ATLANTIS PRESS

Advances in Health Sciences Research, volume 22

4th International Symposium on Health Research (ISHR 2019)

New Research Agenda: Potentials Use of Machine Learning for Public Health

Kemal N. Siregar [1,2*], Tris Eryando[1,2], Martya Rahmaniati[1,2], Indang Trihandini[1,2], Retnowati[1,3]

[1]*Health Informatics Research Cluster, Faculty of Public Health, Universitas Indonesia, Indonesia*
[2]*Department of Biostatistics and Population Studies, Faculty of Public Health, Universitas Indonesia, Indonesia*
[3]*Hospital of Universitas Indonesia, Indonesia*
**Corresponding author. Email: nazarudin.kemal51@gmail.com*

ABSTRACT

In the era of Industrial Revolution 4.0, the question is to what extent this development has been utilized in the field of public health? In fact, in this era there are a number of technological developments, especially information technology, among others, Big Data Analytics and Machine Learning that can be utilized in the field of public health. Public health surveillance that traditionally uses structured data, with the development of information technology today, the surveillance can use textual data. Using Machine Learning, the surveillance would be able to assess public opinion that are negative about a public health program, for example the immunization program. Machine Learning could also assist public health researchers to develop a model to predict the occurrence of certain ill health, which can then test the model with available data, regarding the accuracy of the model in predicting the event. The objectives of this paper are, firstly to provide some of the Faculty of Public Health University of Indonesia (FPHUI) achievements in utilizing the potentials of Machine Learning. Secondly, to identify the future Health Informatics research agenda, especially in the use of Machine Learning.

Keywords*: public health, health informatics, machine learning, sentiment analysis, risk prediction*

1. INTRODUCTION

In the era of Industrial Revolution 4.0, the question is to what extent this development has been utilized in the field of public health? In fact, in this era there are a number of technological developments, especially information technology [1], among others, Big Data Analytics and Machine Learning that can be utilized in the field of public health [2].

Public health surveillance that traditionally uses structured data, with the development of information technology today, the surveillance can use textual data. Using Machine Learning, the surveillance would be able to assess public opinion that are negative about a public health program, for example the immunization program [3]. Machine Learning could also assist public health researchers to develop a model to predict the occurrence of certain ill health, which can then test the model with available data, regarding the accuracy of the model in predicting the event [4].

The objectives of this paper are, firstly to provide some of the Faculty of Public Health University of Indonesia (FPHUI) achievements in utilizing the potentials of Machine Learning. Secondly, to identify the future Health Informatics research agenda, especially in the use of Machine Learning.

2. METHOD

FPHUI has been formed the Health Informatics Research Cluster (HIRC) two years ago at the end of 2017, whose main objective is to strengthen Health Informatics research, which actually has been running for a long time in the FPHUI both by lecturers and students. After the HIRC establishment, a number of research articles were produced during 2018-2019. Among the research articles, a number of articles were identified, of which Machine Learning has been integrated into health informatics researches. This study was conducted to review the research articles that have deliberately used Machine Learning.

The review has been conducted to see the extent of the development of health informatics research at the FPHUI in the use of Machine Learning. Ten articles were selected purposively, which were categorized as preliminary studies before using Machine Learning. Of these 10 articles, there are 3 things reviewed, namely a) main study results before using Machine Learning, b) status of research, c) other information, including funding support for research. Afterwards 10 articles were selected that had used the

Machine Learning method, in which three things reviewed, namely a) as to what extents the use of Machine Learning method, b) status of research, c) other information, including funding support for research.
The results of the review of the 20 articles can be seen in Table 1 and Table 2. Based on the results of the review, we will try to see lessons learned for the new research going forward agenda, especially in the use of Machine Learning.

3. RESULTS AND DISCUSSION

Results

The first result obtained is that with the formation of research clusters, much needed collaboration can be quickly carried out, both with units under the UI, and with institutions outside the UI. The main sources of researches within the FPHUI are lecturers and students. These human resources can apparently be strengthened through collaboration with other researchers from research centers outside the FPHUI. This is evident from the authors of the reviewed research articles. The collaboration that has been carried out has produced a number of publications in indexed journals.
The collaboration between researchers that has been carried out seems to be intentionally integrated into the Machine Learning approach into public health research methods. The research team members come from FPHUI and from outside FPHUI, with different scientific backgrounds. There are a number of research topics carried out, as well as journal publications and intellectual property rights produced in the last 2 years.
But it appears that the use of Machine Learning that has been done is still limited. The level of accuracy that has been obtained from the model developed to make predictions is not optimal. Furthermore, the level of Technology Readiness achieved is still low. Literature studies have been carried out to develop the idea of utilizing Machine Learning. A number of models (algorithms) have been developed, which are then continued to be implemented by coding, which then uses data to train the machine and to test the machine's ability to recognize existing data patterns. The prototype of the application has been generated but nothing has been tested at the user level.

Table 1. Preliminary Studies before Using Machine Learning in Health Informatics Researches in FPHUI, 2018-2019

No	Topic	Main Study Results Before Using Machine Learning	Status of Research	Other Information, incl. Support for Research
1	Roles of Personal Health Record (PHR) in the Management of Type 2 Diabetes Mellitus: A Systematic Literature Review	Risk factors for type 2 DM; Android-based PHR model to record health status for screening DM	Manuscript is submitted for national journal publication	Post graduate research program, UI funding support (PITTA B scheme)
2	The low visit of the Productive Age Group to the Health Post for Non-Communicable Diseases and the Needs to Use a Personal Health Record in the Depok City	Feasibility study on the use of Android-based PHR for improving program performance	The manuscript will be presented at an international scientific meeting, IMOPH Sept 2019	Post graduate research program, UI funding support (PITTA B scheme)
3	Sentiment Analysis of Drug Control from Twitter: A Systematic Review	Use sentiment analysis to assess public opinion on an issue	The manuscript presented at international scientific meeting, IMOPH Oct 2018	Post graduate research program, UI funding support (PITTA B scheme)
4	Factors Influencing Coverage of Elimination of Mother to Child Transmission of HIV (EMTCT): A Systematic Literature Review	Factors of EMTCT coverage	Manuscript is submitted for international journal publication	Doctoral research program
5	Use of Digital Innovation on Elimination of Mother to Child Transmission of HIV: A Systematic Literature Review	Android-based mobile health model to support EMTCT program effectiveness	Manuscript is being prepared for international journal publication	Doctoral research program
6	Determining Factors for Parental Rejection on Measles and Rubella Vaccination: A Systematic Review	Factors that build public opinion	Manuscript is submitted for international journal publication	Post graduate research program
7	Sentiment Analysis to Recognize Public Opinion on Measles and Rubella Vaccination in Facebook	Use sentiment analysis to assess public opinion on an issue; Utilization of social	The manuscript will be presented at an international scientific	Doctoral research program

No	Topic	Main Study Results Before Using Machine Learning	Status of Research	Other Information, incl. Support for Research
	and Twitter: A Systematic Literature Review	media data	meeting, IMOPH Sept 2019	
8	Use of Personal Health Record (PHR) in the Management of Hypertension: A Systematic Literature Review	Risk factors for hypertension; Android-based PHR model to record health status for screening hypertension	The manuscript will be presented at an international scientific meeting, IMOPH Sept 2019	Post graduate research program, UI funding support (PITTA B scheme)
9	Mobile Health (mHealth) Development for Village Midwives to Improve the Performance of the Maternal Health Program in Babakan Madang Sub-District, Bogor	Android-based mobile health prototype to support maternal health program effectiveness	The manuscript presented at international scientific meeting, IMOPH Oct 2018	UI funding support (scheme for community service), has a certificate of intellectual property rights
10	Social Determinant and Maternal Conditions as the Important Risk Factors for Preterm Birth in Public health: A Systematic Literature Review	Risk factors for preterm birth	The manuscript will be presented at an international scientific meeting, IMOPH Sept 2019	Doctoral research program

Table 2. The Use of Machine Learning in Health Informatics Researches in FPHUI, 2018-2019

No	Topic	As to What Extent the Use of Machine Learning Method	Status of Research	Other Information, incl. Support for Research
1	Prototyping Personal Health Record (PHR) for Type 2 Diabetes Mellitus Prevention	Use decision tree to determine whether someone is at risk or not. α stage app dev't.	Manuscript is submitted for int'l journal publication	Post graduate research prog, UI funding support (PITTA B scheme); TRL 2
2	A Prototype System for the Analysis of Sentiment regarding Drug and Food Issues in Indonesia	Selection of accuracy test between NB, SVM, NN. Data 3 labelling. Web-based app dev't.	Manuscript is submitted for int'l journal publication	Post graduate research prog, UI funding support (PITTA B scheme); TRL 5
3	Improving Performance of Puskesmas' Midwives to Increase the Coverage of Elimination of Mother to Child Transmission of HIV using Mobile Health	Plan: Integration of machine learning potentials into android-based application.	Doctoral research proposal to Kemenristekdikti	Doctoral research program; Planned TRL 6
4	Early Detection of Parental Rejection of Measles and Rubella Vaccines on Social Media through Machine Learning Based Sentiment Analysis	Plan: Integration of machine learning and sentiment analysis into web-based application.	Doctoral research proposal to Kemenristekdikti	Doctoral research program; Planned TRL 6
5	Prototyping Android-based Personal Health Record (PHR) for Managing Hypertension	Plan: Selection of accuracy test between NB, SVM, NN. Data 3 labelling. Android-based app dev't.	Post graduate research proposal	Post graduate research prog, UI funding support (PITTA B sch); Planned TRL 4
6	Online Dynamic Risk Calculation for Early Detection of Stroke	Use of historical data and primary data for validation. Carry out data training and data testing. Accuracy test	Presented at the int'l symposium biomed engineering August 2018.	FPH UI funding support (capacity building scheme); TRL 3

No	Topic	As to What Extent the Use of Machine Learning Method	Status of Research	Other Information, incl. Support for Research
		of the classification made.		
7	Dev't of Personal Health Record (PHR) to Reduce the Incidence of Non-Communicable Diseases in the West Java Community using Risk Predictor Analysis	Plan: Integration of machine learning to determine whether someone is at risk or not into android-based application.	First year of three years of research: study literature, prototype development	UI funding support (international collaboration scheme); Planned TRL 6
8	Risk Calculation of Herbal Medicine Effect on Diabetes Mellitus using Bioenergy Symphony Model	Use of historical data and primary data for validation. Carry out data training and data testing. Accuracy test of the classification made.	Presented at 3rd Int'l Conf on Advance Pharmacy and Pharmaceutical Sciences, Sept 2018	FPH UI funding support (capacity building scheme); TRL 3
9	Prediction on the Cost-effectiveness of Health Service in Hospital using Artificial Intelligent	Plan: Integration of machine learning into web-based application.	Data collection is in progress	Doctoral research program; Planned TRL 4
10	Prototype of Risk Predictor for Preterm Birth with Machine Learning in Indonesia	Plan: Integration of machine learning into web-based application.	Doctoral research proposal to Kemenristekdikti	Doctoral research program; Planned TRL 5

Discussion

What Are The Main Results Of The Studies Before Using Machine Learning?

Tthere have been a number of research articles that even though they have not used Machine Learning but have conducted a literature study which is the basis for using Machine Learning in the next research stage. Literature studies that have been carried out include the use of sentiment analysis to assess public opinions on health issues and policies [5], searching for factors that can determine whether someone is at risk or not. At this stage the study also includes making an android-based prototype which can be a Personal Health Record to be able to collect data that can be input for Machine Learning [6].

Some of the results of studies have been published in national or international journals. The researches consist of several forms, including post graduate research programs, doctoral research programs, where some have received funding from ui.

What is the level of use of Machine Learning methods?

In the further study phase, Machine Learning integration has been carried out into applications both based on android- and web-based. There are a number of techniques or Machine Learning methods used in the research at this stage. A decision tree has been used to determine whether a person is at risk or not. Then, there are machine learning techniques that are used to test accuracy, such as naïve bayesian, support vector machine, neural network [7]. The Machine Learning has been used to predict risk factors for strategic public health conditions [4].

Some of the results of studies have been published in national or international journals. The researches consist of several forms, including post graduate research programs, doctoral research programs, where some have received funding from UI (international collaboration scheme). Technology Readiness Level from 2 to 6.

machine learning has begun to be used in health informatics research at fphui. even though the research results obtained are still limited [5], this research experience can be the basis for the future research agenda in public health field, so that the machine learning could be used further in various management decision-making, programs and service delivery care [8]. It could further be used for national level public health intelligence, which is high level decision making and very strategic [9].

The discussion points above can be summarized into an overview of the use of machine learning in public health research at FPHUI during 2018-2019, which can be seen in table 3

.

Table 3. Overview of the Use of Machine Learning in Public Health Researches At FPH UI, 2018-2019

Parameter	Preliminary Studies before Using Machine Learning in Health Informatics Researches	Use of Machine Learning in Health Informatics Researches
What are the main results of the studies before using machine learning?	• Utilization of sentiment analysis to assess public opinion • Searching for factors that need to be asked to determine someone is risky or not • Prototyping android-based Personal Health Record	
What is the level of use of machine learning methods?		• Integration of machine learning into android- or web-based applications • Use decision tree to determine whether someone is at risk or not • Selection of accuracy test between Naïve Bayesian, Support Vector Machine, Neural Network • Machine Learning used to predict risk factors for strategic public health conditions
Status of Research	Some have been published in national or international journals	Some have been published in national or international journals
Other Information, Including Support for Research and TRL	Several types of research, namely post graduate research programs, doctoral research programs, where some have received funding from UI.	Several type of research, namely post graduate research programs, doctoral research programs, where some have received funding from UI (international collaboration scheme). Technology Readiness Level from 2 to 6.

4. CONCLUSION

Machine learning have begun to be used in health informatics research at FPHUI. Machine learning has been used for textual data, which by using sentiment analysis can assess public opinions on health issues and policies. Machine Learning has been used to predict risk factors for strategic public health conditions. Based on the experience of health informatics research at FPH UI in the last 2 years, machine learning methods are very useful for recognizing patterns and making predictions of outcomes with relatively high accuracy.

5. RECOMMENDATIONS

1. This research experience can be the basis for the future research agenda in public health field, so that the machine learning could be used further in various management decision-making, programs and service delivery care.
2. It could further be used for national level Public Health Intelligence, which is high level decision making and very strategic**.**

ACKNOWLEDGMENTS

Appreciation to the Dean and leadership of FPH UI for supporting and launching the establishment of the Health Informatics Research Cluster (HIRC) in early 2018. Thank you to the Research Fellows of HIRC who were actively involved in various health informatics research activities during 2018-2019. Thank you to the post graduate students from both the master and doctoral degree study programs in public health who have produced together a number of articles for publications.

REFERENCES

[1] Flanders Make. Industry 4.0: hype or reality? The current state of play in Flemish manufacturing; 2017.

[2] Benke, Kurt and Geza Benke. Artificial Intelligence and Big Data in Public Health; 2018.

[3] Benke, Kurt. Uncertainties in Big Data When Using Internet Surveillance Tools and Social Media for Determining Patterns in Disease Incidence; 2017.

[4] Trishan Panch, Jonathan Pearson-Stuttard, Felix Greaves, Rifat Atun. Artificial Intelligence: Opportunities and Risks for Public Health. The Lancet Digital Health, Vol. 1, No. 1, e13–e14; 2019.

[5] Isah, Haruna, Paul Trundle, and Daniel Neagu. Social media analysis for product safety using text mining and sentiment analysis.14th UK Workshop on Computational Intelligence (UKCI). IEEE; 2014.

[6] Poulymenopoulou, Michaela; Flora, Malamateniou; George, Vassilacopoulos. Machine Learning for Knowledge Extraction from PHR Big Data. Studies in Health Technology and Informatics; 2014.

[7] Mooney, Stephen J., &Pejaver, Vikas. Big Data in Public Health: Terminology, Machine Learning and Privacy. Annual Review of Public Health; 2018.

[8] Rao Mala, Kidney David. Public Health Intelligence "Issues of Measure and Method". Springer International Publishing Switzerland; 2016.

[9] Fu, P., Luck, J., & Protti, D. Information systems in support of public health in high-income countries. In R. Detels, R. Beaglehole, M. Lansang, & M. Gulliford (Eds.), Oxford textbook of public health; 2009 (p. 395–426).