A MACHINE LEARNING MODEL FOR PREDICTING PRE-ECLAMPSIA FOR RESOURCE-CONSTRAINED REGIONS.

ABDUSHAKUR JAMWA ARINA

A THESIS SUBMITTED TO THE INSTITUTE OF COMPUTING AND INFORMATICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF MASTER OF SCIENCE IN INFORMATION TECHNOLOGY OF TECHNICAL UNIVERSITY OF MOMBASA

2023

DECLARATION

This thesis is my original work and has not been presented for academic award in any other university.

---------- ----------

ABDUSHAKUR JAMWA ARINA **Date**

MSIT/0005/2020

This thesis has been submitted with our approval as university supervisors.

---------- ----------

DR. MGALA MVURYA **Date**

---------- ----------

DR. ANTONY LUVANDA **Date**

---------- ----------

DR. PAMELA T. KIMETO **Date**

DEDICATION

To my beloved wife Salwa Lali Shee and my daughter Fatma Abdushakur Jamwa.

ACKNOWLEDGEMENT

First, I would like to express my gratitude to the divine power of God, who is the ultimate creator of everything visible and invisible, the supreme ruler of life's precious gift, and the driving force behind my passion for seeking knowledge.

Secondly, I would like to convey my sincere gratitude to my supervisors, namely Dr. Mgala Mvurya, Dr. Antony Luvanda, and Dr. Pamela Kimeto, for their guidance and support throughout all the phases of this thesis. Their assistance has been crucial and significant in facilitating the completion of this work.

Thirdly, my sincere gratitude goes to my classmates and friends Titus, George, Joseph, Ester Asenath, Athman, Amran, Michael, and Peter for always motivating my thought process by always challenging me to be my best.

Lastly, I wish to acknowledge all those who gave their input either through prayers, words of wisdom, and any kind of contribution. I cannot thank you enough for the sacrifices you afforded me throughout this period. Only God can repay you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS AND ABBREVIATIONS

ACOG	American College of Obstetricians and Gynecologists
AI	Artificial Intelligence
ANC	Ante Natal Care
ANN	Artificial Neural Network
ICT	Information Communication & Technology
KDH	Kenya Demographic and Health Survey
KNBS	Kenya National Bureau of Statistics
LDA	Linear Discriminant Analysis
LR	Logistic Regression
MDG	Millennium Development Goals
ML	Machine Learning
MCH	Maternal and Child Health
MoH	Ministry of Health
NB	Naïve Bayes
PE	Pre-eclampsia
RF	Random Forest
SDG	Sustainable Development Goals
SVM	Support Vector Machines

UNCTAD United Nations Conference on Trade and Development

WB World Bank

WHO World Health Organization

ABSTRACT

Pre-eclampsia is globally recognized by the World Health Organization as a significant contributor to high rates of morbidity and mortality among infants and mothers. It accounts for approximately 3% to 5% of all reported pregnancy-related complications worldwide. However, in developing nations like Kenya, particularly in the sub-Saharan region, the prevalence of pre-eclampsia is notably higher, ranging from 5.6% to 6.5% of reported pregnancies. Key risk factors associated with pre-eclampsia include sudden elevation in blood pressure, increased protein levels in urine, chronic kidney disease, and the presence of either Type 1 or Type 2 diabetes. This research developed a predictive model for pre-eclampsia utilizing supervised machine learning techniques on socio-demographic data gathered from Kilifi County. A total of 500 secondary data records gathered from Kilifi county hospital were pre-proposed to train and test the machine learning models. To train and test the models, the study employed five supervised machine learning algorithms, namely Logistic Regression, Random Forest, Naïve Bayes, Linear Discriminant Analysis, and Support Vector Machines. Maternal age, marital status, gravida, education level, and ANC attendance were identified as the optimal extracted features using PCA. The logistic regression model outperformed other supervised machine learning models in the study, achieving a high accuracy rate of 0.96 in predicting pre-eclampsia. The results show that Logistic Regression can accurately predict pre-eclampsia within the first trimester of pregnancy. Future research will involve collecting more data from different regions to improve performance and building a mobile application that will improve MCH accessibility in resource-constrained regions in the country.