Machine Learning Driven Approach for Waterborne disease Warning System

Introduction

Waterborne diseases are diseases that are being found in contaminated water which causes by variety of microorganisms, pathogenic microbes, biotoxins contaminants which lead to diseases such as schistosomiasis, cholera and other gastrointestinal problems (Patel, 2019). The transmission or spread of these pathogenic microbes happens due to bathing, drinking water, washing, eating food exposed to contaminated water among other things (National Institute of Environmental Health Sciences (n. d.). Vomiting and diarrhea are commonest symptoms of the waterborne diseases; other symptoms include ear, skin, eye problems, and repository among others and the majority of the waterborne diseases are life threatening ((Reynolds, Mena, Gerba,2008). Children are the majority victims of the waterborne diseases due to poor hygienic water and weak immunity especially in developing countries. Many developing countries do not have good and proper water treatment plans particularly in rural areas (Reynolds, Mena, Gerba, 2008). According World Health Organization (WHO), about 884 million people around the world lack basic drinking water while 159 million people source their drinking water on surface. More than 2 billion people also use contaminated drinking water where the water source transmit waterborne disease which have been linked to about more than 502,000 diarrheal death every year (WHO). United State of America has the best quality water in world, but however waterborne diseases outbreaks continue to occur (Reynolds, Mena, Gerba,2008) and most of the outbreaks are primarily occurred as the result of technological failure in the course of the water treatment (Calderon, 2004).

Machine Learning (ML) is an artificial intelligence concept has been used to develop the models or systems that reason with the existing dataset to predict the future event (Muhammad, Algehyne, Usman, 2020). The Machine learning systems are able to read and modify its structure based on set of observed data in order to achieve desired objectives. Like other artificial intelligence techniques such as fuzzy logic, data mining and deep learning; Machine Learning are being used for diagnosis and prediction of many diseases such as Diabetes, COVID-19, coronary artery diseases, waterborne diseases (Muhammad and Algehyne, 2021).

Many research scholars used machine learning techniques for prediction of waterborne diseases in order to manage the diseases. Nastaran, Masoumeh, Doina, David, (2014) developed a machine learning model to simulate waterborne disease outbreaks of varying magnitude and duration and the model accurately predict the characteristics of the disease outbreaks with an accuracy of more than 80%. Sandeep and Kuljit (2018) in their study of a hybrid forecasting ANFIS genetic algorithm based model for cholera waterborne disease was implemented. Now Casting System with Machine Learning Techniques for prediction of fecal contamination level of water at Recreational Beaches in Korea has been developed with Artificial Neural Network and support vector regression using Haeundae and Gwangalli Beaches in Busan City dataset in the work of (Yongeun et al., 2018).

In this regards, the main motivation in this study is to develop machine learning driven approach for waterborne disease warning system for efficient prediction of waterborne diseases.

**References**

* World Health Organization (2017). Diarrhoeal disease. Retrieved from https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease, accessed date 4th March, 2021
* Patel H. H. (2019). Water-Borne Diseases, News Medical Life Sciences. Retrieved from <https://www.news-medical.net/health/Water-Borne-Diseases.aspx> accessed date 4th March, 2021
* National Institute of Environmental Health Sciences (n. d.) Waterborne Diseases. Retrieved <https://www.niehs.nih.gov/research/programs/geh/climatechange/health_impacts/waterborne_diseases/index.cfm> accessed date 4th March, 2021
* Reynolds K.A., Mena K.D., Gerba C.P. (2008) Risk of Waterborne Illness Via Drinking Water in the United States. In: Whitacre D.M. (eds) Reviews of Environmental Contamination and Toxicology. Reviews of Environmental Contamination and Toxicology, vol 192. Springer, New York, NY. <https://doi.org/10.1007/978-0-387-71724-1_4>
* Calderon RL (2004) Measuring benefi ts of drinking water technology: ten years of drinking water epidemiology. NEWWA Water Quality Symposium, May 20, 2004, Boxborough, MA.
* Nastaran J., Masoumeh I., Doina P., David L. B. (2014). Quantifying the determinants of outbreak detection performance through simulation and machine learning, Journal of Biomedical Informatics; http://dx.doi.org/10.1016/j.jbi.2014.10.009
* Yongeun P., Minjeong K., Yakov P., Seoung-Hwa C., Jeong-Goo C., Junho J., and Kyung H. C. (2018). Development of a Nowcasting System Using Machine Learning Approaches to Predict Fecal Contamination Levels at Recreational Beaches in Korea, Journal of Environmental Quality, Volume 47, Issue 5, Pages 1094-1102
* Sandeep K. and Kuljit K. C. (2018). Hybrid ANFIS-genetic algorithm based forecasting model for predicting Cholera-waterborne disease. International Journal of Intelligent Engineering Informatics, Volume 8, Issue 4.
* L. J. Muhammad, E. A. Algehyne, S. S. Usman (2020). Predictive Supervised Machine Learning Models for Diabetes Mellitus”. Springer Nature Computer Science, Vol. 1 no. 240, 2020.
* Muhammad, L.J., Algehyne, E.A. (2021). Fuzzy based expert system for diagnosis of coronary artery disease in nigeria. *Health Technol.* **11,** 319–329 (2021). <https://doi.org/10.1007/s12553-021-00531-z>
* Ahmed, U.; Mumtaz, R.; Anwar, H.; Shah, A.A.; Irfan, R.; García-Nieto, J. Efficient Water Quality Prediction Using Supervised Machine Learning. Water **2019**, 11, 2210. https://doi.org/10.3390/w11112210