ANFIS MODELLING ON DIABETIC KETOACIDOSIS FOR UNRESTRICTED FOOD INTAKE CONDITIONS

GALUH WILUJENG SARASWATI

MASTER OF COMPUTER SCIENCE
(SOFTWARE ENGINEERING AND INTELLIGENCE)

2017

Faculty of Information and Communication Technology

ANFIS MODELLING ON DIABETIC KETOACIDOSIS FOR UNRESTRICTED FOOD INTAKE CONDITIONS

GALUH WILUJENG SARASWATI

Master of Computer Science in Software Engineering and Intelligence

2017

ANFIS MODELLING ON DIABETIC KETOACIDOSIS FOR UNRESTRICTED FOOD INTAKE CONDITIONS

GALUH WILUJENG SARASWATI

A thesis submitted
in fulfillment of the requirements for the degree of Master of Computer Science
(Software Engineering and Intelligence)

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2017

DECLARATION

I declare that this thesis entitled "Anfis Modeling on Diabetic Ketoacidosis for Unrestricted Food Intake Conditions" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : ……..

Supervisor Name : ……..

Date : ……..……

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in term of scope and quality for the award of Master of Computer Science (Software Engineering and Intelligence).

Signature	:	……..
Supervisor Name	:	……..
Date	:	……..……

DEDICATION

This thesis are dedicated to the Almighty God

To my beloved parents, Mr. Sukadi and Mrs. Eko Rahayuningsih

To my lovely sister and brother, Dyah Asih sulistyawati, S.E and Ismu Darmawan K.

To my admired partner, Praditya Swastikha Ardhy, S.T

To my precious country, Indonesia

ABSTRACT

Diabetic ketoacidosis is a complication of diabetes that occurs when body cannot produce insulin necessarily to convert glucose into energy, instead fat is used as energy source and produce ketone as a byproduct. Ketones can be detected in urine compounds, especially when there is a large number of ketones that produce a distinctive smell of acetone. Odor sensors assembled into Electrical Nose (E-nose) system is used as self-diagnostics pre-test for diabetic's analysis. However, diabetic's analysis often required a subject to fast before sample testing. Currently, different prediction model for diabetic ketoacidosis are used depending on fasting or non-fasting conditions. This is inconvenience for diabetic's analysis to be done at any time anywhere. This project aim to propose an adaptive prediction model capable to diagnose diabetic ketoacidosis in unrestricted food intake conditions. The adaptive Neuro-fuzzy Inference System (ANFIS) is proposed to build the diabetic ketoacidosis classifier. The fuzzy inference model will be used to capture both fasting and non-fasting membership functions before feeding the results for classification to the neural network model. Two sets of experimental data involving 20 diabetic patients and 20 healthy subjects were collected from CITO laboratory Semarang Central Java, Indonesia. Ethics consents were informed and agreed by the subjects before starting the data collection. This project follows the experimental methodology in verifying the hypothesis drawn. The experimental paradigm was designed to simulate fasting and non-fasting conditions. Samples data were recorded in the morning before food intake and two hours after food intake using four MQ 2, MQ 5, MQ 6 and MQ 8 sensors, in previously built Electronic Nose prototype system. A 5-fold cross-validation testing was implemented for classification results analysis. The results are highly promising with at least 90% accuracy in all testing. The proposed model has achieved 96% average accuracy in unrestricted food intake conditions. The prediction results on non-fasting and fasting data samples were recorded as 98% and 96% of average accuracy respectively. This has proven that the proposed ANFIS model is good to detect diabetic's cases through ketoacidosis regardless of food intake. It has better performance in normal food intake as compare to fasting condition, since insulin inefficiency happened in diabetics patients will resulted in obvious acetone secretion in non-fasting condition. The project has also implemented the optimization process onto the proposed ANFIS model through the hybrid of Genetic Algorithm on the fuzzy membership function of the ANFIS model. The proposed GA-ANFIS approach provides excellent classification in accuracy, precision and recall. However, the results are only a minor improvement from the non-optimized ANFIS model since the predecessor has achieved good classification accuracy. In conclusion, diabetic ketoacidosis in unrestricted food intake conditions can be predicted using the proposed ANFIS and GA-ANFIS model. Future work should be focusing on data collection of the E-Nose sensors and the improvement of the learning algorithm robustness towards environmental noise during data acquisition, such as evaporation and contamination of odor samples.

ABSTRAK

Ketoasidosis diabetes adalah komplikasi apabila badan tidak menghasilkan insulin untuk menukarkan glukosa kepada tenaga, sebaliknya, lemak digunakan sebagai sumber tenaga dan menghasilkan ketone sebagai hasil sampingan. Keton boleh dikesan di dalam sebatian air kencing, terutama apabila terdapat sebilangan besar keton yang menghasilkan bau aseton. Sensor bau di dalam sistem Hidung elektrik (E-hidung) digunakan sebagai diagnostik di dalam pra-ujian untuk analisis kencing manis. Analisis diabetes sering memerlukan subjek berpuasa sebelum melakukan pensampelan. Kini, penggunaan model ramalan ketoasidosis kencing manis yang berbeza diperlukan berdasarkan keadaan puasa atau sebaliknya. Ini menimbulkan kesulitan bagi membolehkan analisis diabetes dilakukan pada bila-bila masa dan di mana sahaja. Projek ini bertujuan untuk mencadangkan satu model ramalan kencing manis berasaskan ketoacidosis tanpa sekatan keadaan pengambilan makanan. Sistem Inference Neuro-kabur (ANFIS) Adaptif dicadangkan untuk membina ketoacidosis pengelas kencing manis. Model inferens kabur akan digunakan untuk menggambarkan kedua-dua fungsi keahlian sebelum dan selepas makan, sebelum keputusan disalurkan ke dalam pengkelas rangkaian neural. Dua set data eksperimen yang melibatkan 20 orang pesakit kencing manis dan 20 orang yang sihat telah dikumpulkan dari makmal CITO Semarang Jawa Tengah, Indonesia. Etika persetujuan telah dimaklumkan dan dipersetujui oleh subjek sebelum memulakan pengumpulan data. Projek ini mengikut kaedah eksperimen dalam mengesahkan hipotesis yang digariskan. Paradigma eksperimen direka untuk mensimulasikan keadaan puasa dan sebalinya. Data sampel dicatatkan pada waktu pagi sebelum pengambilan makanan dan juga dua jam selepas pengambilan makanan menggunakan empat sensor MQ 2, MQ 5, MQ 6 dan MQ 8, pada sistem prototaip hidung elektronik yang dibina sebelumnya. Keputusan pengkelasan didapatkan dengan menggunakan 5-fold Cross-validation. Keputusan adalah sangat baik dengan minima 90% ketepatan di dalam semua ujian. Model yang dicadangkan telah mencapai 96% ketepatan purata tanpa mengambilkira keadataan pemakanan. Keputusan purata 98% dan 96% telah dicatatkan untuk keadaan tanpa puasa dan puasa masing-masing Ia adalah lebih baik dalam keadaan tidak puasa kerana ketidakcekapan insulin pada pesakit kencing manis menyebabkan rembesan aseton yang jelas dalam keadaan tidak berpuasa. Proses optimasi juga telah diuji ke atas model ANFIS melalui hibrid algoritma Genetik ke atas fungsi keahlian kabur. Pendekatan GA-ANFIS sangat cemerlang dari segi accuracy, precision dan recall. Walau bagaimanapun, peningkatan prestasi tidak ketara kerana model ANFIS telah berkemampuan mencapai klasifikasi ketepatan yang baik. Kesimpulannya, dianostik kencing manis melalui ketoacidosis tanpa sekatan keadaan pengambilan makanan boleh diramal menggunakan model ANFIS dan GA-ANFIS yang dicadangkan. Kerja masa depan harus memberi tumpuan kepada pengumpulan data sensor E-Nose dan peningkatan keteguhan algoritma pembelajaran ke arah gangguan alam sekitar semasa perolehan data, seperti penyejatan dan pencemaran sampel bau.

AKCNOWLEDGEMENTS

Alhamdullilahi Robbil'Alamnin, first of all the researcher would like to express my deepest praise and gratitude to Allah SWT who has given blessing and merciful to complete this thesis. Second, I would like to express my deepest appreciation to my beloved supervisor Associate Professor Dr. Choo Yun Huoy for her guidance, suggestion, advice and encouragement throughout this study. I would also like to express my appreciation to Dean Faculty Information and Communication Technology, Prof. Dr. Burairah Bin Hussin and also Prof. Nanna Suryana Herman for their support and help towards my postgraduate studies.

I am especially grateful to Ministry of Education and Culture Indonesia, Ministry of Research Technology and Higher Education Indonesia and also Dian Nuswantoro Universisty (UDINUS) for providing scholarship Beasiswa Unggulan, without the scholarship, I could not have stated this master study in UTeM Malaysia.

Last but not least, thanks to my all friends Syaifur, Naufal, Tosan , Maulana, Wendy, Masduki, Aditya, especially Nuur Cahaya, Anisa, Safira and Erba and others for their support.

Finally, my deepest gratitude goes to my parents, Bapak Sugiyo M.Kom, Bapak Agung, Ibu Reni and also Praditya Swastika Ardhy for their love, prayers and encouragement. To those who indirectly contributed to this research, my proud thanks for their kindness.

Melaka, 15 June 2017

Galuh Wilujeng Saraswati

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

Diabetes is one of the world's oldest diseases. *Diabetes Mellitus* is defined as a diseases group metabolic which have characteristic of an increase in the level of blood glucose (*hyperglycemia*) that occur due to abnormal insulin secretion, the activity of insulin increasingly or both of them (Barakat et al., 2010). Diabetes is a metabolic illness that is affected by many factors such as lack of insulin and the body's inability to use insulin (Insulin Resistance) with a sign of chronic hyperglycemia, fats, decreased metabolism of carbohydrates and proteins as a result of a deficiency in the secretion of the hormone insulin, glucose transporter deficiency or both. Diabetes is now a major worldwide public health problem, the epidemiologic transition is evidenced by the tendency of increase in the prevalence of Diabetes mellitus annually. In 2013, the statistical report of *International Diabetes Foundation* (IDF) has estimated that the prevalence of Diabetes Mellitus is around230 million people. Based on data from the *World Health Organization* (WHO), at least 171 million people suffer from diabetes. This is increasing quickly and expected by 2030 this number will increase twice. Additionally, this disease is one of the biggest killers in Southeast Asia and the Western Pacific (WHO, 2011). Diabetes as the biggest killer can rise because most people do not aware of characteristics of diabetes. As the result, they do not know the indicators of a person who has diabetes. There are several symptoms of diabetes which can be an indicator of the emergence of health problems in the form of diabetes include high levels of glucose in the blood and in the urine, performance impaired kidney,

white blood cells are not able to kill bacteria in the body, as frequent urination (Polyuria) for Increase in thirst level (Polydipsia), increase of hunger (Polyphagia) and weightless if the diagnosis is not done at the right time (Siyang et al., 2012). Detection of ketones in the urine have an important role to examine the metabolic condition of the diabetes patient.

Nowadays, the availability of many commercial computer-based systems that patient's onshore processing information in the form of dynamic glucose level. This system is very suitable for self-monitoring and consultation with a physician. The computer-based system actually more suitable if they are able to intelligent to analyze the information early so that patients can take action before there were complications due to diabetes (Ajmalahamed et al., 2014).

1.2 Research Background

In Indonesia, diabetes is a public concern since Indonesia becomes the fourth country in the world has highest diabetes rate increased up to 14 million people it based on the report of WHO. According Central Data and Information Health Ministry of Republic Indonesia, the prevalence of diabetes are 0.6% people aged 15 years and above or around 1 million people actually fell the symptoms of *diabetes mellitus* but the peoples do not test to know more about their condition due to various factors including the high cost of checking blood sugar and scared of needles. ("information-diabetes indonesia.pdf," n.d.). The increasing number of diabetes is caused by delayed diagnosis. There are many factors that caused the delay that influences the choices available or the variety of variables (Pangaribuan, 2014).

Moreover, *diabetes mellitus* is a disease marked by insulin production in pancreas produces insufficient for glucose metabolism in our body as known *diabetic ketoacidosis*. In healthy individuals, the body uses carbohydrate metabolism mostly for fuel cells. Uncertainty adequate carbohydrates are not available, such as during extreme famine, the

body begins fat metabolism into a ketone body to provide the necessary fuel. High level of ketone can be very dangerous and become of toxic. This serious condition is known as *ketoacidosis* and if untreated, it can lead to a diabetic coma or even death (Pradhan and Sahu, 2011). Another condition that will result in elevated levels of ketone bodies is Type I diabetes. People with diabetes mellitus cannot metabolize glucose efficiently, because of insufficient insulin or insulin resistance. Their bodies will begin fat and protein metabolism to make up for the lack of glucose available for energy. Monitoring of ketone levels becomes significant for the diabetic patient so they can keep track of their glucose level in the body (Sharma, 2014)

The high levels of ketone in the urine, these condition called *ketonuria*, indicates that the body uses most of its fat for energy (Luethi et al., 2016). In their study, Lekha.S and Suchetha.M said that urine, sweat, saliva, tears and breath contains traces of glucose in their body, and this trail is different from the levels of glucose in the blood. Therefore some human serum has gained recognition as an alternative measurement of glucose (S. and M., 2015). Urine is one of the most widely biological fluid analysis for medical purposes because it is non-invasive and widely available. Ketones themselves are actually easily recognizable in urine compounds, especially when in the patient's blood there are already a large number of ketones. The ketone produces a distinctive aroma of acetone. This smell is sometimes smelled of the breath and urine of patients who have high levels of ketone cases (Makaram et al., 2014). Determination of ketone in urine has a very important role for the primary detection diabetes, because through our urine able to get immediate information about the pathology of a disease in the body. One tool that is used to determine the level of glucose in urine is used urine test strip. The urine test strips are often used in the hospitals to identify in the early diagnose of patients by investigating a large group of the population. Based on paper (Siyang et al., 2012) states that the diagnosis of diabetes in urine can generally be

carried out using three methods: Fasting plasma glucose test, oral test and random plasma glucose test for diabetes diagnosis.

Inspired by the major global increasing mortality of diabetes patients the main question that whether we able to predict before it comes true. The advance of computer technology, the use of intelligent methods (neural networks, fuzzy logic, and genetic algorithms) have an important role in medical concern such as diagnosis disease. As in the journal (Ziasabounchi and Askerzade, 2014) computer applications for the tools to predict, diagnose and treat patients has attracted some interest of researchers. Furthermore, medical practitioners also usess a computerized technology to assistance diagnosis and prediction (give a high opinion) as the medical diagnosis has uncertain nature(Jain and Raheja, 2015).

The examination of diabetes hopefully can be predicted before it comes true. By using the advance of computer technology, the examination can be achieved used intelligent model. Adaptive Neuro-Fuzzy Inference System (ANFIS) is one of the expert system inference Neuro-Fuzzy work with Takagi-Sugeno type that was developed by Jyh-Shing Roger Jang in 1993. As reported in (Ziasabounchi and Askerzade, 2014) these models are often used for the smart predictive model for healthcare and medical data diagnosis. Many research has highlighted the effectiveness of development ANFIS as well as learning models for predicting. Adaptive ANFIS in building a predictive system, it is hoped in the future the patient will predict dynamic glucose levels in the body so that the patient can take action before complication.

1.3 Problem Statement

Screening of *ketonuria* is often performed for patients with acute diabetes and pregnant women. This test should be done using urinary ketones to determine glucose levels in diabetic urine. Additionally, Monitoring of ketone in the diabetic patient must be check

while diabetes treatment changes from insulin to oral hypoglycemic agents. Diabetic patients should suffer regular urine tests for glucose and ketone checks, due to *ketonuria* that will develop within 24 hours after withdrawal of insulin. In conditions related with acidosis, Monitoring treatment response is needed to test ketones before became the serious illness of acidosis. Urine ketone performs first before there is a significant increase in blood ketones (Spanou et al., 2015). Therefore, measurement of urine ketone is helpful in emergency situations. During pregnancy, early detection of *ketonuria* is important because ketoacidosis is a factor associated with intrauterine death.

Since ketone measurement has unentertained that changes unpredictable, the prediction model is necessary to warn every hour so that the patient makes it possible to take preventive measures before their ketone levels increasing. Considering the fact of ketone has uncertainty whether the control of liberal glucose in diabetic patients leads to increased production of ketone and ketoacidosis. This study propose prediction model solution uses the Computational intelligent methods to compare the prevalence of dynamic sample (fasting and non-fasting) *ketonuria* in GDM (Gestational Diabetes Mellitus) following the Institute of Medicine diet instructions.

Considering the fact that most of the case of prediction proposed solution uses the Artificial Intelligence methods (Karabpour, 2015). Several studies have been conducted to predict various diseases using computer technology with intelligent methods and various algorithms, which is ANFIS method. To obtain the best performance Swain & Sample (2016) states that the performance of ANFIS can be improved by varying the type of input membership function, partitioning techniques, varying methodology and type of ANFIS system. While the best performance can be realized, the determination of the parameters and setting proper rules of the fuzzy membership function is an obstacle in the ANFIS

algorithm. Therefore the lack of using ANFIS algorithm can be improved with Genetic Algorithm (GA) to find the optimal value for modeling parameters.

1.4 Research Question

Based on background and the problem statement described above, the research question in this study are as follows:

1. How to design odor analysis for intelligent diabetic ketoacidosis prediction involving unrestricted food intake conditions.
2. What is the impact of optimization process toward the diabetes ketoacidosis prediction result?

1.5 Research Objective

In this study, there are two objectives to be achieved based on the research problems.

1. To propose an adaptive Neuro-Fuzzy inference model for diabetes ketoacidosis prediction using urine odor analysis
2. To optimize the proposed adaptive neuro-fuzzy inference model for diabetes ketoacidosis analysis.

1.6 Research Contribution

The contribution of this research are:

1) An approach of a hybrid genetic algorithm–adaptive network-based fuzzy inference system for diabetic ketoacidosis prediction under unrestricted food intake conditions using urine odor analysis.
2) A hybrid genetic algorithm–adaptive network-based fuzzy inference system with diabetic ketoacidosis modeling method.

1.7 Scope of Study

This project study concentrations of performing Adaptive Neuro Fuzzy-Inference in are of a medical part which is diabetic ketoacidosis using Artificial Intelligent approach. For this project studies, Data form in the odor of diabetic and normal people will be extracted from Electronic Nose as cases of study. The data consist of 40 specimens divided into two group, they are fasting and non-fasting, also it has complete standardization of medical ketone testing proses in CITO laboratory Semarang Central Java, Indonesia. However, the data transfers from odor to binary in Electronic Nose need to process in Arduino microcontroller then show the output data in Excel file. In addition the output of data in excel examination on artificial intelligent efficiency algorithm as compare to original ANFIS algorithm. Performance and quality measure are used to evaluate the quality, efficiency, and effectiveness of the proposed approach ANFIS and other approaches.

1.8 Project Report Overview

This study follows six chapter of the project report. The report structure is as follow:

Chapter 1 is the introduction of the scope of research by providing information about the authenticity of the highlights of the research include the problem statement, followed by the purpose of the research conducted to answer the problem statement and research question. In this chapter also presents information about the contribution of research in the field of informatics and communication technology. Finally, everything is summarized in the project report overview.

Chapter 2 is the literature review about diabetes prediction system using machine learning approach in the domain urine analysis, it focused on studies, review and examines on diabetes prediction using machine learning algorithm based urine odor analysis.

Chapter 3 is the methodology. It is mentioned in the most suitable method to conduct research and determine effective processes to answer research problem, including detail of

proposed methodology, proposed hybrid GA-ANFIS algorithm, research design, experimental paradigm design that will be used for data collection and analysis.

Chapter 4 is ANFIS prediction model. It's brief step by step in order to build the prediction model. Those steps are constructed structure of ANFIS contain, first is the model prediction building, the second step is explained the generation of FIS then is proposed method Genetic Algorithm in order to optimization ANFIS

Chapter 5 is the Result comparison and analysis. This chapter will continue from the last chapter to analyze the result from the experiment. Result analysis is an important step in the research which is a process to observe and estimate the relevant result in order to select the best method used in the experiment.

Chapter 6 is conclusion that provide the summarize for each chapter This chapter consisst of research background and purpose of study, literature review and proposed method are used in the project, experimental result and future work.

1.9 Summary

The problem of this research is to predict the dynamic blood glucose level from the odor of urine using Adaptive Neuro Fuzzy Inference model. Since blood glucose measurement has a natural temporal ordering glucose level changes have erratic nature, the prediction model is necessary to warn every hour so that the patient makes it possible to take preventive measures before the change will happen in their blood glucose levels. The ANFIS model is use to optimize issue on prediction diabetes.