

An Overview on the Application of Artificial Intelligence in Infectious Disease Epidemiology

J. Suriakumar¹, R. Gopinathan¹, Gopinath Ramalingam², T. Murugalakshmi^{3*}

¹Department of Microbiology, Govt. Medical College, Dindigul, Tamil Nadu, India.

²Department of Microbiology, Govt. Theni Medical College, Theni, Tamil Nadu, India.

^{3*}Department of Pharmacology, Govt. Medical College, Dindigul, Tamil Nadu, India.

Corresponding Authors: T. Murugalakshmi^{3*}

^{3*}Department of Pharmacology, Govt. Medical College, Dindigul, TamilNadu

Abstract: Today, digital technologies are required to forecast, prevent, and control new infectious diseases. Infectious disorders can be transmitted either directly or indirectly by microorganisms such as viruses, bacteria, parasites, or fungus. Artificial intelligence advancements can assist scientists in forecasting infectious diseases to restrict pandemic spread, understanding the behavior of microbes, and accelerating medicine research to contain the sickness. Artificial intelligence not only gives scientists more confidence in their ability to predict, prevent, and control the spread of infectious diseases, but it also assists them in conducting faster clinical trials and identifying successful medication discoveries for treating the diseases. Artificial intelligence is on the verge of changing the healthcare system through disease-specific analysis and interventions to encourage faster, more reliable and cost-effective healthcare solutions for human well-being. This article addresses prospective AI applications that can help health institutions and the global population by fighting the spread of infectious illnesses.

Keywords: Artificial intelligence (AI), Healthcare, Infectious diseases, COVID-19, Prediction.

1. Introduction

Infectious diseases have a massive impact on the worldwide population's and economy's health. The readiness of healthcare system capacities to deal with pandemic risks remains critical. While the world waits for an effective vaccination against the COVID-19 virus, digital technologies such as Artificial Intelligence (AI) are playing an important role in global research investigations, biological tests, and clinical trials. Artificial intelligence techniques such as Machine Learning and Deep Learning can decode the virus's protein structures and forecast the efficacy of vaccines, which can build antibodies in the human body to combat the infection. AI models can be used to analyze hidden patterns in massive amounts of data in order to do hard tasks like categorization and prediction faster. Machine Learning models and artificial neural networks (ANN) can be trained on huge samples of structured and unstructured data from afflicted cells to determine the best feasible path for virus elimination [1].

Artificial intelligence is assisting global influenza tracking platforms by predicting novel influenza outbreaks across different locations and providing real-time insights on disease spread by analyzing social media communications to track potential outbreak events. Artificial intelligence applications can assist humanity in exercising preventative behavior in order to proactively avoid and manage the spread of infectious illnesses. Artificial intelligence has the potential to dramatically benefit the global healthcare system by playing an important role in epidemic forecasting, clinical diagnostics, and drug discovery. Investment in healthcare technologies, with a focus on the most vulnerable populations, can assist global societies prepare for disease outbreak identification, surveillance, and containment [2]. This article will address the function of artificial intelligence in healthcare and how it can aid the human population in the fight against infectious diseases.

Healthcare using artificial intelligence

Artificial intelligence (AI) has been acknowledged as the most powerful and promising analytical instrument for humans [3]. AI is the result of the input resource: large amounts of data that must be cleaned, processed, and

integrated. Big data is defined by volume, velocity, variety, variability, truthfulness, and complexity. These phrases allude to the amount of data, the speed with which data is received and transmitted, the variety of data types and sources, and the accuracy and correctness, respectively. However, most of the amount and velocity of data in health care today are insufficient to necessitate big data. The majority of health-related studies do not require the assistance of data scientists, but rather of bioinformaticians and statisticians. However, in the context of omics, where hundreds of thousands of data points are generated for gene polymorphism, gene expression, metabolomics, lipidomics, and proteomics, better techniques to detect individual cases from the overall orientation of the mass of data are required. The detection of weak signals allows for the early detection of trends before they become large and significant. This is very common in the realm of cybersecurity. In health care, this would imply discovering a signature in a small group of people and forecasting the clinical trajectory of the remainder of the population. Various data sets have been ingeniously used to forecast infectious disease epidemics. As previously said, the problem with infectious diseases is their unpredictability, as well as the numerous elements that influence the process of infection and transmission. AI is a type of computation that enables machines to act or react to stimuli in the same way that people do by executing cognitive tasks. Traditional computing, on the other hand, reacts to data, but the output must be hand coded to do so. Because no cognitive function is performed, independent intelligence is absent. Traditional computers cannot respond to an unforeseen scenario. In summary, AI systems constantly adapt their behavior to changes and adjust their responses accordingly. In one of its most prominent applications, machine learning, the computer learns once how to behave or react to a specific result and then understands how to act in the future.

Recent studies have demonstrated the potential of machine learning in image processing where traditional technologies failed to detect early indicators of sickness [4]. This is especially true for cancer [5,] where AI methods are frequently used to aid in diagnosis and therapy. This is applicable even in developing countries where resources, health-care costs, and other constraints prevent optimal care from being provided. A recent study demonstrated the feasibility of developing a low-cost point of care for cancer detection based on basic imaging and deep learning [6]. Several studies [7, 8] proposed using a Bayesian network (BN) to depict statistical interdependence. A BN is a graph-based model of joint multivariate probability distributions that captures the attributes of variable conditional independence [8]. Such databases, as well as traditional health-care datasets, are being utilized to better understand infectious illnesses, infection processes, treatment resistance, transmission, and vaccine creation (Fig. 2).



Fig.1: The role of the AI ecosystem in this endeavor is central

Healthcare solutions based on smartphones, as well as the use of emerging health technologies and digital practices in health care, such as AI, telemedicine or telehealth, mobile health, e-health, and electronic health records, are effective tools in combating pandemic-prone infectious diseases such as COVID-19 and dengue [9]. During the COVID-19 pandemic, electronic medical records (EMR) were successfully deployed for tracking suspected or confirmed COVID-19 patients [10].

Intelligent disease surveillance system

Sign-tracking devices are the most regularly used. AI-powered tracking tools, such as bracelets and rings, can be used to locate infected individuals. To detect epidemics and effective control actions, different monitoring apps (corona tracer bd, corona identifier, etc.) and GPS can be developed to identify COVID-19 and dengue patients. By browsing an internet address or database record, radio frequency identification (RFID) can be utilized to identify a tagged patient. Emerging AI-powered solutions dramatically improve capacity for disease and vector surveillance, as well as monitoring the environmental and societal risks of COVID-19 and dengue. The frequency-based detection of major dengue vector female *Aedes* mosquitoes using surface acoustic wave technology is also useful for eradicating *Aedes* mosquitoes as soon as they are identified and evacuating the area until they are treated [11].

Disease identification

Using neural network and fuzzy clustering approaches to classify patients at higher risk for COVID-19 and dengue can be a major step forward for the advanced diagnosis sector. Three popular machine learning approaches, Artificial Neural Network (ANN), Decision Tree (D.T.), and Naive Bayes (N.B.), are being tested for developing a diagnostic model for early detection of dengue cases and exploiting available dengue datasets [12].

Disease prediction

The application of AI-based emerging technologies can be used to anticipate the likelihood of infection, identify patterns of viral infections such as COVID-19 and dengue, and so minimize the suffering caused by the outbreak. The created AI-based monitoring system assists in tracking sick individuals from previous data and isolates them from non-infected individuals. The use of these technologies in pandemic situations in healthcare applications has demonstrated their effectiveness in tracking and preventing the spread of pandemic disease [13]. The early warning and response system (EWARS) is another useful toolset that uses outbreak and alarm indicators to generate prediction models that may be used to forecast a future dengue outbreak at the district level [14]. Forecasting weather using a cloud computing framework is utilized to detect dengue epidemics early and sanitize public locations using a drone to raise public safety and social awareness.

Pandemic Containment Using Artificial Intelligence

Infectious diseases represent significant threats to public health. According to recent flu outbreaks, increased population density and mobility have an important role in the transmission of new infectious diseases, perhaps leading to pandemics. The congregation of people in a small space, as well as connecting relationships with the sick patient, can increase the likelihood of disease spread. The pathogen's reproductive quantity and illness transmission rate will fluctuate across different geographies due to differences in population density, demographics, habitats, and behavior patterns. Artificial intelligence can not only provide real-time insights into disease propagation by processing massive amounts of data, but it can also predict new pandemic epicenters. AI algorithms can examine contact tracing and population movements to enable proactive detection, identification, and comprehension of pandemic epidemics. Infectious disorders, such as COVID 19, SARS, are more contagious when the susceptible group has poor immunity and is more vulnerable, such as when they are unwell from other diseases or are hospitalized. In this day and age, AI models are widely used to analyze massive amounts of data from infectious disease sources such as national surveillance systems, sentinel reporting systems, genome databases, outbreak investigation reports, vaccine reports, human dynamics information, social media platforms, internet search queries, vaccine reports, and human dynamics information. The massive inflow of data, data integration under master data management, and knowledge extraction enable the AI application to reveal latent tendencies. Artificial intelligence is also assisting in pandemic modeling and simulation of disease propagation information, which allows policymakers to implement appropriate healthcare policies [15].

AI applications in infectious diseases

The concern of infectious disease transmission has prompted authorities to implement procedures to identify persons who are at risk. As a result, temperature checks are undertaken systematically throughout Singapore airport terminals using a thermal camera to identify individuals with high temperatures. This minimum check is one of the actions taken ahead to prevent illness transmission. This type of surveillance is being improved by new approaches that use mathematical modeling. A similar approach was devised to detect infected individuals using vital signs classification [16]. As a result, the neural network and fuzzy clustering methods were utilized to successfully classify persons at higher risk for influenza using respiration rate, heart rate, and facial temperature. The addition of membership values (degree of belongingness to a cluster based on edge/centroid position in the said cluster) and the fuzzifier distinguishes fuzzy clustering methods from k-means clustering. As a result, unlike nonfuzzy clustering algorithms, each point can belong to several clusters. This highlights the ability to create effective strategies for identifying vulnerable populations. Even in the event of acute infectious diseases, where efforts must be prioritized, triage is vital and part of the process. Machine learning technologies can be applied in increasingly advanced scenarios. A combination of the support vector machine (SVM) learning algorithm, Matlab, the leave one out cross-validation (LOOCV) approach, and nested one-versus-one (OVO) SVM, for example, was utilized to better separate gene sequences from bacteria than other methods such as high-resolution melt (HRM). The combination of SVM and HRM could identify isolated bacteria with excellent accuracy (100%) [17]. The accuracy of real-life biological

samples, blood samples from patients, was compromised, demonstrating the limits of building tools using data obtained in a controlled setting (laboratory). It is unknown if this was related to poor biological sample quality or to the interactions of bacteria in a natural setting. Nonetheless, this demonstrates that the mathematical methods produced should take into account practical aspects such as sample quality and laboratory process duration. This was addressed in the example of tuberculosis diagnosis, the world's second highest cause of death from infection [18]. Existing systems for disease diagnosis exist, such as the artificial immune recognition system (AIRS). AIRS was created by utilizing a property of the immune system. The immune system's function is to sense and remember threats. Immunological memory is perhaps the most essential element of immunity since it helps us to respond more effectively when the threat (infectious agent) is met again. This is consistent with the development of AI tools based on human cognitive function; the main difference is that intelligence is delegated to the periphery (blood). As a classifier, AIRS employs k-nearest neighbor (kNN). There are a few drawbacks with kNN in machine learning: (i) it detects data patterns without requiring an exact match to known patterns, which results in low accuracy; and (ii) if k is too little or high, there may be issues with noise and loose neighborhood. The AIRS that employs supervised machine learning algorithms [19] has demonstrated high accuracy [20].

Transmission and epidemiology

Epidemiological research can be conducted at the population or patient level (clinical epidemiology). Epidemiological investigations should be conducted in a specified order, with infection-related data collected longitudinally. Mathematical models, when used correctly, can estimate the scale of developing infectious diseases. There are large datasets and prediction models for noncommunicable diseases (NCD). Over a 5-year period, a recent study collected data from 50 American states on a variety of NCDs such as diabetes, cardiovascular disease, hypertension, and others [21]. Data from 30 states was used for training, and the remaining 20 states were tested. This massive amount of data, combined with machine learning modeling, allowed for near-realistic results. However, the lack of transfer from patient to patient due to proximity or shared environment in a short period of time is what defines NCD. Infectious diseases are defined by their ability to spread from one person to another in relatively short periods of time. The severity of symptoms and death related with the virus are pushing the necessity of predicting the epidemic's future size and location. Many of the machine learning methods presented in this chapter, in addition to providing indications on size and location, which is frequently communicated information, are primarily used by mathematicians to estimate infection-related variables (e.g., incubation time, transmission mode, symptoms, resistance to treatments). The input data for epidemiological investigations is quite diversified, allowing it to use various AI assets. Epidemiological studies have shown that we may forecast an epidemic from a very tiny incidence, as demonstrated in a recent study on Kyasanur forest illness, a tick-borne viral infectious disease [22]. The team of scientists demonstrated a high prediction rate and proposed localization data to be used in future databases in order to better manage transmission using an extremal optimization adjusted neural network. Recent deadly outbreaks, such as Ebola, have compelled the world to innovate in the field of prediction. Scientists created an ensemble of predictors using machine learning, a single-layer artificial neural network (ANN), logistic regression (LR), decision tree (DT), and SVM classifiers, which could be applied to diverse combinations of Ebola-related data. A significant challenge in such health crises is the lack of rapid response and the low quality of data from the earliest episode of infection. The history of transmission, like forensic science, is heavily reliant on the early steps. Colubri et al. demonstrated how to deal with missing information and/or limited sample size difficulties when using machine learning approaches: a composite of machine learning approaches rather than a single model [23].

The Difficulties of Using Artificial Intelligence for Infectious Disease Applications

It is difficult to apply traditional statistical approaches to large-dimensional medical data (such as textual and image data). When dealing with high dimension data with complexity, the basic assumptions of classic statistical-based methods are frequently violated. Most traditional statistical methods are incapable of dealing with Big Data concerns because emerging infectious disease outbreak data comes in a variety of forms, each with its own set of restrictions and assumptions. Big Data, AI, and machine learning technologies have so far yielded promising results in a variety of business and industrial settings, revealing hidden patterns and forecasting future possibilities. Deep learning artificial neural network techniques, for example, have demonstrated promising results in retrieving highly nonlinear features from large datasets. Recent research has used AI approaches to track down rodent reservoirs of future zoonotic diseases [24], anticipate Extended-spectrum β -lactamase (ESBL) generating organisms [25], and

manage tuberculosis (TB) and gonorrhea disease outbreaks [26]. The public's reaction to illness epidemics can be unpredictable. However, with the introduction of Big Data and AI technologies, we are becoming increasingly capable of correlating population behavior with disease epidemics. For example, devoted researchers [27] analyzed infectious disease outbreaks linked digital activity patterns through web search patterns (such as Google Trend) during the outbreak time using a behavioral informatics and analytics strategy. It is hoped that the growth of AI approaches for infectious illness Big Data analytics would increase our ability to dynamically watch public reactions to disease outbreaks and properly predict disease spread, allowing authorities to implement timely response measures to infectious diseases. Overall, the machine learning applications detailed in table 1 have considerably enhanced the management of infectious diseases. While this demonstrates AI's immense promise, there are still many aspects that require improvement in order to properly harness AI's capacity to assist remove undesirable infections, minimize the burden of seasonal viruses, and better comprehend pathogen-human interactions.

Table.1 AI technologies, their applications, and the probable outcomes

Artificial Intelligence	Infectious diseases	Outcomes
• Bayesian networks • Artificial neural network • Fuzzy clustering • k-Nearest neighbor • Decision forest • Random forest • ARIMA • Unsupervised learning • Super learner	• Pathogen mutation • Diagnosis • Zoonosis • Outbreak • Source of infection • Pandemic prediction • Drug discovery • Host genetic • Host pathogen interaction • Missing data	• Decision support • Reducing time for diagnosis, epidemic prediction, drug discovery • Identification of strategies for blocking transmission • Improving health • Saving life • Saving costs • Better be prepared • Forensic approach

AI applications in COVID-19

At the start of the Pandemic, governments and the general public are focused on the following issues: the duration and peak of the COVID-19 Pandemic, the number of infected people, the influencing variables, and the measures to limit infection, among others. Many works were done by a community of applied mathematicians, virologists, epidemiologists, and AI experts to assist governments in developing appropriate policies in the fight [28]. This section focuses on AI applications in various epidemiological areas, such as developing models to predict COVID-19 spread and severity, identifying and tracing infection cases to control the infectious rate, and researching the related influencing factors to prevent the spread [29].

AI techniques for infectious illness data

AI and machine learning approaches can be divided into two categories depending on learning strategies: supervised learning and unsupervised learning. Supervised learning is the task of inferring a function from labeled training data, and it includes Support Vector Machine (SVM), Decision Tree, Random Forest, Naive Bayes (NB), Artificial Neural Network (ANN), Bootstrap Aggregating, and AdaBoost, which can efficiently handle classification and regression problems in medical data [30]. All of these strategies could be effective in increasing diagnosis accuracy and recommending appropriate treatment for patients. Furthermore, the predictions from these methodologies might be used to alert authorities and the general public of impending dangers and suggest relevant prevention and control strategies. Unsupervised learning approaches such as Principal Component Analysis (PCA) could be used to minimize the dimensions of data, allowing researchers to find some essential characteristics connected to infectious disease [31]. Other unsupervised learning methods, such as K-means, could categorize patients into categories and spot anomalous patients, directing researchers to these medical instances. Topic models like Latent Dirichlet allocation (LDA) could also be used to extract topics from medical textual records. Deep learning architectures have recently seen widespread application in prediction and classification, social network filtering, and bioinformatics, and are regarded as powerful tools for infectious disease analytics.

2. Conclusion

Artificial intelligence systems are becoming a necessary part of our daily lives as they revolutionize the healthcare industry for the benefit of global populations. It can assist medical researchers in developing novel medical vaccines and treatments while assuring the medicine's effectiveness, accuracy, and reliability for patient safety. From a single sickness to multiple diseases, from a single piece of information to numerous pieces of information, more and more AI medical procedures are focusing on integrity and integration. AI technology will be used throughout the health-care life cycle, including pre-, during-, and post-diagnosis. Along with the growth of biological research, AI applications aid in the speedier analysis of large amounts of infectious illness data, allowing policymakers, medical professionals, and healthcare institutions to respond to any future diseases.

3. References

- [1] da Silva Motta, et al., "Use of Artificial Intelligence on the Control of Vector-Borne Diseases." *Vectors and Vector-Borne Zoonotic Diseases*. (2018).
- [2] Davies SE. Artificial intelligence in global health. *Ethics & International Affairs*. 2019;33(2):181-92.
- [3] Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y. Mastering the game of go without human knowledge. *nature*. 2017 Oct;550(7676):354-9.
- [4] Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. *The New England journal of medicine*. 2017 Jun 6;376(26):2507.
- [5] Boon IS, Au Yong TP, Boon CS. Assessing the role of artificial intelligence (AI) in clinical oncology: utility of machine learning in radiotherapy target volume delineation. *Medicines*. 2018 Dec 11;5(4):131.
- [6] Im H, Pathania D, McFarland PJ, Sohani AR, Degani I, Allen M, Coble B, Kilcoyne A, Hong S, Rohrer L, Abramson JS. Design and clinical validation of a point-of-care device for the diagnosis of lymphoma via contrast-enhanced microholography and machine learning. *Nature biomedical engineering*. 2018 Sep;2(9):666-74.
- [7] Xu J, Wickramaratne TL, Chawla NV. Representing higher-order dependencies in networks. *Science advances*. 2016 May 20;2(5):e1600028.
- [8] Belle A, Kon MA, Najarian K. Biomedical informatics for computer-aided decision support systems: a survey. *The Scientific World Journal*. 2013 Jan 1;2013.
- [9] Ye J. The role of health technology and informatics in a global public health emergency: practices and implications from the COVID-19 pandemic. *JMIR medical informatics*. 2020 Jul 14;8(7):e19866.
- [10] Pryor R, Atkinson C, Cooper K, Doll M, Godbout E, Stevens MP, Bearman G. The electronic medical record and COVID-19: Is it up to the challenge?. *American journal of infection control*. 2020 Aug 1;48(8):966-7.
- [11] Salim ZT, Hashim U, Arshad MM, Fakhri MA, Salim ET. Frequency-based detection of female Aedes mosquito using surface acoustic wave technology: Early prevention of dengue fever. *Microelectronic Engineering*. 2017 Jul 5;179:83-90.
- [12] Gambhir S, Malik SK, Kumar Y. The diagnosis of dengue disease: An evaluation of three machine learning approaches. *International Journal of Healthcare Information Systems and Informatics (IJHISI)*. 2018 Jul 1;13(3):1-9.
- [13] Sitharthan R, Rajesh M. RETRACTED ARTICLE: Application of machine learning (ML) and internet of things (IoT) in healthcare to predict and tackle pandemic situation. *Distributed and Parallel Databases*. 2022 Dec;40(4):887-.
- [14] Rahman MS, Pientong C, Zafar S, Ekalaksananan T, Paul RE, Haque U, Rocklöv J, Overgaard HJ. Mapping the spatial distribution of the dengue vector Aedes aegypti and predicting its abundance in northeastern Thailand using machine-learning approach. *One Health*. 2021 Dec 1;13:100358.
- [15] Jain K. Artificial intelligence applications in handling the infectious diseases. *Prim Health Care*. 2020;10(5):351.
- [16] Sun G, Matsui T, Hakozaiki Y, Abe S. An infectious disease/fever screening radar system which stratifies higher-risk patients within ten seconds using a neural network and the fuzzy grouping method. *Journal of Infection*. 2015 Mar 1;70(3):230-6.
- [17] Fraley SI, Athamanolap P, Masek BJ, Hardick J, Carroll KC, Hsieh YH, Rothman RE, Gaydos CA, Wang TH, Yang S. Nested machine learning facilitates increased sequence content for large-scale automated high resolution melt genotyping. *Scientific reports*. 2016 Jan 18;6(1):19218.

- [18] Saybani MR, Shamshirband S, Hormozi SG, Wah TY, Aghabozorgi S, Pourhoseingholi MA, Olariu T. Diagnosing tuberculosis with a novel support vector machine-based artificial immune recognition system. *Iranian Red Crescent Medical Journal*. **2015** Apr;17(4).
- [19] Watkins A, Boggess L. A new classifier based on resource limited artificial immune systems. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)* **2002** May 12 (Vol. 2, pp. 1546-1551). IEEE.
- [20] Cuevas E, Osuna-Enciso V, Zaldivar D, Perez-Cisneros M, Sossa H. Multithreshold segmentation based on artificial immune systems. *Mathematical Problems in Engineering*. **2012** Jan 1;2012.
- [21] Luo W, Nguyen T, Nichols M, Tran T, Rana S, Gupta S, Phung D, Venkatesh S, Allender S. Is demography destiny? Application of machine learning techniques to accurately predict population health outcomes from a minimal demographic dataset. *PloS one*. **2015** May 4;10(5):e0125602.
- [22] Majumdar A, Debnath T, Sood SK, Baishnab KL. Kyasanur forest disease classification framework using novel extremal optimization tuned neural network in fog computing environment. *Journal of medical systems*. **2018** Oct;42:1-6.
- [23] Colubri A, Silver T, Fradet T, Retzepi K, Fry B, Sabeti P. Transforming clinical data into actionable prognosis models: machine-learning framework and field-deployable app to predict outcome of Ebola patients. *PLoS neglected tropical diseases*. **2016** Mar 18;10(3):e0004549.
- [24] Han BA, Schmidt JP, Bowden SE, Drake JM. Rodent reservoirs of future zoonotic diseases. *Proc Natl Acad Sci United States Am* **2015**;112(22):7039e44.
- [25] Goodman KE, Lessler J, Cosgrove SE, Harris AD, Lautenbach E, Han JH, et al. A clinical decision tree to predict whether a bacteremic patient is infected with an extended-spectrum beta-lactamase-producing organism. *Clin Infect Dis Off Pub Infect Dis Soc Am* **2016**;63(7):896e903.
- [26] Blumenthal A. Artificial Intelligence to fight the spread of infectious diseases. *Phys Org* **2018** 20 Feb [cited 2018 25 Oct]. Available from: <https://phys.org/news/2018-02-artificialintelligence-infectious-diseases.html>.
- [27] Bragazzi NL, Alicino C, Trucchi C, Paganino C, Barberis I, Martini M, et al. Global reaction to the recent outbreaks of Zika virus: insights from a Big Data analysis. *PloS One* **2017**; 12(9), e0185263.
- [28] Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* **2020**;382:1708-20.
- [29] Chan JF, Yuan S, Kok KH, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* **2020**;395:514-23.
- [30] Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. *Artif Intell Med* **2001**;23(1): 89e109.
- [31] Wang Y, Fan Y, Bhatt P, Davatzikos C. High-dimensional pattern regression using machine learning: from medical images to continuous clinical variables. *Neuroimage* **2010**;50(4): 1519e35.