraying molecules as mathematical objects, in which binary digits (bits) are utilised to locate and/or distinguish chemical substructures.These descriptors and fingerprints must be used together to create a quantitative structure–activity relationship (QSAR) for the chemical substances under investigation [5]. These descriptors are critical since they have been utilised in earlier studies to predict inhibitors against a variety of pathogenic pathogens [4]. Our created prediction models were used to repurpose the medications against the SARS-CoV-2, SARS, and MERS coronaviruses.We used the best performing SVM models in all three categories to predict the repurposed medicines.The "Approved" category of pharmaceuticals was retrieved from the DrugBank repository [5] for repurposing the drug categories. The PaDel software was used to calculate the descriptors and fingerprints of all 2468 authorised medicines.Furthermore, the approved drug descriptors were utilised to forecast extremely effective medications against all three types of viruses.



Comparison of all machine learning algorithms.

Based on the over all result comparison KNN shows the best result when compared with all other algorithms to give result for repurposing drug prediction.

6. Conclusion

The SARS-CoV2 (COVID-19) virus has taken over the world. Early detection of the transmission can aid in the implementation of appropriate measures. In anticipating the COVID-19 transmission, the results demonstrate that polynomial regression (PR) produced the lowest root mean square error (RMSE) score when compared to other methodologies. The drugs predicted is compared with machine learning algorithms such as KNN, SVM, Naïve Bayes, Logistic Regression and Constant in which KNN shows better result predicting the prescribed drug needed for Corona virus.

REFERNCES

1. de Wit E, van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. Nat Rev Microbiol 2016;14(8):523–34.

2. Barcena M, Oostergetel GT, Bartelink W, Faas FGA, Verkleij A, Rottier PJM, et al. Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirion. Proc Natl Acad Sci U S A 2009;106(2):582–7.
3. Neuman BW, Adair BD, Yoshioka C, Quispe JD, Orca G, Kuhn P, et al. Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy. J Virol 2006;80 (16):7918–28.
4. Woo PCY, Lau SKP, Lam CSF, Lau CCY, Tsang AKL, Lau JHN, et al. Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. J Virol 2012;86(7):3995–4008.
5. Leppard, Gary G. "Evaluation of electron microscope techniques for the description of aquatic colloids." In Environmental particles, pp. 231-289. CRC Press, 2019.
6. Jia, L., Li, K., Jiang, Y., Guo, X. (2020). Prediction and analysis of Coronavirus Disease2019. arXiv preprint arXiv:2003.05447.
7. Mrs.G.Revathy et.al , “Escalating quality of services with channel assignment and traffic scheduling in wireless mesh networks”, Cluster computing, Jan 2018.
8. Mrs G.Revathy et.al , “Route maintenance using tabu search and priority scheduling in wireless mesh networks”, Journal of advanced research in dynamical and control systems, vol 9,sp-6, 2017.
9. Dr G Revathy et.al , “ Of Derm Sickness Using Naïve Bayes Classification With Bipolar Fuzzylogic Set” , LINGUISTICA ANTVERPIENSIA, 2021 Issue-3, May 2021.
10. Kalipe, G., Gautham, V., Behera, R. K. (2018, December). Predicting malarial outbreak using Machine Learning and Deep Learning approach: A review and analysis. In 2018 International Conference on Information Technology (ICIT) (pp. 33-38). IEEE.
11. Singh, R., Singh, R., Bhatia, A. (2018). Sentiment analysis using Machine Learning technique to predict outbreaks and epidemics. Int. J. Adv. Sci. Res, 3(2), 19-24.
12. Ting, Daniel Shu Wei, Lawrence Carin, Victor Dzau, and Tien Y. Wong. "Digital technologyand COVID-19." Nature Medicine (2020): 1-3.
13. Benvenuto, Domenico, Marta Giovanetti, Lazzaro Vassallo, Silvia Angeletti, and MassimoCiccozzi. "Application of the ARIMA model on the COVID-2019 epidemic dataset." Data in brief (2020): 105340.