

ARTIFICIAL INTELLIGENCE – PAVING THE PATH FOR PREDICTIVE MODELLING IN PUBLIC HEALTH

Abstract

As the 21st century unfolds, artificial intelligence (AI) has emerged as a transformative force that is reshaping industries, augmenting human capabilities, and facilitating awe-inspiring innovations. At its core, AI seeks to replicate human intelligence within machines, enabling them to learn, reason, and make decisions, paving the way for a future that was once confined to the realms of science fiction. As societies around the world grapple with complex health issues, AI is emerging as a potent tool that not only aids in diagnosis and treatment but also enables proactive measures through predictive modelling (PM). A pivotal aspect of AI, PM involves creating algorithms that learn from historical data to forecast future trends, enabling health authorities to proactively address potential crises. This chapter seeks to provide a basic understanding of the PM types, its working principle and its role in addressing the issues of public health. In addition, this chapter also addresses the ethical concerns associated with AI.

Keywords: Artificial intelligence, predictive modelling, public health, outcomes

Authors

Dr. Rashmi Agarwalla

Assistant Professor
Department of Community and Family
Medicine
All India Institute of Medical Sciences
(AIIMS)
Guwahati
rashmiagarwalla@gmail.com

Dr. Himashree Bhattacharyya

Associate Professor
Department of Community and Family
Medicine
All India Institute of Medical Sciences
(AIIMS)
Guwahati
himashreebhattacharyya@gmail.com

Dr. Subhrajeet Chakraborty

Senior Resident
Department of Community and Family
Medicine
All India Institute of Medical Sciences
(AIIMS)
Guwahati
dr.subhrajeet93@gmail.com

Dr. Ankur Khandelwal

Associate Professor
Department of Anaesthesiology, Critical
Care and Pain Medicine
All India Institute of Medical Sciences
(AIIMS)
Guwahati
ankurchintus@gmail.com

I. INTRODUCTION

As time keeps progressing it keeps getting more and more evident that digital transformation of healthcare is not just an option rather a mandate. The goal of Predictive Modelling (PM) is to identify and predict future events and their likelihood [1]. PM has been used in various fields such as predicting deterioration of patient's health, end-of-life care, chronic care management, etc. Apt utilization of PM in the field of public health can revolutionize health care delivery for the mass.

The term Artificial Intelligence (AI) was coined by John McCarthy in 1956. It is a wide branch of computer science that is involved in creation of machines and systems which can conduct tasks that are otherwise too complicated for a machine [2]. The core philosophy behind it is to mimic human intelligence in a way, so as to enable the machine or system to recognize patterns, perform tasks or predict outcomes on the basis of data acquired through multiple sources such as user databases [3]. Although, the term AI invokes a lot of worry and anxiety amongst a certain section of society, it is already in full-fledged use in the form of Self Driving Cars, Chatbots, Digital Assistants, etc. PM on the other hand is a typical data science method that analyses previous data to generate future predictions [2]. Techniques including data mining, modelling, machine learning, AI and statistics are used to do this. It is frequently used as a tool to assist in recognition of impending hazards and opportunities.

II. AI & PM IN PUBLIC HEALTH

A large number of functions of AI and PM can be used in the field of public health. With the rapidity of progress that AI is making, it is expected that AI will be integrated in almost every aspect of public health.

III. EARLY WARNING SYSTEMS FOR EPIDEMICS

Numerous unfiltered, open-source data sources, such as news articles and social media posts, record the issues and discussions of the community. These data, if consistently mined, may reveal early warning signs of epidemics before they are discovered by health officials [4]. The traditional methods of detection of outbreak depend upon reporting by healthcare and laboratory personnel, which is not rapid enough to enable early detection and subsequent prevention of an epidemic. During the recent COVID-19 outbreak in China, the first cases of COVID-19 presenting with severe pneumonia of unknown etiology were formally reported on December 8, 2019. However, a retrospective analysis employing open-source intelligence data identified another COVID-19 case in China around mid-November 2019 [5]. Similarly, even in Guinea, which has a relatively low smartphone usage, the Ebola epidemic could have been identified in late 2013, three months before the World Health Organization (WHO) was alerted, using quick social media-based intelligence and surveillance [6]. There are multiple web-based warning systems which are currently being used for the surveillance of public health events. The three major ones are discussed here –

1. **Pro MED-Mail:** The International Society for Infectious Diseases created the Pro MED-mail system in 1994 to track down uncommon health occurrences that affect people, animals, and plants [7]. With a network of employees from at least 30 countries working

in different time zones and nearly 80,000 subscribers from around 200 countries, Pro MED-mail operates 24/7 and has helped identify a number of significant outbreaks [8].

- 2. Epidemic Intelligence from Open Sources (EIOS):** EIOS was developed in collaboration between WHO and the Joint Research Commission of the European Commission [9]. EIOS is estimated to have a system capacity of at least 40 million articles from 12,000 web sources, including social media in multiple languages. The system includes native language processing recognition technology, classified articles and preferential algorithms to identify, tag and categorize reports [9]. System access is granted exclusively to the WHO and specific agencies or countries.
- 3. Health Map:** Health Map is a fully automated system that is not subjected to human intervention. It reports all health events including non-communicable diseases and is therefore not specific to infectious epidemics [10]. The system contains a number of modules including a data collection tool, a classification engine, a backend web application, a database and a backend web application that enables the system to function properly [11].

IV. DISASTER MANAGEMENT

Predictive analytics can help identify areas of particular vulnerability to future catastrophic events. It can also be used to identify patterns and trends of natural disasters. The effectiveness of a catastrophe response can be tracked with the help of predictive analytics. Authorities can identify areas in need of more resources or support and modify their response by examining the data. Recently, even machine learning has been employed to analyse satellite imagery to identify areas particularly damaged by the calamity for prioritisation of response to those areas.

- 1. Screening and Diagnosis of Diseases:** Machine learning has enabled systems to learn autonomously by analyzing training data and experience without explicit programming. Furthermore, its performance improves with time [12]. In outpatient clinics, a deep neural network trained on more than 37,000 head computed tomography scans of cerebral bleeding examined 9,500 uncovered cases, cutting the time to diagnosis by 96% with an accuracy rate of 84% [13].
- 2. Simulation of Public Health Measures:** Various forms of PM can be utilized to estimate the cost benefit ratio of the public health measures which are being implemented across the nation. However, to generate such models, one need to have a robust data collection system in place which can curate data from sources such as social media, medical records etc. AI can be a vital tool in this process of data collection.
- 3. Predictive Models of Health Outcomes:** AI can assist in generation of PMs of public health outcomes by providing data-driven tools to identify and monitor etiological risk factors for diseases and other health-related issues. For instance, one National Institutes of Health study discovered that AI chatbot systems could predict the likelihood of type 2 diabetes among Spanish-speaking people with up to 90% accuracy. Data from patient records, including age, sex, and lifestyle, as well as information from social media,

including dietary preferences and physical activity, were gathered and analysed by the AI chatbot system [14].

V. TYPES OF PREDICTIVE MODELS IN PUBLIC HEALTH

Creating a PM in public health is a multi-step process. First, managers set goals for their model, such as defining what they intend to measure or predict. Thereafter, the individual or group begins to collect and enter the data. During this phase, they can also sort the data, classifying it as needed for use in the model. The main models used in predictive analysis are as follows:

- 1. Classification Model:** It is one of the most basic types of modelling which produces either ‘Yes’ or ‘No’ responses to questions. It uses historical data to produce answers to a query.
- 2. Forecast Model:** It is a very popular model and it works on anything with numerical value based on learning from historical data. For example, by answering how many cases of influenza requiring admission might report to the health centre next week or the number of calls a telemedicine centre can handle per day or week.
- 3. Clustering Model:** The clustering model separates data into different categories based on similar attributes. It then uses data from each cluster to determine their large-scale outcomes. This model works by using two types of clustering; hard clustering classifies data by determining whether each point completely belongs to a certain cluster. Soft clustering assigns probabilities to each data point instead of separating them into separate clusters.
- 4. Outliers Model:** An outliers model identifies unusual or outlying information within a dataset. It can analyse individual instances of unusual data or connections to other categories and numbers.
- 5. Neural Network:** A neural network, much like the human brain is a complex model. A neural network is a complex model that resembles the human brain. It involves multiple algorithms working together to identify patterns, group data, and create categories for different data sets. Neural networks usually consist of three layers. The input layer puts the data in the next layer, which is the hidden layer. The hidden layer includes predictive generation functions. Data from various predictors are gathered by the output layer, which then creates a comprehensive, final conclusion. Policy makers can use these networks in conjunction with other PMs, such as time series or clustering.

VI. ETHICAL ISSUES AND BARRIERS PERTAINING TO AI IN PUBLIC HEALTH

Ever since AI has staked a claim in public sphere, it has faced a great deal of criticism, of which a few are indeed genuine and thought provoking. One of the most pertinent criticisms of AI has been that it has the potential of biases. Underlying reasons for biases are often complicated and technically tough to explain, but since these AI applications “learn” from data produced in biased societies, they are shaped by both information biases and societal biases [15]. Algorithmic bias against a marginalized community can exist even if

that group's social identity or position is not provided directly to the algorithm, since AI methods readily identify latent constructs reflected in combinations of other variables [15]. Another concern that AI systems raise pertains to data security and privacy. Since health records are important and vulnerable, hackers often target them during data breaches [16]. Availability of big data and its accessibility to programmers is another major roadblock. Since patient data is regarded as confidential, there is understandably a great deal of reluctance amongst institutions to exchange data [17]. The question of accountability is another area of concern as because if a prediction goes wrong, subsequently adversely affecting the medical or administrative management of an issue, it shall be hard to hold the clinician accountable since he or she had no role in developing the system. Moreover, it shall be difficult to hold the developer responsible since their experience is not related to the clinical set up [18].

VII. CONCLUSION

With the rapid advancements in science and technology, increasingly sophisticated PM methods have been used in public health. The primary focus of PM techniques should be to strengthen health systems and empowering health care providers, policy makers and consumers to achieve the vision of health for all. The technologies used should take into consideration the issues of ethics, patient safety, equity, data security, affordability, sustainability and cost effectiveness. The outcomes of AI and use of PM should focus on achieving people centric health solutions, development of health infrastructure, prevention of disease and promotion of health and well-being of the people.

REFERENCES

- [1] Ranapurwala SI, Cavanaugh JE, Young T, Wu H, Peek-Asa C, Ramirez MR. Public health application of predictive modeling: an example from farm vehicle crashes. *Injury Epidemiology*. 2019 Jun 17;6(1):31.
- [2] AI versus ML versus predictive analytics | Tableau [Internet]. [cited 2023 Aug 2]. Available from: <https://www.tableau.com/data-insights/ai/AI-ML-predictive-analytics>.
- [3] Yang CC. Explainable Artificial Intelligence for Predictive Modeling in Healthcare. *J Healthc Inform Res*. 2022 Feb 11;6(2):228–39.
- [4] Yan SJ, Chughtai AA, Macintyre CR. Utility and potential of rapid epidemic intelligence from internet-based sources. *Int J Infect Dis*. 2017 Oct;63:77–87.
- [5] Kpozehouen EB, Chen X, Zhu M, Macintyre CR. Using Open-Source Intelligence to Detect Early Signals of COVID-19 in China: Descriptive Study. *JMIR Public Health Surveill*. 2020 Sep 18;6(3): e18939.
- [6] Joshi A, Sparks R, Karimi S, Yan SLJ, Chughtai AA, Paris C, et al. Automated monitoring of tweets for early detection of the 2014 Ebola epidemic. *PLoS One*. 2020;15(3): e0230322.
- [7] Madoff LC. ProMED-mail: an early warning system for emerging diseases. *Clin Infect Dis*. 2004 Jul 15;39(2):227–32.
- [8] Madoff L, Brownstein J. ProMED and HealthMap: Collaboration to improve emerging disease surveillance. *International Journal of Infectious Diseases*. 2010 Mar 1;14: e184.
- [9] The Epidemic Intelligence from Open Sources Initiative [Internet]. [cited 2023 Aug 3]. Available from: <https://www.who.int/initiatives/eios>
- [10] Freifeld CC, Mandl KD, Reis BY, Brownstein JS. HealthMap: global infectious disease monitoring through automated classification and visualization of Internet media reports. *J Am Med Inform Assoc*. 2008;15(2):150–7.
- [11] HealthMAP - A Model-Based Framework for On-Line Prognostics and Health Management (PHM) | Infotech@Aerospace Conferences [Internet]. [cited 2023 Aug 3]. Available from: <https://arc.aiaa.org/doi/abs/10.2514/6.2010-3500>

- [12] Wang X, He X, Wei J, Liu J, Li Y, Liu X. Application of artificial intelligence to the public health education. *Frontiers in Public Health* [Internet]. 2023 [cited 2023 Aug 3];10. Available from: <https://www.frontiersin.org/articles/10.3389/fpubh.2022.1087174>
- [13] Arbabshirani MR, Fornwalt BK, Mongelluzzo GJ, Suever JD, Geise BD, Patel AA, et al. Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. *NPJ Digit Med*. 2018; 1:9.
- [14] Jungwirth D, Haluza D. Artificial Intelligence and Public Health: An Exploratory Study. *Int J Environ Res Public Health*. 2023 Mar 3;20(5):4541.
- [15] Caliskan A, Bryson JJ, Narayanan A. Semantics derived automatically from language corpora contain human-like biases. *Science*. 2017 Apr 14;356(6334):183–6.
- [16] Baowaly MK, Lin CC, Liu CL, Chen KT. Synthesizing electronic health records using improved generative adversarial networks. *J Am Med Inform Assoc*. 2019 Mar 1;26(3):228–41.
- [17] Khan B, Fatima H, Qureshi A, Kumar S, Hanan A, Hussain J, et al. Drawbacks of Artificial Intelligence and Their Potential Solutions in the Healthcare Sector. *Biomed Mater Devices*. 2023 Feb 8;1–8.
- [18] Lee SI, Celik S, Logsdon BA, Lundberg SM, Martins TJ, Oehler VG, et al. A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia. *Nat Commun*. 2018 Jan 3;9(1):42.