ChatFDA: Medical Records Risk Assessment

Charlie Sun
Language Technologies Institute
Carnegie Mellon University
cssun@andrew.cmu.edu

Minh Tran
Robotics Institute
Carnegie Mellon University
minht@cs.cmu.edu

Abstract

In the realm of healthcare, patient safety and the reduction of medical errors are paramount. Despite well-intentioned efforts, healthcare systems, particularly in low-resource regions, often lack the robustness to prevent these errors effectively. This research introduces a novel mobile application designed to mitigate this issue by aiding caregivers in assessing potential risk from medical notes. Utilizing multiple input modalities and data from openFDA to provide real-time, actionable insights on prescriptions. Preliminary results on MIMIC-III dataset indicate a proof of concept to reduce in medical errors and an enhancement in patient safety. This application has the potential to significantly improve healthcare outcomes in resource-poor settings. For reproducibility and further research, the complete code for our method is made available on Github. DISCLAIMER: We used ChatGPT as an initial step explore research ideas for this project.

1 Introduction

Every year in the United States only, there are 250000 death due to medical errors, these could come from diagnostics error, surgery, patient care infection, and largely (44%) from medication. And this is the situation in a country will well-developed healthcare system with advanced medical facilities, well-organized and digitalized records with careful maintenance. In low-resource regions, particularly rural areas in third-world countries, the challenge of ensuring patient safety is exacerbated. There are regions where only one nurse need to take care a whole tribe village, and she/he doesn't have the access to a standardized medical records, sometimes it is just hand-written notes or even from parole of the patient from memory.

2 Related Works

2.1 Medical Records Correction

The correction of medical records has been a topic of interest in healthcare informatics for several years. Various studies have explored the use of Electronic Health Records (EHRs) to improve the accuracy and reliability of medical data. Previously probabilistics models have been use to identify and correct inconsistencies in medical records, particularly in the context of medication prescriptions. Recently, large language models (GPT4, PaLM) have been used for error correction in general, but there is no benchmark exist for medical data. These works lay the foundation for the importance of accurate medical record-keeping, which our research aims to extend into evaluating the risk from the records.

2.2 Language Models for Medical Data

The application of language models in the medical domain is a burgeoning field, especially with the advent of more sophisticated models like Med-PaLM 2 and Med-Bert. These models have been

Preprint. Under review.

employed for various tasks, such as medical text summarization, diagnosis prediction, and drug interaction identification. However, there is no existing research on evaluating the efficiency of using LLMs in increasing patient safety, especially for multi-language settings where the information is limited. Our research builds upon these advancements by integrating language models to interpret and verify medical notes, particularly in settings where expert oversight is limited. This work gives an overview about potential research and application directions of LLMs in medical settings.

Both of these areas—medical records correction and language models for medical data—provide valuable insights and foundational knowledge that inform and support the objectives of our research. By combining elements from these two domains, we aim to create an application that can significantly improve patient safety and reduce medical errors in low-resource settings.

3 Proposed Approach

3.1 Pipeline Design

The architecture of our application is organized into a pipeline consisting of several interconnected modules. The first module collects and process medical data, our input could be text, voice, or image. For this experiment, we focused on medical notes since this is a reliable source of information for the caregivers. With the MIMIC-III dataset, we need a module for data processing to standardizes the collected data for analysis. Subsequently, from the raw data, we use GPT4 block to process the medical records in to prescription and medical history part. This step reduce the human error in writting notes and the extract the type of prescriptions of each patient. The prescription then will be sent to openFDA, providing actionable insights like medication interactions and treatment guidelines. The final block of the pipeline take the insights from openFDA and medical history of the patient and return a risk evaluation, as well as save the medical records to the database for the future use.

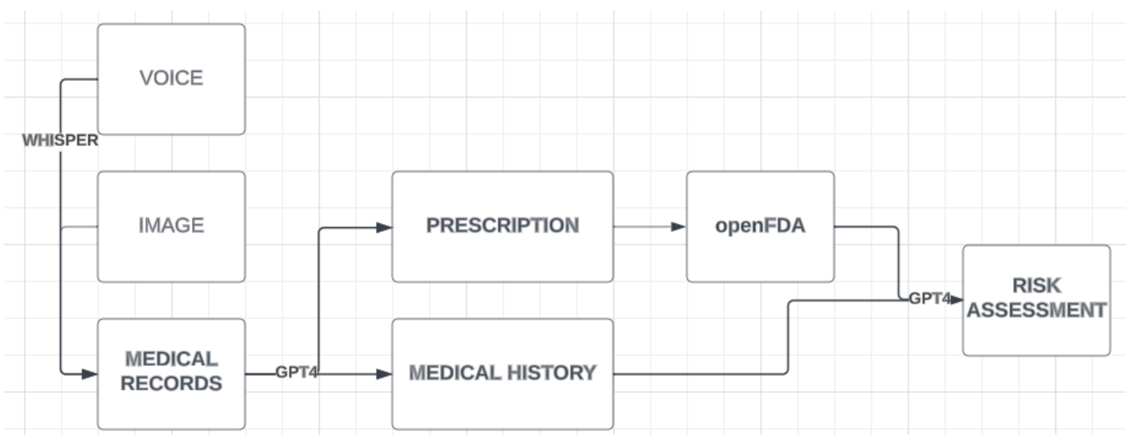

Figure 1: Pipeline Diagram

3.2 Prompt Design

Designing a useful prompt for a language model is critical, especially in high-stakes situations like medical evaluations. In this study, we introduce two separate prompts:

1. The first is to extract relevant medical information from a doctor's notes. This information pertains to a patient's pre-existing conditions, symptoms, and prescribed medications.
2. The second prompt aims to assess the potential risks related to the extracted medical information. In this section, we elaborate on our approach to designing this risk-assessment prompt.

Prompt Structure: The risk assessment prompt comprises three primary sections:

- **Parsed Notes:** This section integrates the parsed information from the doctor's notes.
- **Drug Information:** Here, we incorporate drug interactions and warnings retrieved from the FDA database.
- **Answer Section:** The language model is tasked to analyze the prescribed treatment, identifying potential drug interactions and assessing potential patient reactions based on their pre-existing conditions. The culmination of this analysis is the evaluation of the treatment's dangerousness on a scale: LOW, MEDIUM, or HIGH. An answer template is also provided to ensure consistency in responses.

Actual Prompt:

```
"I am a doctor, and I need you to evaluate my prescription:
{parsed_notes}

Drug contexts:
{drug_info_string}

Please answer the following in a concise point format, considering the provided drug context:
- Possible interactions between the prescribed drugs?
- Specific adverse effects of the drugs that relate to the patient's pre-existing conditions and

Conclude your response by assessing the treatment's dangerousness based on interactions and adver

Your answer should adhere to this format:
* INTERACTIONS:
- <interaction 1>
- <interaction 2>
- ...

* ADVERSE EFFECTS:
- <adverse effect 1>
- <adverse effect 2>
- ...

* DANGEROUSNESS: <LOW / MEDIUM / HIGH>

Include only necessary interactions or adverse effects in your response."
```

3.3 User Interface

The user interface of the application is intentionally designed to be intuitive and user-friendly. It features a series of prompts that guide healthcare workers through the verification process. Initially, users are prompted to select their preferred method of input—image, voice, or text. Based on this selection, the application provides an interface for capturing the image, recording the voice, or typing the text. Once the data is processed and analyzed, a summary report is displayed, and users are prompted to confirm its accuracy. They are also prompted to review the actionable insights generated from openFDA data, allowing for more informed decision-making.

By synergizing an efficient pipeline with user-friendly prompts, our proposed approach aims to offer a robust and accessible application capable of significantly reducing medical errors and improving patient safety, particularly in low-resource settings.

4 Results Analsysis

We tested our application on a public sample of the MIMIC-III dataset that contains de-identified health data associated with over 40,000 patients who stayed in critical care units of the Beth Israel Deaconess Medical Center. For this scope of this project, we only test on a small public sample, and focused on medical notes only.

Yesterday

admission date discharge date date birth sex service medicine allergy known allergy adverse drug reactio st name un chief complaint shortness breath major surgical invasive procedure esophagogastroduodenoscopy endoscopic clipping intubation extubation history present illness yo male history recent admission av block presumed lyme disease htn dm prior imi w systolic chf ef h gib unclear etiology presented acute onset dyspnea lying bed home lasting two hour patient discharged cardiology service following admission dyspnea found new heart block elevated troponin st change ekg concern new onset av block secondary lyme disease patient discharged home ceftriaxone also restarted aspirin admission tonight patient started dyspnea home rest pt denies chest pain complains nausea episode non bloody emesis looked dark brown got go bathroom felt lightheaded fell hit head lose consciousness pt endorses dark stool noticed since starting iron pt denies fever cough abdominal pain ed initial v ra exam significant pale conjunctiva guaiac positive dark stool lab notable hct previous hct day ago wbc potassium bicarb creatinine lactate inr ekg significant sinus rhythm street address elevation ii iii avf twi avl felt consistent prior two gauge iv placed patient transfused unit prbcs additional unit crossmatched ng lavage marroon return clear patient given protonix bolus gtt potassium patient received calcium chloride insulin albuterol amp sodium bicarb emergency department patient noted worsening dyspnea telemetry became bradycardiac three four epidoses lasting approximately one minute patient heart rate improved spontaneously require atropine transfer patient sinus tach sbp arrival micu patient feel comfortable endorses intermittent dyspnea chest pain ng tube place draining dark brown red fluid denies abdominal pain nausea vomiting diarrhea past medical history schf ef reported dm complicated neuropathy ckd ac htn hl ckd baseline cr chronic anemia uncertain etiology baseline high chronic leukocytosis chronic gi bleed uncertain etiology barretts esophagus prior sbo adhesion p loa social history life hospital wife bedbound m name stitle full time caretaker one son life home retired former pack year smoker quit year ago former beer drinker denies illicits family history mother died name ni father died cancer grandfather died mi dm physical exam admission general appearance acute distress eye conjunctiva perrl lymphatic cervical wnl cardiovascular normal normal peripheral vascular bilateral dp pulse respiratory chest crackle base bilaterally abdominal soft non tender bowel sound present skin warm pertinent result lab admission wbc rbc hgb hct mcv mch mchc rdw glucose urea n creat sodium potassium chloride total co anion gap lactate k pt ptt inrpt hematocrit blood hct blood hct blood hct pm blood hct blood hct pm blood hct lactate blood lactate k blood lactate blood lactate microbiology blood culture x ngtd urine culture growth lyme serology antibody b burgdorferi detected eia imaging tte left atrium mildly dilated left ventricular wall thickness normal posterior wall thin fibrotic akinetic left ventricular cavity size normal overall left ventricular systolic function moderately depressed lvef secondary akinesis inferior posterior wall tissue doppler imaging suggests increased left ventricular filling pressure pcwp mmhg right ventricular free wall thickness normal right ventricular chamber size normal depressed free wall contractility aortic valve leaflet mildly thickened minimally increased gradient consistent minimal aortic valve stenosis mitral valve leaflet mildly thickened mitral valve prolapse moderate mitral regurgitation seen tricuspid valve leaflet mildly thickened moderate pulmonary artery systolic hypertension trivial physiologic pericardial effusion echocardiographic sign tamponade compared finding prior study image reviewed left ventricular ejection fraction reduced secondary extensive inferior posterior wall dysfunction cxr chin elevated tip endotracheal tube upper margin clavicle le cm carina probably acceptable position tube could advanced mm secure seating pulmonary edema mild atelectasis left base new mild cardiomegaly stable pleural effusion pneumothorax brief hospital course mr known lastname year old male history av nodal blockade htn dm prior imi systolic chf ef h gib unclear etiology presenting dyspnea hematemesis secondary upper gib well myocardial ischemia setting gib gi bleed upper gi bleed demonstrated hematemesis ng lavage bloody fluid initially given unit prbcs l ivf improvement hemodynamics patient evidence active end organ ischemia given troponin elevation st change ekg elevated lactate patient received total unit prbc well one ffp platelet transfusion gi saw patient performed endoscopy twice first provide adequate visualization due significant bleeding second endoscopy visualized vascular lesion consistent dieulafoy lesion clipped post procedure patient remained hemodynamically stable stable hct require transfusion hematocrit remained stable floor ppi transitioned iv po repeat endoscopy day discharge showed barretts biopsy taken repeat egd week myocardial ischemia patient likely demand ischemia setting gib without chest pain patient troponin elevation prior hospitalization setting renal failure repeat tte performed compared finding prior study image reviewed left ventricular ejection fraction reduced secondary extensive inferior posterior wall dysfunction atrius cardiology evaluated patient beta blocker initially held acute gi bleed restarted stable heart rhythm stable occasional nd degree block similar previous hospitalization asa restarted need restarted discretion pcp cardiologist restarted home dos lisinopril hctz restarted mg metoprolol succinate follow atrius cardiology lyme carditis av block patient presented osh new onset high grade av block narrow complex junctional escape rhythm patient currently undergoing empiric treatment lyme disease given history tick exposure initial lyme serology negative repeated still negative continued ceftriaxone project day course end cardiology feel pacer indicated time given improvement treatment hyperkalemia unclear etiology improved ed following administration calcium bicarb insulin likely secondary ckd potassium normalized wnl time transfer floor remained stable chf tte ef hospitalization history ef prior tte repeat tte ckd creatinine increased baseline possibly setting poor perfusion setting hemorrhage patient cr remained elevated time transfer slowly trended back toward baseline leukocytosis baseline elevated wbc additional elevation felt secondary inflammatory state created gi bleed myocardial ischemia dm continued home dose lantus insulin sliding scale transitional issue need asa restarted need lab checked pcp follow visit need uptitration bb tolerated code status full communication son name ni telephone fax wife name ni telephone fax follow appts gi cardiology id pcp medication admission ceftriaxone g iv qh course complete simvastatin mg daily insulin glargine unit qhs omeprazole mg hospital ferrous sulfate mg hospital aspirin mg daily discharge medication atorvastatin mg tablet sig one tablet po daily daily disp tablet refill insulin glargine unit ml solution sig eighteen unit subcutaneous bedtime omeprazole mg capsule delayed releasee c sig one capsule delayed releasee c po twice day ceftriaxone dextroseiso o gram ml piggyback sig two g intravenous qh every hour day last day completed disp q vial refill sodium chloride syringe sig see ml injection qh every hour needed line flush sodium chloride flush ml iv qh prn line flush peripheral line flush ml normal saline every hour prn heparin porcine pf unit ml syringe sig see ml intravenous prn needed needed line flush heparin flush unit ml ml iv prn line flush picc heparin dependent flush ml normal saline followed heparin daily prn per lumen order filled pharmacy dosage form syringe strength unit ml metoprolol succinate mg tablet extended release hr sig one tablet extended release hr po day disp tablet extended release hr refill lisinopril mg tablet sig one tablet po daily daily hydrochlorothiazide mg capsule sig one capsule po daily daily discharge disposition home service facility year digit discharge diagnosis primary diagnosis upper gastrointestinal bleed dieulafoys lesion non st elevation myocardial infarction secondary demand ischemia secondary diagnosis chronic systolic congestive heart failure pulmonary hypertension chronic kidney disease barretts esophagus hypertension hyperlipidemia discharge condition mental status clear coherent level consciousness alert interactive activity status ambulatory requires assistance aid walker cane discharge instruction dear mr known lastname pleasure caring hospital admitted serious gastrointestinal bleed required endoscopic procedure intensive care unit procedure bleeding controlled blood test following intervention stable heart trouble previous hospitalization kept close eye well made following change medication continue ceftriaxone g iv daily stopped simvastatin start atorvastatin mg instead changed metoprolol mg daily continue take med prescribed weigh every morning name md md weight go lb followup instruction department infectious disease monday first name namepattern name md md telephone fax building lm hospital unit name hospital campus west best parking hospital ward name garage name last name lf first name lf location location un university college primary care address hospital university college numeric identifier phone telephone fax appt thursday also need seen cardiologist one month follow please call make appointment also need repeat endoscopy week called gi department schedule hearsd e week need call telephone fax schedule

Figure 2: Processed Medical Notes

The results indicate that the our application is effective in verifying medical notes, integrating real-time data for informed decision-making, and improving the user experience for healthcare workers. Most importantly, the application shows promise in its primary objective—reducing medical errors and enhancing patient safety.

5 Limitation and Conclusion

In this project, we introduce an application aimed at reducing medical errors and enhancing patient safety, particularly in low-resource settings. We demonstrate the app's capability to verify medical notes through various input methods and leverage openFDA data for informed decision-making.

However, our work has limitations. The application's effectiveness is currently tested on a limited dataset, questioning its generalizability. Additionally, the reliance on real-time data integration could pose challenges in regions with poor internet connectivity.

Future directions should focus on expanding the dataset for more robust testing and exploring offline capabilities to make the application more versatile. This work serves as a stepping stone for leveraging technology to improve healthcare outcomes in resource-poor settings.

References