SentiCare

Voice Based Mental Health Support System

Project Code

Project Advisor

Dr. Saad Razzaq (internal)
Dr. Mohsin Atta (external)

Project Manager

Dr. Muhammad Ilyas

Project Team

Esha Gulzar	BSCS51F22S047	\<Team leader\>
Wajeeha Ijaz	BSCS51F22S032	\<Team member\>
Muhammad Zain Ul Abidin	BSCS51F21S054	\<Team member\>

Submission Date

6/10/2025

Table of Contents

1.Abstract

Mental health issues remain critically underserved due to stigma, a lack of professionals, and limited awareness, leaving millions without access to timely help. Existing digital mental health tools are mostly English-based and text-driven, excluding Urdu-speaking individuals who prefer natural, verbal communication. To address this gap, we will develop an AI-powered bilingual (English and Urdu) voice chatbot designed to provide evidence-based Cognitive Behavioral Therapy (CBT) support through speech-based interaction. We will integrate Speech-to-Text (STT), voice biomarker extraction, emotion detection, and Reinforcement Learning (RL) to recognize emotional cues in user speech and deliver context-aware, personalized CBT interventions. The system will support users dealing with anxiety, depression, and stress by offering structured CBT exercises, therapeutic guidance, and adaptive learning that improves with each interaction. By combining language inclusivity, evidence-based therapy, emotional intelligence, and accessible AI, the system seeks to bridge the mental health service gap in Pakistan while contributing to AI research for low-resource languages.

2.Background and Justification

Pakistan faces an escalating mental health crisis, with around 50 million individuals experiencing psychological distress and less than one psychiatrist per 600,000 citizens [1]. Stigma, cost, and lack of professionals prevent timely care. Most mental health applications, such as Woebot or Youper, are English-only and text-based, making them inaccessible to the majority of Pakistani users who prefer Urdu or bilingual communication and natural voice interaction.

AI-driven conversational agents have proven effective in delivering CBT interventions and reducing psychological distress [2]. Research highlights that culturally adapted, language-specific systems increase user trust and engagement [3]. However, Urdu remains a low-resource language for speech and emotion recognition, and existing systems rarely combine voice biomarker analysis with adaptive reinforcement learning for personalized therapy [4].

We will address these gaps by developing a bilingual (English and Urdu) voice-based chatbot that integrates speech emotion recognition through voice biomarkers, natural language understanding using BERT and Sentence-BERT, and reinforcement learning for adaptive CBT intervention selection. The system will follow evidence-based CBT protocols including structured exercises for anxiety, depression, and stress management. We will design the system to act as a supportive therapeutic companion rather than a diagnostic tool, focusing on accessibility, cultural sensitivity, language inclusivity, empathy, and ethical design. By supporting both English and Urdu, we will maximize accessibility for Pakistan's diverse linguistic population while delivering personalized, stigma-free mental health support.

3.Project Methodology

The project follows an iterative and modular approach, divided into three layers: **Frontend**, **Backend**, and **AI/ML Core**. Each layer will be improved through continuous feedback and evaluation.

Frontend Layer:

We will develop a bilingual voice-enabled interface for web and mobile use, supporting real-time speech input in English and Urdu with text-to-speech output. The design will prioritize simplicity and accessibility for users with varying literacy levels.

Backend Layer:

We will implement a secure server framework using **Flask/FastAPI** to handle API requests and manage conversations. We will use **PostgreSQL** for structured data storage with secure authentication, HTTPS encryption, and data anonymization to ensure user confidentiality.

AI/ML Core:

Speech-to-Text (STT): we will use **OpenAI Whisper** for accurate English and Urdu speech-to-text conversion.

Voice Biomarker Extraction: we will use **Librosa** to extract language-independent acoustic features (pitch, speech rate, pause duration, energy, jitter) for emotional context analysis.

Natural Language Understanding (NLU):

- *English:* **Sentence-BERT** (all-mpnet-base-v2) for semantic understanding; fine-tuned **BERT** on existing mental health samples for anxiety, depression, and stress classification
- *Urdu:* Multilingual **Sentence-BERT** (LaBSE) or **XLM-RoBERTa**; translation bridge using **Helsinki-NLP models** (Urdu ↔ English) for processing

Emotion Detection: project will combine voice biomarkers with text-based emotion models to detect anxiety, sadness, stress, and calmness for intervention selection.

Dialog Management: we will create 5-10 CBT templates per disorder (anxiety, depression, stress) in both languages, including grounding techniques, breathing exercises, and cognitive reframing. We will use **PHQ-9** and **GAD-7** questionnaires for initial assessment.

Reinforcement Learning (RL): project will implement **contextual bandits** (NumPy/SciPy) initially, advancing to **PPO** (Stable-Baselines3) to learn optimal CBT technique selection based on engagement, sentiment improvement, and voice biomarker changes.

Text-to-Speech (TTS): project will use **gTTS** or **Coqui TTS** for bilingual therapeutic response delivery.

4.Project Scope

The chatbot's goal is to provide immediate, evidence-based CBT support for English and Urdu-speaking individuals in Pakistan using AI-driven voice interaction and therapeutic intervention.

In Scope:

- Bilingual voice-based conversations (English and Urdu) with language selection
- Clinical assessment using PHQ-9 and GAD-7 , PSS-4 (stress) questionnaires in both languages
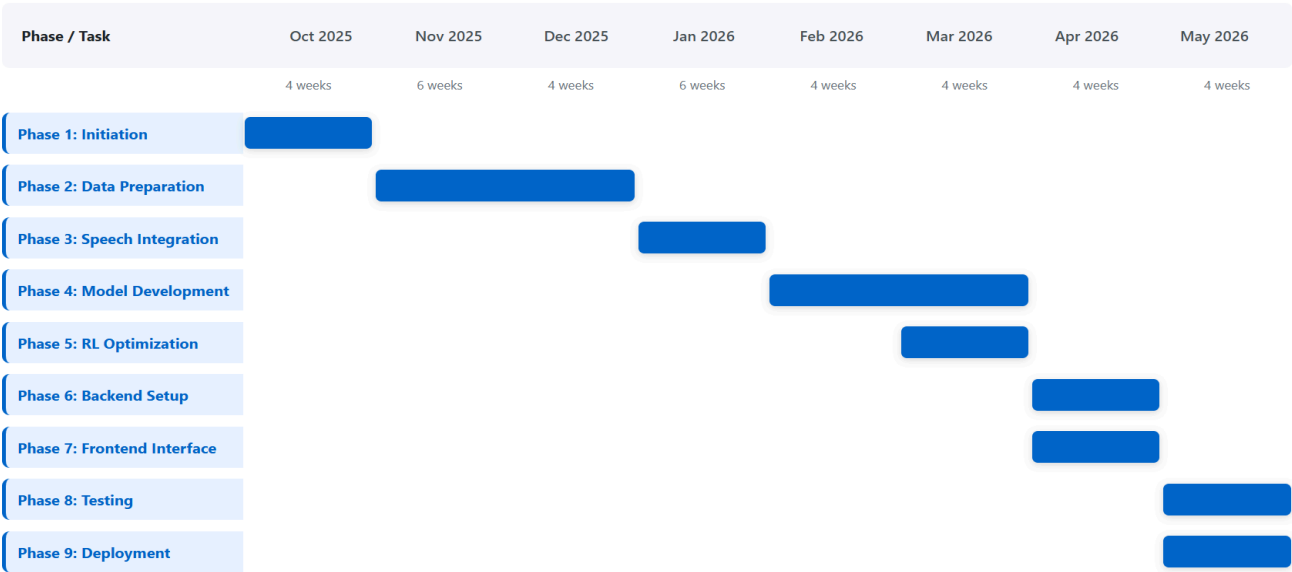- Voice biomarker extraction for emotional analysis

- Emotion and sentiment recognition through speech and text
- Structured CBT interventions (grounding, breathing exercises, cognitive reframing)
- Reinforcement learning for personalized therapy selection
- Secure data storage with encryption and authentication
- Culturally adapted content for Pakistani users

Out of Scope:

- Medical diagnosis or medication prescription
- Replacement of professional therapy
- Languages beyond English and Urdu
- Video communication
- Clinical trials or large-scale deployment
- Real-time human therapist intervention
- Integration with hospital system

5.High-Level Project Plan:

The project spans approximately 36 weeks, divided into phases with clear tasks, descriptions, durations, timelines, and dependencies. This plan ensures iterative development and risk mitigation

Phase / Task	Oct 2025	Nov 2025	Dec 2025	Jan 2026	Feb 2026	Mar 2026	Apr 2026	May 2026
	4 weeks	6 weeks	4 weeks	6 weeks	4 weeks	4 weeks	4 weeks	4 weeks
Phase 1: Initiation								
Phase 2: Data Preparation								
Phase 3: Speech Integration								
Phase 4: Model Development								
Phase 5: RL Optimization								
Phase 6: Backend Setup								
Phase 7: Frontend Interface								
Phase 8: Testing								
Phase 9: Deployment								

6.References:

[1] British Asian Trust, "Mental health in Pakistan," 2024. [Online]. Available: https://www.britishasiantrust.org/our-work/mental-health/mental-health-in-pakistan/ [2] A. Ankomah and R. Turkson, "Emotion-Aware AI Chatbots for Mental Health Support in Low-Resource Public Health Systems," *World Journal of Public Health*, vol. 10, no. 3, pp.

265–272, 2025.[Online]. Available:
https://www.sciencepublishinggroup.com/article/10.11648/j.wjph.20251003.17

[3]   S. Akhtar et al., "mPareshan: Technology-Assisted Mental Health Intervention in Rural Pakistan," *BMC Psychiatry*, vol. 25, art. 16, 2025.[Online]. Available:
https://bmcpsychiatry.biomedcentral.com/articles/10.1186/s12888-024-06459-8

[4]   H. Mehmood and S. A. Rauf, "Human-Evaluated Urdu-English Speech Corpus for Low-Resource Languages," *Proc. Intl. Workshop on Spoken Language Translation (IWSLT)*, pp. 138–144, 2025. [Online]. Available: https://aclanthology.org/2025.iwslt-1.12.pdf [5] A. Asghar, S. Sohaib, M. Shafi, and K. Fatima, "An Urdu Speech Corpus for Emotion Recognition," *PeerJ Computer Science*, vol. 8, p. e954, 2022.[Online]. Available:
https://www.sciencedirect.com/science/article/pii/S2352340925003580