PROCEEDING
2st International Conference of Health Polytechnic of Jambi 2023
icon@poltekkesjambi.ac.id
http://journal.poltekkesjambi.ac.id/index.php/ICoHPJ
Transforming Future Health Research : Current Challenge Technology and Inovation
ISSN 2988-7356

NON-COMMUNICABLE DISEASES IN PREGNANT WOMEN BASED ON A BEHAVIORAL APPROACH

Hesti Kurniasih[1], Wanodya Hapsari[1], Katrin Dwi Purnanti[1]
[1]Poltekkes Kemenkes Semarang, Indonesia
Corresponding author: hestikurniasih27@gmail.com

ABSTRACT

Background: Early detection of pregnant women as a prevention of the risk of non-communicable diseases can be done with routine health checks. The general aim of this research is to produce a program from an artificial intelligence system to detect non-communicable diseases early and provide WhatsApp-based recommendations to pregnant women.
Method: The implementation of this research began by creating a dataset obtained from the Medical Records Installation, namely data on pregnant women for 3 years from 2019 to mid-2023. Then the data obtained was coded, processed, and classified according to research needs, resulting in 9,289 data. The data is entered into machine learning to be processed by the machine to determine the mean risk factors, which will then produce prediction data. The first stage in data processing required is a machine learning application which will be used to process big data into predictions.
Result: In this research, the application used is Google Collab, which is a default application from Google and can be used with various devices. In this study, the dataset used by researchers is a dataset that predicts heart disease, hypertension, preeclampsia, and eclampsia and recommendations for pregnant women that provide good performance on each accuracy test. After the first process of data sharing, the training data is 90% and the 10% data is called testing data.
Conclusion: The data obtained from pregnant women is then processed to obtain quality data by applying data cleaning using a scaler, namely data whose attribute values will be empty so that the data becomes more accurate. A pregnant woman dataset of 9289 records with complete attributes of 9289 records will be used in the experimental process.

Keywords: pregnant; non communicable diseases; behavioural

INTRODUCTION

Globally, the world is faced with an increasing problem of chronic diseases in the adult population. For example, around 250 million people in the world suffer from type 2 diabetes mellitus. Until now, efforts to overcome this problem have been dominated by efforts to improve healthy lifestyles aimed at adults. It cannot be denied that these efforts are important, but forgetting one of the root causes of chronic disease and shifting the direction of coping strategies is ineffective. Various evidence from studies on both small and large scales increasingly clarifies the importance of the period of the first thousand days of life. The period of fetal development in the womb and during the first two years of life influences various aspects of the quality of human resources. This does not only include physical quality but also cognitive quality and the risk of chronic disease. (Ajay Aggarwal, Preeti Patel, 2020)

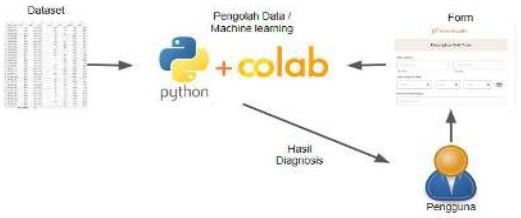
The current pattern of disease incidence has undergone changes marked by an epidemiological transition. Changes in disease patterns that were previously dominated by infectious diseases have shifted to non-communicable diseases (NCDs). World attention to non-communicable diseases is increasing along with the increase in their frequency of occurrence. Two of the ten main causes of death in the world are caused by non-communicable diseases, stroke and ischemic heart disease. it is even the second top cause in both developed and developing

countries. Non-communicable diseases have become the main cause of death globally today (Shilton, 2013). WHO data shows that as many as 57 million (63%) of the deaths that occur in the world and 36 million (43%) of the morbidity rates are caused by non-communicable diseases. All deaths due to NCDs occur in people aged less than 60 years, 29% in developing countries, while in developed countries it is 13%. (Aida Budreviciute, Samar Damiati et al., 2020)

Non-communicable diseases that are related to the impact of nutritional status on the growth path at the beginning of life, not only start from the mother and stop in themselves until adulthood but are also multigenerational or transgenerational. Various evidence supports the phenomenon that this has been happening for at least 100 years. The growth of the fetus or baby is not solely determined by the nutritional status of the mother, but also by the size and shape of the placenta and the condition of the fetus or baby itself. A child develops from an egg released by the grandmother for the mother, so apart from the genes of the father, the grandmother determines the child's genes. The mother releases eggs to be fertilized by the father's sperm (genes) providing nutrition and influencing the placenta. While in the womb, the fetus's nutritional needs are obtained through the placenta, the shape and size of which also influence the smooth delivery of nutrients from the mother. In addition, the placenta will produce hormones that are important for fetal growth and development. After birth, the mother's parenting style influences the baby's growth and development. Thus, it is clear that a person's health status in adulthood has been determined since the previous two generations and the same cycle will occur for the next two generations. (Ajay AggarwalID1, 2*, Preeti Patel3, 4 et al., 2018; Inderawati et al., 2022)

METHOD

The research design used in this study was descriptive observational. This research goes through development using content-based. In this study, the application used was Google Collab, then connected with the Python application for big data analysis. The population in the study were pregnant women who came to the gynecology clinic during the last 3 years. The sample in this study were pregnant women with non-communicable diseases. This study received ethical approval from the Ethics Committee of Prof. Dr. Margono Soekarjo Purwokerto Hospital No. 420/13103 of 2023: 420/13103 year 2023.

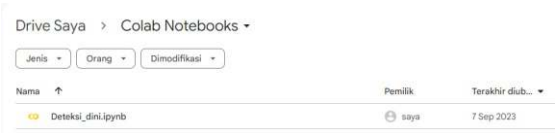


Picture 1. System Overview

RESULT AND DISCUSSION

The implementation of this research began by creating a dataset obtained from the Medical Records Installation, namely data on pregnant women for 3 years from 2019 to mid-2023 totaling 21,345 data. Then the data obtained was coded, processed, and classified according to research needs, resulting in 9,289 data. Data is entered on machine learning to process the data by a machine to determine the mean risk factor which will then produce prediction data. The first stage in data processing required is a machine learning application which will be used to process big data into predictions. In this research, the application used is Google Collab, which is a default application from Google and can be used with various devices. Google Colaboratory or Google Colab is an executable document that allows you to write, edit, and share programs that have been saved on the

drive or that you have just created. One of the benefits of Google Colab is that it has complete machine learning built in and is more memory efficient. Then we connect this Google Collab with the Python application. Python has become a mainstay in data science. This programming language allows data analysts to perform complex statistical calculations, and create data visualizations and machine learning algorithms. Python is a programming language widely used in web applications, software development, data science, and machine learning (ML). (Dey A, 2016; Michie et al., 2017).



Picture 2. Google Collab Application

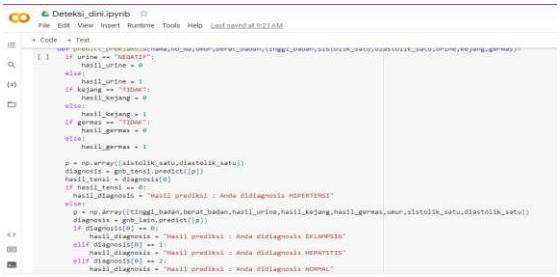
The first stage in training is preprocessing. The dataset is divided into three parts, consisting of classification data, recommendation data, and classification reference data. Especially for recommendation datasets, empty data/cells must be replaced with zeros so that they can be calculated. Then the standard dataset is scaled which usually uses a scaler (used because the scales are not the same). Classification data is processed using a classification algorithm which consists of 6 types of algorithms, namely: R-Nearest Neighbors, Linear SVM, RBF SVM, Gaussian Process, Neural Net, and Ada Boost. In this paper, the researcher chose Ensemble Soft Voting. An ensemble is an algorithm that combines or improves each existing classification to produce a better or at least similar unit. (Dey A, 2016; Michie et al., 2017; The Royal Society, 2017).

Picture 3. Google Collab Application

The algorithm testing results that obtained the best accuracy was MLP. MLP, which is a neural network, is the weak classifier chosen because it produces the best results and automatically includes the following neural networks. Then this model will be implemented and tested in the system. After conducting training, it will be saved as a file with the soft voting extension. Meanwhile, the recommended dataset uses content based on cosine similarity. The data will be stored as data in the form of a constant variable. So both types of algorithms (classification and recommendation) are treated differently. (Dey A, 2016; Michie et al., 2017).

Picture 4. Google Collab Application

The output files of classification and recommendations will be sent to the server to create a REST API. The API is used so that it can be called by other people because in this program the consumers of the training results are in the mobile application. Falcon-app.py is the bridge between the mobile app and the training results. In the same testing process as training, there is also preprocessing. In the training process, it is part of the logic/process of the mobile application as well. Researchers enter a new dataset and then the preprocessing stage, first entering the classification stage. Not yet included in the recommendation because the user profile has no known diagnosis. After identifying which class the pregnant woman belongs to, immediately enter the recommendation stage. The user will automatically output the final classification and recommended results. (Paryati & Krit, 2022; Simbolon et al., 2020).



Picture 5. Diagnose

In this study, the dataset used by researchers is a dataset that predicts heart disease, hypertension, preeclampsia, and eclampsia and recommendations for pregnant women that provide good performance on each accuracy test. After the first process of data sharing, the training data is 90% and the 10% data is called testing data. The data obtained from pregnant women is then processed to obtain quality data by applying data cleaning using a scaler, namely data whose attribute values will be empty so that the data becomes more accurate. A pregnant woman dataset of 9289 records with complete attributes of 9289 records will be used in the experimental process. From the results of data processing carried out using Python, the classification accuracy was obtained.

In this research, an artificial intelligence program has been developed to carry out the early detection of preeclampsia and recommendations for pregnant women based on Android. Then this research will be compared with previous research related to early detection of preeclampsia. The process of prediction or early detection is one of the prevention efforts in the form of early warning about health situations or diseases that will occur in the future and can facilitate health workers to be able to provide health care. Several studies have explored preeclampsia screening in early pregnancy using various variables and technologies. Preeclampsia can be identified by analysis of proteinuria in pregnancy for more than 20 weeks in pregnant women with hypertension. Hypertensive conditions such as preeclampsia during pregnancy cause maternal and fetal death. Preeclampsia is categorized as a hypertensive disorder in pregnancy which is diagnosed by an increase in the pregnant mother's blood pressure (systolic and/diastolic blood pressure greater than or equal to 140 mmHg and 90 mmHg), usually after the 20th week of pregnancy.

Several screening and detection systems for preeclampsia have been developed using statistical learning methods (Eskandar H, Land M, Arnold V, Pujari S, 2015), multivariate Gaussian distribution (Keisling K, 2014), logistic regression (Labrique AB, Vasudevan L. Kochi E, 2013), artificial neural network (Lopez. L, Resser L, 2016), and Bayesian network (Ray R, Sewell AA, 2017). However, detection accuracy remains a priority concern. Today, smartphone technology has improved processing power, storage space, wireless networking, high-quality cameras, and various sensors. Additionally, the reduced cost of mobile phones and smaller sizes have made smartphones available everywhere, even in low-income countries. Mobile phones can be used to alert pregnant women and can also be used by medical practitioners remotely to monitor symptoms of preeclampsia. Many researchers have proposed the development of mobile phones for monitoring preeclampsia.

In this study, researchers developed an android application, which has two unique contributions, namely: Predicting the risk of preeclampsia using soft selector-based ensemble learning, built by six individual classifiers. Then it can provide a recommendation system for women who are at high risk of preeclampsia using the cosine similarity technique.

CONSLUSION

Based on the results of the research and discussions that have been explained in the previous chapters up to the implementation stage of the WhatsApp application, it can be concluded that this research has the following contributions: Ability to improve and produce high accuracy for early detection of PTM

using the Ensemble Learning technique with soft voting, where the accuracy results program is more than 98% and can produce high recommendation accuracy for pregnant women using content-based techniques, where the accuracy of the recommendation program is 98.66%.

ACKNOWLEDGMENT

The author would like to thank the Poltekkes Kemenkes Semarang for providing funding for research activities. The researcher also does not forget to thank all parties who have helped with research activities starting from the RSUD Medical Records Installation which has permitted to collection of data as well as research enumerators.

CONFLICT OF INTEREST

There are no conflicts of interest in conducting this research.

REFERENCES

Aida Budreviciute, Samar Damiati, D. K. S., Kamil Onder, Schuller-Goetzburg6, P., 1, G. P., Katileviciute1, A., And, S. K., Rimantas Kodzius 1, 7, 8, & *. (2020). Management and Prevention Strategies for Non-communicable Diseases (NCDs) and Their Risk Factors. *Frontiers in Public Health*, *8*, 1–11. https://doi.org/10.3389/fpubh.2020.574111

Ajay Aggarwal, Preeti Patel, G. L. dkk. (2020). The Profile of Non-Communicable Disease (NCD) research in the Middle East and North Africa (MENA) region: Analyzing the NCD burden, research outputs, and international research collaboration. *PLOS ONE*. https://doi.org/10.1371/journal.pone.0232077

Ajay AggarwalID1, 2*, Preeti Patel3, 4, G. L., , Abdulkarim EkzayezID3, 4, Coutts5, A., Fouad6, F. M., ShamiehID7, O., Giacaman8, R., Kutluk9, T., Rima, Khalek6, A., & , Mark Lawler10, Peter Boyle11, Diana Sarfati12, Richard Sullivan1, 4. (2018). The Profile of Non-Communicable Disease (NCD) research in the Middle East and North Africa (MENA) region: Analyzing the NCD burden, research outputs, and international research collaboration. *PLOS ONE*, *15*(4).

Dey A. (2016). Machine Learning Algorithms: A Review. *Int J Comput Sci Inf Technol*, *7*(3), 1174–1179.

Eskandar H, Land M, Arnold V, Pujari S, P. M. dan S. R. (2015). *Mobile Technology in Cancer Control for Emerging Health Systems: Metode Statistical Learning Control* (pp. 65–70).

Inderawati, D., Mirawati, M., Aryahnyani, N. P., & Yantina, D. (2022). Peningkatan Keterampilan Kader dalam Skrining Penyakit Tidak Menular (PTM) Warga Kelurahan Jatiwarna Pondok Melati Bekasi. *Jurnal Pemberdayaan Komunitas MH Thamrin*, *4*(1), 84–95. https://doi.org/https://doi.org/10.37012/jpkmht.v4i1.866

Keisling K. (2014). *Multivariate Gaussian Distribution Field Guide for New Born Health*.

Labrique AB, Vasudevan L. Kochi E, F. R. and M. G. (2013). Logistic Regression Innovations as Health System Strengthening Tools: 12 Common Applications and Visual Framework. *Global Health : Science and Practice*, *1*(160), 71.

Lopez. L, Resser L, R. N. and R. D. (2016). Leveraging Artificial Neural Network to Increase Breastfeeding Rates in Los Angeles Country (LAC). *Obstetric Gynaecology PubMed*, *127*(23).

Michie, S., Thomas, J., Johnston, M.,

Aonghusa, M., P., Shawe-Taylor, J., & Etc. (2017). The Human Behaviour-Change Project: harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. *Implementation Science*, *12*(1), 1–12.

Paryati, & Krit, S. (2022). Expert System for Early Detection and Diagnosis of Central Nervous Diseases in Humans with Forward Chaining and Backward Chaining Methods Using Interactive Multimedia. *The International Conference on Artificial Intelligence and Engineering 2022 (ICAIE'2022)*. https://doi.org/https://doi.org/10.1051/itmconf/20224301016

Ray R, Sewell AA, G. K. and R. J. (2017). Missed Opportunity Leveraging Bayesian Network to Reduce Racial Health Disparities. *Journal of Health Politics, Policy and Law*, *42*(901), 24.

Simbolon, O., Widyawati, M. N., Kurnianingsih, K., Kubota, N., & Ng, N. (2020). Predicting the Risk of Preeclampsia using Soft Voting-based Ensemble and Its Recommendation. *2020 International Symposium on Community-Centric Systems (CcS)*, 1–6.

The Royal Society. (2017). Machine learning: the power and promise of computers that learn by example. *Report by Royal Society*, *66*, 125p.