

A Hybrid AI Chatbot Framework for Scalable Mental Health Assistance

Parth Vyas

SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

Abstract:

The growing global burden of mental health disorders has exposed significant limitations in existing healthcare systems, particularly in providing accessible and scalable support. Artificial intelligence-based conversational agents have emerged as a promising solution; however, current approaches face a fundamental trade-off between safety and empathy. Rule-based systems ensure reliability but lack conversational flexibility, while generative models enable humanlike interactions but introduce risks in high-stakes situations. This paper presents a hybrid AI chatbot framework that integrates a BERT-based risk classification module, a rule-based safety system, and a fine-tuned large language model for empathetic response generation. The system dynamically routes user inputs based on real-time risk assessment, ensuring appropriate handling of both crisis and non-crisis interactions. Experimental evaluation on 500 simulated conversations shows that the proposed approach reduces unsafe outputs by 94.3% compared to standalone generative models, while maintaining high levels of user satisfaction and engagement. These findings demonstrate the effectiveness of hybrid architectures in delivering safe, scalable, and context-aware mental health support.