# AI Agent And NLP Based Medical Differential Diagnosis And History Analyzer

Abhi Akshat[1], Kritika Aggarwal[2], Sheenam Naaz[3]
*Department of Computer Science Engineering[1,2,3]*
*Sharda University[1,2,3]*

***Abstract*— In the era of rapid advancements in Artificial Intelligence (AI) and Large Language Models (LLMs), these technologies are increasingly being applied in the medical field to enhance efficiency and accuracy. One area where LLMs are making a significant impact is in automating tasks like medical history recording and analysis. By leveraging Natural Language Processing (NLP), AI models can streamline the process, reducing the time spent on mundane tasks and improving the overall quality of care provided to patients. This paper proposes a system that utilizes LLMs powered AI Agents, specifically models like Llama, to automate the collection and analysis of patient medical history. The system functions as a virtual assistant, interacting with patients to gather information on their symptoms and medical background. It then provides doctors with a summarized history and a differential diagnosis, ranked in order of likelihood, along with potential causes. The model also considers regional factors, improving diagnostic accuracy by incorporating location-based insights. The goal is to ease the workload of healthcare providers, enabling them to focus more on patient care. The system uses fine-tuning techniques to continuously improve its performance over time. This paper details the system's architecture, training methodologies, and the results of its implementation. By automating medical history recording and analysis, this AI-driven approach has the potential to enhance diagnostic accuracy, speed up decision-making, and ultimately improve patient outcomes in clinical settings.**

***Index Terms*—Large Language Models, Natural Language Processing, Artificial Intelligence, Machine Learning, LLaMA**

## I. INTRODUCTION

The integration of Artificial Intelligence (AI) and Natural Language Processing (NLP) into healthcare has the potential to reshape the way medical professionals handle patient data, diagnose conditions, and deliver care. One of the most exciting developments in recent years is the rise of Large Language Models (LLMs), such as GPT, BERT, and LLaMA, which have demonstrated remarkable capabilities in understanding and generating human language. These models have found applications in a wide array of fields, but their potential impact on the medical domain is especially significant. LLMs can play a vital role in automating processes that are otherwise time-consuming, tedious, and prone to human error. One such application lies in the automation of medical history recording and analysis, which is an essential, yet often burdensome task for healthcare professionals. The introduction of LLMs into this process promises to improve efficiency, reduce errors, and enhance the overall quality of patient care.

Medical history recording is a critical component of patient care, as it involves gathering comprehensive information about a patient's past medical conditions, surgeries, medications, family health history, and current symptoms. Traditionally, this process is manual and time-consuming, requiring healthcare professionals to conduct in-depth interviews with patients. In many cases, valuable time is spent on data entry and documentation, tasks that divert attention from more critical aspects of patient care such as diagnosis and treatment. Furthermore, this manual process is susceptible to errors or omissions, particularly when clinicians are under time pressure. The margin for human error in medical history documentation can have significant consequences, potentially leading to misdiagnosis or delayed treatments. This inefficiency highlights the need for automated solutions that can streamline medical history intake without sacrificing accuracy or detail.

Large Language Models present a promising solution to this challenge. With their ability to process and generate natural language at a near-human level, LLMs can be leveraged to automate the task of gathering patient information. These models, trained on vast datasets, can understand and interpret the complexities of medical language, making them well-suited for use in clinical environments. They can engage in conversational exchanges with patients to gather relevant information, cross-reference historical data, and even detect subtle cues that may otherwise be missed in a rushed consultation. By deploying LLMs for medical history recording, healthcare providers can not only save time but also ensure that patient data is recorded consistently and accurately, reducing the likelihood of human error.

Large Language Model (LLM)-powered AI agents are specialized, language-processing systems that interact with patients to gather symptoms, medical history, and demographic information. They interpret and synthesize this data to generate a prioritized differential diagnosis and offer relevant treatment suggestions. Leveraging LLMs enables these agents to understand complex medical language and patient narratives, providing context-aware and nuanced responses. Their natural language processing capabilities also facilitate seamless communication, making it easy for patients to describe their symptoms while the agents dynamically adapt to varied medical scenarios, ultimately supporting doctors in delivering faster, more accurate care.

Beyond the collection of medical history, LLMs can assist in the analysis of patient data. The ability of these models to contextualize and summarize large amounts of information can prove invaluable to clinicians, who are often tasked with synthesizing vast amounts of patient data within short periods. An AI-driven system could provide doctors with a summarized version of the patient’s medical history, highlighting the most pertinent details such as chronic conditions, previous surgeries, or ongoing treatments. This summary can serve as a decision-support tool, allowing physicians to make faster, more informed decisions. Additionally, LLMs can offer suggestions for differential diagnoses by analyzing the patient’s symptoms in conjunction with their medical history and geographic factors that may influence certain conditions. This capability is particularly useful in areas with limited access to specialists, where primary care providers must rely on a broad base of knowledge to make diagnoses.

One of the key strengths of using LLMs in medical settings is their capacity for continuous learning and improvement. As these models are exposed to more data, they can become more accurate in their predictions and more adept at handling complex medical scenarios. Furthermore, LLMs can be fine-tuned for specific medical applications, such as certain specialties or types of patient interactions. This adaptability makes them highly versatile tools that can be customized to fit the specific needs of a hospital or clinic. For instance, a system designed to work in a pediatric setting could be fine-tuned to better understand and engage with child patients, while a system used in an oncology department could be trained to handle the complexities of cancer diagnosis and treatment.

However, the implementation of LLMs in the medical domain is not without challenges. Concerns surrounding patient privacy and data security must be addressed to ensure that sensitive health information is protected. Furthermore, the development of these models requires vast amounts of data and computational resources, raising questions about accessibility and scalability, especially for smaller healthcare providers. Despite these hurdles, the potential benefits of integrating LLMs into healthcare workflows are immense, particularly when it comes to improving the speed and accuracy of medical history recording and diagnosis.

The integration of Large Language Models into the medical field offers the opportunity to automate key processes such as medical history recording and analysis, ultimately leading to enhanced patient care. By reducing the burden of administrative tasks on healthcare professionals, LLMs enable clinicians to focus more on patient care and clinical decision-making. As these models continue to evolve and improve, their role in healthcare is likely to expand, transforming not only how patient data is recorded but also how diagnoses are made and treatments are administered. This paper aims to explore the architecture, training techniques, and potential outcomes of such a system, demonstrating how LLMs like LLaMA can be leveraged to improve efficiency and accuracy in medical history recording and diagnosis.

## II. RELATED WORK

The rapid integration of artificial intelligence (AI) and machine learning (ML) into healthcare has resulted in significant advancements across diagnostic systems, clinical decision support tools, and treatment optimization. The reviewed literature highlights major innovations in conversational AI, large language models (LLMs), natural language processing (NLP), deep learning in medical imaging, reinforcement learning (RL) for treatment strategies, and more specialized applications like telemedicine and genomics.

A notable contribution is the development of a conversational AI system that employs self-play to simulate clinical diagnostic scenarios. The Articulate

Medical Intelligence Engine (AMIE) demonstrated improved accuracy in 149 simulated scenarios [1]. However, limitations in clinician interaction with the system's synchronization protocols pointed to the challenges of real-world adoption. Another prominent study developed a differential diagnosis (DDx) model based on an optimized large language model, which was tested on over 300 real-world cases. While the model showed high diagnostic accuracy, the rigidity of the clinical case formats limited flexibility in diagnosis [2].

In terms of data security and prediction, one study utilized a crossover-based multilayer perceptron (CMLP) to enhance prediction accuracy and security in healthcare, achieving 97% accuracy and 93% precision [3]. Meanwhile, zero-shot and few-shot learning techniques have emerged as viable solutions for addressing the scarcity of labeled healthcare data, particularly in rare disease diagnosis [4,9]. These techniques were found to improve diagnostic outcomes, though their effectiveness remains contingent on the availability of quality training data [9].

Several works focus on the role of natural language processing (NLP) in clinical data extraction and decision support. For example, one study applied Named Entity Recognition (NER) to automate chart review processes, reducing the time healthcare professionals spend on manual tasks [5]. NLP techniques have also been used to analyze electronic health records (EHRs) and extract structured information from unstructured clinical data, improving diagnostic precision [16]. However, the reliance on labeled datasets continues to limit the scalability of NLP systems in healthcare applications [5,16,32].

Deep learning, particularly convolutional neural networks (CNNs), has proven particularly effective in medical imaging, especially in radiological image analysis. A study applying CNNs to detect abnormalities in radiological images reported accuracy rates exceeding 95% [6]. Similarly, deep learning has been used for cancer detection and diagnosis, with models demonstrating high performance in identifying cancerous lesions [11,19]. Despite these successes, challenges related to dataset biases and the lack of interpretability in AI-driven diagnostics remain critical concerns for clinicians [6,15,19].

In addition to image analysis, several studies have explored real-time clinical decision support systems (CDSS). A hybrid approach integrating machine learning models into a rule-based CDSS demonstrated improved diagnostic speed and accuracy in emergency care scenarios [7]. These systems excelled under high-pressure conditions but encountered difficulties in real-time data integration and model interpretability, which are essential for widespread clinical adoption [7,14,17].

Reinforcement learning (RL) has also been applied in healthcare, particularly in optimizing treatment strategies. One RL-based model was used to dynamically adjust treatment plans for chronic diseases based on real-time patient data, significantly improving patient outcomes [10]. However, real-world application remains a challenge due to the complexity of clinical environments and the need for ongoing adaptation [10,29,33].

AI’s role in mental health diagnosis has expanded, with several studies using machine learning to predict mental health disorders based on patient behavior and clinical data. These models have been successful in early detection of conditions like depression and anxiety, providing clinicians with valuable insights [20,21]. Nonetheless, concerns over patient data privacy and model transparency limit their broader application [20,21].

In the field of genomics, AI has facilitated personalized treatment strategies, particularly in oncology. By analyzing genomic data, one study demonstrated how AI could tailor cancer treatments to individual patients, enhancing the effectiveness of therapies [22]. Other research has focused on integrating AI with large-scale genomic datasets to identify disease markers and improve diagnostic accuracy [22,23]. However, challenges related to data privacy, scalability, and the interpretability of AI decisions remain significant [22,23,35].

Wearable technology and mobile health (mHealth) applications have also gained traction as tools for chronic disease monitoring and real-time health data collection. AI algorithms integrated into wearable devices provide predictive insights by continuously monitoring patient vitals, enabling early intervention in conditions like cardiovascular diseases and diabetes [24,27]. These innovations have the potential to revolutionize patient care, though issues related to data security, device interoperability, and patient adherence remain hurdles [24,27,30].

AI has also been applied to forecast disease outbreaks and improve public health management. One study utilized time-series data and AI models to predict epidemic trends, offering critical early warning systems for health authorities [12]. AI-enhanced telemedicine platforms are similarly growing in importance, particularly in underserved areas where remote diagnostic tools and consultations can bridge healthcare access gaps [28]. These systems have demonstrated their utility, but limited infrastructure and privacy concerns still impede widespread implementation [25,28].

In the realm of rare disease diagnosis, few-shot learning approaches have proven effective in generalizing across limited and underrepresented datasets. A study employing this technique demonstrated improved diagnostic accuracy for rare diseases, which traditional models often struggle with due to the scarcity of data [9]. Data augmentation and adversarial training have also been used to enhance model robustness in diagnosing rare conditions [18,26,36].

Lastly, AI's potential in drug discovery has attracted significant attention. One study leveraged AI to analyze molecular structures and clinical trial data, identifying potential drug candidates and accelerating the drug discovery process [31]. This application has the potential to reduce both the time and cost associated with developing new therapies [31].

In summary, the body of work on AI in healthcare spans a wide range of applications, from conversational diagnostic systems and LLMs for differential diagnosis to deep learning models for medical imaging and NLP for clinical data extraction. AI systems have shown remarkable improvements in diagnostic accuracy, patient outcomes, and workflow automation. However, challenges such as data availability, model interpretability, privacy concerns, and real-world scalability remain. Future research must focus on addressing these limitations to fully realize the potential of AI in transforming healthcare.

## III. DATASET USED

For the development and evaluation of the AI agent-based architecture for medical differential diagnosis and medical history recording and analysis, a variety of publicly available healthcare datasets were utilized. These datasets provide comprehensive and diverse medical data, allowing for effective model training, validation, and performance evaluation.

One of the key datasets used is the MIMIC-III (Medical Information Mart for Intensive Care) database[37], which is widely recognized for its extensive collection of de-identified health records from patients admitted to the intensive care unit (ICU). This dataset includes detailed information such as demographics, vital signs, laboratory results, medications, and clinical notes. The MIMIC-III database is invaluable for developing models that handle real-time, continuous data and for training the AI system to understand complex patient histories and recognize potential diagnoses.

For symptom-to-diagnosis prediction, the SymCAT [38] dataset was employed. SymCAT contains detailed symptom-based data collected from various patients, including their reported symptoms and corresponding diagnoses. This dataset is particularly useful in training the system to make accurate predictions based on patient-reported symptoms, which is a key component in medical decision-making.

Together, these datasets provide a rich source of structured and unstructured data across multiple modalities, ensuring that the AI architecture is well-equipped to handle various aspects of medical diagnosis and history analysis. Data pre-processing techniques, including de-identification, normalization, and augmentation, were applied to ensure privacy and consistency across datasets, improving the robustness and generalizability of the model.

## IV. PROPOSED METODOLOGY

This paper proposes a comprehensive multi-agent AI architecture that collects and analyses detailed patient information—ranging from current symptoms to demographic details—to generate a prioritized differential diagnosis and suggest potential treatment options. The architecture is built around multiple specialized agents, each with distinct roles, working together to gather, interpret, and process patient data. This approach aims to provide a holistic view of the patient's condition, supporting healthcare providers with efficient, data-driven insights that are both timely and relevant.

The system begins by interacting with patients to collect input on their symptoms, medical history, lifestyle factors, and demographic information. By incorporating demographic details such as age, gender, geographic location, and other context-specific factors, the model is able to tailor the diagnostic output, offering an individualized assessment that reflects the patient’s unique profile. This level of customization is especially valuable in medical diagnosis, where demographics often play a crucial role in determining disease risk factors and probable conditions.

The AI model then processes this input to generate a ranked differential diagnosis, providing healthcare providers with a list of the top three most likely diagnoses. These diagnoses are prioritized based on the probability of each condition given the patient's specific symptoms, history, and demographic information. For each diagnosis, the system presents relevant background data derived from the patient’s medical history and demographic context, enabling the provider to quickly assess the reasoning behind each diagnostic possibility. By presenting diagnoses in order of likelihood, the system aids the healthcare provider in efficiently narrowing down potential causes, ensuring that the most probable conditions are considered first in the clinical decision-making process.

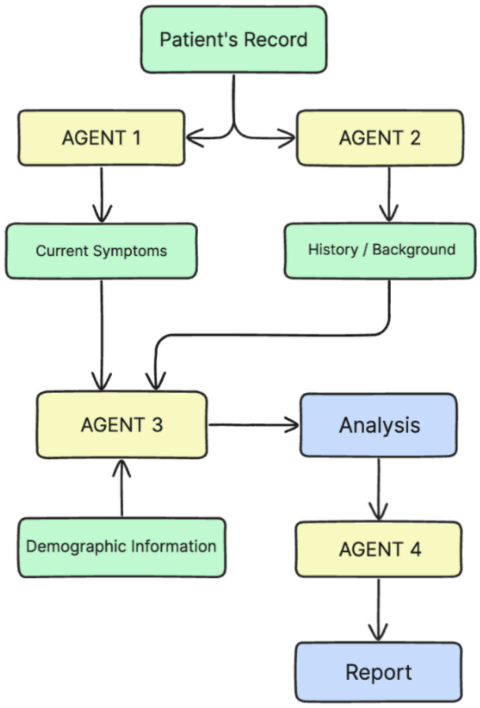


*Figure 1: Proposed methodology and flow*

In addition to the differential diagnosis, the model offers treatment suggestions for each listed condition. These suggestions are drawn from medical guidelines and tailored to the individual’s demographic profile, considering factors such as age and regional healthcare standards. The integration of treatment options provides a valuable starting point for healthcare providers, helping them make informed decisions about patient management and potentially accelerating the pathway from diagnosis to treatment.

This multi-agent system is designed not only to reduce the burden of repetitive tasks on healthcare providers but also to enhance the accuracy and relevance of diagnostic insights. By leveraging large language models and fine-tuning techniques, each agent within the architecture is optimized to handle its specific role—whether it's data collection, interpretation, or synthesis—ensuring that the system as a whole functions with high accuracy and efficiency. This paper provides a detailed overview of the architecture, training methodologies, and performance evaluation of this multi-agent approach, demonstrating its potential as a powerful tool in modern healthcare. The Flow of system can be seen in Figure 1.

Overall, this AI-driven, multi-agent approach to differential diagnosis and treatment recommendations aims to improve clinical workflows, support evidence-based medical decisions, and ultimately enhance the quality of patient care.

### *A. Agent 1*

This AI agent is designed to capture patient information in natural language and interprets details about symptoms through conversational input. By leveraging natural language processing (NLP) and LLMs like GPT, the agent can accurately understand and highlight critical information from the patient’s descriptions. It then organizes and outputs a clear summary of current symptoms for use by healthcare providers, effectively translating complex or ambiguous patient language into structured, clinically relevant data. This approach supports providers by simplifying the intake and diagnosis process, enhancing both speed and accuracy in patient care.

### *B. Agent 2*

This AI agent designed to take conversational input from a patient and extract relevant medical history listens to patient narratives, capturing details about past conditions, treatments, medications, allergies, and family medical background. Using natural language processing (NLP) techniques and large language models, the agent identifies and organizes key information into structured medical history fields, allowing for clear and concise data output. By understanding complex language, including casual phrasing and non-standard medical terms, the AI agent can accurately interpret patient statements and distill relevant details. This enables healthcare providers to quickly access an organized summary of the patient’s medical background, improving the efficiency and accuracy of diagnosis and treatment planning.

### C. Agent 3

This AI agent processes a patient's current symptoms, lifestyle information, medical history, and demographic details to provide a structured output that assists healthcare providers in clinical assessments. First, it generates a clear summary of relevant lifestyle factors—such as exercise habits, smoking, and dietary patterns—alongside essential medical history, offering a comprehensive snapshot of the patient's health background. Then, the agent analyzes the reported symptoms, identifying patterns or relationships that may indicate specific health issues. Based on this thorough evaluation, the AI delivers a prioritized list of the top three likely diagnoses, considering probabilities that reflect the patient's overall profile. Additionally, for each diagnosis, the agent suggests actionable recommendations, including potential treatments, further tests, or referrals to specialists, thereby guiding doctors toward well-informed, efficient decision-making for optimal patient care.

### D. Agent 4

This AI agent is specifically designed to take the detailed analysis provided by Agent 3 and transform it into a comprehensive, structured report. This report synthesizes the patient's symptoms, medical history, lifestyle factors, and demographic context, presenting key insights and the top differential diagnoses in a clear, accessible format. The report includes a summary of probable diagnoses with accompanying rationale, organized in order of likelihood, along with recommended next steps or treatment options tailored to each diagnosis.

By providing a concise yet thorough overview, the report helps doctors quickly grasp essential patient information, enabling them to make more accurate and timely diagnostic decisions. For the patient, this document serves as a clear summary of their health assessment, promoting understanding and encouraging active involvement in their care journey. Ultimately, the agent's ability to convert complex data into a structured, actionable report enhances communication between patients and doctors and supports more personalized, effective medical care.

This multi-agent AI system enhances patient care by streamlining data collection, analysis, and reporting for healthcare providers. Agent 1 gathers patient information in natural language, interpreting and summarizing current symptoms. Agent 2 captures relevant medical history, extracting past conditions, treatments, and family background into structured fields. Agent 3 synthesizes current symptoms, lifestyle factors, medical history, and demographics, providing an organized output with a prioritized list of the top three probable diagnoses and tailored recommendations for each. Agent 4 converts this analysis into a comprehensive, structured report, detailing symptoms, medical history, lifestyle insights, and suggested diagnoses in an accessible format. This report helps doctors make faster, well-informed decisions, while also giving patients a clear summary of their health, encouraging understanding and active involvement in their care. Overall, the system improves diagnostic accuracy and enhances communication between providers and patients.

## V. RESULTS

This paper proposed a Multi Agent system that take in a patient's record and provides a analysis to doctor. This was not meant to replace doctors but to act as an assistant. This system aims to be help health care worker by saving time and helping in provide better diagnosis. To understand how this system works, this section shows and entire flow of the system with key details about each. The patient records could be seen in Figure 2. A patients record include their current symptoms, medical history and Lifestyle information

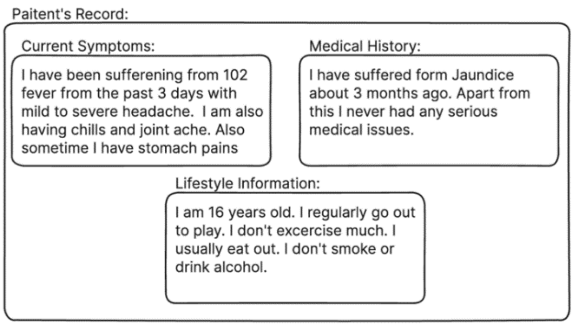


*Figure 2: Patient's Record*

Agent 1 is responsible for extracting current symptoms from the patient's conversational input. It identifies and organizes key symptoms from the patient's descriptions, ensuring that critical information is captured accurately. As shown in Figure 3, the agent efficiently processes natural language input to summarize the reported symptoms. The performance of Agent 1 has been highly effective in accurately extracting symptoms from text, even when presented in diverse or informal language. This capability enhances the diagnostic process by providing healthcare providers with a clear and concise summary of the patient's symptoms, improving both the speed and accuracy of diagnosis.

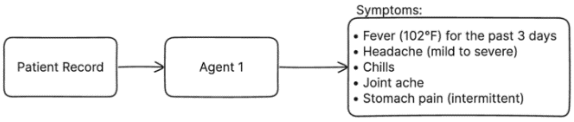


Figure 3: Working of Agent 1

Agent 2 is responsible for extracting key medical history from the patient's conversational input. It identifies important details such as past conditions, treatments, medications, allergies, and family history, organizing them into structured data points. As illustrated in Figure 34, the agent efficiently processes natural language input to summarize the medical history. Its performance in summarizing this information has been highly effective, consistently capturing relevant details even in informal language. This capability allows healthcare providers to quickly access a patient's medical background, improving the speed and accuracy of diagnosis and treatment planning.

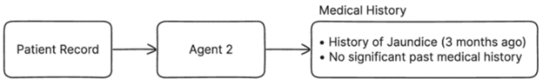


Figure 4: Working of Agent 2

Agent 3 serves as the core of the multi-agent system, integrating information provided by Agent 1 and Agent 2, along with demographic details, to perform comprehensive differential diagnosis and in-depth analysis. Acting as the central processing unit of the system, it synthesizes patient-reported symptoms, medical history, lifestyle factors, and demographic data to generate an insightful and prioritized list of potential diagnoses. Agent 3's primary function is to provide a thorough evaluation of the gathered information, identifying patterns and correlations that may not be immediately apparent. Additionally, it offers critical assistance to healthcare providers by suggesting possible treatment options, next steps, or referrals based on the analysis.

As the "brain" of the system, Agent 3 leverages advanced algorithms and machine learning models to weigh the relevance and significance of each data point, ensuring that the most likely diagnoses are highlighted, and decision-making is informed by accurate and contextually rich insights. By providing this level of analysis, Agent 3 not only supports faster and more accurate clinical decision-making but also contributes to the overall efficiency and effectiveness of patient care. Figure 5 shows the working of this agent.

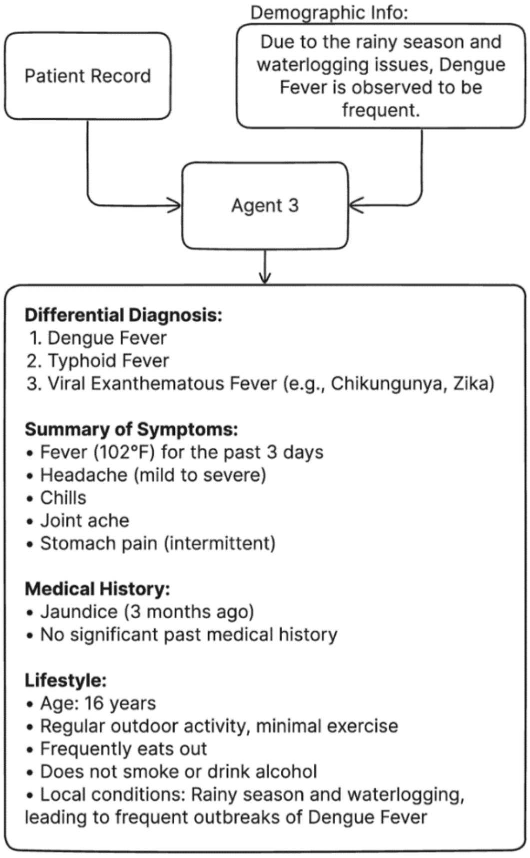


Figure 5: Working of Agent 3

Agent 4 is responsible for taking the analysis provided by Agent 3 and enhancing it by adding suggested courses of action, such as potential treatments, further tests, or referrals. It then compiles this information into a structured report for the doctor. This report includes a summary of the patient's symptoms, medical history, lifestyle factors, and the top differential diagnoses, along with recommended next steps tailored to each diagnosis. Agent 4's role is crucial in ensuring that the healthcare provider receives a clear, actionable summary, facilitating informed decision-making and enhancing patient care.

## VI. CONCLUSION

This paper introduces a multi-agent system that functions as a medical assistant, streamlining essential aspects of patient care by automating the collection of data, symptom analysis, and diagnostic support. Utilizing natural language processing (NLP) and large language models (LLMs), the system captures patient-reported symptoms, medical history, and lifestyle details in a conversational format, ensuring accurate interpretation and organization of even complex or vague patient input. It then synthesizes this data to generate a prioritized list of possible diagnoses and personalized treatment recommendations, aiding in informed clinical decision-making. Additionally, the system creates a well-structured report summarizing the patient’s health profile, improving communication between healthcare providers and patients.

This solution not only saves time by automating routine tasks but also enhances diagnostic accuracy and enables more personalized care. The results highlight the system’s potential to improve patient-provider interactions and support better decision-making. While challenges in handling unclear or incomplete input persist, future developments will focus on refining these capabilities. Ultimately, this multi-agent medical assistant has the potential to transform healthcare workflows, promoting faster, more accurate, and efficient patient care.

## VII. FUTURE SCOPE

The future scope of this research involves enhancing the multi-agent system's effectiveness in real-world healthcare settings. Key improvements include refining natural language understanding to better handle ambiguous or informal patient input, and incorporating advanced machine learning techniques to interpret diverse communication styles. Integrating real-time data from electronic health records (EHRs) and wearable devices could further enrich the system’s analysis.

Expanding the system’s diagnostic capabilities by incorporating specialized medical knowledge and implementing continuous learning will enable more accurate diagnoses. Additionally, addressing data privacy and security concerns, through robust encryption and compliance with regulations like HIPAA, will be crucial for maintaining patient confidentiality. These advancements will increase the system's utility, accuracy, and adoption in healthcare environments.

## REFERENCES

[1] Tu, T., Palepu, A., Schaekermann, M., Saab, K., Freyberg, J., Tanno, R., ... & Natarajan, V. (2024). Towards conversational diagnostic ai. arXiv preprint arXiv:2401.05654.

[2] McDuff, D., Schaekermann, M., Tu, T., Palepu, A., Wang, A., Garrison, J., ... & Natarajan, V. (2023). Towards accurate differential diagnosis with large language models. arXiv preprint arXiv:2312.00164.

[3] Abidi, M. H., Alkhalefah, H., & Aboudaif, M. K. (2024). Enhancing Healthcare Data Security and Disease Detection Using Crossover-Based Multilayer Perceptron in Smart Healthcare Systems. CMES-Computer Modeling in Engineering & Sciences, 139(1).

[4] Nazary, F., Deldjoo, Y., Di Noia, T., & di Sciascio, E. (2024). XAI4LLM. Let Machine Learning Models and LLMs Collaborate for Enhanced In-Context Learning in Healthcare. arXiv preprint arXiv:2405.06270.

[5] Ahmed, U., Iqbal, K., & Aoun, M. (2023). Natural language processing for clinical decision support systems: A review of recent advances in healthcare. Journal of Intelligent Connectivity and Emerging Technologies, 8(2), 1-16.

[6] Panagoulias, D. P., Virvou, M., & Tsihrintzis, G. A. (2024). Evaluating LLM--Generated Multimodal Diagnosis from Medical Images and Symptom Analysis. arXiv preprint arXiv:2402.01730.

[7] Ahmad, N., Du, S., Ahmed, F., ul Amin, N., & Yi, X. (2023). Healthcare professionals satisfaction and AI-based clinical decision support system in public sector hospitals during health crises: a cross-sectional study. Information Technology and Management, 1-13.

[8] Desai, M., & Shah, M. (2021). An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). Clinical eHealth, 4, 1-11.

[9] Galitsky, B. A. (2024). LLM-Based Personalized Recommendations in Health.

[10] Chen, X., Mao, X., Guo, Q., Wang, L., Zhang, S., & Chen, T. (2024, August). RareBench: Can LLMs Serve as Rare Diseases Specialists?. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 4850-4861).

[11] Andrew, J. J., Vincent, M., Burgun, A., & Garcelon, N. (2024, May). Evaluating LLMs for Temporal Entity Extraction from Pediatric Clinical Text in Rare Diseases Context. In Proceedings of the First Workshop on Patient-Oriented Language Processing (CL4Health)@ LREC-COLING 2024 (pp. 145-152).

[12] Shoham, O. B., & Rappoport, N. (2023). Cpllm: Clinical prediction with large language models. arXiv preprint arXiv:2309.11295.

[13] Liu, F., Zhu, T., Wu, X., Yang, B., You, C., Wang, C., ... & Clifton, D. A. (2023). A medical multimodal large language model for future pandemics. NPJ Digital Medicine, 6(1), 226.

[14] Nayak, S. K., Garanayak, M., Swain, S. K., Panda, S. K., & Godavarthi, D. (2023). An Intelligent Disease Prediction and Drug Recommendation Prototype by Using Multiple Approaches of Machine Learning Algorithms. IEEE Access.

[15] Krishna, C. L., & Reddy, P. V. S. (2019, February). An efficient deep neural network multilayer perceptron based classifier in healthcare system. In 2019 3rd International Conference on Computing and Communications Technologies (ICCCT) (pp. 1-6). IEEE.

[16] Jin, M., Yu, Q., Shu, D., Zhang, C., Fan, L., Hua, W., ... & Meng, Y. (2024). Health-llm: Personalized retrieval-augmented disease prediction system. arXiv preprint arXiv:2402.00746.

[17] Dou, Y., Huang, Y., Zhao, X., Zou, H., Shang, J., Lu, Y., ... & Peng, S. (2024). ShennongMGS: An LLM-based chinese medication guidance system. ACM Transactions on Management Information Systems.

[18] Safar, A. A., Salih, D. M., & Murshid, A. M. (2023). Pattern recognition using the multi-layer perceptron (MLP) for medical disease: A survey. International Journal of Nonlinear Analysis and Applications, 14(1), 1989-1998.

[19] Jeong, D. P., Lipton, Z. C., & Ravikumar, P. (2024). Llm-select: Feature selection with large language models. arXiv preprint arXiv:2407.02694.

[20] Leonardi, G., Portinale, L., & Santomauro, A. (2023). Enhancing medical image report generation through standard language models: leveraging the power of LLMs in healthcare. In In2nd AIxIA Workshop on Artificial Intelligence For Healthcare (pp. 41-45).

[21] Amin, M. Z., & Ali, A. (2017, February). Application of multilayer perceptron (MLP) for data mining in healthcare operations. In 2017 3rd International Conference on Biotechnology, Lahore, Pakistan (p. 2e11).

[22] Li, L., Zhou, J., Gao, Z., Hua, W., Fan, L., Yu, H., ... & Ma, S. (2024). A scoping review of using Large Language Models (LLMs) to investigate Electronic Health Records (EHRs). arXiv preprint arXiv:2405.03066.

[23] Jiang, Y., Irvin, J. A., Ng, A. Y., & Zou, J. (2023). VetLLM: Large Language Model for Predicting Diagnosis from Veterinary Notes. In PACIFIC SYMPOSIUM ON BIOCOMPUTING 2024 (pp. 120-133).

[24] Omar, M., Brin, D., Glicksberg, B., & Klang, E. (2024). Utilizing natural language processing and large language models in the diagnosis and prediction of infectious diseases: A Systematic Review. American Journal of Infection Control.

[25] Ahsan, H., McInerney, D. J., Kim, J., Potter, C., Young, G., Amir, S., & Wallace, B. C. (2024). Retrieving evidence from ehrs with llms: Possibilities and challenges. Proceedings of machine learning research, 248, 489.

[26] Elsborg, J., & Salvatore, M. (2023). Using LLMs and Explainable ML to Analyze Biomarkers at Single-Cell Level for Improved Understanding of Diseases. Biomolecules, 13(10), 1516.

[27] Bhuyan, H. K., Ravi, V., Brahma, B., & Kamila, N. K. (2022). Disease analysis using machine learning approaches in healthcare system. Health and Technology, 12(5), 987-1005.

[28] Cui, H., Shen, Z., Zhang, J., Shao, H., Qin, L., Ho, J. C., & Yang, C. (2024). LLMs-based Few-Shot Disease Predictions using EHR: A Novel Approach Combining Predictive Agent Reasoning and Critical Agent Instruction. arXiv preprint arXiv:2403.15464.

[29] Mashor, M. I. (2024). Developing a data-driven clinical decision support system for global healthcare: A us-Nigeria collaborative project.

[30] Wang, H., Zhao, S., Qiang, Z., Xi, N., Qin, B., & Liu, T. (2024). Beyond Direct Diagnosis: LLM-based Multi-Specialist Agent Consultation for Automatic Diagnosis. arXiv preprint arXiv:2401.16107.

[31] Sharan Prabahar Balasubramanian, N., & Dakshit, S. (2024). Can Public LLMs be used for Self-Diagnosis of Medical Conditions?. arXiv e-prints, arXiv-2405.

[32] Cardenas, L., Parajes, K., Zhu, M., & Zhai, S. (2024, January). AutoHealth: Advanced LLM-Empowered Wearable Personalized Medical Butler for Parkinson's Disease Management. In

2024 IEEE 14th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 0375-0379). IEEE.

[33] Hong, S., Xiao, L., Zhang, X., & Chen, J. (2024). ArgMed-Agents: Explainable Clinical Decision Reasoning with Large Language Models via Argumentation Schemes. arXiv preprint arXiv:2403.06294.

[34] Li, Q., Yang, X., Wang, H., Wang, Q., Liu, L., Wang, J., ... & Group, G. Z. A. (2023). From beginner to expert: Modeling medical knowledge into general llms. arXiv preprint arXiv:2312.01040.

[35] Zhou, H., Gu, B., Zou, X., Li, Y., Chen, S. S., Zhou, P., ... & Liu, F. (2023). A survey of large language models in medicine: Progress, application, and challenge. arXiv preprint arXiv:2311.05112.

[36] Singhal, K., Azizi, S., Tu, T., Mahdavi, S. S., Wei, J., Chung, H. W., ... & Natarajan, V. (2023). Large language models encode clinical knowledge. Nature, 620(7972), 172-180.

[37] Johnson, A., Pollard, T., Shen, L. et al. MIMIC-III, a freely accessible critical care database. Sci Data **3**, 160035 (2016). https://doi.org/10.1038/sdata.2016.35

[38] A. R. Inc., “Symcat: Symptom-based, computer assisted triage,” http://www.symcat.com/, 2020.