SCHOOL OF COMPUTING
Faculty of Engineering

Project Proposal Form MCSD 6215

Sem:1 Session:2024/2025

SECTION A: Project Information.

Program Name:	**Masters of Science (Data Science)**
Subject Name:	**Project 1 (MCSD 6215)**
Student Name:	ZHU QIAN
Metric Number:	MCS241017
Student Email & Phone:	zhuqian@graduate.utm.my +60-1123132056
Project Title:	Topic-Based Analysis of Social Media Posts Using RNN and LSTM
Supervisor 1:	Dr. Mohd Shahizan Othman
Supervisor 2 / Industry Advisor(if any):	

SECTION B: Project Proposal

Introduction:

Social media has emerged as the most important platform for public opinion. Accompanied by the improvement of mobile internet, traditional mass media has lost much of its influence over the public. Under more relaxed audit conditions, social media reflects the public's genuine perspectives on various topics. It is essential to discern valuable information from posts that are presented of text and emojis. In the following sections, the deatils will explain how to filter posts by topics to identify related content. The project will annalyze and interpret genuine reactions of the posts. The analysis results will be summarized to draw conclusions.

Problem Background:

Unlike traditional media, social media serves as a more open platform. However, the rise of fake users

and internet trolls has become a significant issue that is threatening data quality. The platform is often inundated with irrelevant data, while the internet is filled with non-compliant, emotional, repetitive, and meaningless content. Furthermore, some users distort their own views of the perspectives of others. Meanwhile, web surveys hardly verify whether respondents are real individuals, and the results are easy to teamper. In contrast, social media serves as a more reliable and genuine source of data. Real pulic reactions to various topics provide insigths that can help enterprises and governments make more informed decisions in response to real-world changes and challenges. At the same time, these organizations can decrease the costs with traditional questionnaires, which benefits their financial standing.

The weaknesses of the current working model highlight the need for new data sources and a more efficient, sophisticated approach. Utilizing data science and machine learning techniques to analyze topic-based social media posts will help enterprises and governments identify valuable data to address these challenges.

Problem Statement:

Social media posts are presented in various data formats such as text and emojis often mixed with irrelevant content and influenced by emotions. Traditional survey methods frequently produce false or unreliable results, which often deviate significantly from the actual situation.

Aim of the Project:

Collecting and filtering social media posts by topic, and using meachine learning techniques to analyze them, yield valuable and authentic data about the topic.

Objectives of the Project:

1. To identify significant relationships between the content of posts and the topic.
2. To build and develop analytical models that capture the topic inclination of posts.
3. To measure public reactions to the topic by summarizing the analysis results.

Scopes of the Project:

1. Creating and implementing an analycal model for post content. The model can identify valuable, quality data from the conllected posts. The data sources are mostly concentrate on soial media.
2. Testing and validating the model to ensure its efficiency and effectiveness in a controlled setting.

Expected Contribution of the Project:

1. Replace the outdated and unreliable traditional survey system to enhance organizational efficiency.
2. Make better decisions based on genuine data to benefit real-world outcomes.

Project Requirements:

Software:	PyTorch, Python, CUDA
Hardware:	CPU, GPU

Technology/Technique/
Methodology/Algorithm: Recurrent Neural Networks, Long Short-Term Memory

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:

Signature **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** .....