Health-care systems generate massive amounts of unstructured data such as clinical notes, radiology reports, discharge summaries, prescriptions, and laboratory observations. Manual interpretation of this data is time-consuming, error-prone, and often leads to missed critical information. Natural Language Processing (NLP) provides the ability to automatically extract valuable insights from clinical text, enabling faster decision-making and improved patient outcomes. This research focuses on developing an NLP-based system to extract key medical information and support clinicians in delivering efficient and accurate care.

Problem Statement

A large portion of Electronic Health Records (EHRs) is unstructured, making it difficult for physicians to access meaningful insights quickly. Important details such as symptoms, diagnoses, medication, and allergies are often buried inside lengthy notes. There is a critical need for an automated system that can analyze clinical text, extract relevant entities, and provide real-time decision support.

Objectives

1. To design an NLP model capable of identifying and extracting medical entities such as diseases, symptoms, medications, and lab values.
2. To build a classification module for detecting potential clinical issues (e.g., drug interactions, abnormal symptoms)
3. To integrate the extracted information into a clinical decision support interface.
4. To evaluate the system using real or publicly available medical datasets.