Capstone Project Concept Note and Implementation Plan

Project Title: **MindCare**

Team Members:
1. Pann Su Khine
2. Pau En Piang
3. Shwe Sin Moe
4. Wanna Myo Min
5. Ye Bhone Lin

Concept Note

1. Project Overview

In Myanmar, mental health challenges are growing among youths and middle-aged workers due to economic hardship, political instability, and social pressure. These situations cause stress, fear, and hopelessness. Access to professional mental health services is limited by stigma, a lack of awareness, internet shutdowns, and a shortage of qualified psychologists.

This project is relevant to SDG 3 (Good Health and Well-being) by promoting mental health awareness and early detection of mental health risks. It also addresses SDG 10 (Reduced Inequalities) by providing accessible and affordable digital mental health support.

The project's solution is "MindCare," an AI-powered chatbot. It will function as a Mental Health Risk Indicator and a safe, local, digital support tool that understands Burmese users' voices, faces, and words.

2. Objectives

The primary goal is to develop an AI-powered chatbot that serves as a Mental Health Risk Indicator and Personalized Well-being Roadmap for users in Myanmar.

Specific objectives include:

- Using AI and Machine Learning (ML) to analyze user responses and detect early signs of mental distress.
- Supporting people's mental well-being in the Burmese language.
- Creating a system that understands a user's emotions from their voice, face, and text.
- Gently guiding users toward awareness, relaxation, or professional help when needed.

3. Background

People in Myanmar are facing significant challenges from political crisis, civil war, and climate disasters, which has led to widespread stress and hopelessness. It is difficult for people to seek help due to factors like stigma, limited education, internet shutdowns, and a shortage of professional support.

Existing solutions include AI chatbots like Woebot and Wysa, which have demonstrated that AI can support mental health. Furthermore, the WHO Mental Health Gap Action Programme (mhGAP) advocates for using digital tools and AI-assisted assessments to scale up services in low-income countries.

A machine learning approach is necessary because no similar system currently exists for the Burmese language. ML is required to develop a tool that can understand Burmese text and speech, as well as detect emotions from voice tone, facial expressions, and text sentiment. This will enable early detection and provide culturally relevant support.

4. Methodology

The proposed system utilizes a combination of Instruction Fine-Tuning and Retrieval-Augmented Generation (RAG) to enhance the language model's ability to follow complex user queries and provide contextually accurate responses.

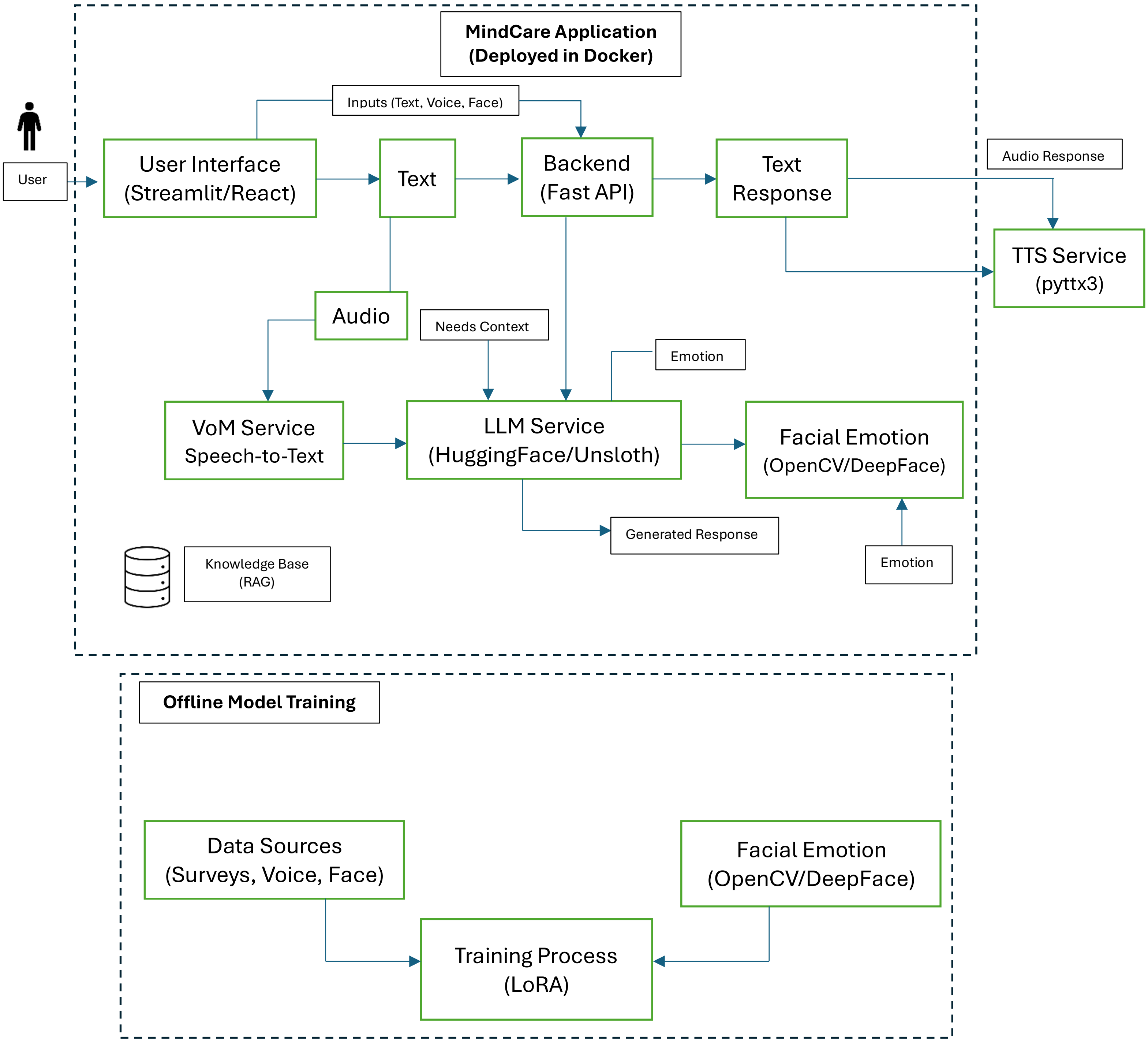
Instruction Fine-Tuning will be applied to adapt a pre-trained Large Language Model (LLM) to follow domain-specific prompts and respond in a structured, human-like manner. This process involves fine-tuning the base model on a dataset containing instruction–response pairs, enabling the model to better understand user intentions and generate more relevant outputs.

In addition, the RAG (Retrieval-Augmented Generation) technique will be integrated to improve the model's factual accuracy. RAG retrieves relevant external documents or database entries in real time and combines them with the model's generative capabilities, ensuring that the responses are both knowledge-grounded and up to date.

For user interaction, the system will support multilingual voice input through the Speech Recognition library, which converts spoken language into text. This allows users to communicate naturally using a microphone. To enhance accessibility, pyttsx3 will be employed for text-to-speech (TTS) functionality, enabling the system to provide audible responses.

The overall pipeline integrates these components into a cohesive system that supports multilingual voice-based conversation with contextual retrieval and fine-tuned natural language understanding.

5. Architecture Design Diagram



6. Data Sources

The project will utilize primary data collected via online Google Form surveys and secondary data from public datasets such as WHO and Kaggle. The data collected consists of survey responses (CSV), voice samples (WAV/MP3) from local Burmese speakers, and facial images/videos collected with consent. This data will be used to understand emotional trends, fine-tune text models, and train emotion detection models for voice and facial expressions. While an initial 20 responses have been collected, the target dataset size is 5,000–10,000 records. Required preprocessing steps include data cleaning (removing duplicates and missing values), language translation from Myanmar to English, and feature extraction for AI/ML analysis.

7. Literature Review

Existing research shows that digital mental health tools like Woebot and Wysa can effectively support users, and WHO reports confirm the need for such tools in crisis-afflicted, low-income countries. Studies also demonstrate that AI can accurately recognize emotions by combining text sentiment, voice tone, and facial expressions, and can incorporate standard screening tools like PHQ-9 and GAD-7. For low-resource languages like Burmese, models such as mBERT and XLM-R can be adapted. This project builds upon this work by using transfer learning and LoRA fine-tuning to create what aims to be the first local, safe, and emotional support system that combines these features for the Burmese language.

Implementation Plan

1. Technology Stack

- Python programming languages
- FastAPI
- LLM Models
- Hugging Face Transformers
- PEFT (Parameter-Efficient Fine-Tuning)
- Instruction Fine-Tuning
- LoRa (Low-Rank Adaptation)
- QLoRa
- T4 GPU
- OpenCV/DeepFace
- Docker
- Streamlit/React
- Unsloth LLM
- PyTorch
- bitsandbytes or DeepSpeed

2. Timeline

Phase	Task	Start Date	End Date
Phase1: Project Initiation & Research	Project kick-off and brainstorming	1-Sep-2025	7-Sep-2025
	Literature and Data Research (Initial)	8-Sep-2025	21-Sep-2025
	Define project scope and objectives	22-Sep-2025	25-Sep-2025
	Technology review (Tools & Models)	5-Oct-2025	15-Oct-2025
Phase2: Data Collection & Preprocessing	Initial data collection (Survey setup)	26-Sep-2025	30-Sep-2025
	Data collection (Survey distribution and monitoring)	1-Oct-2025	20-Oct-2025
	Data preprocessing (Cleaning initial 20 responses)	21-Oct-2025	25-Oct-2025
	Data collection (Voice & Facial samples)	1-Nov-2025	20-Nov-2025
	Data preprocessing (Voice & Text translation)	21-Nov-2025	30-Nov-2025
Phase3: Model Development, Training & Evaluation	Initial model setup (Environment setup)	16-Oct-2025	25-Oct-2025
	Prototype1: Basic LLM fine-tuning (Text)	26-Oct-2025	31-Oct-2025
	Prototype1: Evaluation & Feedback	1-Nov-2025	5-Nov-2025
	Model development (Emotion detection-Voice)	10-Nov-2025	25-Nov-2025

	Model development (Emotion detection-Face)	15-Nov-2025	30-Nov-2025
	Integration – Phase 1 (Text + Voice Models)	26-Nov-2025	30-Nov-2025
	Prototype2: Multi-modal Chatbox (Text + Voice + Face)	1-Dec-2025	10-Dec-2025
	Testing and evaluation (Prototpype2)	11-Dec-2025	15-Dec-2025
Phase4: Deployment & Finalization	Deployment (Setup FastAPI Backend)	5-Dec-2025	15-Dec-2025
	Deployment (Built Streamlit/React Frontend)	5-Dec-2025	18-Dec-2025
	Final system integration & Testing	16-Dec-2025	20-Dec-2025
	Prepare final report & presentation	18-Dec-2025	23-Dec-2025
	Final project submission	24-Dec-2025	24-Dec-2025
	Final presentation	25-Dec-2025	25-Dec-2025

3. Milestones

- Fine-tune a small bilingual transformer on the collected Burmese data.
- Build an emotion detection pipeline that utilizes text, voice, and facial inputs.
- Design a simple interface to provide emotional feedback to the user (e.g., text feedback, emoji, or a color scale).
- Gradually integrate all components into the final MindCare AI application.

4. Challenges and Mitigations

Challenge	Mitigation Plan
Data Availability: Burmese datasets for training are small.	Use transfer learning and LoRA fine-tuning, which allow large models to be adapted using small, locally collected datasets (like surveys and voice samples).
Technical Constraints: Training large models requires significant computing power.	Employ LoRA fine-tuning, which makes training possible even on a laptop or free GPU. Use optimization tools like bitsandbytes or DeepSpeed for efficient GPU memory utilization. The project will also use open-source tools that are free and accessible.

5. Ethical Considerations

- Data Privacy: Mental health data must be kept private.
- Mitigation: All data will be collected only with user consent and will be anonymized. Data will be stored safely.
- User Safety and Impact: Users must feel safe talking about their feelings without fear of exposure or judgment.
- Mitigation: The chatbot will focus on empathy and safe conversation, not on providing medical advice. The model will be fine-tuned for culturally sensitive communication. Clear crisis-handling plans will be implemented, such as showing hotline numbers if serious distress is detected.

6. References

Woebot Health. (n.d.). *Woebot: Your mental health ally.* Retrieved November 3, 2025, from https://woebothealth.com

Wysa Ltd. (n.d.). *Wysa: Your AI mental health support.* Retrieved November 3, 2025, from https://www.wysa.io

World Health Organization. (n.d.). *Mental health: Strengthening our response.* World Health Organization. Retrieved November 3, 2025, from https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response

World Health Organization. (2016). *mhGAP Intervention Guide for mental, neurological and substance use disorders in non-specialized health settings (Version 2.0).* World Health Organization. https://www.who.int/publications/i/item/mhgap-intervention-guide—version-2.0

Kaggle. (n.d.). *Kaggle: Your Machine Learning and Data Science Community.* Retrieved November 3, 2025, from https://www.kaggle.com

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). *The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine, 16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). *A brief measure for assessing generalized anxiety disorder: The GAD-7. Archives of Internal Medicine, 166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019*, 4171–4186. https://arxiv.org/abs/1810.04805

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., … & Stoyanov, V. (2020). *Unsupervised cross-lingual representation learning at scale. Proceedings of ACL 2020*, 8440–8451. https://arxiv.org/abs/1911.02116