

# Multimodal Cognitive Reframing System: A Multi-Agent Approach for Mental Health Support

Das Projekt beschäftigt sich mit der Entwicklung eines multiagentenbasierten Systems, das darauf abzielt, Menschen dabei zu helfen, negative Gedanken in positivere und konstruktivere Perspektiven umzuwandeln. Dabei wird die Methode des kognitiven Reframens, eine zentrale Technik der kognitiven Verhaltenstherapie (CBT), verwendet, um ungesunde Denkmuster herauszufordern und eine gesündere Denkweise zu fördern. Das System erlaubt multimodalen Input wie Text, Sprache und Bilder, die jeweils von autonomen Agenten analysiert werden. Ziel des Projekts ist es, eine schnelle und einfache Unterstützung für Personen zu bieten, die sich davor meiden, professionelle Hilfe in Anspruch zu nehmen, aufgrund von Stigmatisierung oder Scham.

Die Methode des kognitiven Reframens spielt die zentrale Rolle im System. Diese Technik basiert auf der Idee, dass Gedanken, Gefühle und Verhaltensweisen miteinander verbunden sind und sich gegenseitig beeinflussen. Negative Denkmuster können zu emotionalem Stress und problematischen Handlungen führen. Alternative Denkweisen, wie das Wechseln zu einer positiveren Perspektive, kann das psychische Wohlbefinden verbessern.

Die Hauptarchitektur ist ein Multi-Agenten-System. Dieser Ansatz eignet sich besonders für komplexe Aufgaben, bei denen unterschiedliche Kompetenzen erforderlich sind, welche zu komplex für einen einzigen Agenten wären. Das System besteht aus mehreren autonomen Agenten, die jeweils auf eine spezifische Aufgabe spezialisiert sind. Diese Agenten kollaborieren miteinander und nutzen Large Language Models, um fortgeschrittene Analysen durchzuführen. Jeder Agent bearbeitet eine spezifische Eingabemodalität: Der Text-Agent identifiziert kognitive Verzerrungen in den Texteingaben. Der Sprach-Agent verwendet die automatische Spracherkennung von OpenAIs Whisper, um Audioaufnahmen zu transkribieren und emotionale Inhalte aus Audioaufnahmen zu interpretieren. Der Bild-Agent untersucht visuelle Daten wie Zeichnungen oder Fotos und analysiert visuelle Elemente wie Farben, Komposition und Symbole, um emotionale Zustände und unbewusste Gedanken des Benutzers zu identifizieren. Alle Agenten schlagen basierend auf ihrer Analyse Optionen vor, wie diese erfassten Gedanken umgedacht werden können. Der Multimodale-Agent fokussiert sich auf das Gespräch mit dem User und integriert gleichzeitig die Ergebnisse der einzelnen Agenten, um eine Antwort zu dem Input des Users zu geben, mit dem Fokus auf das kognitive Reframens. Das System nutzt darüber hinaus Techniken des Prompt-Engineerings, um die Leistungsfähigkeit der Sprachmodelle zu maximieren. Durch gezielt gestaltete Eingabeaufforderungen werden die Modelle darauf ausgerichtet, präzisere und therapeutisch relevante Antworten zu generieren. Rollenspiel des Large Language Models und Empathisches Paraphrasieren fördern dabei die Interaktion und ermöglichen dem Benutzer, seine Gedanken und Gefühle tiefer zu reflektieren.

Ein Hauptziel des Projekts ist es, eine realistische therapeutische Konversation zu simulieren. Zusammenfassend stellt das Projekt eine innovative und zugängliche Lösung für die Unterstützung der psychischen Gesundheit dar. Es bietet einen neuen Ansatz, um Barrieren abzubauen und Menschen auf ihrem Weg zu einem besseren mentalen Wohlbefinden zu begleiten, indem es ihnen hilft, negative Gedanken in konstruktive Perspektiven umzuwandeln.