Capstone Project Concept Note and Implementation Plan

Project Title: Mental health detection from social media posts

Team Members

1. Aung Thu Phyo
2. Su Mon Hlaing
3. Moe Thet Hmue
4. Nang Mon Kham
5. Phyo Ko Ko

Concept Note

1. Project Overview

- Our capstone project, "Mental Health Detection from Social Media Posts," is a system designed to analyze social media text to identify potential signs of mental health issues such as stress, anxiety, or depression. This project directly supports **Sustainable Development Goal 3 (SDG 3): Good Health and Well-being** by creating a tool for early detection and mental health awareness. It also aligns with **SDG 9 (Industry, Innovation and Infrastructure)** by applying an innovative, AI-driven approach to a public health challenge.
- The problem we aim to address is that many individuals, including in Myanmar, express emotional distress online, but these signs often go unnoticed. A system that can automatically detect and classify posts showing signs of depression or stress can provide significant positive impact. The potential solution would enable early awareness, help NGOs or health organizations identify at-risk individuals, and promote timely intervention, ultimately contributing to the promotion of mental well-being for all.

2. Objectives

- The specific objective of this project is to develop a system capable of automatically detecting and classifying social media posts that show signs of depression or stress.
- This system aims to achieve a significant positive impact by addressing the problem of unnoticed emotional distress online. It will contribute by:
    - Enabling early awareness of potential mental health issues.
    - Assisting NGOs and health organizations in identifying at-risk individuals.
    - Promoting timely intervention, which ultimately contributes to the broader promotion of mental well-being for all.

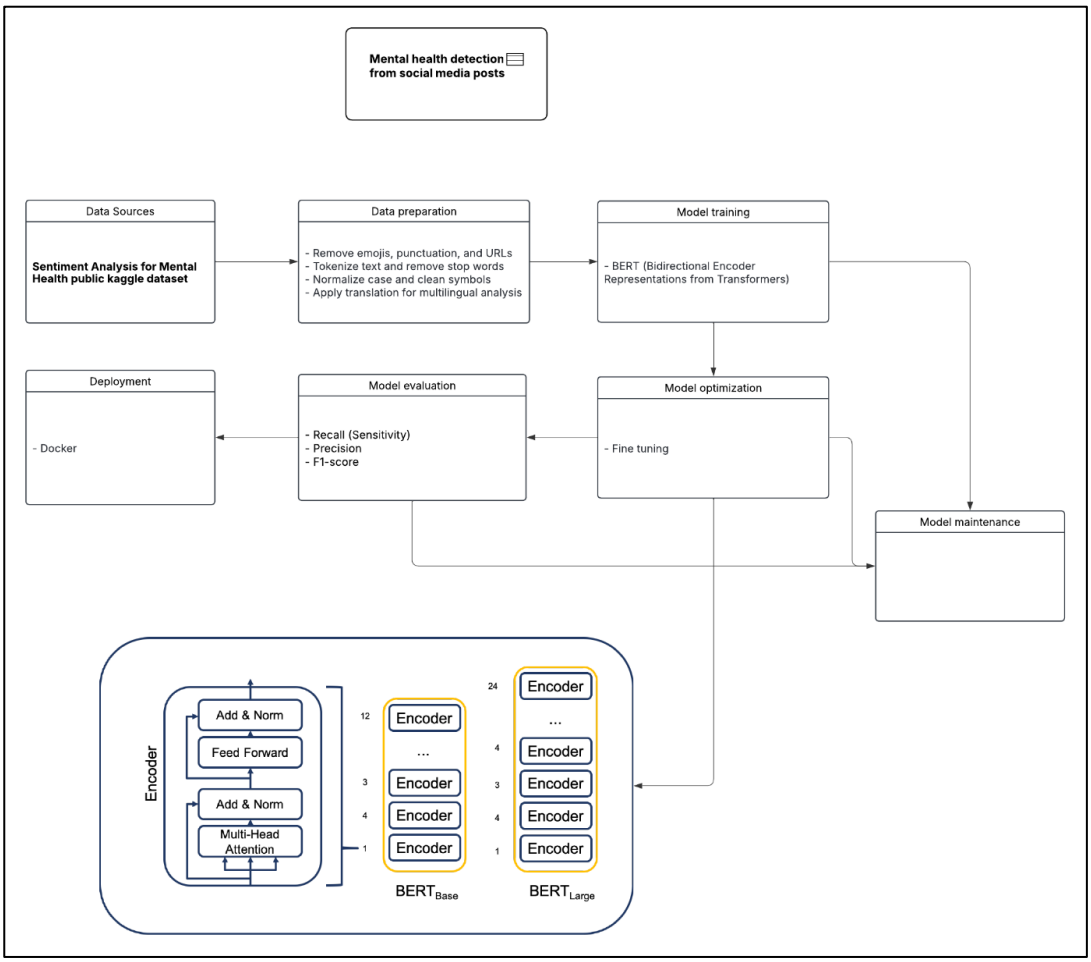
3. Background

- Mental health has been rightfully recognized as a critical component of global well-being, as emphasized by Sustainable Development Goal 3. In parallel, platforms like Facebook and X have become primary forums for personal expression, where billions of users, including many in Myanmar, share their daily emotions and struggles. This creates a massive, real-time data stream where clear indicators of stress, anxiety, and depression are often lost. While traditional support from NGOs and reactive platform tools like post-reporting are essential, they are not sufficient. These methods cannot manually sift through the millions of daily posts and often only address a crisis after it has become apparent, leaving many who are quietly struggling unidentified.
- This is precisely where a machine learning (ML) approach becomes necessary. The sheer volume of text data on social media makes manual analysis impossible. A system built on Natural Language Processing (NLP) can analyze text at a massive scale, identifying the subtle linguistic patterns and semantic cues correlated with mental distress. Such a system is not intended to diagnose individuals but to serve as a large-scale, automated early-warning tool. It can help bridge the gap by proactively identifying individuals who may be at risk, enabling social media and health organizations to offer support more effectively and contributing to a more innovative, responsive public health infrastructure.

4. Methodology

- Preprocessing Steps: Remove emojis, punctuation, and URLs
    - Tokenize text and remove stop words
    - Normalize case and clean symbols
    - Apply translation for multilingual analysis
- Data training using
    - BERT (Bidirectional Encoder Representations from Transformers)
- Model evaluation
    - Recall (Sensitivity)
    - Precision
    - F1-score

5. Architecture Design Diagram



6. Data Sources

- This project will utilize the "**Sentiment Analysis for Mental Health public dataset**" comprehensive, structured collection of textual statements provided in CSV format in Kaggle. This dataset was curated by aggregating multiple publicly available mental health datasets and applying extensive cleaning procedures, making it ideal for Natural Language Processing. For the initial model training phase, we will use the primary English-language data. Subsequently, to ensure the system's scalability and adaptability for Myanmar's multilingual context, we will apply a translation algorithm (such as the Google Translate API or MarianMT) to expand the dataset to Burmese. The dataset provides statements tagged with one of seven mental health statuses—Normal, Depression, Suicidal, Anxiety, Stress, Bi-Polar, and Personality Disorder—which will form the basis for our classification model.

7. Literature Review

- In recent years, the application of natural language processing (NLP) and machine learning to detect mental health signals from text has grown rapidly. A comprehensive

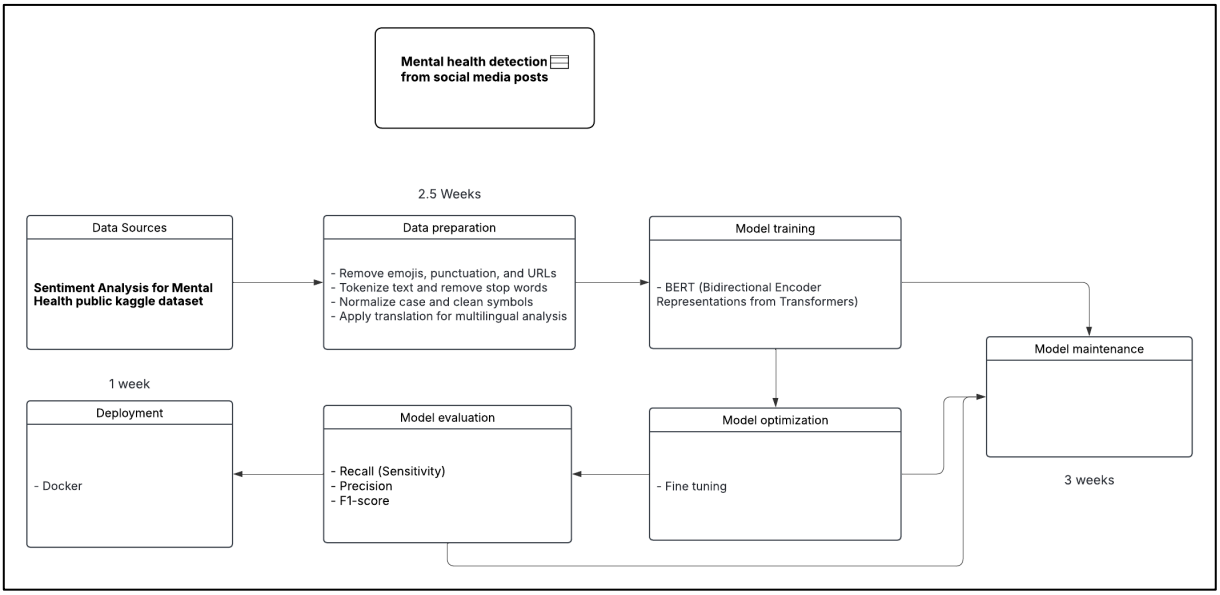
narrative review **[1]** identifies key trends across a decade of research, noting a significant shift from traditional feature-based classifiers toward deep-learning architectures and an increasing focus on varied data sources like social-media posts. This trend is demonstrated empirically by **[2],** who used a deep learning model on user content from online communities (such as Reddit) to achieve high accuracy in building classification models for disorders like depression, anxiety, and bipolar disorder. More specifically, in the critical area of suicide risk, another review **[3]** compiles evidence on machine-learning approaches for detecting suicidal ideation, highlighting both specific algorithmic methods (like neural networks) and the crucial ethical imperatives of deploying AI in mental-health contexts.

Implementation Plan

1. Technology Stack

Component	Tools / Libraries
Programming Language	Python
Data Handling	pandas, NumPy
Baseline Models	scikit-learn, BERT
Deep Learning	Hugging Face Transformers
Text Processing	Hugging Face Tokenizers
Evaluation	scikit-learn
Visualization	Matplotlib, Seaborn
Deployment	Docker, Streamlit / Flask

2. Timeline



3. Milestones

- Completed
    - Data collection

    - Choose the right model, data handling
- Pending
    - Model evaluation
    - Optimizing to get better accuracy
- Target
    - Social media companies and health care can detect the early signs of mental health and supports mental health by taking care of them in person or by showing the contents to make them relieve.

4. Challenges and Mitigations

- Data quality
    - Myanmar → English mistranslation can reduce the quality.
- Model performance
    - If user has 100 posts, only 5 detected depression and 95 shows as happy stage.
    - The model can decide that the user is happy.
- Technical constraints.
    - Resource requirements (GPU) for the training and model maintenance.

5. Ethical Considerations

- Our project carries significant ethical responsibilities. The first is data privacy. We are analyzing sensitive personal text, and while we will use an anonymized, public dataset to protect individuals, we acknowledge users did not explicitly consent to this analysis.
- The second is algorithmic bias. Our primary dataset is English-language and Western-centric, which risks creating a model that is unfair or inaccurate for the cultural and linguistic nuances of our community in Myanmar. Our plan to adapt the model for Burmese is a critical step to mitigate this.

6. References

1. Zhang, T., Schoene, A.M., Ji, S. & Ananiadou, S. (2022) 'Natural language processing applied to mental illness detection: a narrative review', *npj Digital Medicine*, 5, p. 46. doi: 10.1038/s41746-022-00589-7. https://www.nature.com/articles/s41746-022-00589-7

2. Kim, J., Lee, J., Park, E., Han, J. & Sungkyunkwan University (2020) 'A deep learning model for detecting mental illness from user content on social media', *Scientific Reports*, 10(1),

p. 11846. doi: 10.1038/s41598-020-68764-y. https://www.nature.com/articles/s41598-020-68764-y

3. Sheng, Z. (n.d.) 'Suicidal ideation detection on social media using machine learning: A review'. [online] Available at: https://arxiv.org/abs/2201.10515

4. Sarkar, S. (n.d.) *Sentiment analysis for mental health* [dataset]. Kaggle. Available at: https://www.kaggle.com/datasets/suchintikasarkar/sentiment-analysis-for-mental-health/data