MEDCOACH

An LLM-powered learning tool that helps medical students sharpen their diagnostic skills through interactive clinical case simulations.

Project Proposal

Mai Werthaim & Maya Kimhi

Repository: projectProposal/MedCoach_ProjectProposal.pdf

Problem statement

Project Motivation

Medical students and interns often rely on static teaching methods that lack the complexity of real-world clinical scenarios. They need an interactive, case-based platform to sharpen their diagnostic reasoning and decision-making skills.

By harnessing the power of large language models, we can simulate authentic patient interactions, generate diverse medical questions, and provide immediate feedback—ultimately bridging knowledge gaps and enhancing the learning experience.

Problem statement

Problem Definition

⭐ Inputs

Named Entity Recognition (NER)
Medical questions or patient case descriptions from a dataset (MedQuad).

Question Answering, Semantic Evaluation
Student responses or queries about diagnoses and treatment.

⭐ Outputs

Medical Text Generation
A clinical case scenario generated by an LLM based on a medical QA dataset (includes age, symptoms, etc.).

Question Answering
Feedback comparing student answers to reference answers.

Problem statement

Problem Challenges

- Evaluating open-ended responses which requires semantic and medical understanding beyond keyword matching.

- Accurately interpreting specialized medical terminology and context

- Generating medically accurate and coherent clinical scenarios.

- Providing meaningful feedback.

Training and Test Data

Diseases and their Symptoms Dataset (Kaggle): This dataset provides a comprehensive collection of disease names and associated symptoms, encoded in a one-hot manner. Each row in the dataset represents a single instance

Each symptom can be 0/1 for each disease, where:
- 1 indicates the symptom is relevant to the disease
- 0 implies no relevance

Synthetic Data Generation: Using LLM to create case scenarios from the Disease-Symptom Dataset.

Training and Test Data

Example

MedCoach: Mr. John Miller, a 42-year-old male, presented to the emergency department with a 12-hour history of progressively worsening neurological symptoms. Approximately 24 hours prior, he had attended a family picnic where he consumed home-canned vegetables that were later suspected to be improperly sterilized. Initially, Mr. Miller noticed a subtle blurring of his vision and a sensation of heaviness around his eyes. Within a few hours.
symptoms: double vision, drooping eyelids, and a noticeable difficulty in speaking as his speech became slurred. difficulty swallowing, a persistently dry mouth, and generalized muscle weakness. What is Mr. Miller's diagnosis?

Student: Did Mr. Miller undergo a physical examination?

MedCoach: On physical examination, the clinical team observed bilateral ptosis, impaired cranial nerve responses, and decreased muscle strength in both his upper and lower limbs. What Mr. Miller has?

Student: Botulism

MedCoach: Correct

Evaluation

Student Answer Assessment

Comparison with Baseline

Evaluating model performance by comparing it to a simple baseline (TF-IDF + cosine similarity).

Classification Metrics

Accuracy, Precision, Recall, F1-score

Generated Clinical Cases

Medical Plausibility & Completeness

BERTScore between generated cases and original data to gauge linguistic quality.

Cases Diversity

Embedding distance to measure semantic differences between generated cases

THANK YOU

Do you have any questions for us?