

## PAPER

# medLLM: Fine-Tuning a Medical language Model for Clinical Decision Support

First Author,<sup>1,\*</sup> Second Author,<sup>2</sup> Third Author,<sup>3</sup> Fourth Author<sup>3</sup> and Fifth Author<sup>4</sup><sup>1</sup>Department, Organization, Street, Postcode, State, Country, <sup>2</sup>Department, Organization, Street, Postcode, State, Country, <sup>3</sup>Department, Organization, Street, Postcode, State, Country and <sup>4</sup>Department, Organization, Street, Postcode, State, Country

\*Corresponding author. email-id.com

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

## Abstract

**Key words:** keyword1, Keyword2, Keyword3, Keyword4

## Introduction

In the rapidly evolving field of medicine, doctors are required to process vast amounts of information to make well-informed decisions. However, accessing relevant medical literature can be time-consuming and challenging. This project aims to address this issue by developing a large language model designed to assist doctors in gathering crucial patient-related information efficiently. The proposed model allows doctors to input a patient description as a prompt, after which it retrieves and summarizes relevant medical research papers. This process provides physicians with critical insights derived from medical literature. This approach enhances work efficiency, reduces cognitive load, and ensures that important information is readily accessible when needed. To achieve this, our system employs a two-stage approach: retrieval and summarization. For the retrieval phase, we utilize RoBERTa, a robust Transformer-based model, to process the input prompt and identify the most relevant medical research articles from the PMC-Patients database. Once the relevant articles are retrieved, the summarization phase is handled using either Llama 3.2 1b instruct or Llama 3.3. We experimented with both models separately to determine which one provided the most accurate and useful summaries for clinical decision-making.

Through comparative analysis, we assessed which model performed better in summarizing medical literature effectively for our use case. This report outlines the development, training, and evaluation of the model, detailing its capabilities and limitations. Specifically, we discuss the dataset utilized for training, the preprocessing techniques employed, the symptom extraction process, the machine learning architecture, and the evaluation methods used to assess summarization quality. Furthermore, we examine the model's performance and highlight areas for improvement.

## Methods

## Results

## Discussion

## contributions

## References

1. Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.