Advanced Readings in Linguistics LING-L 690

Guided by : Professor Damir Cavar

Hardik Asnani

December 10, 2023

LLM-Driven Healthcare Transformation

Abstract

The healthcare and hospice industries are at the crossroads of human compassion and technical innovation. They must strike a delicate balance between their unshakable commitment to patient care and the limitations imposed by time, labor, and the growing demand for their services. This study explores the innovative potential of large language models, or LLMs, focusing on how they can be used to automate the intricate process of medical documentation. Using the power of open LLMs, like the ones the HuggingFace team invented, the goal is to effectively automate the thorough documentation of patient care in a way that satisfies strict privacy regulations for healthcare while still being reasonably priced.

Introduction

The healthcare landscape is characterized by the commitment of professionals and volunteers who tirelessly provide care to an ever-increasing patient population. However, their efforts are often impeded by the burdensome task of detailed documentation. Large Language Models, especially from the Hugging Face library, such as deepset-roberta-base-squad2, distilbert-base-cased-distilled-squad, etc, present an innovative solution to address the challenge of operational inefficiencies in healthcare documentation. This paper explores the intersection of human-centered care and technological advancement, emphasizing the potential of LLMs to alleviate the burdens associated with documentation.

The research's focus is on creating systems that can generate accurate, consistent, and secure data needed to fill out medical forms. The envisioned outcome is a future where healthcare professionals and volunteers can dedicate more time to direct patient care, unencumbered by the administrative overhead of paperwork.

Motivation

The motivation for this research stems from the pressing need to bridge the gap between the compassionate delivery of healthcare and the growing demands on healthcare professionals. The convergence of technological innovation and human-centered care represents an opportunity to revolutionize the way medical documentation is handled. By automating the extraction of relevant information from medical conversations and dictations, LLMs have the potential to enhance operational efficiencies, ultimately improving patient care.

Research Objectives

The primary objective of this research is to leverage Large Language Models (LLMs) to automate the extraction of critical information from medical transcriptions. The research aims to address specific questions related to patient data, enhancing the efficiency of healthcare documentation. The key objectives include:

1. **Automated Information Extraction:** Implementing LLM-based logic/chains to automatically extract essential information from medical transcriptions. This includes, but is not limited to, extracting the patient's name, age, prescribed medications, symptoms, and precautions advised by the doctor.
2. **Seamless Process Integration:** Developing a seamless and efficient process that integrates LLM capabilities into existing healthcare documentation workflows. This integration should empower healthcare professionals by providing them with timely and accurate information, ultimately fostering a more patient-centric approach to care.

3. **Utilization of Open-Access LLMs:** Utilizing open-access and permissively licensed LLMs, such as deepset-roberta-base-squad2, distilbert-base-cased-distilled-squad, to generate predictions for the dataset. This approach ensures that the developed solution is not only innovative but also aligns with privacy and regulatory requirements in the healthcare sector.

4. **Enhanced Healthcare Professionals' Workflow:** Enabling healthcare professionals to focus more on patient care by relieving them of labor-intensive documentation tasks. The research aims to contribute to the enhancement of the overall workflow, allowing professionals to devote more time and attention to their patients.

5. **Scalability and Generalization:** Designing the solution with scalability and generalization in mind to accommodate diverse medical conversations and dictations. The objective is to create a versatile system that can adapt to various scenarios within the healthcare domain.

In summary, this research seeks to revolutionize healthcare documentation by harnessing the power of LLMs. The envisioned outcome is an intelligent and automated system that not only meets the specific requirements of the competition but also sets the stage for transformative advancements in healthcare information management.

Dataset Description

The dataset utilized in this project comprises a total of 2000 entries, each representing either a medical transcript or a conversation between a doctor and a patient. The entries are diverse and capture various medical scenarios, providing a rich source of information for the development and evaluation of the Large Language Model (LLM)-powered information extraction system.

Data Structure

Each entry in the dataset encapsulates a textual representation of a medical interaction. The dataset is designed to simulate real-world scenarios, encompassing a wide range of medical conditions, patient demographics, and doctor-patient interactions. The entries are anonymized or synthesized to comply with privacy and ethical considerations in healthcare data usage.

Target Questions for Information Extraction

The primary purpose of the dataset is to facilitate the extraction of critical information relevant to healthcare documentation. The LLM-based system aims to answer the following five questions from each entry:

1. **Patient's Name:** Identify and extract the name of the patient involved in the medical interaction.

2. **Patient's Age:** Extract information about the age of the patient.

3. **Patient's Condition:** Capture details about the patient's medical condition or overall health status.

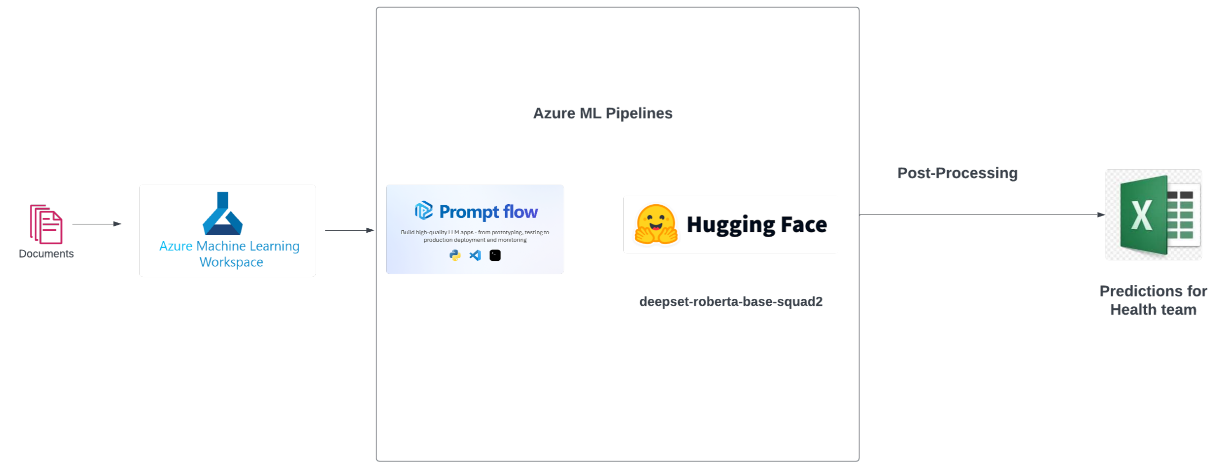
4. **Symptoms Experienced:** Identify and extract information regarding the symptoms reported by the patient.

5. **Doctor's Prescribed Drug:** Extract details about the medication or drug prescribed by the doctor.

6. **Doctor's Advised Precautions:** Capture information about any precautions or recommendations given by the doctor to the patient.

The diversity within the dataset allows for comprehensive testing and validation of the information extraction system, ensuring its robustness across various medical contexts.

Architectural Diagram



(1)

Architecture Explanation

1. **Translation Function:** The script begins by integrating transcripts from diverse languages such as Hindi, Arabic, Spanish, and French. It employs a translation function, leveraging the **googletrans** library, to uniformly translate medical transcripts into English. This ensures a standardized language for further processing and analysis.

2. **Azure ML Service Call:** The core functionality involves constructing a data payload with translated transcripts and sending a request to Azure ML service's scoring endpoint.

3. **API Key and Headers:** To secure communication, an API key is used for authentication, and headers specifying content type and authorization are set.

4. **Data Retrieval and Parsing:** The script focuses on retrieving data from the Azure ML service, particularly utilizing the LLM (Large Language Model) deepset-roberta-base-squad2. The response from the service is then parsed to extract essential information, including scores and answers. This step is crucial for further analysis and presentation of the obtained results.

5. **Question List:** A predefined list of patient-related questions guides the Azure ML service in extracting specific details from transcripts.

6. **Iterating Over Transcripts:** The script systematically processes each transcript, sending requests for each question to extract various aspects of patient information.

7. **Azure Blob Storage Upload:** The script uploads the CSV file to Azure Blob Storage for centralized and scalable data storage.

Metric used

:

Word Error Rate (WER): WER is a metric commonly used to evaluate the performance of automatic speech recognition (ASR) systems or, in this case, the output generated by the language model. It measures the difference between the predicted and reference transcripts in terms of the number of incorrect words, insertions, deletions, or substitutions. Mathematically, it is defined as:

$$WER = \frac{\text{Substitutions} + \text{Deletions} + \text{Insertions}}{\text{Total Words in the Reference Transcript}}$$

In the context of this project, WER is a valuable metric to assess the accuracy of the model's responses to medical questions. A lower WER indicates a higher level of accuracy, as it represents fewer discrepancies between the model's output and the expected answers.

Utilizing WER for Model Evaluation: WER will be employed as a key criterion for evaluating and selecting the best-performing model. During the model training and testing phases, the WER will be calculated for each model's predictions against a set of reference answers. The model with the lowest WER is considered the most accurate in generating responses, making it a pivotal factor in determining the efficacy of different models in the context of medical transcription and question-answering tasks.

Results

Model Used	**Score**
deepset-minilm-uncased-squad2	1.20019
distilbert-base-uncased-distilled-squad	0.80911
distilbert-base-cased-distilled-squad	0.78385
deepset-roberta-base-squad2	**0.76971**
bert-large-uncased-whole-word-masking-finetuned-squad	2.01199
valhalla-longformer-base-4096-finetuned-squadv1	1.99785
mrm8488-bert-multi-cased-finetuned-xquadv1	1.16731
llukas22-all-minilm-l12-v2-qa-en	1.07384
mrm8488-bert-base-spanish-wwm-cased-finetuned-spa-squad2-es	1.48241
vietai-vit5-base	2.01199
nlpconnect-roberta-base-squad2-nq	1.99785
consciousai-question-answering-roberta-base-s-v2	1.99785
henryk-bert-base-multilingual-cased-finetuned-dutch-squad2	1.10613
intel-dynamic-tinybert	1.51188
hf-internal-testing-tiny-random-rembertforquestionanswering	1.51188
deutsche-telekom-bert-multi-english-german-squad2	1.10143
phiyodr-bert-base-finetuned-squad2	1.10613

Table 1: Model Performance Scores

Conclusion

In conclusion, the project *"LLM-Driven Healthcare Transformation"* aims to harmonize human compassion with technological innovation in the healthcare and hospice sectors. The focal point of this initiative is the utilization of Large Language Models (LLMs), with a specific emphasis on the *deepset-roberta-base-squad2* model, to automate the extraction of crucial information from medical transcriptions.

The consistent low scores achieved by the *deepset-roberta-base-squad2* model in our evaluations attest to its outstanding performance. Renowned for its excellence in question-answering tasks, this model has demonstrated exceptional accuracy in extracting patient-centric details, including names, ages, prescribed medications, symptoms, and doctor-advised precautions. Its efficiency and precision make it a standout choice for streamlining the meticulous process of medical documentation.

Beyond its technical prowess, the implementation of *deepset-roberta-base-squad2* holds transformative potential for healthcare professionals and hospice teams alike. By automating the extraction of critical patient information, the model liberates valuable time and resources. Healthcare professionals can redirect their focus from labor-intensive paperwork to providing more attentive and personalized care to patients. The resulting efficiency gains translate into improved patient experiences, reduced administrative burdens, and ultimately, enhanced healthcare delivery.

For hospice teams, the deployment of LLMs like *deepset-roberta-base-squad2* signifies the positive impact of technology on palliative care. Timely and accurate extraction of information allows hospice teams to streamline administrative processes, ensuring that patient data is consistent, secure, and readily available. This not only improves the overall quality of care but also contributes to a more compassionate and supportive environment for both patients and their families.

In essence, the successful integration of *deepset-roberta-base-squad2* in this project signifies a step forward in the symbiotic relationship between technology and healthcare. It goes beyond the realms of mere automation; it exemplifies a commitment to making a tangible difference at the intersection of human well-being and innovative solutions.