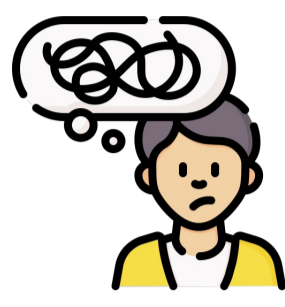


Motivation & Background



Mental Health

- Highly intricate, multi-factorial processes.
- Prevalence, Under-treatment, Stigma

50% individuals encounter a mental health disorder at some point in their lives. **25%** individuals with mental health are lack access to adequate care [1].

Key Limitations of Existing Systems

- Reliance only on subjective text—leading to low accuracy and engagement.
- Lack of context and perception of the real state.
- Limited personalization, poor interpretability.
- Cloud dependency raises privacy risks.

Can we turn every mobile phone into a **personal psychological assistant** that create **accessible, stigma-free, scalable** mental health care solution that can understand you, remember you and take the initiative to help you?

Aspect	Personal Psychological Assistant	AI-Chatbot
Personalization	Tailors care based on history, personality, and context	Often responds generically; limited memory
Reasoning	Uses clinical knowledge, causal reasoning, and intuition	Primarily reactive; lacks causal understanding or reasoning
Adaptability	Can adjust strategy based on how a session unfolds or emotional cues	Usually follows pre-set flows; struggles with unexpected input
Accountability & Trust	Operates under professional ethics; trusted relationship	No accountability; perceived as a tool, not a partner
Multimodal Awareness	Observes tone, body language, behavioral patterns	Relies mostly on text input; lacks full context of the user's state

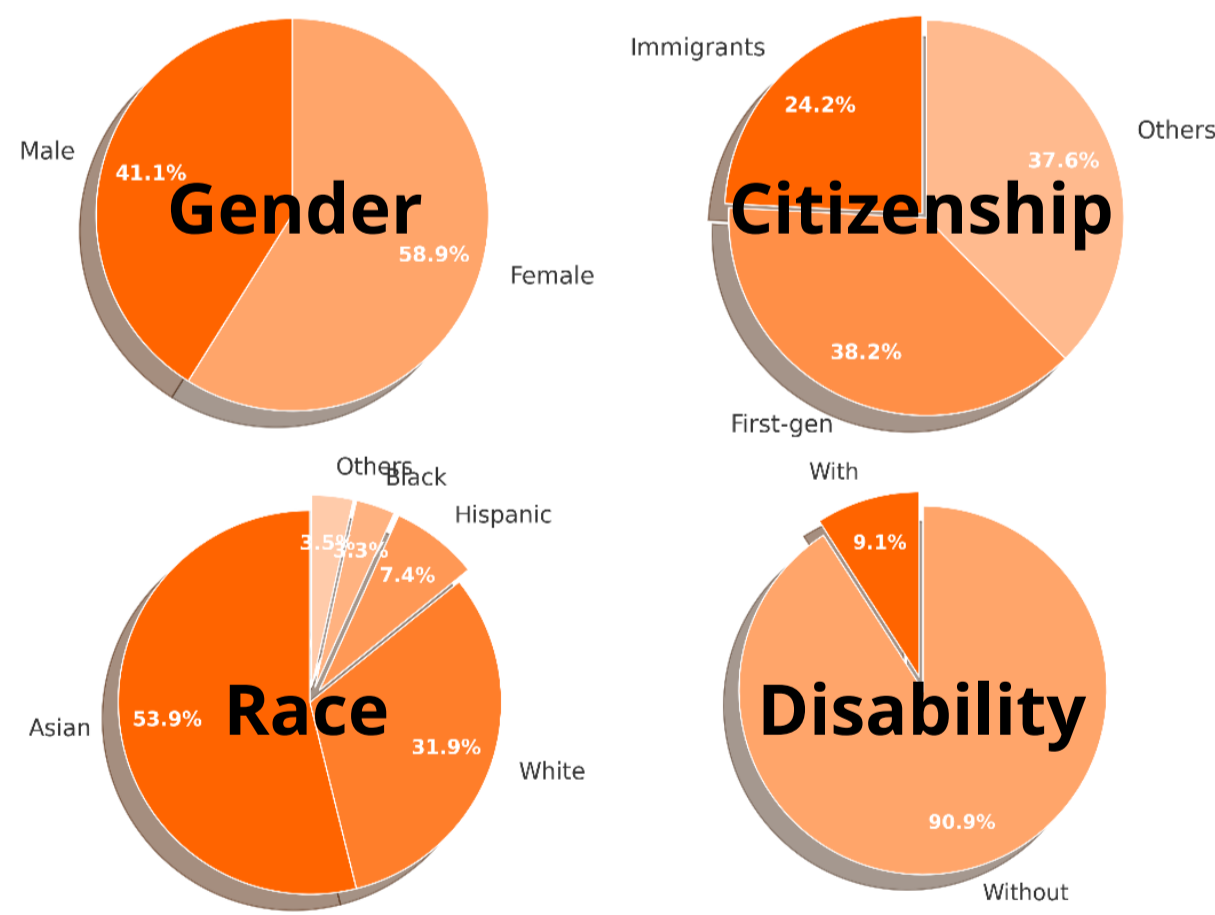


ProMind-LLM - Autonomous LLM agents enhanced with wearable sensor data and causal reasoning - **just like a trained counselor right from your pocket**

Evaluation

Dataset:

 Globem	4 years	~500
 PMData	>5 months	16
 IRB-Pilot	2 weeks	20



Ablation Study:

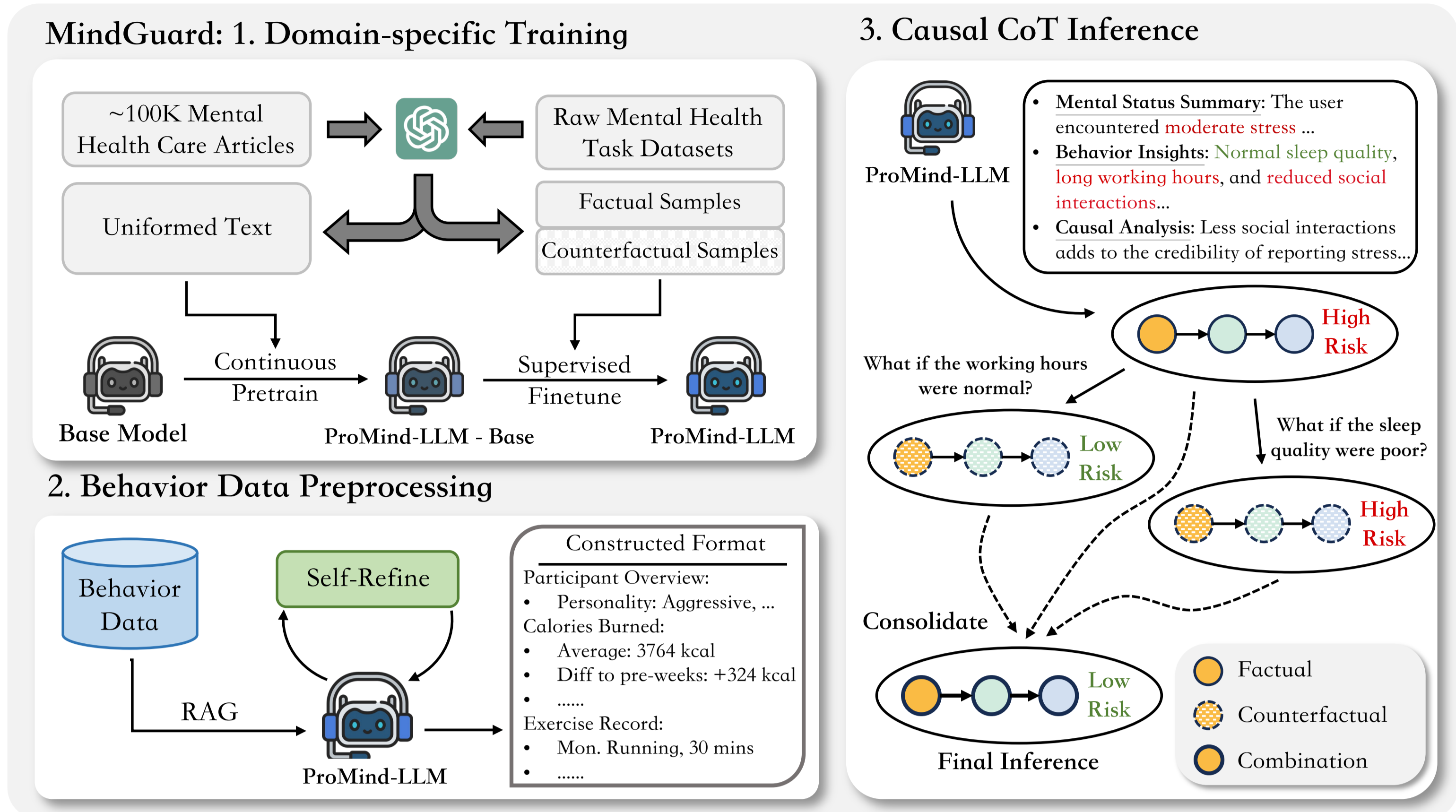
❖ Self-refine Mechanism and causal CoT:

Model	Method		PMData				Globem			
	Self-Refine	Causal CoT	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
GPT-4o	✗	✗	0.938	0.734	0.726	0.729	0.803	0.575	0.818	0.675
	✓	✗	<u>0.952</u>	<u>0.737</u>	<u>0.800</u>	<u>0.767</u>	0.814	0.586	<u>0.951</u>	<u>0.725</u>
	✗	✓	0.933	0.687	0.769	0.726	<u>0.829</u>	<u>0.615</u>	0.848	0.713
	✓	✓	0.956	0.781	0.821	0.800	0.867	0.663	0.955	0.783
Qwen2-chat-72B	✗	✗	0.837	0.382	0.693	0.495	0.761	0.516	0.742	0.609
	✓	✗	<u>0.903</u>	<u>0.514</u>	0.514	<u>0.514</u>	0.769	0.523	0.879	<u>0.655</u>
	✗	✓	0.859	0.423	<u>0.635</u>	0.507	<u>0.781</u>	<u>0.541</u>	<u>0.798</u>	0.645
	✓	✓	0.918	0.674	0.547	0.604	0.819	0.608	0.783	0.684
ProMind-LLM	✗	✗	0.845	0.369	0.504	0.426	0.734	0.475	0.641	0.546
	✓	✗	0.940	0.792	<u>0.543</u>	<u>0.644</u>	<u>0.844</u>	<u>0.640</u>	<u>0.902</u>	<u>0.748</u>
	✗	✓	0.849	0.376	0.479	0.422	0.766	0.525	0.672	0.589
	✓	✓	<u>0.938</u>	<u>0.765</u>	0.667	0.712	0.859	0.646	0.939	0.765

Overall Performance:

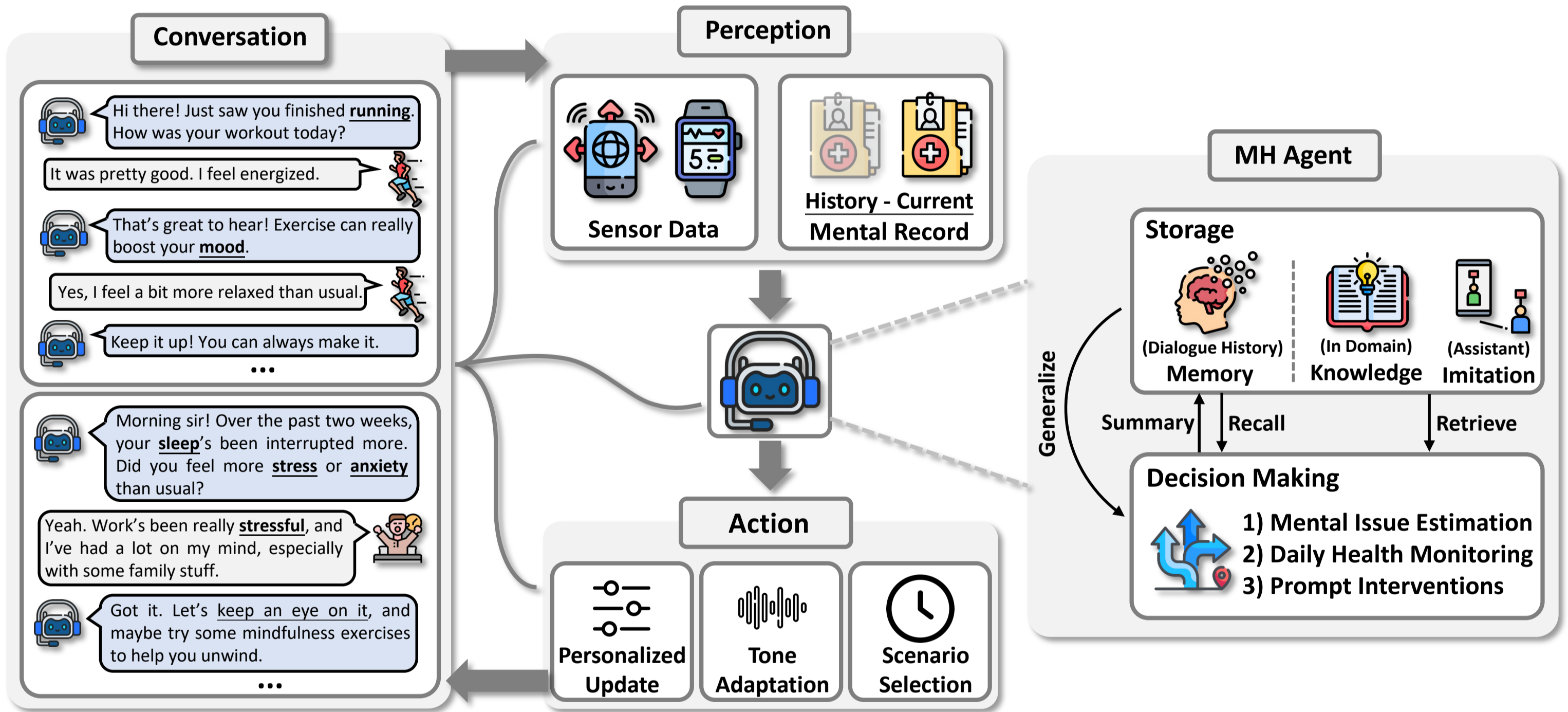
Category	Model	PMData				Globem			
		Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
Comercial	GPT-4o1	0.956	0.781	<u>0.821</u>	0.800	0.867	0.663	<u>0.955</u>	0.783
	GPT-3.51	0.858	0.385	0.714	0.500	0.747	0.505	0.918	0.651
	Claude-3.51	0.923	0.786	0.314	0.449	0.789	0.552	0.951	0.699
Open-sourced	LLaMA3-chat-70B	0.795	0.297	0.771	0.429	0.807	0.579	0.833	0.683
	QWen2-chat-72B	0.918	0.674	0.547	0.604	0.819	0.608	0.783	0.684
	InternLM2-chat-7B	0.634	0.191	0.828	0.310	0.367	0.289	1.000	0.449
ProMind-LLM	InterLM2-base-7B	<u>0.938</u>	<u>0.765</u>	0.667	<u>0.712</u>	<u>0.859</u>	<u>0.646</u>	0.939	<u>0.765</u>

System Architecture

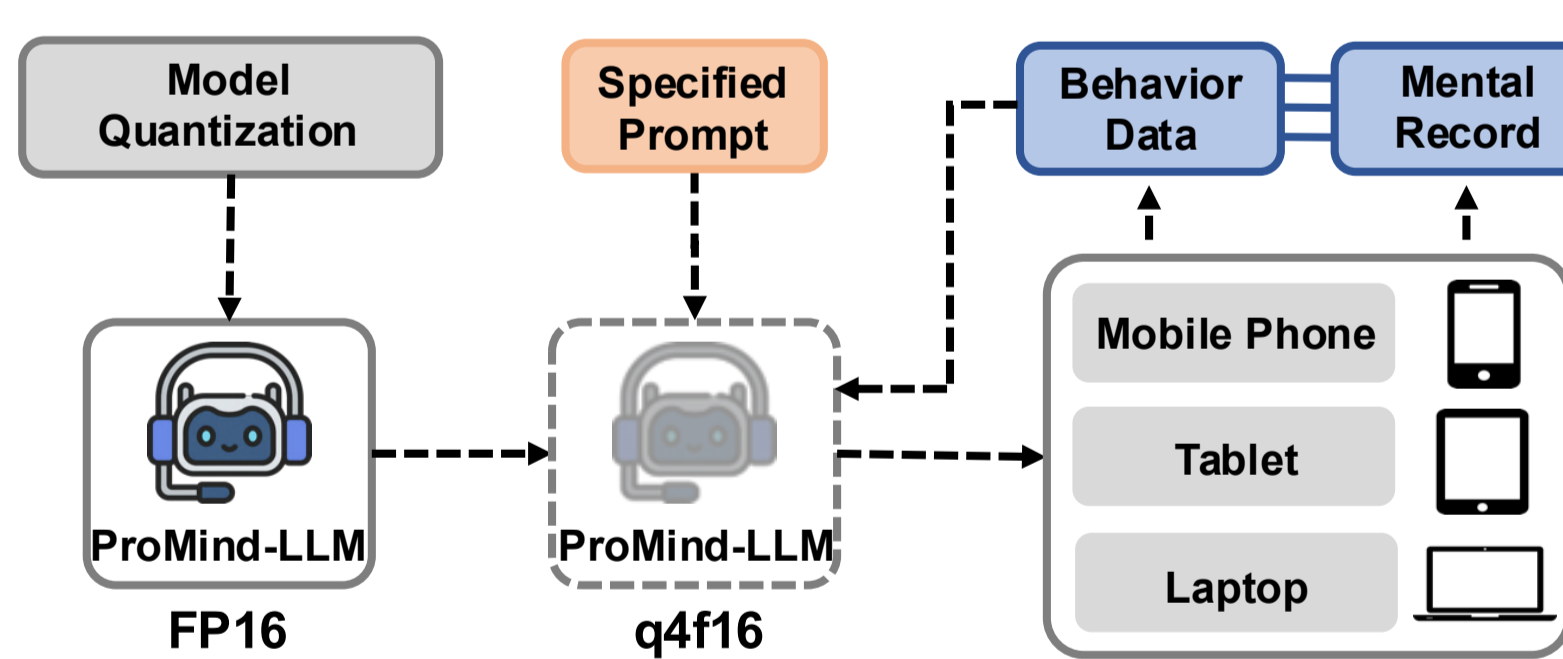


Key Techniques:

- 📖 Domain Pretraining + SFT (Counterfactual Augmentation)
- 🧠 Causal Chain-of-Thought (CoT) Reasoning
- 🔄 Self-refine Mechanism
- 📱 MLC-LLM Mobile Deployment (4GB)



Implementation



Device	Chip	Memory	Decode (tok/s)
GPT-4-turbo	-	-	20
iPhone 14 Pro	A16	6GB	5
iPhone 15 Pro	A17	8GB	7
iPad Pro 2021	M1	8GB	7
MacBook Pro 13	M1	16GB	11

Contributions

Key Impact

- ✓ 130%↑ over base LLMs
- ✓ Mobile-ready (4GB)
- ✓ Human-aligned, explainable
- 🌐 Anyone. Anytime. Anywhere
- ✓ Support personalized, interpretable, cause-driven psychological intervention
- ✓ More resource-efficient than commercial LLMS
- ✓ Transition from cloud inference to edge intelligence
- 🌐 A future for mental health services open to all

References

- [1] Xu, Xuhai, et al. "GLOBEM dataset: multi-year datasets for longitudinal human behavior modeling generalization." *Advances in neural information processing systems* 35 (2022): 24655-24692.
- [2] Zheng, Xinzhe, et al. "ProMind-LLM: Proactive Mental Health Care via Causal Reasoning with Sensor Data." *arXiv preprint arXiv:2505.14038* (2025).
- [3] Ji, Sijie, et al. "Mindguard: Towards accessible and sitgma-free mental health first aid via edge llm." *arXiv preprint arXiv:2409.10064* (2024).