CATEGORY: RESEARCH

ARGMED-AGENTS: EXPLAINABLE CLINICAL DECISION REASONING WITH LLM VIA ARGUMENTATION SCHEMES

Anushka Joseph, Figo Fernandez, Jess John
Department of Computer Engineering,
The Bombay Salesian Society's, Don Bosco Institute of Technology, Mumbai- 400070

2024-2025

ABSTRACT

- Traditional AI models *lack transparency*, limiting reliability in clinical decision-making.
- LLMs struggle with *precise, explainable* medical reasoning.
- A multi-agent argumentation framework improves **precision, coherence,** and **explainability**.
- **Generator, Verifier,** and **Reasoner** agents collaborate to refine clinical arguments.
- An **argumentation graph** maps *reasoning*, allowing clinicians to **trace decisions**.
- Generates **clinically meaningful** decisions when supported by domain knowledge.
- Challenges remain in handling *complex medical cases*.
- Enhances **trust** and **accountability** in AI-driven clinical decision support systems.

SCOPE

Develop a **multi-agent argumentation framework** to *simulate clinical reasoning* by *generating* and *evaluating arguments* from ***different perspectives***.

Integrate large language models (***LLMs***) with *domain-specific knowledge* for *natural language processing* and *evidence synthesis* in clinical contexts.

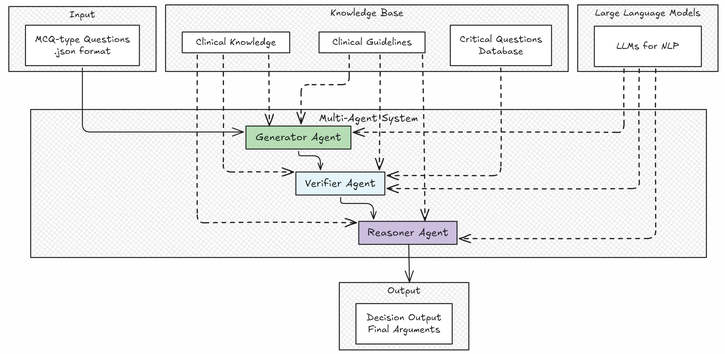
Create a **computational model** to **evaluate** the *coherence, plausibility*, and *relevance* of arguments for ***clinical decision-making***.

Conduct thorough evaluation and **validation** for the *framework's performance* against ***traditional*** *clinical decision support systems*.

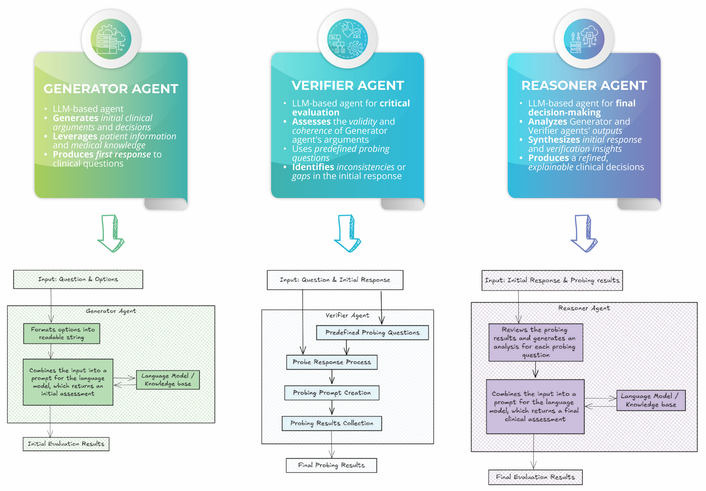
PROBING QUESTIONS

1. What are the key factors that led you to this conclusion by ruling out the other options?
2. Can you identify any potential weaknesses in your reasoning?
3. What alternative explanations might exist for this scenario?
4. How might your answer change if [insert relevant detail] was different?
5. On a scale of 1-10, how certain are you of this answer, and why?
6. What additional information would help you be more confident in your answer?
7. What are the potential consequences if this answer is incorrect?
8. How does your answer align with standard medical practices or guidelines?

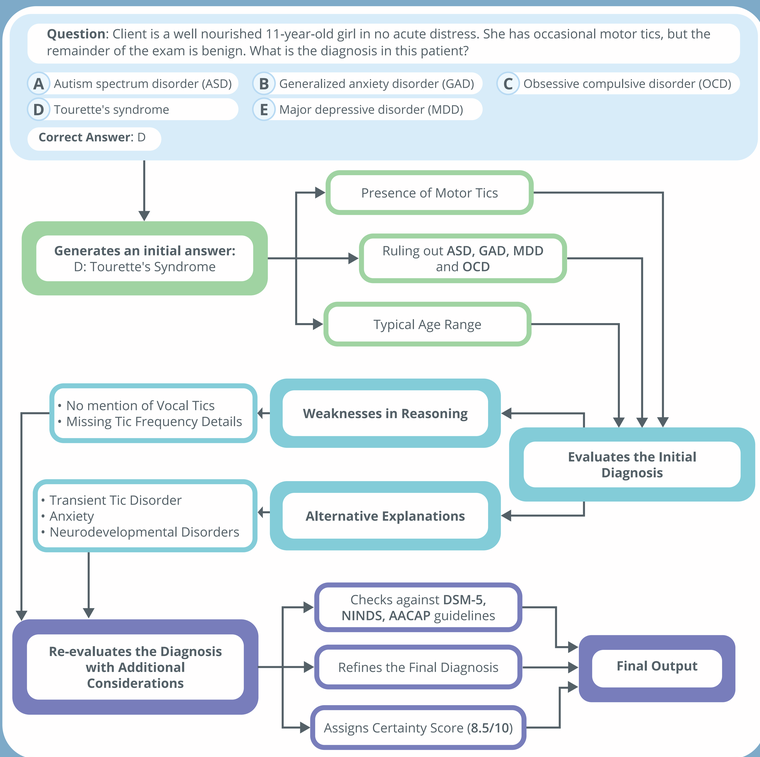
SYSTEM ARCHITECTURE



FRAMEWORK DESIGN



WORKFLOW EXAMPLE



REFERENCES

1. D. Saraswat et al., "Explainable AI for Healthcare 5.0: Opportunities and Challenges," in IEEE Access, vol. 10, pp. 84486-84517, 2022, doi: 10.1109/ACCESS.2022.3197671.
2. Y. Jia, J. McDermid, N. Hughes, M. Sujan, T. Lawton and I. Habli, "The Need for the Human-Centred Explanation for ML-based Clinical Decision Support Systems," 2023 IEEE 11th International Conference on Healthcare Informatics (ICHI), Houston, TX, USA, 2023, pp. 446-452, doi: 10.1109/ICHI57859.2023.00064.
3. C. Comito, D. Falcone and A. Forestiero, "AI-Driven Clinical Decision Support: Enhancing Disease Diagnosis Exploiting Patients Similarity," in IEEE Access, vol. 10, pp. 6878-6888, 2022, doi: 10.1109/ACCESS.2022.3142100.
4. C. Comito, A. Forestiero and G. Papuzzo, "A clinical decision support framework for automatic disease diagnoses," 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Vancouver, BC, Canada, 2019, pp. 933-936, doi: 10.1145/3341161.3343509.
5. S. Lee, S. Kim, J. Lee, "Explainable Artificial Intelligence for Patient Safety: A Review of Application in Pharmacovigilance," in IEEE Access, vol. 11, pp. 50830-50840, 2023, doi: 10.1109/ACCESS.2023.3271635.