Statement of Purpose (SoP)
DSL501: Machine Learning Project

Team Name: Ashutosh Kumar Jha

October 9, 2025

1. Project Details

- **Project Title:** Self-State Identification and Classification in Mental Health Data using Agentic AI Systems
- **Code Repo Link (if available):** Not Available
- **If Own Idea:** Yes
  This project proposes a novel approach to self-state identification and classification (adaptive vs. maladaptive) in mental health data using agentic AI architectures rather than conventional LLM-only systems.

2. Problem Statement

- The analysis of self-states and their classification is a new and under-explored field, with little technical literature, no stable open-source code implementations, and no mature architectural patterns available. [2]
- The motivation of this project is to design agentic AI architectures and pipelines for identifying and classifying self-states in Reddit mental health posts, going beyond conventional LLM usage.
- This problem is important in the ML domain because self-state prediction is an under-attended challenge, absent in high-tier (A*) conferences and standard benchmarks. Solving it contributes to the computational domain of mental health linguistics and responsible AI for well-being. [4]
- Sub-problem focus is kept on addressing bias and performance degradation when filtering documents, and detecting deeply embedded states in sentences via hybridization of reasoning and action systems in agentic environments. [3]

3. Methodology

- **Proposed ML Models/Architectures:** Open-source LLMs such as 4-bit quantized Gemma 2 9B [2], modular agentic AI architecture, Retrieval-Augmented Generation (RAG), and reasoning-based agents.
- **Key Techniques, Algorithms, Frameworks:** Agentic AI, LangChain, LangGraph, State Graphs, ReAct Agents, RAG pipelines.
- **Difference from Existing Methods:** Current methods rely mainly on direct LLM classification with sentence splitting and chunking. Our approach integrates reasoning-enabled agents, modular pipelines, and iterative decision-making, making the system more robust to hidden adaptive states and deeper semantic granularity.

- **Components to be Implemented:** End-to-end agentic pipeline with data loaders, preprocessing modules, context filters, annotators, reasoning-action agents, and final classification generators.

4. Dataset Details

- **Name and Source:** Reddit mental health forums (self-collected dataset) as a proxy to the CLPsych 2025 Shared Task dataset.
- **Size and Structure:** Target size ∼300–400 posts across ∼30 user timelines, consistent with the CLPsych 2025 baseline dataset structure (annotated with adaptive/maladaptive spans and summaries).
- **Preprocessing/Augmentation:** Sentence splitting (spaCy), paraphrasing for privacy, importance filtering, and context windowing (2–3 sentence chunks).
- **Justification:** Suitable because it closely resembles the shared task dataset while ensuring access and reproducibility; annotations reflect real-world expressions of adaptive/maladaptive states in online discourse.

5. Required Resources

- **Hardware:** GPU with ≥16GB memory (e.g., NVIDIA V100/A100 or whatever is available) for inference of quantized LLMs; high-RAM CPU for preprocessing tasks.
- **Software:** Python (obviously), PyTorch, Hugging Face Transformers, spaCy, BERTScore, LangChain, LangGraph.
- **Additional Tools:** Weights & Biases for experiment tracking, APIs for Reddit data collection, open-source LLM checkpoints (Gemma 2 9B, as used in the original paper).

6. Novelty of the proposed Approach

- Unlike conventional LLM-only methods that treat the task as sentence or span classification, our approach leverages agentic AI, enabling reasoning-based exploration of self-states, modular pipelines, and hybrid action-reasoning workflows. [2, 1]
- We aim to improve recall for adaptive states (often under-detected) while maintaining balance with maladaptive detection.
- Additional novelty lies in the architectural design: building reusable agentic components for mental health NLP, which are currently absent in literature and code repositories.

7. Team Composition and Individual Contributions

- **Member 1:** Ashutosh Kumar Jha, 12340390, Responsible for the end-to-end project pipeline including data collection, preprocessing, system implementation, experimentation, and documentation.

Note: Single-member project.

8. Expected Outcomes

- **Performance Metrics:** Target recall ≥ 0.60 for adaptive and maladaptive self-state detection; weighted recall improvements over sentence-only baselines; better balance between adaptive vs. maladaptive detection.
- **Final Outputs/ Deliverables:**
  - Trained agentic AI pipeline for self-state classification.
  - Comparative study against baseline sentence/LLM-only methods.
  - Research paper/report documenting methodology and findings.
  - (Optional) Demo application for real-time classification of self-states in social media text.

References

[1] Callum Chan, Sunveer Khunkhun, Diana Inkpen, and Juan Antonio Lossio-Ventura. Prompt engineering for capturing dynamic mental health self states from social media posts. In Ayah Zirikly, Andrew Yates, Bart Desmet, Molly Ireland, Steven Bedrick, Sean MacAvaney, Kfir Bar, and Yaakov Ophir, editors, *Proceedings of the 10th Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2025)*, pages 256–267, Albuquerque, New Mexico, May 2025. Association for Computational Linguistics. ISBN 979-8-89176-226-8. doi: 10.18653/v1/2025.clpsych-1.22. URL https://aclanthology.org/2025.clpsych-1.22/.

[2] Laerdon Kim. A baseline for self-state identification and classification in mental health data: Clpsych 2025 task, 2025. URL https://arxiv.org/abs/2504.14066.

[3] Heinrich Peters and Sandra Matz. Large language models can infer psychological dispositions of social media users, 2024. URL https://arxiv.org/abs/2309.08631.

[4] Talia Tseriotou, Jenny Chim, Ayal Klein, Aya Shamir, Guy Dvir, Iqra Ali, Cian Kennedy, Guneet Singh Kohli, Anthony Hills, Ayah Zirikly, Dana Atzil-Slonim, and Maria Liakata. Overview of the CLPsych 2025 shared task: Capturing mental health dynamics from social media timelines. In Ayah Zirikly, Andrew Yates, Bart Desmet, Molly Ireland, Steven Bedrick, Sean MacAvaney, Kfir Bar, and Yaakov Ophir, editors, *Proceedings of the 10th Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2025)*, pages 193–217, Albuquerque, New Mexico, May 2025. Association for Computational Linguistics. ISBN 979-8-89176-226-8. doi: 10.18653/v1/2025.clpsych-1.16. URL https://aclanthology.org/2025.clpsych-1.16/.

Disclaimer

Since the idea has not been approved by the course instructor, the contents of this document are to be inferred as tentative.