Final Project Proposal

Project Title

AI-Powered Early Detection and Risk Assessment of Mental Health Issues

---

Project Description

This project aims to develop an AI-driven system for the early detection and risk assessment of mental health challenges such as depression, anxiety, and burnout.
The system integrates both **survey-based data** and **textual data** (such as user reflections or social media-like statements) to assess emotional well-being. By combining structured and unstructured data, the model will generate an interpretable mental health risk score and highlight the main contributing factors.
The final output will be an **interactive dashboard or web app** that predicts the user's risk level (Low, Moderate, High) and visualizes the top stress indicators.

Group Members & Roles

Name	Role	Responsibilities
[Sagda Esmat]	**Team Leader / Data Scientist**	Coordinate the team, design ML pipeline, manage integration.
[Rawan Ahmed]	**NLP Engineer**	Develop and fine-tune the text classification model (BERT/LSTM).
[Marwa ashraf]	**Data Engineer**	Data cleaning, preprocessing, and feature engineering for survey data.
[Nada nasser]	**Frontend Developer**	Build Streamlit/Flask dashboard and integrate prediction APIs.
[Monika Ibrahim]	**Evaluator / Analyst**	Model evaluation, documentation, and visualization.
[shereen nasser]	**MLOPS Engineer**	Deploy the model, manage APIs, and ensure system integration.

Team Leader

Sagda Esmat

Objectives

1. To predict the mental health risk level of individuals using both survey responses and textual reflections.

2. To integrate multiple datasets (students and professionals) for a diverse and representative model.
3. To ensure interpretability of predictions through SHAP or LIME visualization.
4. To deploy an interactive dashboard for real-time mental health screening and visualization.
5. To promote ethical AI practices with data privacy and transparency.

Tools & Technologies

- **Programming Languages:** Python
- **Libraries & Frameworks:** Pandas, Scikit-learn, TensorFlow/PyTorch, Transformers (BERT), SHAP, LIME
- **Visualization & App:** Streamlit, Matplotlib, Seaborn, Plotly
- **Deployment:** Flask API / Streamlit Cloud / Hugging Face Spaces
- **Dataset Sources:**
  - Mental Health in Tech Survey
  - Student Mental Health Dataset
  - Sentiment Analysis for Mental Health Dataset

Milestones & Deadlines

Milestone	Description	Deadline
Week 1	Data collection and understanding	Oct 20, 2025
Week 2	Data cleaning and preprocessing	Oct 27, 2025
Week 3	Train tabular models (Random Forest, XGBoost)	Nov 3, 2025
Week 4	Train NLP model (BERT / LSTM)	Nov 10, 2025
Week 5	Fusion model and explainability (SHAP/LIME)	Nov 17, 2025
Week 6	Build Streamlit dashboard and integrate models	Nov 24, 2025
Week 7	Testing, evaluation, and documentation	Dec 1, 2025
Final Submission	Final report + live demo	Dec 5, 2025

KPIs (Key Performance Indicators)

1. Data Quality

Metric	Target
Percentage of missing values handled	**> 98%**
Data accuracy after preprocessing	**> 95%**
Dataset diversity (representation of categories)	**≥ 3 distinct demographic groups**

2. Model Performance

Metric	Target
Model accuracy (F1-Score)	**≥ 85%**
Model prediction speed (Latency)	**< 200 milliseconds**
Error rate (False Positive / False Negative)	**< 10%**

3. Deployment & Scalability

Metric	Target
API uptime	**> 99%**
Response time per request	**< 500 milliseconds**
Real-time processing (if applicable)	**5 FPS (for streaming inputs)**

4. Business Impact & Practical Use

Metric	Target
Reduction in manual screening effort	**≥ 60%**
Expected cost savings for institutions	**≥ 40%**
User satisfaction (survey-based)	**≥ 85%**

Expected Deliverables

- Cleaned and preprocessed datasets (CSV + documentation).
- Machine learning and NLP models (trained + evaluated).
- Streamlit dashboard with real-time prediction and explainability.
- Final report summarizing model design, evaluation, and ethical considerations.

Ethical & Privacy Considerations

- All data anonymized and used only for research purposes.
- No collection of identifiable information (names, emails, etc.).

- Clear disclaimer: “This tool is not a diagnostic system, only a screening support.”
- Results are interpreted cautiously and reviewed by professionals.

Conclusion

This project aims to empower organizations and universities with an AI-based mental health screening tool that is **accurate, interpretable, and ethical**. By fusing survey and text data, it provides deeper insights into emotional well-being, helping institutions intervene early and support mental wellness proactively.