# Medical Compliance Checker

## Approach:

I have used **llama-3.1-8b-instant** with **instruction-based prompting** technique to the best results. Used **Langchain** and **Grok API** for that.

### How I crafted the prompt:

I have gone through multiple guidelines online sources and collected data.

Then I summarized all the guidelines information in few bullet points.

Based on summarized information, I have crafted the prompt by setting rule for compliant or non-compliant claims.

For structured output, I have implemented pydantic with structure output instruction tuning.

## Why not other approaches:

* Firstly, LLMs can understand subtle differences in phrasing, tone, and intent while ruled based approach required require hundreds of hard-coded keywords combinations to catch phrase.
* Secondly, LLMs are pretrained on vast amount of medical and regulatory data, giving them broad knowledge out of the box.
* I used instruction prompting over fine-tuning LLM because it would be time consuming to collect and process data and training LLM is computationally expensive.

## Other Thing I was trying to do.

For example, like: **"Clinical studies show this knee surgery has a 95% success rate."**

For cases like these where LLM also need to know about the studies for fact check.

I was trying to integrate an external search tool for fact checking. My idea was that the agent calls the tool that will perform a search of the claim like proven studies and once it gets information from reliable websites like **PubMed Central**, the agent performs classification.

## Evaluation:

Out of 20 test examples, 19 are classified correctly.

All test data used for evaluation is synthetically generated via GPT using few short promptings.

**(If more prompt engineering was done it could classify all.)**

**Report:**

Accuracy: 0.96

Precision (Compliant): 0.92

Recall (Compliant): 1.00

F1 Score (Compliant): 0.96

Classification Report:

precision recall f1-score support

Compliant 0.92 1.00 0.96 11

Non-Compliant 1.00 0.93 0.96 14

accuracy 0.96 25

macro avg 0.96 0.96 0.96 25

weighted avg 0.96 0.96 0.96 25

A diagram of a confusion matrix

AI-generated content may be incorrect.