SCHOOL OF COMPUTING
Faculty of Engineering

Project Proposal Form MCST1043
Sem: 2 Session: 2024/25

SECTION A: Project Information.

Program Name:	**Masters of Science (Data Science)**
Subject Name:	**Project 1 (MCST1043)**
Student Name:	Lock Chun Hern
Metric Number:	MCS241047
Student Email & Phone:	lock@graduate.utm.my 0167965911
Project Title:	Prediction of Health Expenditure in Malaysia using Machine Learning
Supervisor 1:	
Supervisor 2 / Industry Advisor(if any):	

SECTION B: Project Proposal

Introduction:

Health expenditure is defined as all the money spent on health goods and services, including preventative measures, promotion and provision of health services, nutrition, pharmaceuticals, and research. It includes funding from various sources such as government, household out-of-pocket expenditures, private health insurance, and other non-government organisations.

Recently, machine learning approaches have been used to deepen understanding and provide insight into healthcare spending. Researchers from Jordan predicted total healthcare expenditure for their country by using two neural network strategies: Adaptive Neuro-Fuzzy Inference System and Hybrid Neural Fuzzy Inference System (Saleh et.al., 2023). Artificial intelligence techniques to estimate total health expenditure per capita in Turkey are shown by Ceylan and Atalan (2021), which compares five different models and proposes an effective hybrid model of genetic algorithm-based feature selection and random forest for the prediction. In this study, we intend to apply machine learning model to the health expenditure data from Malaysia and World Health Organisation to shed light on the future health expenditure in Malaysia.

Problem Background:

In Malaysia, total health expenditure has been increasing from 2011 to 2022, from RM 36.9 billion to RM 78.9 billion, and as % GDP (Gross Domestic Product) from 3.94% to 4.41% (Malaysia National Health Accounts, 2023). Malaysia government has allocated RM 45.3 billion in Budget 2025 for the Ministry of Health for spending on healthcare, which is the second highest after education (Ministry of Finance, 2025). In addition, the rise in health spending by 38 countries

participating in Organisation for Economic Co-operation and Development (OECD) is predicted to peak at 11.8% GDP in 2040. By that time, the growth in healthcare expenses from public sources is estimated to be twice government revenues. (OECD, 2024). With the medication price inflation and the aging populations in Malaysia, the growing pressure on healthcare budgets remains a challenge. However, limited academic research has been done to predict future health expenditures in Malaysia. Therefore, there is a need for an accurate predictive model to be developed to aid in the planning of future healthcare budgets

Problem Statement:

This project intends to find out what the health expenditure of Malaysia will be from the year 2026 to the year 2040 using machine learning. Health expenditure can be further classified into Total Health Expenditure (TEH), Current Health Expenditure (CHE), which excludes health-related expenditure (e.g. personnel training, research and development), General Government Health Expenditure (GGHE), and household Out-Of-Pocket health expenditure. (OOP).
This study will also investigate what are the determinants of health expenditure (e.g. life expectancy, population aged 65 years old and above) that can be selected as parameters to apply to the machine learning model.

Aim of the Project:

This project aims to predict health expenditure in Malaysia using machine learning method to aid in health financing and planning.

Objectives of the Project:

1. To analyse and select the determinants of health expenditure to use as input parameters
2. To apply machine learning method using selected input parameters to predict the health expenditure in Malaysia
3. To evaluate and improve the performance metrics of the machine learning model

Scopes of the Project:

1. The source of data will be from Ministry of Health Malaysia, Department of Statistics Malaysia and World Health Organisation (WHO).
2. The data collected will only involve health expenditure and population data related to health economics. No individual data that reveals an individual's medical and medication history will be used.
3. The time frame for data used will be between 2000 to 2022, which provides relevant data to the study and is able to reveal insights for the recovery from the most recent pandemic (COVID-19).

Expected Contribution of the Project:

This project can provide insights for policymakers in the country in planning health expenditures and allocation of budgets for other expenses. This helps in ensuring the long-term sustainability of health financing. The findings of this project are also expected to provide insights for other countries with similar healthcare system or income levels. Together, these can contribute to better health outcomes for the patients and people in Malaysia.

Project Requirements:

Software: Python, Jupyter Notebook

Hardware: Laptop(AMD Ryzen 5 4600H with Radeon Graphics 3.00 GHz, 16GB RAM)

Technology/Technique/ Methodology/Algorithm: Machine Learning

Type of Project (Focusing on Data Science):

[ ] Data Preparation and Modeling

[ ] Data Analysis and Visualization

[ ] Business Intelligence and Analytics

[/] Machine Learning and Prediction

[ ] Data Science Application in Business Domain

Status of Project:

[/] New

[ ] Continued

If continued, what is the previous title?

SECTION C: Declaration

I declare that this project is proposed by:

[/] Myself

[ ] Supervisor/Industry Advisor ( )

Student Name: Lock Chun Hern

Signature **Date** 7/4/2025

SECTION D: Supervisor Acknowledgement

The Supervisor(s) shall complete this section.

I/We agree to become the supervisor(s) for this student under aforesaid proposed title.

Name of Supervisor 1:

Signature **Date**

Name of Supervisor 2 (if any):

Signature **Date**

SECTION E: Evaluation Panel Approval

The Evaluator(s) shall complete this section.

Result:

[ ] FULL APPROVAL
[ ] CONDITIONAL APPROVAL (Major)*
[ ] CONDITIONAL APPROVAL (Minor)
[ ] FAIL*

***** Student has to submit new proposal form considering the evaluators' comments.

Comments:

Name of Evaluator 1: .....

Signature **Date**

Name of Evaluator 2: .....

Signature **Date**