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:	**Project1 (MCST1043)**
Student Name:	Muhammad Haziq bin Mohamad
Metric Number:	MCS241036
Student Email & Phone:	muhammadhaziqmohamad@graduate.utm.my & 601110667854
Project Title:	Identifying Patterns in Salary Disparities Among Malaysians Using Machine Learning
Supervisor 1:	
Supervisor 2 / Industry Advisor(if any):	

SECTION B: Project Proposal

Introduction:

Salary disparity is one of the major economic issues around the world. In Malaysia, the salary disparity also has raised concern among the workers. This issue will become disadvantage to the country economic health and growth. "High-income inequality could be detrimental to the economic growth of a country" (Masseran, N., & Ibrahim, K. 2019). According to Yusof (2019), high-income inequality can lead to negative impact for the country's economic growth as the hard access to the education and economic opportunities. For instance, the high-income inequality only focused on the small portion of population and will not benefit the others. It will become crucial as the income is unevenly distributed among marginalized groups such as the poor or minorities. Therefore, the benefits such as the quality education, healthcare and job opportunities will become hard to access. The chain will continue as without proper education and economic opportunities, these groups will not contribute fully to the country's economy. This is because, they will remain stuck in the low paying jobs or be underemployed. Thus, Malaysia will be lost out the potential talent and productivity that can slow down the overall economic progress. Therefore, this project aims to identifying patterns in salary disparity among Malaysian workers by utilizing machine learning approach such as classifying and regression. The insights obtain from this project will better understanding the key features that contribute to the salary disparity that will benefit the society.

Problem Background:

The unfairness of salary disparity among workers has become the main issue in this country. This can be seen, despite Malaysia's continued attempts to establish high-income economy. The studies have indicated that the income inequality can lead to the stunted growth of the country's economy. "This situation can lead to stunted economic growth, political instability, and social unrest" (Anyanwu et al., 2016; Anyanwu, 2016). Furthermore, salary disparity does not just affect the country economy but also will negatively affects the society happiness. For instance, Ferrer-i-Carbonell and Ramos (2014) found that the income inequality is negatively associated with the happiness of the society in Latin America and Western countries. For example, a big gap between the rich and poor will lead to feelings of injustice that can affect the happiness of the society. This cause to the lower life satisfaction that can affect the government's goal to increase the standard of living in this country. According to the Department of Statistics Malaysia (DOSM), there are several factors that contributed to the income inequality such as education level, ethnicity, profession type, and geographical location. Therefore, this project will use machine learning approach that will help to identify the patterns. The insights obtain will provide the key features that contribute to the salary disparity the most.

Problem Statement:

Government of Malaysia has set the pivotal goals which are to achieve high-income status and improving living standards. However, these objectives can be not achieved by the persistent issues of salary disparity. Income inequality will hamper the economic growth and progress. Furthermore, it also not only affects the wealth distribution but can negatively impact on various societal dimensions such as healthcare, education and economic opportunities. Moreover, there is lack of discovery on how the numerous demographics affects the salary disparity. Besides that, the traditional statistical method often lacks in capturing the non-linear and complicated big set of data. According to Coursera (2025), the machine learning provides benefits such as enormous and encompass methods that can help the companies to identify the patterns within the big data far more quickly than the conventional statistical techniques. Therefore, this project aims to identify the pattern of salary disparity among Malaysian worker and to know the key features that contribute to the income inequality in Malaysia that will

Aim of the Project:

This project aims to identify patterns in salary disparities among Malaysians and to understand the key features that contribute to income inequality through the application of machine learning methods.

Objectives of the Project:

1. To identify the key features affecting salary disparity among employees in Malaysia.
2. To utilize machine learning techniques for examining and predict trends in income disparity.
3. To visualize and quantify salary disparities using data analytics tools

Scopes of the Project:

1. The data is gathered from public resources such as Jobstreet, Glassdoor, DOSM and may include self-administered surveys to fill in any gaps in existing datasets.
2. This project use python as the programming language.
3. The analysis takes into account elements like gender, education background, industry type, years of experience, and job location.

Expected Contribution of the Project:

The expected contribution of this project is to benefit the society by improving the understanding of how the numerous demographics and professional features affect the income differences. Therefore, the important key features and patterns identified through analysis will then be visualized for easier interpretation. The project will give insight into the nature and extent of salary disparity in Malaysia. The results will aim to reduce income inequality and lead to a fairer labor market in Malaysia.

Project Requirements:

Software:	Python
Hardware:	Intel i5, 16GB Ram
Technology/Technique/ Methodology/Algorithm:	Machine learning: regression, classification, and clustering

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: MUHAMMAD HAZIQ BIN MOHAMAD

Haziq — **Signature**

7/4/2025 — **Date**

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**