Status and Prospects of Applying Machine Learning for Diagnosis and Treatment in African Healthcare Industry

Fan Wu*

Master of Science in Business Analytics, University of California San Diego, La Jolla, United States

* Corresponding Author Email: wufan9702@hotmail.com

Abstract. There is a big difference on healthcare industry between Africa and advanced countries, which represent in lacking medical equipment and backward skills of treatment. The United Nation are thinking about how to eliminate this variance with not only give medical resources but also apply technology to improve the efficiency of using supplies. To train skilled healthcare workers require long-term and expensive investment, but when the pandemic outbreak, there are few talents who can handle complex and extensive work. Therefore, the scientists start to use computer science to replace with manual work, which is more precise and efficient than human beings. Meanwhile, the information technology has developed to the new stage that using big data and advanced algorithms for predicting future potential results which aim to guide better intervention and treatment work. The United States and other advanced countries have the number of effective cases about using these technologies, but the revolution of healthcare is still processing in Africa. The early stage of applying computer technologies in Africa means that there is a big potential market for both scientists and businessmen to explore more solutions for Africa since the adaptive transformation is the key step in the implement.

Keywords: Machine Learning, Africa, Healthcare Industry.

1. Introduction

The Africa takes the about 20% of the global population that leads to the healthcare problem in Africa cannot be ignored. While the Africa has lacked public health services for long time which caused Africans are suffering infectious diseases and chronic. Implementing effective diagnosis and treatment projects is important to optimize the public health system [1].

Takes the nearest pandemic happened in the world, COVID-19 as the example, there is only 16% vaccination rate for 1.3 billion African, which was not enough to cover most people. Also, Africa owns very limited resources to build up local factories for producing vaccines and medical supplies, which caused lots of African countries keep relying on the global medical supplies include drugs, medical equipment and so on. The other reason comes from talent management, which shows from African health workers chose to work out of Africa after finishing education and stay outside which leads to lacking few doctors and nurses in Africa. Because of the backward economy situation, scientists are not willing to stay in Africa and teach for long time. Long-term investment seems necessary and urgent.

ML is one of the technologies that could advance healthcare resolutions through enhancing medical decisions with data and mathematics. For now, the computer technologies have been used in several industries and brought number of benefits for normal life. The ML is one of advanced technologies that aims to find the relationship between past data and future results to make reliable and trustable decisions. The common usages include detecting how fast the pandemic widespread, the association between physical indicators with possibility of being sick. There are many countries have applied the ML in healthcare industry, for example, the scientists applied deep learning, one of the ML methods, to predict and detect for providing early treatment for lung cancer patients in the United States. But there are injustices between Africa and advanced countries which limit the African medical development. The gap is that there are less successful cases to apply ML in Africa than in high-income countries, which needs to be solved as soon as possible. This research paper is to discuss the status and prospects of applying Machine Learning (ML) technologies for diagnosis and treatment

in African healthcare industry and aims to provide theoretical knowledge for African healthcare further research.

2. Current Situation of the African Healthcare

The current situation in Africa is not ideal, not only in penetration rate of vaccine and supply of medical instruments, but also in high medical costs which cannot be handled by developing countries.

For example, the human papillomavirus (hereafter denoted as 'HPV') is the culprit causes cervical, neck, chest cancers. For preventing the HPV for teenagers, a lot of countries started to vaccine them to versus risky kinds of HPV. However, there are not many people in developing counties know about the HPV vaccines, some of them do not trust the validity of vaccinations, or even query the safety of vaccines. In Southwest Nigeria, there is about 79% people have never heard about the vaccine for HPV, and lots of parents refused to let their children to take the vaccinations for HPV [2]. The reasons are complicated, but the main one was that the people in South Nigeria have no chance to get the information from outside resources. Therefore, to improve the public basic understanding of medicine is necessary and urgent.

Also, the availability of medical devices in African countries is not ideal due to the backward manufactory power and market forces. The advanced countries have taken the most part of market of medical device, which made the standards of medical equipment cannot satisfied easily by other countries, especially in advancing countries [3]. The African manufacturers have no skill to produce and design medical devices which fit these standards or need to spend more funding than expected. The diagnosis step requires medical devices to achieve the rapid analysis, otherwise most of diagnosis work need to be done by the doctors' experience which has high possibility to make mistakes. Moreover, the workers who know how to use these advanced medical devices are rare in African medical institutions. The truth is that lacking talents reduce the success rate of using medical equipment in diagnose and treatment usage even they have enough devices.

The supply chain of medicine damaged by the recent pandemic is also one of the severe problems does harm to the basic public health in Africa. There is more than 70% of agreed medications are not produced locally, since produced in Chinese and Indian cost less [4]. The Africa holds big populations which is deeply hurt by fragile supply chain of medications, so building up own factories to produce medicines seems important. However, the cost to improve productive capacity to produce enough medicine for such a large population of African people is not easy. The limitations come from the weak organization, researchers, and technology, which cannot be solved rapidly because of the long-term lagging level of economy and manufacturing. Lacking medicines caused the treatment in healthcare lost the basic support to treat and lost opportunity to cure patients. These are the difficulties pushed Africa to find the solution to improve whole healthcare level, and ML, as the new technology seems supportive in this healthcare revolution.

3. Application of Machine Learning in Africa

Take the COVID-19 as the first example, the way applying ML in Africa to detect COVID-19 has been applied by researchers to help countries to find infected travelers who enter Africa. The principle of ML in detection is to analyze the connotation between travel records with being infected and give out the list of countries which have high possibility to import diseases. The other case in Africa using ML to detect COVID-19 with the smartphone soundtrack function. By using the data collected in South Africa, the researchers developed the models to identify the coughs of COVID-19 patients and normal control group because there are about 20% difference on the length of cough between COVID and not COVID [5]. This is the advantage for ML and computer science technology since human being could not figure out the difference of this tiny variance of length of cough. Also, by connecting with smartphone which has the soundtrack tool, the scientists could collect on-time data and allowed the ML to output the conclusion of positive or negative as soon as possible. Moreover, the research

about using ML to predict the survival rate of suspicious COVID patients with the various indicators was arranged in Africa [6]. The doctors could cure more patients with limited medical supplies with this technology. The success of these cases means that the ML has been convenient in rebuilding the diagnose system in Africa which improve the accuracy and speed of the diagnosis.

The other field of using ML is to detect cancers. The successful case in sub-Saharan Africa was to apply computers to analyze the tiny potential Colorectal cancer area through models of ML [7]. The first step is the image preprocessing which isolated entire picture into small pieces, and second step is to use K-means, one of the ML models, to calculate the rate of each crisis shows up, and then using SVM, one of the ML models, to find the characteristics of Colorectal cancer virus with the similarity and difference of neighboring pieces. These features are impossible to be recognized by human eyes because the cancers belong to the cell field. The normal tool used to screen lung cancer is the CT images which represent whether there is lung cancer or not inside the patients. As the eyes cannot notice the small symptoms of lung cancer, the doctors need the CT images to assist in diagnose. The researchers find that the ML could improve the accuracy of using CT to detect the lung cancer in Africa [8]. These successful cases are the reason why scientists and researchers need to use ML tech to assist in diagnosis step. Most of the previous detection methods were relied on the professional healthcare workers who are rare in Africa.

Except for the cancers, the HIV/AIDS is one of the killable diseases in sub-Saharan Africa, which not only do harm to adults but also has about 7 percent possibility to pass virus to new-born child according to the research [9]. The HIV can also be detected by ML technologies in many research. The first step was to collect population data through healthcare survey which required citizens to answer several intended questions, then used these data which include the domestic, prosperity, education level, married status, and geographical information of people to train models of ML to find the relationship between each parameter and being affected by HIV/AIDS [10]. Besides, the ML models are used in analyzing the exposure of HIV risk to female in South Africa and some neighborhood countries. The preexposure prophylaxis (hereafter denoted as 'PrEP'), is the inhibition for AIDS can be calculated through mathematics models. The scientists could conclude the specific age range women has lowest PrEP which can adjust the policies of intervention to rely less on the government funding and strength the cost-efficiency level [11]. Reducing the costs of prevention is urgently necessary in Africa so that to fit its economy situation. The other research supports that the ML used by Balzer et al could help to determine the PrEP candidates in rural Africa to prepare early intervention for HIV [12]. The other common application is the intelligent robot to detect potential diseases with symptom indexes that patients type into the chat box. The patients answer several questions to fit the potential diseases symptoms, the algorithm could calculate the percentage of being infected by flu virus or HIV.

There are over 95% of dead baby and early newborn deaths happened in African sub-Saharan part and Asian southern part, and ML technology can be used is childbirth [13]. The ML was used to predict the death percentage based on the indexes of pregnancies, and doctors could use these data to apply one of the early interventions – Induction of labour (hereafter denoted as 'IOL'). The ML applied by scientist team in Tanzania was to find the important indexes which determine the IOL success rate [14]. The results showed that the age of pregnant women, age of fetus and indexes of body are essential when doing prediction for success proportion of IOL. This could be useful when doctors improve the survival rate of newborn child because they have already known that what indicators needed to be concerned when facing maternities. This is also a big gap of low birth weight (hereafter denoted as 'LBW') represented the weight of babies between developing countries and developed countries. The LBW could increase the 25 times of death rate of baby based on the research. The researchers in Ethiopia applied the ML to predict LBW of newborns with the data collected from survey. The consequence of this research in Ethiopia was to conclude the gender of baby, the period between marriage and birth, age and job of pregnancy are the four key parameters affected the LBW [15]. The medical institutions could investigate relevance indicators to reduce childbirth mortality because they have already known the possibility of death rate and prepared the IOL or other methods.

There are also some branches of ML can be used in healthcare industry. The Natural Language Processing (hereafter denoted as 'NLP'), mainly used in translating unstructured language into structure language, especially in computerized text and communication evaluation. In the healthcare area, NLP is used to grab the key words from the sentences, which has been used in Free Open Access Medical Education which is the good public recourses [16]. There are many unstructured posts or articles in this website, but too many words are not easily be recognized by readers who want to read materials specific on healthcare industry. The NLP helps websites like Free Open Access Medical Education to classify the contents by shorten long sentence into key words. For example, "This is a very good article that talks about how age affects vision in the most of African countries.", the NLP could translate this sentence into key words, like "ages", "affect", "vision", "African" automatically when authors upload sentence into the website. The readers could understand the main content of this sentence by only read four key words instead of reading whole sentence. The other branch of ML is called Neural Network (hereafter denoted as 'NN'), which is the used in Zambia for the diabetic retinopathy screening program which aims to find the correlation between vision and diabetic retinopathy [17]. In the meantime, for more convenient use, the new NN model called AfriBERTa15 created for African usage because this model made particularly for African languages. Also, NN is used to develop medicines with virtual simulations like brains of people which reduce development costs for medical companies.

4. Limitation

Although ML has been used in healthcare industry for several years, the limitations still exist in not only algorithms but also policies about data safety. The first limitation comes from the algorithm or ML internal design. Because of the basic information technology instructions in Africa is far behind advanced countries, like United States and European countries, which causes the difficulties of collecting data from people in Africa. The normal way to collect data of African people was from communities, which lack of timeliness and accuracy compared with the office for National Statistics [18]. Moreover, the models of ML require as many as possible data to get the best parameters, so the correctness for African ML algorithm is lower than average level. The design of ML was made for big data when doing analyzing process, which ignored the trace data situation happens in areas like Africa. For ML itself, more data means better prediction results. Also, the algorithms of some models in ML required the programmers to decide the categories before building up models. There are about sixty countries in Africa which cannot be recognized as the separate country when doing categorical preparations [19]. When scientists want to analyze the healthcare data of specific country, for example, Zimbabwe, the algorithm does not have enough data to figure out the results. The programmers who build up models are willing to use broad group facing big data resources.

The second limitation comes from the data policy related to compliance, privacy, and security. The development of Artificial Intelligence or ML technology pushed governments to promulgate more and more policies to protect the sensitive data of their citizens. Based on the related policy, the data from special patients or medical institutions are required to do regulatory analysis before using. These kinds of policies lead to the data lack of diversity which lower the accuracy of ML models. Meanwhile, the people worry about giving out their personal information to institutions or companies, which limit the scope of real data. The technical way to solve this question is to create and use synthetic data for protecting these special data. However, the drawback of using too much synthetic data is obvious because the synthetic data is made by programmers who made it, which caused the unreality of data and high risk caused by individuals.

5. Conclusions

This research paper talks about the current situation in Africa and the application of ML. And then talks about the advantages and limitations about using ML to improve the healthcare level. The

functionalities of ML could increase the speed of diagnose and treatment of infectious diseases by predicting the spread and lethality, and the accuracy of detecting cancers by recognizing the cancer cells. Also, the ML assists in childbirth work in Africa to increase the survival rate of new-born child. The limitations include the accuracy of ML models affected by data volume collected in Africa and the data policies limits the use of data collected from people.

Based on the benefits and drawbacks of ML applied in Africa, the growing of ML needs to be concerned by both African government and global forces. The revolution of healthcare requires new technologies to provide the long-term support since training skilled healthcare workers to do manual work lack of speed and accuracy. Also, ML technically has strong functionalities in enlightening original efficiency and finding details of illness which cannot be identified by doctors without technical supports. So, to promote ML on healthcare industry is one of the fundamental tasks for science teams and African governments. In the meantime, the governments have responsibility to control abusing data when doing business use. Instead of using synthetic data, the organizations should try data desensitization which is to replace some sensitive information with special characters. The common example is to rewrite patients' names into capital letters of first and last name or using "patient1" to distinguish the difference names of patients. By applying data desensitization, the data used in building models keeps the reality and integrity.

In conclusion, this paper analyzed ML in healthcare industry in Africa comprehensively. The successful cases of applying ML could be learned by other advancing countries and improved when facing more complicated situations. The limitations of ML could give lessons to African countries to think about how to achieve better adaptive transformation on public health area. The ML keeps improving along with the development of computer technologies, which means there will be more efficient solutions in the future.

References

[1] Aborode, A. T., Hasan, M. M., Jain, S., Okereke, M., Adedeji, O. J., Karra-Aly, A., & Fasawe, A. S. (2021). Impact of poor disease surveillance system on COVID-19 response in africa: Time to rethink and rebuilt. Clinical epidemiology and global health, 12, 100841.

[2] Mihretie, G. N., Liyeh, T. M., Ayele, A. D., Belay, H. G., Yimer, T. S., & Miskr, A. D. (2022). Knowledge and willingness of parents towards child girl HPV vaccination in Debre Tabor Town, Ethiopia: a community-based cross-sectional study. Reproductive health, 19(1), 136.

[3] Medenou, D., Fagbemi, L. A., Houessouvo, R. C., Jossou, T. R., Ahouandjinou, M. H., Piaggio, D., Kinnouezan, C.-D. A., Monteiro, G. A., Idrissou, M. A. Y., Iadanza, E., & Pecchia, L. (2019, March 19). Medical devices in Sub-Saharan africa: Optimal assistance via a Computerized Maintenance Management System (CMMS) in Benin - Health and Technology.

[4] The impact of COVID-19 pandemic on medicine security in Africa: Nigeria as a case study. The Pan African medical journal, 35(Suppl 2), 73.

[5] Pahar, M., Klopper, M., Warren, R., & Niesler, T. (2021). COVID-19 cough classification using machine learning and global smartphone recordings. Computers in biology and medicine, 135, 104572.

[6] Ngabo, D., Dong, W., Ibeke, E., Iwendi, C., & Masabo, E. (2021). Tackling pandemics in smart cities using machine learning architecture. Mathematical biosciences and engineering: MBE, 18(6), 8444–8461.

[7] Waljee, A. K., Weinheimer-Haus, E. M., Abubakar, A., Ngugi, A. K., Siwo, G. H., Kwakye, G., Singal, A. G., Rao, A., Saini, S. D., Read, A. J., Baker, J. A., Balis, U., Opio, C. K., Zhu, J., & Saleh, M. N. (2022). Artificial intelligence and machine learning for early detection and diagnosis of colorectal cancer in sub-Saharan Africa. Gut, 71(7), 1259–1265.

[8] Sori, W. J., Feng, J., Godana, A. W., Liu, S., & Gelmecha, D. J. (2020, October 2). DFD-net: Lung cancer detection from denoised CT scan image using Deep Learning - Frontiers of Computer Science.

[9] Tordoff, D. M., Buskin, S., Lechtenberg, R., Golden, M. R., Kerani, R. P., & Herbeck, J. T. (2021). Combining traditional and molecular epidemiology methods to quantify local HIV transmission among foreign-born residents. AIDS (London, England), 35(4), 655–664.

[10] Nichol, A. A., Bendavid, E., Mutenherwa, F., Patel, C., & Cho, M. K. (2021). Diverse experts' perspectives on ethical issues of using machine learning to predict HIV/AIDS risk in sub-Saharan Africa: a modified Delphi study. BMJ open, 11(7), e052287.

[11] Grant, H., Gomez, G. B., Kripke, K., Barnabas, R. V., Watts, C., Medley, G. F., & Mukandavire, Z. (2020). Time to Scale Up Preexposure Prophylaxis Beyond the Highest-Risk Populations? Modeling Insights from High-Risk Women in Sub-Saharan Africa. Sexually transmitted diseases, 47(11), 767–777.

[12] Krakower, D. S., & Marcus, J. L. (2020). Machine Learning for Human Immunodeficiency Virus Prevention in Rural Africa: The SEARCH for Sustainability. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America, 71(9), 2334–2335.

[13] Mboya, I. B., Mahande, M. J., Mohammed, M., Obure, J., & Mwambi, H. G. (2020). Prediction of perinatal death using machine learning models: a birth registry-based cohort study in northern Tanzania. BMJ open, 10(10), e040132.

[14] Tarimo, C. S., Bhuyan, S. S., Li, Q., Mahande, M. J. J., Wu, J., & Fu, X. (2021). Validating machine learning models for the prediction of labour induction intervention using routine data: a registry-based retrospective cohort study at a tertiary hospital in northern Tanzania. BMJ open, 11(12), e051925.

[15] Bekele W. T. (2022). Machine learning algorithms for predicting low birth weight in Ethiopia. BMC medical informatics and decision making, 22(1), 232.

[16] Siddique, S., & Chow, J. C. L. (2021, February 14). Machine learning in Healthcare Communication. MDPI.

[17] Akpanudo, S. (2022). Application of Artificial Intelligence Systems to Improve Healthcare Delivery in Africa.

[18] Abebe, R., Aruleba, K., Birhane, A., Kingsle, S., Obaid, G., Remy, S. L., & Sadagopan, S. (2021, March 1). Narratives and counternarratives on Data Sharing in Africa: Proceedings of the 2021 ACM Conference on Fairness, accountability, and transparency. ACM Conferences.

[19] Porras, A. R., Rosenbaum, K., Tor-Diez, C., Summar, M., & Linguraru, M. G. (2021). Development and evaluation of a machine learning-based point-of-care screening tool for genetic syndromes in children: a multinational retrospective study. The Lancet. Digital health, 3(10), e635–e643.