Project Title: Comprehensive Drug Interaction Analysis and Clinical Decision Support through Sophisticated Natural Language Processing and Multifaceted Data Integration within Electronic Health Records

Abstract

Drug interactions are a critical challenge in healthcare, often leading to adverse drug events, extended hospital stays, and significant healthcare costs. Existing systems rely heavily on static databases, which lack contextual and patient-specific insights from electronic health records (EHRs) or unstructured clinical notes. Leveraging natural language processing (NLP) and structured data integration offers the potential to transform drug interaction analysis and clinical decision-making.

This research aims to integrate structured drug interaction data with patient-specific EHR information to develop an intelligent system capable of real-time clinical decision support. Structured data, such as DrugBank's comprehensive database, will be combined with unstructured clinical notes accessed through the MIMIC-III clinical databases using advanced NLP techniques such as domain-specific entity recognition using BioBERT, sentence parsing with SpaCy, and Transformers for data tokenization and vectorization.

The methodology involves extracting and integrating critical data such as drugs, medical conditions, and observations from multiple sources. NLP models will identify relationships between entities and predict potential adverse outcomes. The unified framework will enable clinicians to access precise, patient-tailored drug interaction insights, enhancing safety and accuracy in prescribing. By bridging gaps, this research could establish a groundbreaking model for personalized clinical support systems, offering a scalable approach to improving drug safety in diverse healthcare settings.