CS 510 Team 8 Project - Development Track

Wangzhe Sun (wangzhe2@illinois.edu)
Chenyu Song (chenyus3@illinois.edu)
Chaobo Cai (chaoboc2@illinois.edu)
Lujiajie Shen (ls34@illinois.edu)

Functions and Users: This project aims to develop a web-based clinical decision support system utilizing text mining techniques and medical data. Users input symptom descriptions, which are analyzed to extract topics, make diagnoses, and provide treatment recommendations. Target users include healthcare professionals and individuals seeking preliminary health information before consulting a medical provider.

Significance: The clinical decision support system streamlines medical decision-making by leveraging technology to efficiently analyze symptoms, providing fast and accurate diagnoses and treatment recommendations. It addresses the pain points of complex medical knowledge and time-consuming diagnosis processes, benefiting both healthcare professionals and patients. By offering a self-diagnosis tool for patients and aiding professionals in recalling complex medical information, it enhances patient care and outcomes, addressing the societal need for accessible and efficient healthcare solutions. This is crucial for reducing healthcare costs and promoting social justice.

Approach: Our application involves three stages. First, we preprocess text using NLTK or spaCy to extract keywords. Next, we apply the Latent Dirichlet Allocation (LDA) algorithm from libraries like Gensim for topic modeling. Then, we train a classification model using the "DISEASE PREDICTION USING MACHINE LEARNING WITH GUI" dataset from Kaggle. Finally, we generate diagnostic results based on the outcomes of the trained model and the classification model.

Evaluation: After implementation, the project can be tested with sample text data as test cases. The data will contain the patient's symptom descriptions and true diagnosis. Then we can compare the project generated result with the true result to evaluate its performance. Additionally, the system would be tested with a diverse set of sample patient cases covering various medical conditions. This ensures that the system performs consistently and reliably across different scenarios.

Timeline: Our timeline is as follows

- Backend & Frontend (2 weeks)
- Integration, Deployment & debugging (1 week)
- Evaluation (1 week)
- Buffer (2 days)

Task Division: Our task division is as follows

- Chenyu Song (Coordinator): Backend algorithms including LLM API and NN.
- Chaobo Cai: Full-stack development with a focus on the backend side
- Wangzhe Sun: Backend framework setup, logic development, and deployment
- Lujiajie Shen: Frontend UI including text/file input and output display