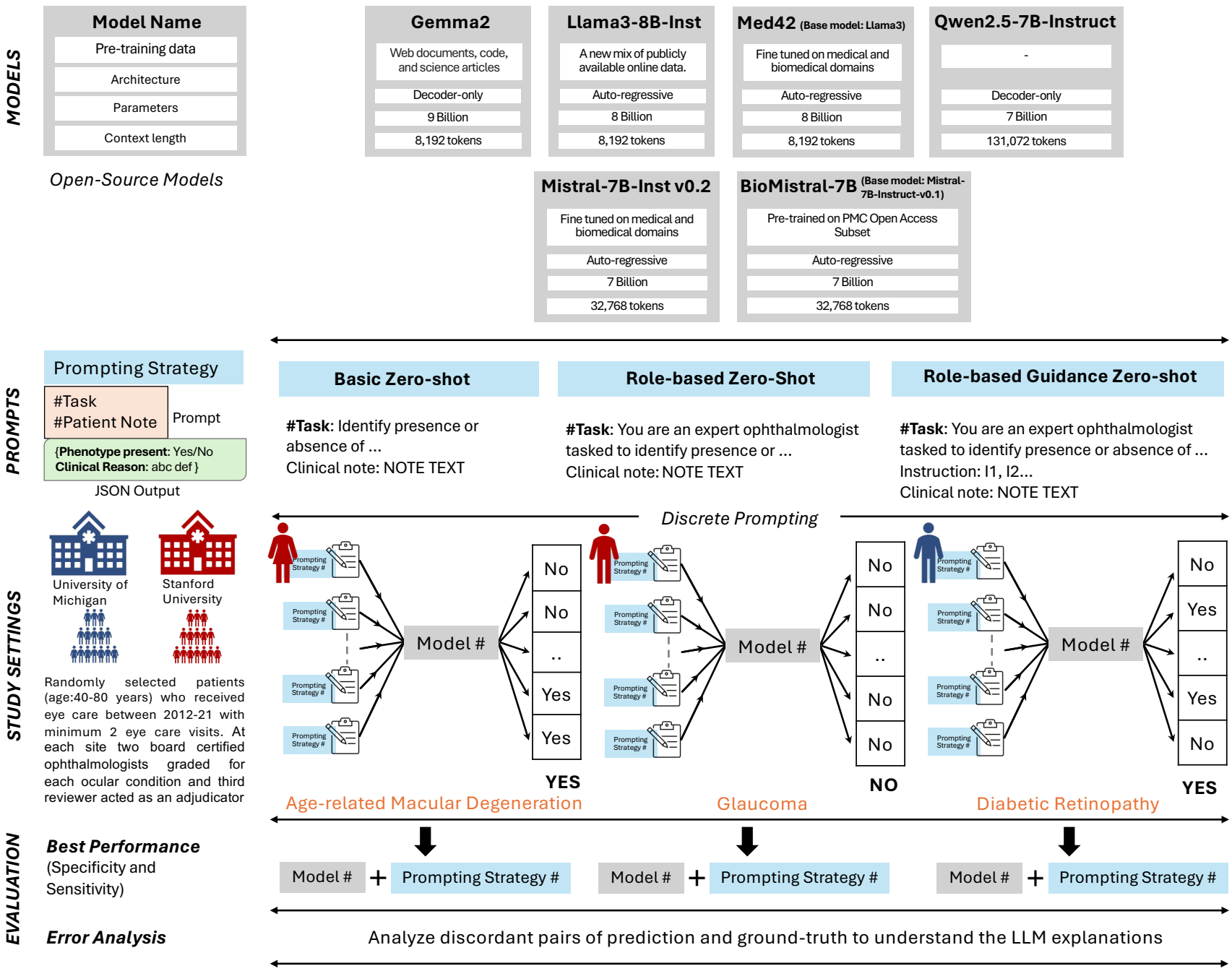
MODELS
Model Name
Pre-training data
Architecture
Parameters
Context length
Open-Source Models
Gemma2
Web documents, code, and science articles
Decoder-only
9 Billion
8,192 tokens
Llama3-8B-Inst
A new mix of publicly available online data.
Auto-regressive
8 Billion
8,192 tokens
Med42 (Base model: Llama3)
Fine tuned on medical and biomedical domains
Auto-regressive
8 Billion
8,192 tokens
Qwen2.5-7B-Instruct
-
Decoder-only
7 Billion
131,072 tokens
Mistral-7B-Inst v0.2
Fine tuned on medical and biomedical domains
Auto-regressive
7 Billion
32,768 tokens
BioMistral-7B (Base model: Mistral-7B-Instruct-v0.1)
Pre-trained on PMC Open Access Subset
Auto-regressive
7 Billion
32,768 tokens
PROMPTS
Prompting Strategy
#Task
#Patient Note
Prompt
{Phenotype present: Yes/No
Clinical Reason: abc def }
JSON Output
Basic Zero-shot
#Task: Identify presence or absence of ...
Clinical note: NOTE TEXT
Role-based Zero-Shot
#Task: You are an expert ophthalmologist tasked to identify presence or ...
Clinical note: NOTE TEXT
Role-based Guidance Zero-shot
#Task: You are an expert ophthalmologist tasked to identify presence or absence of ...
Instruction: I1, I2...
Clinical note: NOTE TEXT
STUDY SETTINGS
University of Michigan
Stanford University
Randomly selected patients (age:40-80 years) who received eye care between 2012-21 with minimum 2 eye care visits. At each site two board certified ophthalmologists graded for each ocular condition and third reviewer acted as an adjudicator
Discrete Prompting
Prompting Strategy #
Model #
No
No
..
Yes
Yes
YES
Age-related Macular Degeneration
No
No
..
No
No
NO
Glaucoma
No
Yes
..
Yes
No
YES
Diabetic Retinopathy
EVALUATION
Best Performance
(Specificity and Sensitivity)
Model # + Prompting Strategy #
Model # + Prompting Strategy #
Model # + Prompting Strategy #
Error Analysis
Analyze discordant pairs of prediction and ground-truth to understand the LLM explanations