05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

PUBLIC HEALTH MAPPING BASED ON CHILDREN'S NUTRITIONAL STATUS PREDICTION USING ENHANCED FRAMEWORK OF MACHINE LEARNING

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

LILIANA SWASTINA

SULTAN IDRIS EDUCATION UNIVERSITY
2025

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

PUBLIC HEALTH MAPPING BASED ON CHILDREN'S NUTRITIONAL STATUS PREDICTION USING ENHANCED FRAMEWORK OF MACHINE LEARNING

LILIANA SWASTINA

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

THESIS PRESENTED TO QUALIFY FOR A DOCTOR OF PHILOSOPHY

FACULTY OF COMPUTING AND META-TECHNOLOGY
SULTAN IDRIS EDUCATION UNIVERSITY

2025

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

UPSI/IPS-3/BO 32
Pind : 00 m/s: 1/1

SULTAN IDRIS EDUCATION UNIVERSITY

Please tick (√)	
Project Paper	
Masters by Research	
Master by Mixed Mode	
PhD	√

INSTITUTE OF GRADUATE STUDIES

DECLARATION OF ORIGINAL WORK

This declaration is made on the 26 (day) September (month) 2025 (year)

i. Student's Declaration:

I, LILIANA SWASTINA (MATRIC NO. P20211002730), Faculty of Computing and Meta-Technology hereby declare dissertation / thesis for Doctor of Philosophy in Artificial Intelligence titled PUBLIC HEALTH MAPPING BASED ON CHILDREN'S NUTRITIONAL STATUS PREDICTION USING ENHANCED FRAMEWORK OF MACHINE LEARNING is my original work, and any sources containing copyrighted material have been properly cited for the permitted purposes. Any quotations, excerpt, reference or re-publication from or any works that has copyright had been clearly and well cited.

Signature of the student

ii. Supervisor Declaration:

I, Profesor Dr. Bahbibi binti Rahmatullah hereby certify that the work entitled PUBLIC HEALTH MAPPING BASED ON CHILDREN'S NUTRITIONAL STATUS PREDICTION USING ENHANCED FRAMEWORK OF MACHINE LEARNING was prepared by the above the named student, and was submitted to the Institute of Graduate Studies as a partial fulfilment for the conferment of Doctor of Philosophy in Artificial Intelligence, and the aforementioned work, to the best of my knowledge, is said student's work.

26/09/ 2025

Date

Professor Dr. Bahbibi Rahmatullah
Pensyarah
Fakulti Komputeran dan Meta-Teknologi
Universiti Pendidikan Sultan Idris
35900 Tanjong Malim, Perak

Signature of the Supervisor

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

UPSI/IPS-3/BO 3'
Pind.: 01 m/s:1/'

UNIVERSITI PENDIDIKAN SULTAN IDRIS
SULTAN IDRIS EDUCATION UNIVERSITY

INSTITUT PENGAJIAN SISWAZAH /
INSTITUTE OF GRADUATE STUDIES

BORANG PENGESAHAN PENYERAHAN TESIS/DISERTASI/LAPORAN KERTAS PROJEK
DECLARATION OF THESIS/DISSERTATION/PROJECT PAPER FORM

Tajuk / *Title*: PUBLIC HEALTH MAPPING BASED ON CHILDREN'S NUTRITIONAL STATUS PREDICTION USING ENHANCED FRAMEWORK OF MACHINE LEARNING

No. Matrik /*Matric's No.*: P20211002730

Saya / *I* : LILIANA SWASTINA

(Nama pelajar / *Student's Name*)

mengaku membenarkan Tesis/Disertasi/Laporan Kertas Projek (Kedoktoran/Sarjana)* ini disimpan di Universiti Pendidikan Sultan Idris (Perpustakaan Tuanku Bainun) dengan syarat-syarat kegunaan seperti berikut:-
acknowledged that Universiti Pendidikan Sultan Idris (Tuanku Bainun Library) reserves the right as follows:-

1. Tesis/Disertasi/Laporan Kertas Projek ini adalah hak milik UPSI.
   The thesis is the property of Universiti Pendidikan Sultan Idris
2. Perpustakaan Tuanku Bainun dibenarkan membuat salinan untuk tujuan rujukan dan penyelidikan.
   Tuanku Bainun Library has the right to make copies for the purpose of reference and research.
3. Perpustakaan dibenarkan membuat salinan Tesis/Disertasi ini sebagai bahan pertukaran antara Institusi Pengajian Tinggi.
   The Library has the right to make copies of the thesis for academic exchange.
4. Sila tandakan (√) bagi pilihan kategori di bawah / *Please tick (√) for category below:-*

☐	**SULIT/*CONFIDENTIAL***	Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub dalam Akta Rahsia Rasmi 1972. / *Contains confidential information under the Official Secret Act 1972*
☐	**TERHAD/*RESTRICTED***	Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan ini dijalankan. / *Contains restircted information as specified by the organization where research was done.*
√	**TIDAK TERHAD / *OPEN ACCESS***	

(Tandatangan Pelajar/ Signature)

Tarikh: 26 September 2025

Professor Dr. Bahbibi Rahmatullah
Pensyarah
Fakulti Komputeran dan Meta-Teknologi
Universiti Pendidikan Sultan Idris
35900 Tanjong Malim, Perak

(Tandatangan Penyelia / *Signature of Supervisor*)
& (Nama & Cop Rasmi / *Name & Official Stamp*)

Catatan: Jika Tesis/Disertasi ini **SULIT** @ **TERHAD**, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan ini perlu dikelaskan sebagai **SULIT** dan **TERHAD**.

Notes: If the thesis is CONFIDENTAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

ACKNOWLEDGEMENT

The journey to this doctoral degree has been challenging yet rewarding, made possible by the support of many individuals. This thesis is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Artificial Intelligence at the Sultan Idris Education University, Malaysia.

I would like to express my gratitude to Jesus Christ, God Almighty, for His blessings, which have enabled me to complete my thesis, "Public Health Mapping Based on Children's Nutritional Status Prediction Using Enhanced Framework of Machine Learning." I am grateful to everyone who has helped and supported me throughout this journey. I am very grateful to my supervisor, Prof. Dr. Bahbibi binti Rahmatullah, for her invaluable guidance, encouragement, and insights throughout this process. Her extensive knowledge and high standards continue to inspire me to strive for excellence. To Prof. Dr. Nazri bin Mohd Nawi, Assoc. Prof. Dr. Rozmie Razif Othman, and Dr. Mad Helmi bin Abd Majid, as evaluator for the suggestions given to improve this thesis. I would like to express my sincere gratitude and deep appreciation to the following individuals. I express my sincere gratitude to Prof. Dr.-Ing. Maizatul Hayati binti Mohammad Yatim, Dean of the Faculty of Computing and Meta-Technology at the Sultan Idris Education University, for providing the opportunity and facilities necessary for the thesis. I am grateful to Assoc. Prof. Dr. Aslina binti Saad, Dr.Aslina binti Ahmad for her facilitation, valuable inputs, and suggestions that improved this research.

I also express my gratitude to Yayasan Indah Banjarmasin and Universitas Sari Mulia, who have allowed me to pursue this doctoral education. May Universitas Sari Mulia continue to advance and develop.

I would like to express my gratitude to all the Heads of Community Health Centers and staff in Banjarmasin City and in South Kalimantan Province who have been very helpful in the process of collecting research data and providing information according to their respective fields of expertise. My deepest heart goes to my parents, beloved husband Agus Priyono, and beloved son Jeremy Christian in heaven, Perkantas Banjarmasin, PRII Banjarmasin, GRII Kertajaya Surabaya, and family that I love and care for. Thank you for all the moral and prayer support that is immeasurable and very meaningful to me. May God Almighty reward everyone with kindness, Amen.

Finally, I express my deepest appreciation for the guidance and support from all parties that I have mentioned here. No word can express my gratitude. I dedicate this work to all of you. Thank you.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

ABSTRACT

Public health monitoring involves community efforts to prevent disease, extend life, and promote health care, considering social, cultural, and economic influences. Surveillance process plays a crucial role in addressing health issues, with children's nutrition serving as a key indicator of community well-being. To identify the priority of intervention by the government, a map indicating the mapping of public health status is needed. Numerous previous studies have explored using machine learning models to analyze children's nutrition status. However, a definitive framework for predicting children's nutritional status using machine learning remains uncertain, which is essential for creating a public health map. Developing and validating an enhanced machine learning framework for predictive analysis is necessary to generate this public health map. This study used various algorithms in the proposed enhanced framework, including Neural Networks, Random Forests, Decision Trees, Logistic Regression, and Extreme Gradient Boosting. The findings showed that the framework that employs a class imbalance handler offers better results. The Neural Networks method demonstrated the highest accuracy rates of 90.5% for wasting, 79.0% for underweight, and 74.8% for stunting, outperforming other methods. Performance comparisons are primarily accuracy-based but consider additional metrics like the F1 score for a comprehensive assessment. The framework for predicting WHZ (weight-for-height z-score) status combines Handler Class Imbalance Bagging (7:3) with Synthetic Minority Over-sampling Technique (SMOTE). Predicting HAZ (height-for-age z-score) status, the framework uses Handler Class Imbalance Bagging (7:3) augmented with Sample 300. The enhanced frameworks were employed to predict WHZ and HAZ status. Based on those predictions, the study proposed various public health maps, including cluster-based and threshold-based maps. This research implies the potential of the proposed machine learning frameworks to enhance public health intervention strategies by accurately predicting children's nutritional status and guiding targeted interventions.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

PEMETAAN KESIHATAN AWAM BERDASARKAN RAMALAN STATUS NUTRISI KANAK-KANAK MENGGUNAKAN RANGKA KERJA PEMBELAJARAN MESIN YANG DIPERTINGKATKAN

ABSTRAK

Pemantauan kesihatan awam melibatkan usaha komuniti untuk mencegah penyakit, memanjangkan hayat, dan mempromosikan penjagaan kesihatan, dengan mengambil kira pengaruh sosial, budaya, dan ekonomi. Proses pemantauan digunakan untuk menangani isu-isu kesihatan, dengan nutrisi kanak-kanak menjadi ukuran penting kesihatan komuniti. Untuk menentukan keutamaan intervensi oleh kerajaan, sebuah peta yang menunjukkan pemetaan status kesihatan awam diperlukan. Banyak kajian terdahulu yang menggunakan model pembelajaran mesin untuk menganalisis nutrisi kanak-kanak telah dilakukan. Walau bagaimanapun, rangka kerja yang muktamad untuk meramalkan status nutrisi kanak-kanak menggunakan pembelajaran mesin masih belum jelas, sedangkan ia penting untuk menghasilkan peta kesihatan awam. Pembangunan dan pengesahan rangka kerja pembelajaran mesin yang dipertingkatkan untuk analisis ramalan adalah perlu bagi menghasilkan peta kesihatan awam ini. Kajian ini menggunakan pelbagai algoritma dalam rangka kerja dipertingkatkan yang di cadangkan, termasuk Rangkaian Neural, Hutan Rawak, Pokok Keputusan, Regresi Logistik, dan Extreme Gradient Boosting. Dapatan kajian menunjukkan bahawa rangka kerja yang menggunakan pengendali ketidakseimbangan kelas menawarkan hasil yang lebih baik. Kaedah Rangkaian Neural menunjukkan kadar ketepatan tertinggi iaitu 90.5% untuk susut berat badan *(wasting*), 79.0% untuk kurang badan rendah *(underweight*), dan 74.8% untuk bantut (*stunting*) mengatasi kaedah-kaedah lain. Perbandingan prestasi adalah berdasarkan ketepatan, di samping mengambil kira metrik tambahan seperti skor F1 untuk penilaian yang lebih menyeluruh. Rangka kerja untuk meramal status WHZ (z-skor berat-untuk-tinggi) menggabungkan Pengendali Ketidakseimbangan Kelas Bagging (7:3) dengan Teknik Sampel Minoriti Sintetik (SMOTE). Untuk meramal status HAZ (z-skor tinggi-untuk-umur), rangka kerja menggunakan Pengendali Ketidakseimbangan Kelas Bagging (7:3) yang diperkuat dengan Sampel 300. Rangka kerja yang dipertingkatkan digunakan untuk meramal status WHZ dan HAZ. Berdasarkan ramalan tersebut, kajian ini telah mencadangkan pelbagai peta kesihatan awam, termasuk peta berasaskan kluster dan peta berasaskan nilai ambang. Kajian ini menunjukkan potensi rangka kerja pembelajaran mesin yang dicadangkan untuk meningkatkan strategi intervensi kesihatan awam dengan meramalkan status nutrisi kanak-kanak secara tepat dan menjadi panduan intervensi bersasar.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

TABLE OF CONTENTS

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

LIST OF TABLES

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

LIST OF FIGURES

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi
05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi
05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

LIST OF ABBREVIATIONS

ACC	Accuracy
AUC	Area Under Curve
BDHS	Bangladesh Demographic and Health Survey
DL	Deep Learning
DT	Decision Tree
E-Net	Elastic Net
HAZ	High for Age Z scores
k-NN	k-Nearest Neighbor
KMS	Kartu Menuju Sehat/ children's health indicators card
LDA	Linear Discriminant Analysis
LR	Logistic Regression
NB	Naive Bayes
NN	Neural Networks
MBL	Mutually Beneficial Learning
ML	Machine Learning
Posyandu	Pos Pelayanan Terpadu/local integrated health services
Puskesmas	Pusat kesehatan masyarakat/public health facility
RF	Random Forest
SMOTE	Synthetic Minority Over-sampling Technique
SVM	Support Vector Machine
WAZ	Weight for Age Z score
WHZ	Weight for Height Z score

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

LIST OF APPENDICES

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

CHAPTER 1

INTRODUCTION

1.1 Introduction

Public health surveillance is key to understanding and addressing health problems affecting communities. This includes improving sanitation, medical services, and care. More broadly, it includes all factors that influence health, such as social, cultural, and economic factors (Azari & Borisch, 2023). Public health is a broad field covering all aspects of population health. In this context, children's nutritional status not only reflects individual health but can also be considered a strong indicator of overall community health (Chiolero & Buckeridge, 2020). Some health problems can be caused by malnutrition. Various diseases can appear, including Hyponatremia, Hypokalemia, Anemia, and various other diseases due to vitamin deficiency (Finn, 2014). A person can be malnourished, despite eating enough. Adequate nutrition in the food consumed is essential. Otherwise, he may suffer from a lack of carbohydrates, proteins, fats, vitamins, and minerals. It can bring permanent damage to the growth and development of children.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

Nutrition plays a crucial role as one of the most significant risk factors in early childhood development. This is supported by research showing that an estimated 250 million children under five in low- and middle-income countries, representing 43% of this demographic, may not reach their full development potential due to extreme poverty (Lu, Black, & Richter, 2016). Citing the 2018 Global Nutrition Report, it is revealed that about a third of children from these countries suffer from malnutrition. Additionally, complaints of hunger account for 20% of total child deaths worldwide (Hawkes, 2018). Even the United Nations Children's Fund (UNICEF) has described underweight, overweight, and hidden hunger (suffering from micronutrient deficiencies) as threats to the survival, growth, and development of a country's children, economy, and society. These three things are called the Triple Burden.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

In addition, a study by the World Bank showed that if a country does not eliminate stunting when their future workers are children, there will be a loss of 7% of GDP per capita (World Bank, 2018). Thus, the effects of malnutrition on the country's economy do not occur suddenly but stem from the neglect of providing sufficient nutrition to the prospective labor force.

On the other hand, parenting in many communities generally prioritizes children. Therefore, monitoring children's health and nutritional status becomes a critical benchmark. If children's nutrition is well cared for, it can positively affect the community's nutrition and health (Finn, 2014). In other words, child malnutrition indicates that the community may have experienced a disruption in the supply of nutrition or lacked food nutrition education, leading to broader public health issues. A

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

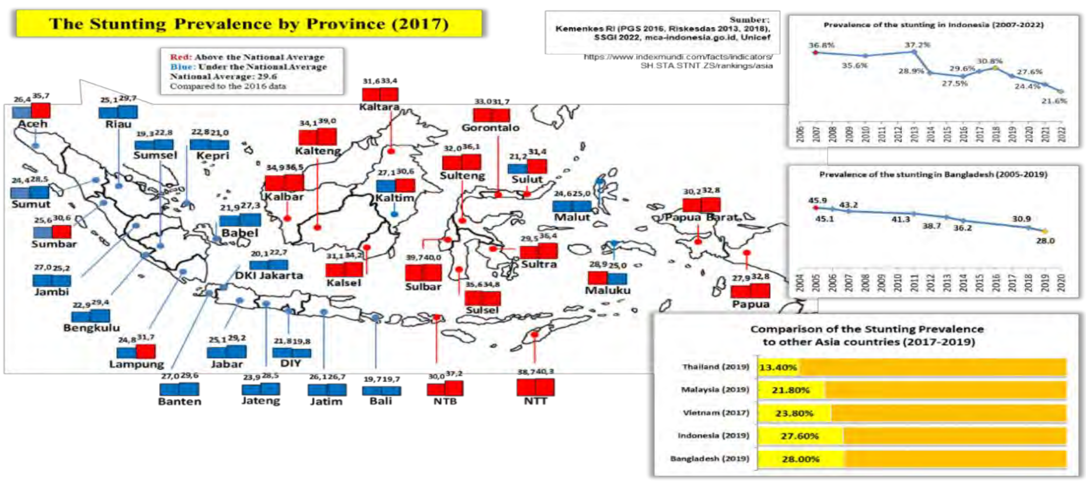
05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

high prevalence of stunting may suggest that the community faces challenges related to reproductive health, nutritional education, parenting, access to a healthy environment, hygiene, water, sanitation, or other health facilities (Chiolero & Buckeridge, 2020; Hong et al., 2020). Thus, the prevalence of nutritional status can serve as a significant public health indicator.

In Indonesia, citing the study of Direktorat Bina Gizi Masyarakat Departemen Kesehatan Republik Indonesia (Directorate of Community Nutrition Development of Health Ministry of Indonesia), changes in weight of under-five children within a certain time is an early indication of changes in nutritional status. In six months, the children whose weight does not increase two times are at risk of experiencing malnutrition 12.6 times compared to the children whose weight continues to increase (Lareno, B., Swastina, L. & Junaidi, H.M., 2018). The Indonesian stunting prevalence is quite large and spread over several areas (Figure 1.1) to the previous year. They are included in the ten highest stunting provinces.

Figure 1.1

The stunting prevalence in Indonesia (Lareno et al., 2020)



05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

As shown in Figure 1.1, although a decline in stunting prevalence has been observed, it has not been as rapid as expected. The end-of-year 2024 target of 14% stunting prevalence remains challenging to achieve uniformly across all regions, particularly in Kalimantan, Nusa Tenggara, and Papua. The South Kalimantan Province had a stunting prevalence of 34.2% in 2017. This reflects a rise of 1.1% compared to the previous data. The Indonesian government has established Posyandu (neighborhood integrated healthcare post) under Puskesmas (public health facility) to monitor maternal, neonatal, and child health (MNCH). Posyandu's primary goal is to monitor the growth of children and record data on the Kartu Menuju Sehat/KMS (MNCH Card). In Banjarmasin, there are 26 Puskesmas, each hosting 8-12 Posyandu. However, despite the extensive data recorded, it does not always directly indicate the overall health condition of the community.

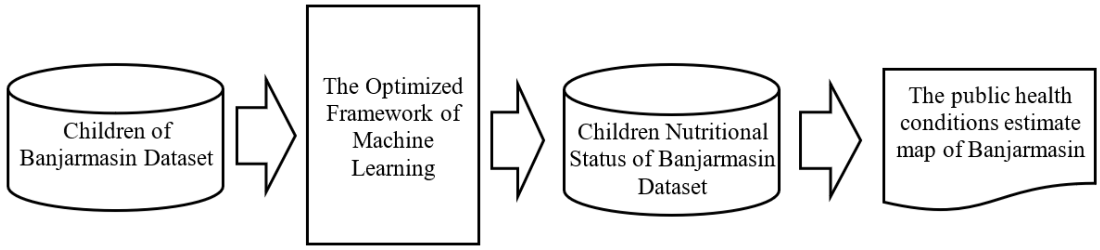
After a temporary halt due to the Covid-19 pandemic, Posyandu activities in Banjarmasin gradually resumed in January 2022. Efforts to reduce stunting necessitate expanded health access and facilities, but limited government budgets require prioritization of urgent areas. Therefore, maximizing the budget involves prioritizing areas of more urgent need. This requires a dataset to produce a map showing the level of need for government intervention to improve public health conditions. Utilizing health data from Posyandu through machine learning offers a promising approach to generating such a map (Figure 1.2). However, data from Posyandu must be pre-processed to be used effectively for machine learning applications.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

Figure 1.2

Schematic diagram to generate a map of health conditions



Thus, public health involves community efforts to prevent disease, extend life, and promote health, considering social, cultural, and economic influences. Surveillance addresses health issues, with children's nutrition being a vital measure of community health. To prioritize, a map indicates where government intervention is most urgently required.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

The existing maps are only available at the city or district level and do not provide more detailed mapping at the sub-district or public health facility (puskesmas) level. Additionally, these existing maps do not connect both child and mother/family data. Building a dataset that connects both child and mother/family data and provides finer geographic detail is essential for accurately capturing the relationships between socioeconomic, environmental, and health-related factors. This enables more effective interventions and strategies to improve public health outcomes in the region.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

1.2 Research Background

Machine learning (ML) is a branch of artificial intelligence that presents a significant opportunity to enhance mathematical modeling, particularly in handling nutrition data and anthropometric measurements, which often exhibit non-linear associations. ML's capacity to extract meaningful insights and patterns from vast datasets is instrumental in predicting nutritional status and formulating effective problem-solving strategies.

Numerous studies have explored the application of ML models in analyzing children's nutrition. Talukder & Ahammed (2020) employed Random Forest (RF), k-nearest Neighbor (k-NN), Logistic Regression (LR), Linear Discriminant Analysis (LDA), and Support Vector Machine (SVM) algorithms to classify children's nutritional status in Bangladesh. They reported that the RF algorithm yielded the highest accuracy. Similarly, Hemo & Rayhan (2021) compared RF and Decision Tree (DT) models, concluding that RF outperformed DT in accuracy. Fenta, Zewotir & Muluneh (2021) also found RF to be the most effective model among Elastic Net (E-Net), LR, Ridge Regression, Lasso, and Neural Network (NN) methods. Conversely, Shahriar et al. (2019) observed that Neural Network (NN) outperformed Support Vector Machine (SVM) and RF models. Ferdowsy et al. (2021) also noted that Logistic Regression (LR) achieved the highest accuracy compared to other models, including k-NN, SVM, RF, DT, Mutually Beneficial Learning, and Naive Bayes.

Many studies utilize Body Mass Index (BMI) or height-for-weight as indicators to classify nutritional status (as shown in Table 2.4). Another anthropometric parameter used is the Weight-for-Height Z score (WHZ), Weight-for-Age Z score (WAZ), and

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

Height-for-Age Z score (HAZ) as indicators for wasting, stunting, and underweight, respectively (Ariyadasa et al., 2012; Rahman, 2021; Shahriar, 2019).

1.3 Problem Statement

Based on the background provided, various studies have explored machine learning (ML) algorithms for analyzing children's nutrition. Talukder & Ahammed (2020) reported that the Random Forest (RF) algorithm achieved the highest accuracy. Similarly, Hemo & Rayhan (2021) and Fenta, Zewotir & Muluneh (2021) favored RF as the most effective model. Conversely, Shahriar et al. (2019) observed that Neural Networks (NN) outperformed RF. Meanwhile, Ferdowsy et al. (2021) highlighted Logistic Regression (LR) as having superior accuracy compared to RF. In addition, Bitew, Sparks & Nyarko (2021) and Pang et al. (2021) claim Extreme Gradient Boosting (XGB) has a better performance compared to LR, NN, RF, and DT. However, the algorithms for predicting the nutritional status of children in Banjarmasin remain uncertain, which is crucial for generating a map of public health conditions in the area.

Therefore, further investigation is warranted, involving rigorous testing of several nominee algorithms including NN, DT, RF, LR, and XGB. These algorithms have been selected due to their well-established track record in classification and prediction within the medical field (Fenta, Zewotir, & Muluneh, 2021; Hemo & Rayhan, 2021; Ferdowsy et al., 2021; Rahman et al., 2021; Zare et al., 2021; Talukder & Ahammed, 2020). Each algorithm undergoes evaluation based on criteria such as accuracy, area under curve (AUC), precision, f1-score, sensitivity, and specificity.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

On the other hand, there are several frameworks to predict the nutritional status of children. The frameworks are namely Rahman’s Framework (Rahman et al., 2021), Ferdowsy’s Framework (Ferdowsy et al., 2021), and Shariar’s Framework (Shariar et al., 2019). Those frameworks employ several algorithms like Neural Network (NN), Decision Tree (DT), Random Forest (RF), and Logistic Regression (LR). When these frameworks were applied to the Banjarmasin dataset, optimal performance was not yet produced.

Differences in the dataset's features or other underlying factors present an opportunity to adapt the frameworks to better align with the Banjarmasin dataset, or customize the dataset features for improved performance. Adding a class imbalance handler is essential to address the common challenge of class imbalance in machine learning, particularly in health-related predictions where some classes are often underrepresented. Without addressing this issue, models may become biased toward the majority class, leading to poor generalization and misleading performance metrics (Khare et al., 2017; Momand et al., 2020; Amin and Novitasari, 2022). Thus, a new framework is needed because existing ones do not effectively handle the class imbalance present in the Banjarmasin dataset and may not be optimized for its unique features. By integrating techniques like SMOTE and Bagging, the new framework addresses these imbalances, ensuring more accurate predictions. Additionally, the new framework is tailored to the local context, enhancing prediction accuracy and reliability, and offers significant improvements such as faster computation times, more efficient utilization, and a user-friendly interface.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

The research problem (RP) can be stated as follows:

1. The absence of connected child and mother/family data limits effective public health interventions and strategies.
2. The best performance algorithms for predicting the nutritional status of children in Banjarmasin remain uncertain.
3. The previous frameworks have not yet addressed class imbalance.

1.4 Research Question

Addressing the complexities of predicting children's nutritional status, it is essential to pinpoint the pivotal questions guiding this research. These research questions will not only drive the investigation but also ensure that the findings can significantly contribute to public health strategies aimed at improving children's nutritional outcomes in Banjarmasin.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

The research question (RQ) can be stated as follows:

1. How to link child and maternal/family data, and identify key features for enhancing analysis?
2. Which algorithm gives the best results in predicting children’s nutritional status?
3. Do frameworks that apply class imbalance steps provide better prediction performance?

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

1.5 Research Objectives

The research questions outlined provide a clear roadmap for investigating the nutritional status of children in Banjarmasin. It is necessary to compile comprehensive information on the nutritional status of children in Banjarmasin. This dataset will serve as the foundation for predictive modeling. Then develop and validate an optimized framework by integrating various machine learning algorithms. The goal is to generate maps of public health conditions based on children's nutritional status. That process is supported by predicting children's nutritional status.

This study aims to:

1. Assess the suitability of the evaluated features of the Banjarmasin dataset for predicting the nutritional status of children.
2. Identify and enhance the algorithms that provide better prediction performance of children's nutritional status.
3. Enhance the framework by using class imbalance to achieve the best predictive performance and generate a comprehensive mapping of public health conditions based on the nutritional status of children.

1.6 Operational Definition

Before diving into the analysis and results of this research, it is crucial to operationally define the key terms. These operational definitions ensure that each concept is identified and measured consistently, providing clarity and reproducibility throughout the study.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

Children's Nutritional Status. Definition: The nutritional status of children will be assessed using anthropometric indicators such as weight-for-height, height-for-age, and weight-for-age, classified according to WHO standards. Measurement Tools: Data was gathered through health surveys or medical records in the Banjarmasin area.

Prediction Using Machine Learning. Definition: The prediction of children's nutritional status was carried out using an enhanced machine-learning framework. Models were trained on historical data, incorporating variables such as child age, gender, maternal age, maternal education, maternal employment, and other relevant features. Measurement Tools: Algorithms like neural networks, random forests, decision trees, extreme gradient boost, or logistic regression were employed. The performance of these models was evaluated using metrics such as accuracy, precision, and recall.

Public Health Mapping. Definition: Public health mapping was conducted based on the predicted nutritional status of children, identifying regions with a high prevalence of malnutrition. Measurement Tools: Spatial Information tools were used to visualize prediction data on maps, facilitating the identification of patterns and spatial distribution.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

1.7 Study Limitations

This research focuses on developing the best performance of the framework for predicting the nutritional status of children so that the dataset can be used to generate a map of the health conditions of children and the community of an area. So, the stakeholders (government in several levels and field focus, including public work, health, demographic affairs, and food security) that appropriate planning and interventions can be made and assist efforts to address problems arising from malnutrition or other nutrition-related health problems.

In addition, focus on Banjarmasin children and use only well-established algorithms that have been previously proven accurate in the medical field and appropriate for this nutritional status prediction. The Banjarmasin dataset will be developed and validated using optimized features derived from established related research datasets. If successful, this approach could be extended to other cities or applied to broader areas.

1.8 Importance of Research

The study outcomes, including a framework for predicting nutritional status and maps of public health estimates, are expected to be valuable resources for lecturers, students, and researchers. Lecturers in information systems or computer science departments can utilize the framework as an instructional tool to enhance teaching and learning processes, particularly in demonstrating how to enhance and compare the performance

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

of machine learning frameworks. Similarly, students can utilize the modified and improved framework as a learning resource in the classroom, engaging in data-based predictive exercises.

The research aims to inspire interdisciplinary research initiatives and real-world applications within communities. It serves as a reference for future studies seeking to evaluate the predictive performance of machine learning frameworks in various contexts.

For stakeholders, particularly government entities at different levels and focused on different fields, the study outcomes can aid in addressing malnutrition-related issues, particularly in monitoring aspects. The framework enables quick identification of areas requiring intervention and facilitates the formulation and implementation of appropriate policies and mitigation plans based on estimated scenarios.

The research emphasizes the importance of the feature selection algorithm and evaluation metrics in improving the nutritional status prediction framework. The dataset preparation process is identified as a crucial step in achieving optimal machine learning algorithm performance. Previous studies (Ferdowsy et al., 2021; Rahman et al., 2021; Shahriar et al., 2019) support this claim, highlighting the significance of appropriately adapting datasets to algorithm characteristics and intended purposes to avoid suboptimal results and misinterpretations. The study contributions are derived from the research objectives in Table 1.1.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

Table 1.1

Relation between research objectives and study contributions

	Research Objectives	Study Contributions
RO1	Assess the suitability of the evaluated features of the Banjarmasin dataset for predicting the nutritional status of children.	• Nominate a widely used open dataset for predicting nutritional status. • Identify relevant features that enable accurate nutritional status prediction. • Capture probabilistic interrelations among selected features rather than considering them independent to improve accuracy. • Propose a new dataset of Banjarmasin with specific features to predict the nutritional status of children.
RO2	Identify and enhance the algorithms that provide better prediction performance of children's nutritional status.	• Identify the suitable and frequently used machine learning algorithms in nutritional status prediction • Nominate several machine learning algorithms that claim the best performance in nutritional status prediction. • Propose the algorithm that achieves the best performance algorithms in nutritional status prediction using the Banjarmasin dataset
RO3	Enhance the framework by using class imbalance to achieve the best predictive performance and generate a comprehensive mapping of public health conditions based on the nutritional status of children.	• Propose an enhanced framework that applies class imbalance to predict the nutritional status of children • Develop and employ the Banjarmasin dataset for children's nutritional status prediction framework • Propose the Banjarmasin map of public health condition estimates based on children's nutritional status

05-4506832 Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah ptbupsi

1.9 Thesis Organization

This thesis is divided into five distinct chapters. The first chapter introduces the nutritional status and the background of the study, the problem statement, the research objectives, the research questions, the scope, and the significance of the study. The second chapter discusses the review of relevant literature. The third chapter details the research methodology utilized in this study. The fourth chapter discusses data analyses,

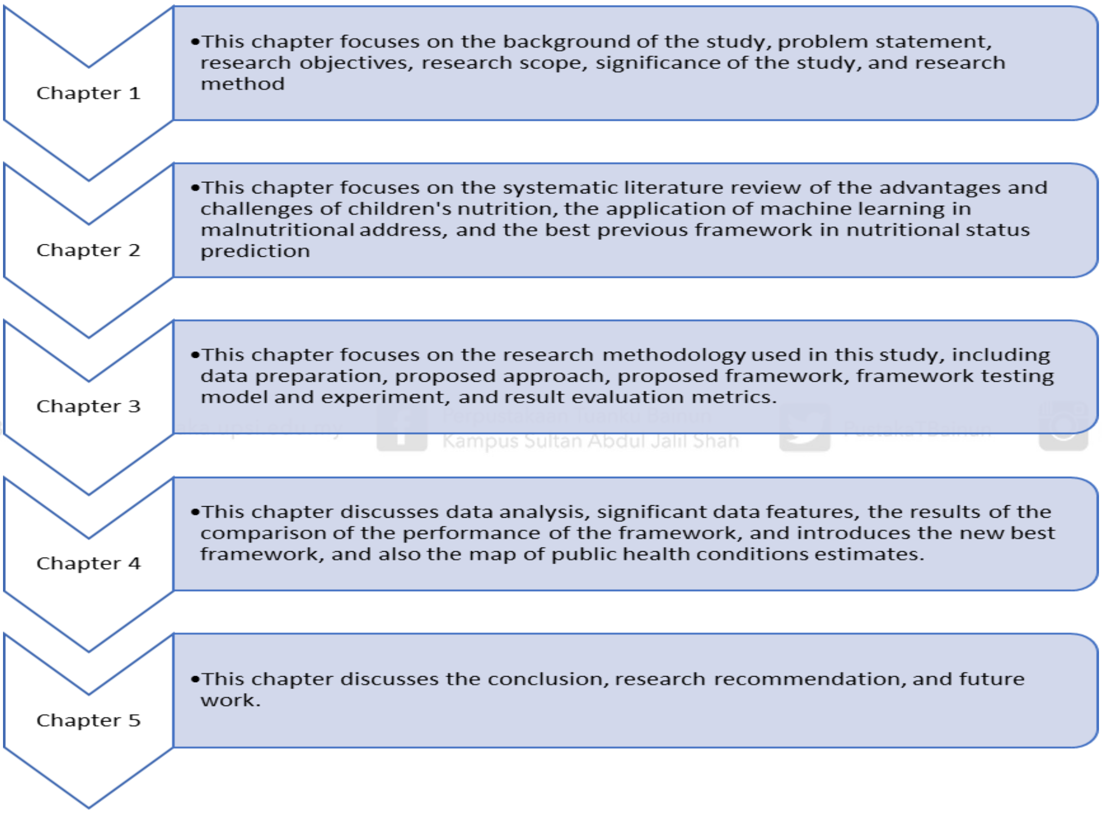
05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

and the result of finding the best framework while the fifth chapter details the discussions of the study conclusion, the recommendation, limitations of the study, and further work. The outline of the thesis is shown in Figure 1.3

Figure 1.3

The outline of the thesis



05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi