PROJECT REPORT

Machine Learning to Predict the onset of Diabetes Disease based on Diagnostic Measures using Two Different Libraries and Algorithms.

EWEN LIAM
17.K1.0052

Faculty of Computer Science
Soegijapranata Catholic University
2021

APPROVAL AND RATIFICATION PAGE

Judul Tugas Akhir	:	Machine Learning to Predict the onset of Diabetes Disease based on Diagnostic Measures using Two Different Libraries and Algorithms.
Diajukan oleh	:	Ewen Liam
NIM	:	17.K1.0052
Tanggal disetujui	:	07 Juli 2021
Telah setujui oleh		
Pembimbing	:	R. Setiawan Aji Nugroho S.T., MCompIT., Ph.D
Penguji 1	:	R. Setiawan Aji Nugroho S.T., MCompIT., Ph.D
Penguji 2	:	Rosita Herawati S.T., M.I.T.
Penguji 3	:	Hironimus Leong S.Kom., M.Kom.
Penguji 4	:	Y.b. Dwi Setianto
Penguji 5	:	Yulianto Tejo Putranto S.T., M.T.
Ketua Program Studi	:	Rosita Herawati S.T., M.I.T.
Dekan	:	R. Setiawan Aji Nugroho S.T., MCompIT., Ph.D

Halaman ini merupakan halaman yang sah dan dapat diverifikasi melalui alamat di bawah ini.

STATEMENT OF ORIGINALITY

I, the undersigned:

Name : EWEN LIAM

ID : 17.K1.0052

Certify that this project was made by myself and not copy or plagiarize from other people, except that in writing expressed to the other article. If it is proven that this project was plagiarizes or copy the other, I am ready to accept a sanction.

Semarang, July 7, 2021

EWEN LIAM

APPROVAL PAGE FOR PUBLICATION OF SCIENTIFIC PAPERS FOR ACADEMIC INTEREST

The undersigned below:

Name : Ewen Liam

Undergraduate Program : INFORMATICS ENGINEERING

Faculty : COMPUTER SCIENCE

Type of work : THESIS

Approved to give Non-Exclusive Royalty Free Right to Soegijapranata Catholic University Semarang for scientific work entitled "Machine Learning to Predict the onset of Diabetes Disease based on Diagnostic Measures using Two Different Libraries and Algorithms." along with the existing tools (if needed). With this Non-Exclusive Royalty Free Right Soegijapranata Catholic University has the right to store, transfer data / format, manage in the form of database, maintain and publish this final project as long as I keep my name as a writer / creator and as a Copyright owner.

This statement I made in truth

Semarang, July 7, 2021

Sincerely

EWEN LIAM

ACKNOWLEDGEMENETS

First of all, all glory belongs to Jesus that allows me to make this thesis and always strengthens me in the making of this thesis. This thesis is dedicated to faith in humanity as the motto of Soegijapranata Catholic University that goes "Talenta Pro Patria at Humanitate". I hope this thesis could be useful for the development of machine learning implementation in medicine and could help people that struggle with diabetes diseases. This thesis also for the completion of the requirement to finish a Bachelor Degree in Informatics Engineering, Soegijapranata Catholic University.

This thesis can be successfully done also with the support of my surroundings, special thanks to:

1. All of the lectures, that always acknowledge me and make a time to discuss this problem with me, especially my supervising lecturer R. Setiawan Aji Nugroho S.T., M CompIT., Ph.D.
2. My dearest family, Law Mie Tjoe, Oei Sien Ling, Gagarin, Bethea, and Hansen. For always supporting me with materials and love.
3. All of my dearest friends (Abigail Anita, Stevanus, Aditya Morning Star, David, Lucas, Dinar) for always be there and encourage me when things get hard. And all of my friends that I cannot mention.
4. Last but not least, I want to thank me, I want to thank me for believing in me, I want to thank me for doing all this hard work, I want to thank me for having no days off, I want to thank me for never quitting, I want to thank me for always being a giver, and trying to give more than I receive, I want to thank me for trying to do more right than wrong, I want to thank me for just being me at all times.

Semarang, July 7, 2021

Sincerely

EWEN LIAM

ABSTRACT

Diabetes is a very serious problem that needs to be solved, the cure has not been found yet, but we still could prevent it by diagnosing people who have diabetes disease. If people positively have a diabetes disease, they can quickly prevent it from a young age by watching their diet. So, it is an important thing to predict the patient who has diabetes disease.This study will show about how technology could help the medic as a milestone in medical study

The technology that used in this project is machine learning based, using logistic regression to predict the onset of diabetes. This project will also be using two methods, the first method is using Naïve Bayes from Sci-Kit learn library and the second method is Logistic Regression without using library at all.

The final result is the Naïve Bayes with Sci-Kit library performs better compare to the logistic regression without library. It could happen because, the more complex the algorithm the more the algorithm has higher accuracy. Although the model of the logistic regression without library does not have good accuracy as the Naïve Bayes with library, the logistic regression also high value which means it can be trusted code.

Keyword: Logistic Regression, Naïve Bayes, Sci-kit Learn Library, Diabetes, Prediction

TABLE OF CONTENTS

ILLUSTRATION INDEX

INDEX OF TABLES