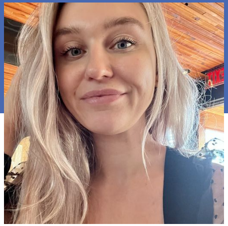
MedyCode Assistant

DATASCI 210 | Summer 2024

August 6th, 2024

Team


Alex Roy
(alexhroy@ischool.berkeley.edu)

Darya Likhareva
(likhareva@ischool.berkeley.edu)

Dylan Daniels
(dylansdaniels@ischool.berkeley.edu)


Tanmay Mahapatra
(tmahapatra@ischool.berkeley.edu)

Quyen Ha
(qha@ischool.berkeley.edu)

Agenda

- Demo
- Problem
- Dataset & EDA
- Technical Discussion
- Evaluation
- Future Work


https://www.pexels.com/search/medical/

[Click here for full product demo]

"The administrative burden in primary care significantly detracts from my ability to care for patients and contributes to burnout."

Dr. Glubok Gonzalez, MD, Family Medicine

30%

national shortage of medical coders.

15 hours

Extra time spent working after hours by physicians, i.e. "pajama time"

$31.2B

lost by Medicare in 2023 from coding errors, causing delays and higher costs.

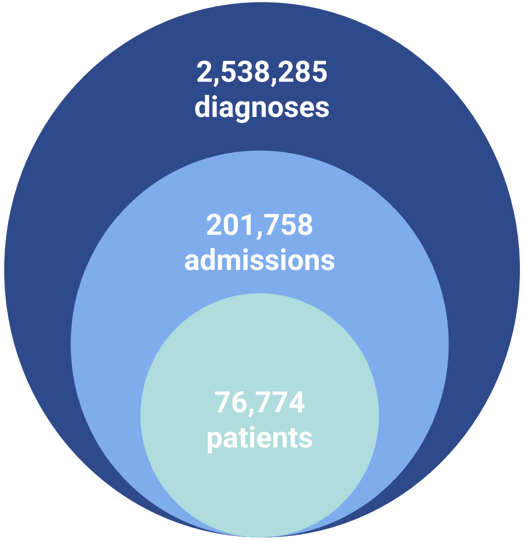
Let’s eliminate pajama time…

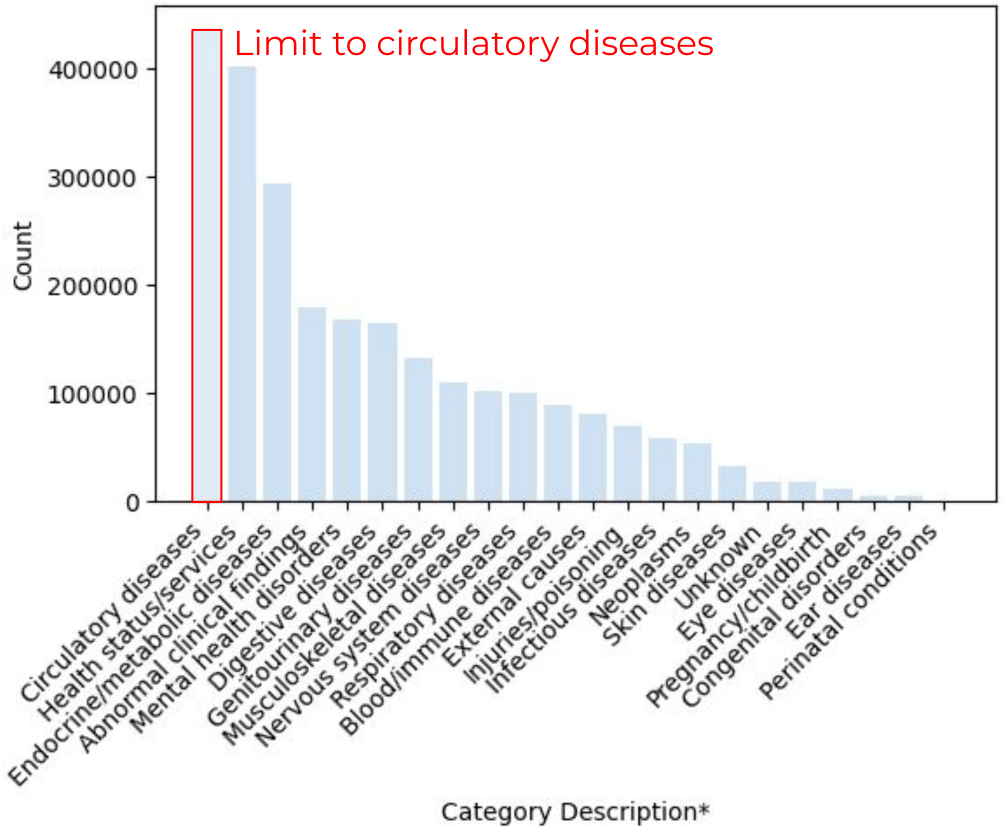
Medycode Assistant will:

- *Improve coding accuracy*
- *Speed up payments*
- *Support under-resourced healthcare providers*

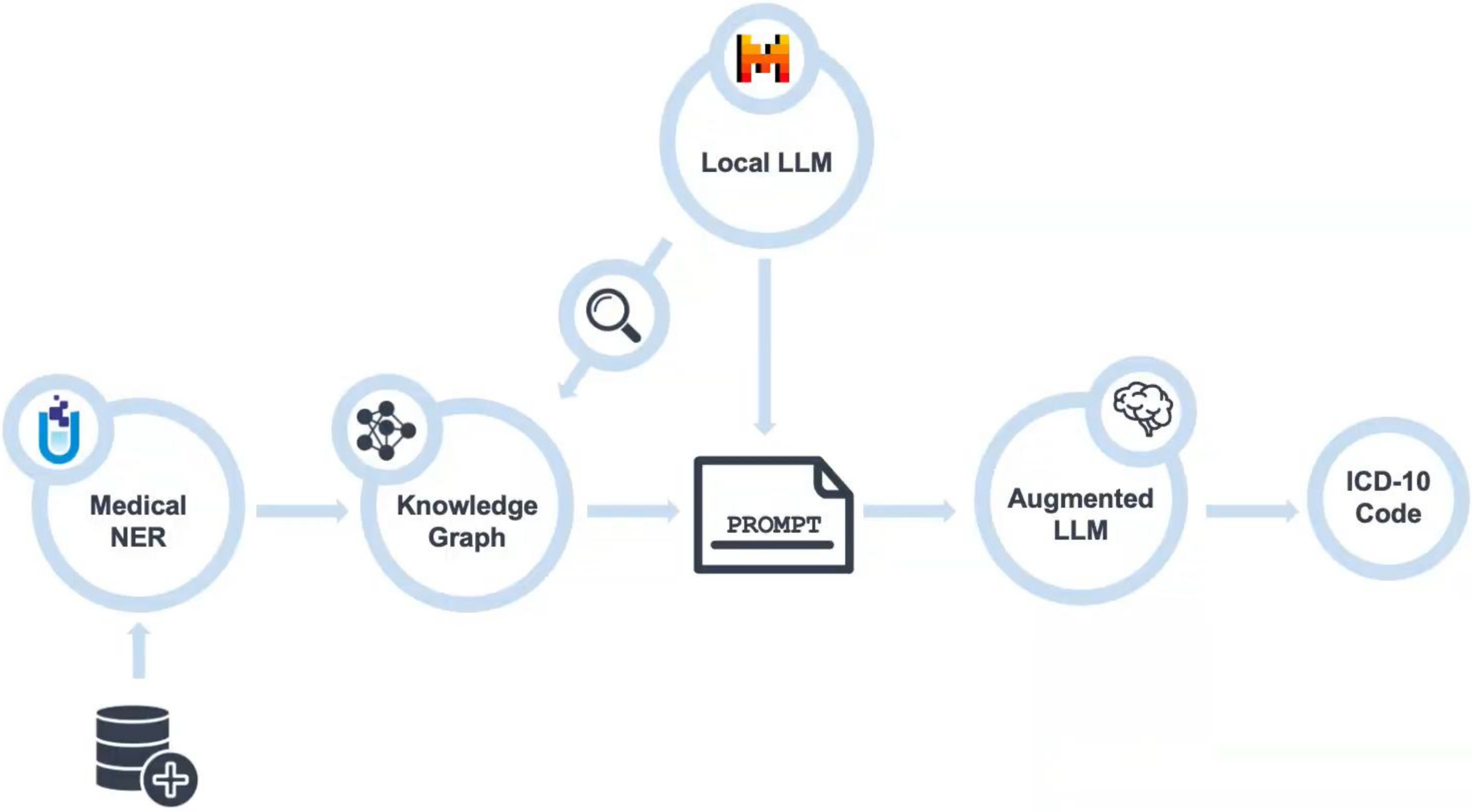
Dataset: MIMIC IV

De-identified health database from Beth Israel Medical Center with clinical notes.





Technical Overview



Evaluation

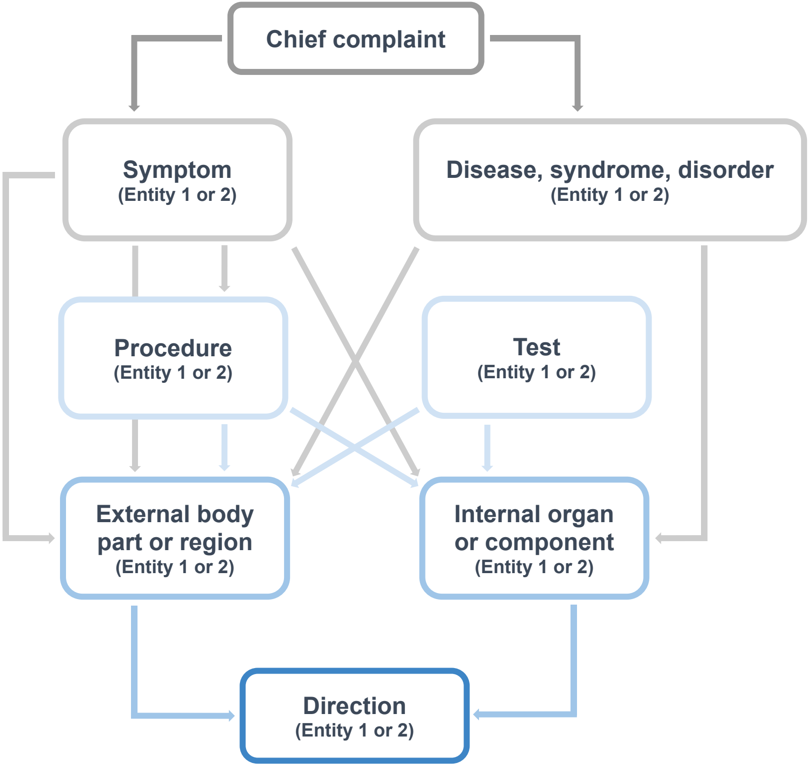
Base Pipeline (without RAG)

Model	NER Prompt	ICD-10 Prompt	Accuracy
Mistral-7b-Instruct	Standard	Updated	14%
Mistral-7b-Instruct	Updated	Updated	28%
Llama3-8b-Instruct	Standard	Updated	10%
Llama3-8b-Instruct	Updated	Updated	13%

MedyCode Pipeline (with RAG)

RAG Model	Embedding Model	NER Model	NER Prompt	Accuracy
Mistral-7b-Instruct	Cohere	Llama3-8b-Instruct	Standard	25%
Mistral-7b-Instruct	Cohere	Llama3-8b-Instruct	Updated	33%
Llama3-8b-Instruct	OpenAI	Llama3-8b-Instruct	Standard	31%
Llama3-8b-Instruct	OpenAI	Llama3-8b-Instruct	Standard	42%
Llama3-8b-Instruct	Cohere	Llama3-8b-Instruct	Standard	33%
Llama3-8b-Instruct	Cohere	Llama3-8b-Instruct	Updated	38%

Knowledge Graphs Improve LLMs



Future Work

Addition of more data and feedback loop

Leverage additional data and incorporate a feedback loop for better ICD-10 predictions.

Inclusion of all ICD-10 codes

Scale up beyond circulatory diseases and include predictions for all ICD-10 codes

Integration into electronic medical record systems

Seamless integration with clients’ existing internal systems

https://www.ischool.berkeley.edu/projects/2024/medycode-assistant

Acknowledgements

Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L. A., and Mark, R., "**MIMIC-IV (version 2.2)**," PhysioNet, 2023.

John Snow Labs: Kocaman, V., & Talby, D. (2021). Spark NLP: Natural Language Understanding at Scale. Software Impacts, 8, 100058. ISSN 2665-9638. https://doi.org/10.1016/j.simpa.2021.100058

What is Medical Coding? Retrieved from https://www.aapc.com/medical-coding/medical-coding.aspx

Minemyer, P. (2023, November 20). **Medicare paid out $31.2B in improper payments for FY23**, CMS says. Fierce Healthcare. Retrieved from https://www.fiercehealthcare.com/payer/medicare-paid-out-31-2b-improper-payments-fy23-cms-says

Feke, T. (2023, July 11). **Learn About Insurance Codes to Avoid Billing Errors.** Verywell Health. Retrieved from https://www.verywellhealth.com/learn-about-insurance-codes-to-avoid-billing-errors-4147371

Verywell Health. (2023, July 11). **You pay more when your healthcare provider picks the wrong diagnosis code.** Retrieved from https://www.verywellhealth.com/you-pay-more-when-your-healthcare-provider-picks-the-wrong-diagnosis-code-4147372

Cure for the Common Code: San Francisco Startup Uses AI to Automate Medical Coding. Retrived from https://blogs.nvidia.com/blog/fathom-ai-medical-coding/#:~:text=Natural%20language%20processing%20AI%2C%20accelerated,coding%20process%20while%20increasing%20accuracy

Nym Health's Clinical Language Understanding (CLU) Technology. Retrieved from https://www.nym.health/technology

Dong, H., Falis, M., Whiteley, W. et al. **Automated clinical coding: what, why, and where we are?**. npj Digit. Med. 5, 159 (2022). https://doi.org/10.1038/s41746-022-00705-7

Template by Slidesgo

A special thank you to Todd Holloway, Fred Nugen, and Mark Butler for their support and guidance on our project. We would also like to express our gratitude to our content advisors and test users: Dr. Yasmin Azar, Dr. Glubok Gonzalez, MD and Dr. Judson Brewer.

MEDYCODE ASSISTANT

Improving outcomes and reducing administrative burdens through enhanced medical coding accuracy and efficiency with AI.

Appendix

Evaluation of NER Extraction

Model	NER Prompt	Evaluation Metrics	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Mistral-7b-Instruct	Baseline	Precision	0..76	0.81	0.83	0.82	0.84	0.86
Mistral-7b-Instruct	Baseline	Recall	0.76	0.82	0.84	0.84	0.86	0.89
Mistral-7b-Instruct	Baseline	F1 score	0.77	0.82	0.84	0.83	0.84	0.86
Mistral-7b-Instruct	Updated	Precision	0.76	0.84	0.85	0.84	0.85	0.88
Mistral-7b-Instruct	Updated	Recall	0.79	0.83	0.84	0.84	0.85	0.90
Mistral-7b-Instruct	Updated	F1 score	0.80	0.84	0.84	0.84	0.85	0.87
Llama3-8b-Instruct	Baseline	Precision	0.74	0.82	0.84	0.83	0.85	0.87
Llama3-8b-Instruct	Baseline	Recall	0.75	0.83	0.85	0.85	0.86	0.89
Llama3-8b-Instruct	Baseline	F1 score	0.77	0.83	0.84	0.84	0.85	0.87
Llama3-8b-Instruct	Updated	Precision	0.80	0.85	0.86	0.85	0.86	0.88
Llama3-8b-Instruct	Updated	Recall	0.80	0.84	0.85	0.85	0.87	0.92
Llama3-8b-Instruct	Updated	F1 score	0.81	0.84	0.85	0.85	0.86	0.89

Evaluation WITH RAG

RAG Model	Embedding Model	NER Model	NER Prompt	Accuracy
Mistral-7b-Instruct	OpenAI	Llama3-8b-Instruct	Baseline	25%
Mistral-7b-Instruct	OpenAI	Llama3-8b-Instruct	Updated	33%
Mistral-7b-Instruct	OpenAI	Mistral-7b-Instruct	Baseline	25%
Mistral-7b-Instruct	OpenAI	Mistral-7b-Instruct	Updated	25%
Mistral-7b-Instruct	Cohere	Llama3-8b-Instruct	Baseline	23%
Mistral-7b-Instruct	Cohere	Llama3-8b-Instruct	Updated	21%
Mistral-7b-Instruct	Cohere	Mistral-7b-Instruct	Baseline	28%
Mistral-7b-Instruct	Cohere	Mistral-7b-Instruct	Updated	27%
Llama3-8b-Instruct	OpenAI	Llama3-8b-Instruct	Baseline	31%
Llama3-8b-Instruct	OpenAI	Llama3-8b-Instruct	Updated	42%
Llama3-8b-Instruct	OpenAI	Mistral-7b-Instruct	Baseline	37%
Llama3-8b-Instruct	OpenAI	Mistral-7b-Instruct	Updated	30%
Llama3-8b-Instruct	Cohere	Llama3-8b-Instruct	Baseline	33%
Llama3-8b-Instruct	Cohere	Llama3-8b-Instruct	Updated	38%
Llama3-8b-Instruct	Cohere	Mistral-7b-Instruct	Baseline	21%
Llama3-8b-Instruct	Cohere	Mistral-7b-Instruct	Updated	37%

Baseline Prompt for NER Extraction:
Please extract and list the medical entities from the following text.

Text:
{note}

Instructions:
Extract and label entities:
In clear and concise language, find and extract medical entities which should include medical symptoms, tests, procedures, diseases, syndromes, or disorders, external body parts or regions, internal organs or components, and direction from the text.

Format
Use the following labels:
'Internal_organ_or_component', 'Symptom', External_body_part_or_region', 'Direction'.
'Disease_Syndrome_Disorder', ‘Procedure’, 'Test'.

Provide the extracted entities as a dictionary with the labels as keys after the words 'Extracted Entities:' .

Example:
Extracted Entities: 'Internal_organ_or_component': ['aortic valve','heart'], 'Symptom': ['dyspnea','edema'],
'External_body_part_or_region': ['neck','leg','abdomen','skin'], 'Direction': ['left', 'lateral aspect','anterior'],
'Disease_Syndrome_Disorder': ['hypertension','hyperlipidemia'],'Test': ['echo']

Final Prompt for NER Extraction:

You are a highly experienced and skilled medical expert. Your task is to find similar entities from given notes and to label them with the right entity type.

Text:
{note}

Instructions:
Extract and label entities:
In clear and concise language, find and extract medical entities which should include medical symptoms, tests, procedures, diseases, syndromes, or disorders, external body parts or regions, internal organs or components, and direction from the text.

Format
Use the following labels:
'Internal_organ_or_component', 'Symptom', External_body_part_or_region', 'Direction'.
'Disease_Syndrome_Disorder', ‘Procedure’, 'Test'.

Provide the extracted entities as a dictionary with the labels as keys after the words 'Extracted Entities:' .

Example:
Extracted Entities: 'Internal_organ_or_component': ['aortic valve','heart'], 'Symptom': ['dyspnea','edema'],
'External_body_part_or_region': ['neck','leg','abdomen','skin'], 'Direction': ['left', 'lateral aspect','anterior'],
'Disease_Syndrome_Disorder': ['hypertension','hyperlipidemia'],'Test': ['echo']

Your task is to follow these instructions and provide the extracted entities in the specified format. Do not provide additional information outside of this format.

Final Prompt for ICD-10 Assignment:

You are a highly experienced and skilled medical coder. Your task is to review the given medical entities and to label them with the correct ICD-10 code for the present illness.

Based on the given medical entities, assign the single best ICD-10 code for the overall present illness.

Instructions:
Assign a single ICD-10 code for the present illness based on the given medical entities. The code should belong to one of the following ICD-10 codes:
['I21', 'I26', 'I34', 'I50', I35', 'I13', 'I48', 'I11', 'I16', 'I25', 'I27', 'I05', 'I08', 'I31', 'I44', 'I09', 'I97', 'I47', 'I12', 'I10', 'I34', 'I33', 'I21', 'I95', 'I51', 'I05', 'I30', 'I71', 'I01', 'I30', 'I97', 'I31', 'I89', 'I82', 'I16']

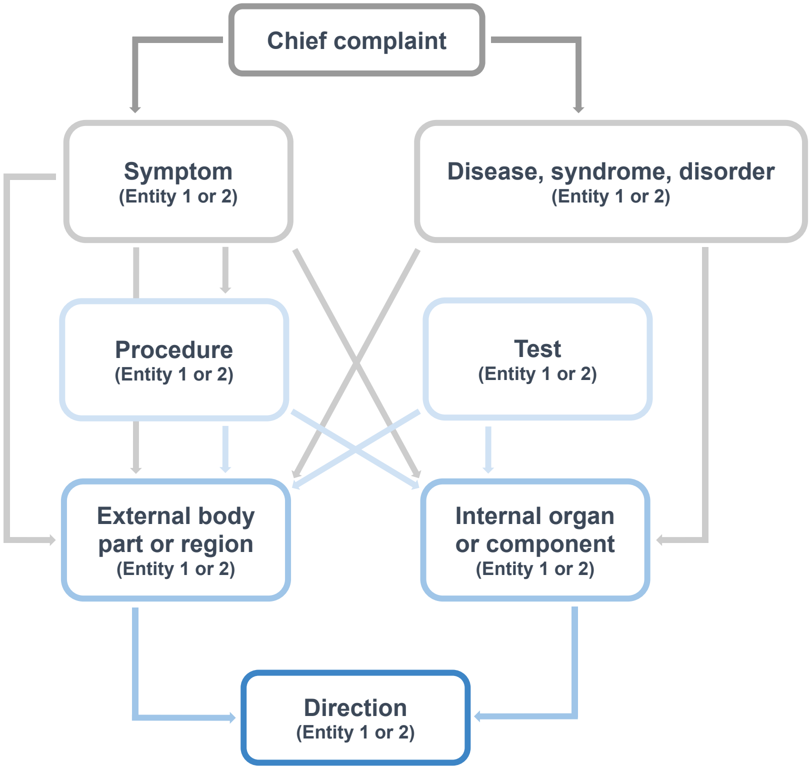
If you are not sure of the specific code, assign a code from one of the 10 categories where the range is the first three digits of the code along with the associated disease categories.

Diseases of circulatory system categories:
1) Codes: I00 - I02, Category: Acute rheumatic fever;
2) Codes: I05 - I09, Category: Chronic rheumatic heart diseases;
3) Codes: I10 - I1A, Category: Hypertensive diseases;
4) Codes: I20 - I25, Category: Ischemic heart diseases;
5) Codes: I26 - I28, Category: Pulmonary heart disease and diseases of pulmonary circulation;
6) Codes: I30 - I5A, Category: Other forms of heart disease;
7) Codes: I60 - I69, Category: Cerebrovascular diseases;
8) Codes: I70 - I79, Category: Diseases of arteries, arterioles, and capillaries;
9) Codes: I80 - I89, Category: Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified;
10) Codes: I90 - I99, Category: Other and unspecified disorders of the circulatory system.

Here are the entities extracted from the clinical note:
{ner}

Please provide your reasoning for why you think the chosen code
is the best.
The output should be in the format:
{'final_output': {'Code_01': [code],'Reasoning': [reasoning]}, 'Code_02': [code],'Reasoning': [reasoning]}, 'Code_03': [code], ‘Reasoning': [reasoning]}

John Snow Lab for Relation Extraction

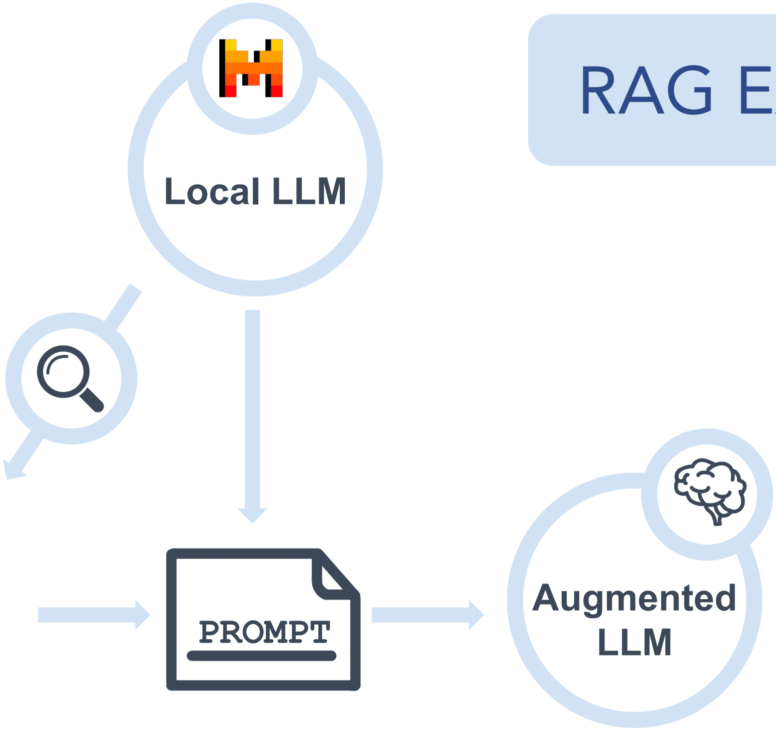


Relation Models

1. Symptoms <--> Procedures
   Symptoms <--> Tests
2. Symptoms <--> Body parts
   Disease/syndrome/disorders <--> Body parts
3. Tests <--> Body parts
   Procedures <--> Body parts
4. Body parts <--> Direction (left v. right)

RAG Example

```
query_text = \
"high blood pressure"
```




```
'name: discrepancy of blood pressure'
metadata={'type': 'Symptom', 'codes': []}

'entity_name: carotid bruits'
metadata={'type': 'Symptom', 'codes': []}

'entity_name: harsh systolic murmur'
metadata={'type': 'Symptom', 'codes': []}

'entity_name: excess fluid'
metadata={'type': 'Symptom', 'codes': []}

'entity_name: thin blood secretions'
metadata={'type': 'Symptom', 'codes': []}
```

Related and Existing Work

- Most existing commercial solutions and academic publications focus on more traditional NLP approaches such as rule-based systems, CNN, and/or BERT.
  - Fathom: SF startup uses BERT.
  - NYM: NY startup uses a rule-based system.
  - A perspective paper in 2022 notes that many papers have focused on CNN and/or BERT-based approaches, stating these models achieve sub-par performance (e.g., < 60% Micro F1 score).
- How we differ:
  - Using RAG + LLM
  - Focusing on a subset of ICD-10 codes

Ethical Considerations and Privacy Concerns

- All data on the server, and in-transit will be encrypted using industry standards
  - Based on a 2048-bit encryption key
  - Only the server can decrypt the data
  - Data will be physically stored on Amazon S3 data center servers protected by armed guards
- The data collection and server storage will be HIPAA compliant.

User Feedback

"The administrative burden in primary care significantly detracts from my ability to care for patients and contributes to burnout. This tool would help alleviate that and make my job more enjoyable while improving work-life balance."

Dr. Glubok Gonzalez, MD, Family Medicine

"The organized information and accurate coding would help me assign diagnosis codes more precisely. This reduces mistakes and saves time. In a busy hospital setting, this kind of efficiency is truly invaluable."

Dr. Yasmin Azar, UNM medical resident

Questions

How secure is MedyCode? Does it comply with HIPAA regulations?

How fast does the ICD-10 code display after inputting the clinical note?

How would this tool integrate with our current systems?

Could we afford it as a smaller practice?