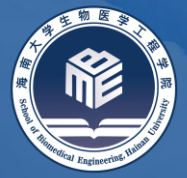


海南大学生物医学工程学院
Hainan University School of Biomedical Engineering

Extracting principal diagnosis from discharge summary using deep learning and LLMs

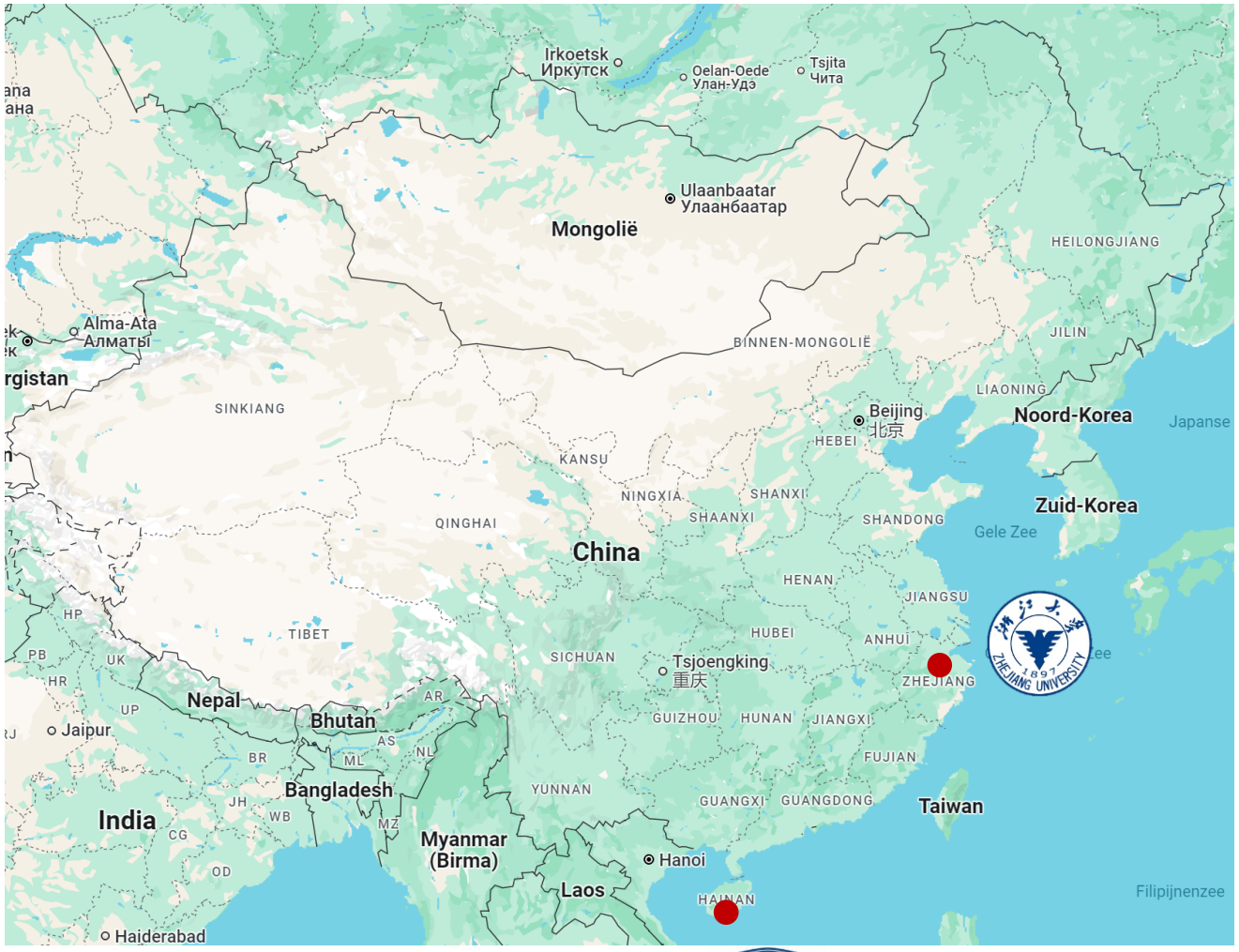
Dr. Shan NAN
nanshan@zju.edu.cn

Me and My Group

Shan Nan Ph.D

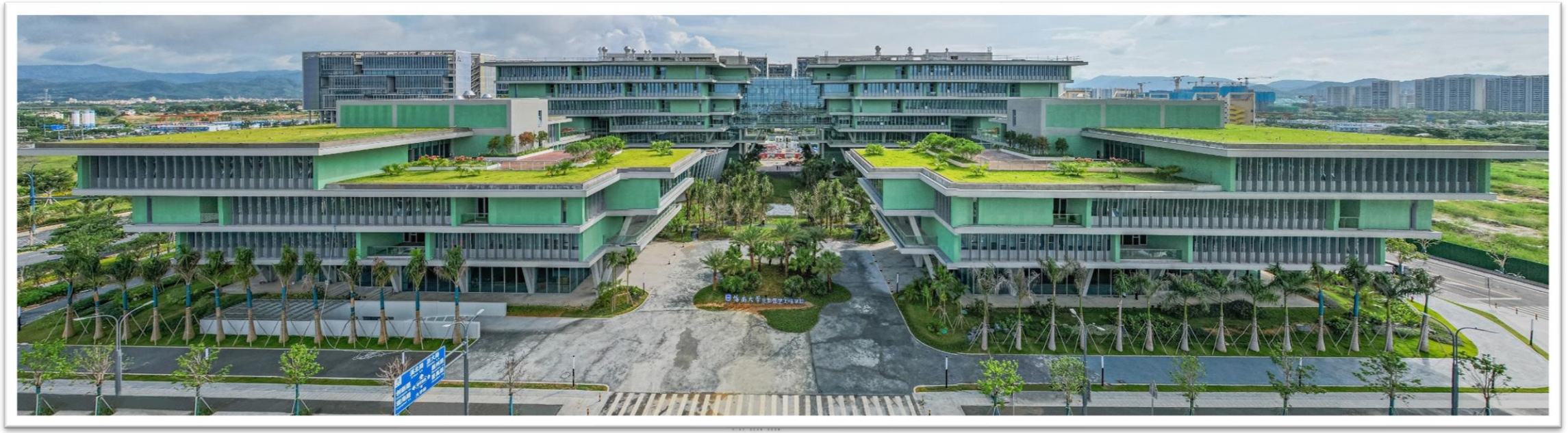
- PhD of Zhejiang University and TU/e
- Associate Prof. of Hainan University
- Working on Clinical Decision Support

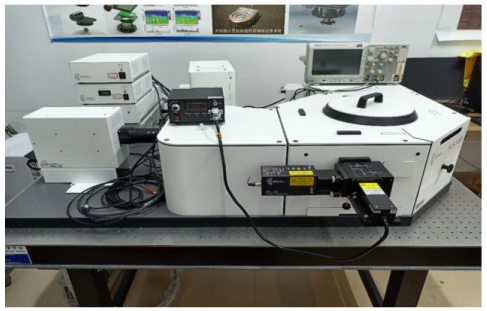
Medical Informatics Group, the earliest medical informatics research group in China

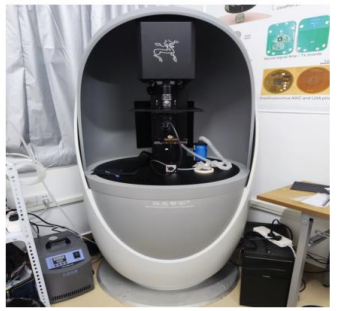


Hainan University

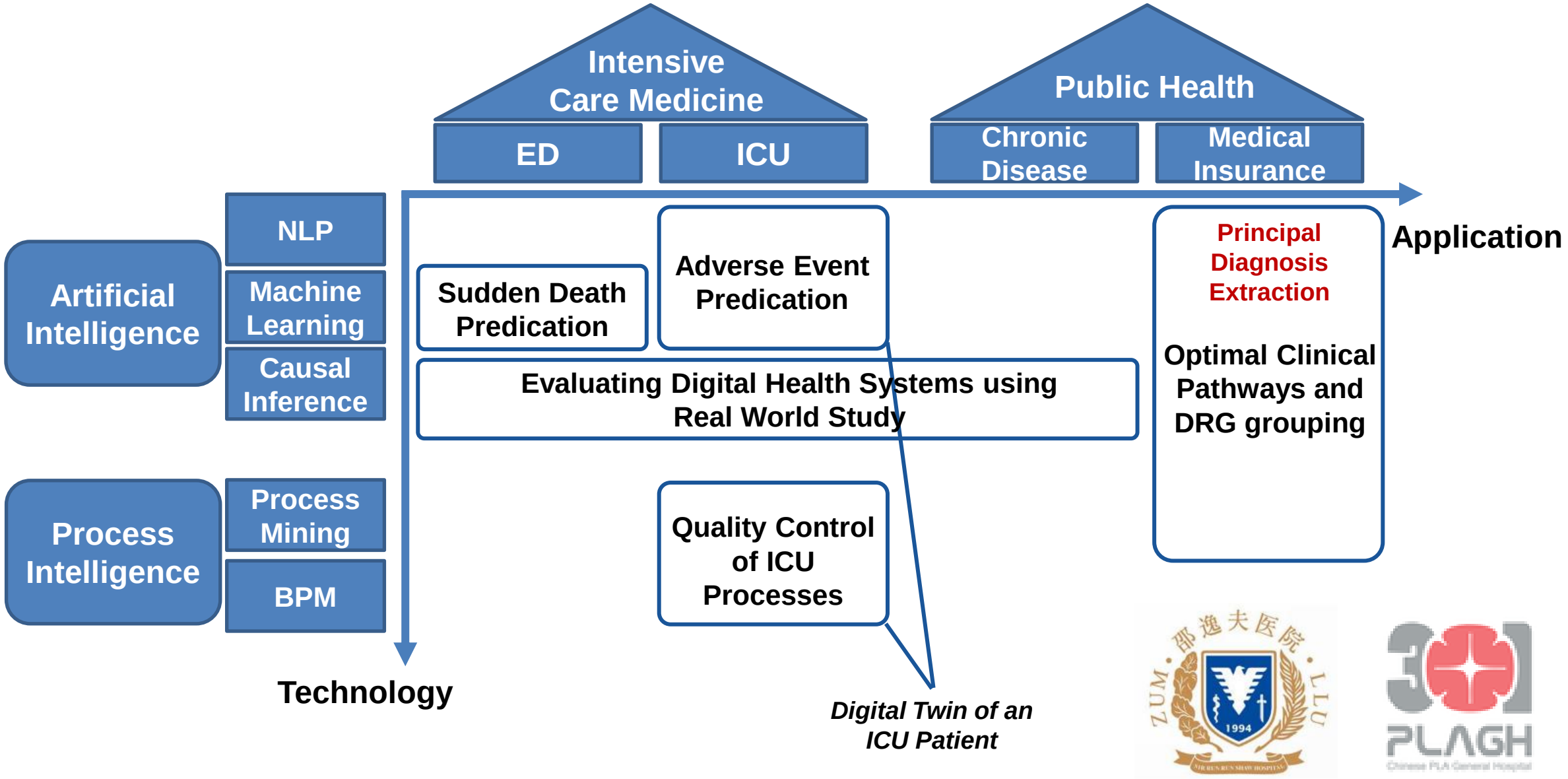
About BME of Hainan University



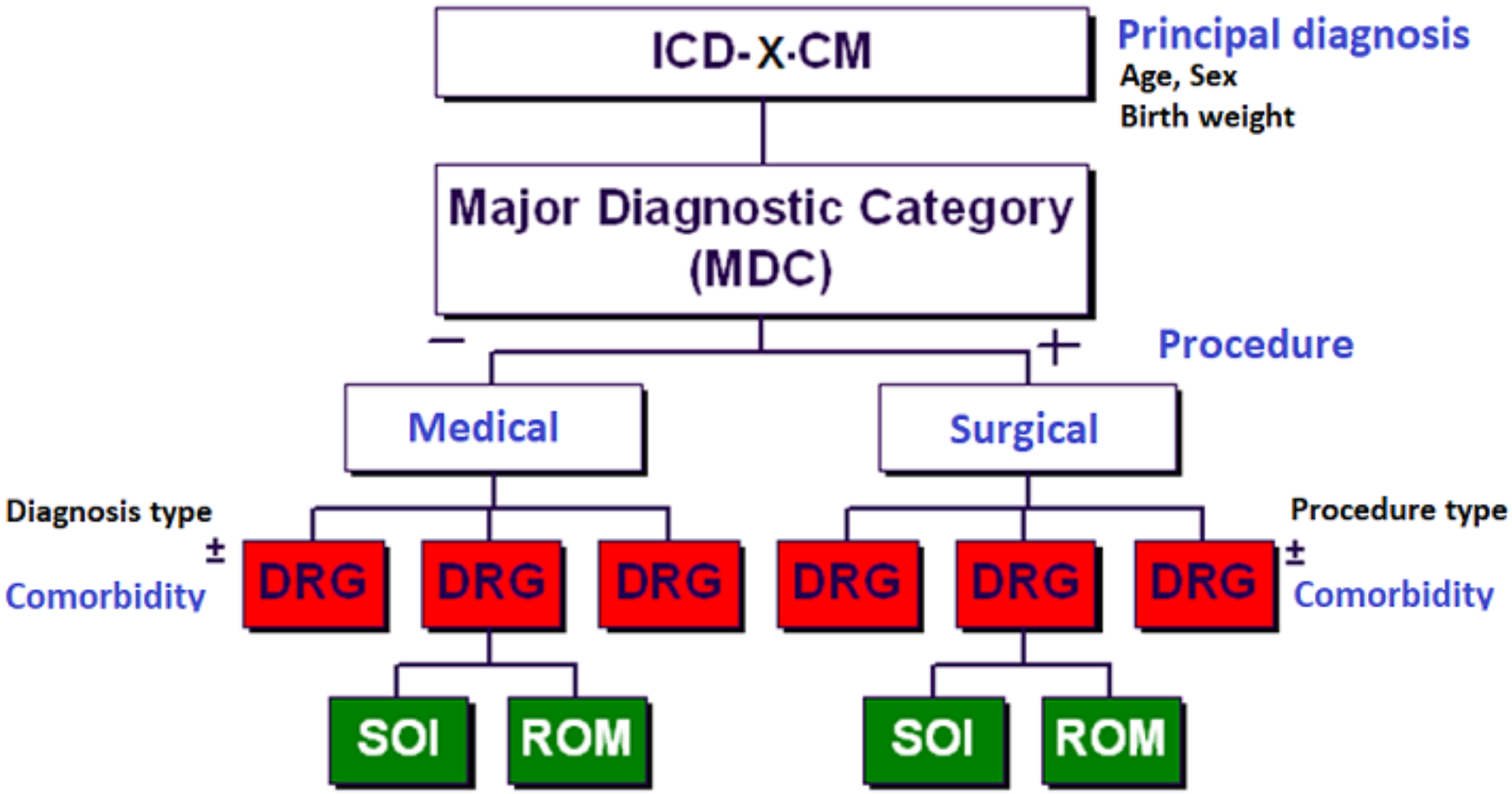




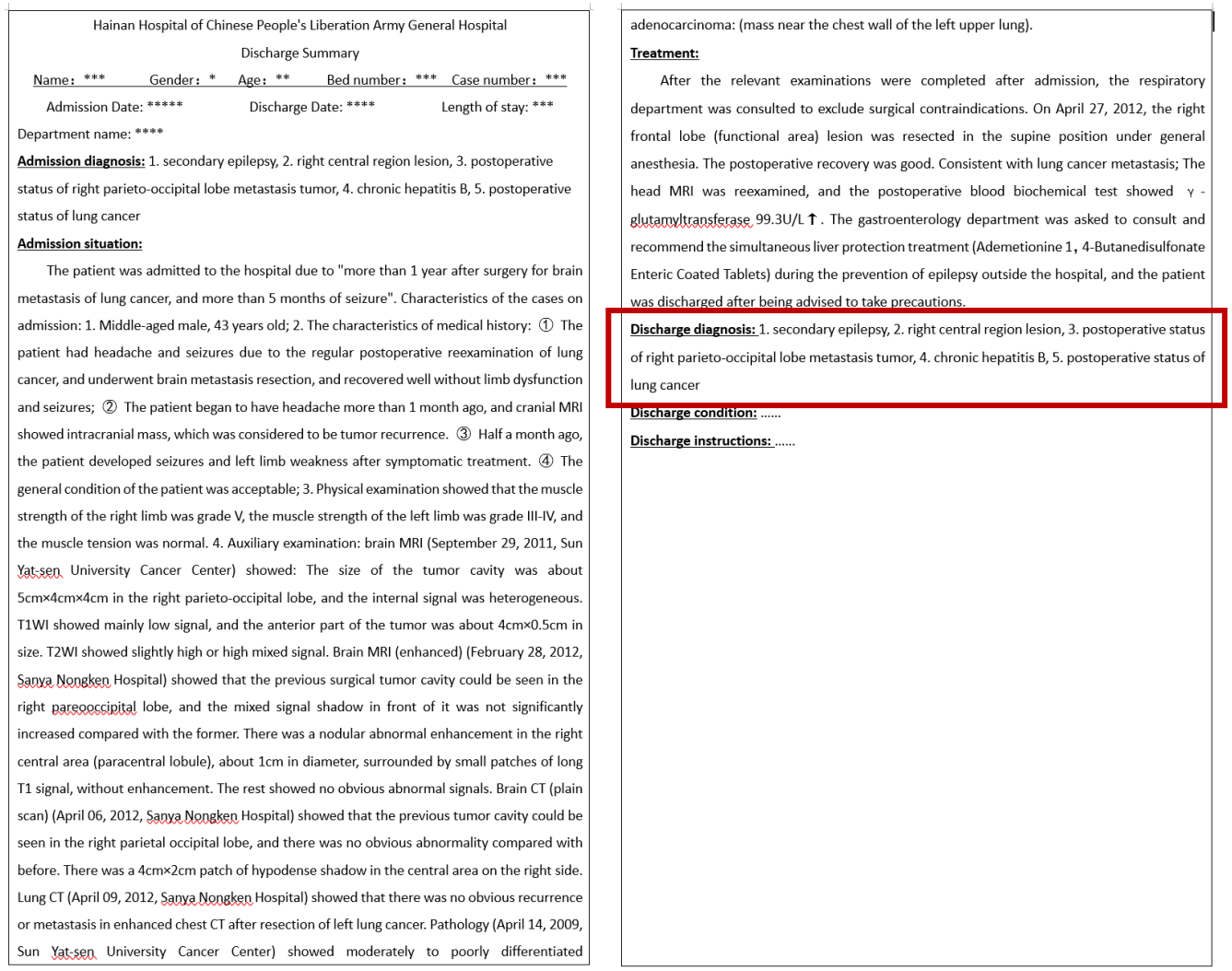
Our Recent Research in Hainan University



Why do we need to extract principal diagnosis?



Principal Diagnosis



Hainan Hospital of Chinese People's Liberation Army General Hospital

Discharge Summary

Name: *** Gender: * Age: ** Bed number: *** Case number: ***

Admission Date: ***** Discharge Date: **** Length of stay: ***

Department name: ****

Admission diagnosis: 1. secondary epilepsy, 2. right central region lesion, 3. postoperative status of right parieto-occipital lobe metastasis tumor, 4. chronic hepatitis B, 5. postoperative status of lung cancer

Admission situation:

The patient was admitted to the hospital due to "more than 1 year after surgery for brain metastasis of lung cancer, and more than 5 months of seizure". Characteristics of the cases on admission: 1. Middle-aged male, 43 years old; 2. The characteristics of medical history: ① The patient had headache and seizures due to the regular postoperative reexamination of lung cancer, and underwent brain metastasis resection, and recovered well without limb dysfunction and seizures; ② The patient began to have headache more than 1 month ago, and cranial MRI showed intracranial mass, which was considered to be tumor recurrence. ③ Half a month ago, the patient developed seizures and left limb weakness after symptomatic treatment. ④ The general condition of the patient was acceptable; 3. Physical examination showed that the muscle strength of the right limb was grade V, the muscle strength of the left limb was grade III-IV, and the muscle tension was normal. 4. Auxiliary examination: brain MRI (September 29, 2011, Sun Yat-sen University Cancer Center) showed: The size of the tumor cavity was about 5cm×4cm×4cm in the right parieto-occipital lobe, and the internal signal was heterogeneous. T1WI showed mainly low signal, and the anterior part of the tumor was about 4cm×0.5cm in size. T2WI showed slightly high or high mixed signal. Brain MRI (enhanced) (February 28, 2012, Sanya Nongken Hospital) showed that the previous surgical tumor cavity could be seen in the right pareooccipital lobe, and the mixed signal shadow in front of it was not significantly increased compared with the former. There was a nodular abnormal enhancement in the right central area (paracentral lobule), about 1cm in diameter, surrounded by small patches of long T1 signal, without enhancement. The rest showed no obvious abnormal signals. Brain CT (plain scan) (April 06, 2012, Sanya Nongken Hospital) showed that the previous tumor cavity could be seen in the right parietal occipital lobe, and there was no obvious abnormality compared with before. There was a 4cm×2cm patch of hypodense shadow in the central area on the right side. Lung CT (April 09, 2012, Sanya Nongken Hospital) showed that there was no obvious recurrence or metastasis in enhanced chest CT after resection of left lung cancer. Pathology (April 14, 2009, Sun Yat-sen University Cancer Center) showed moderately to poorly differentiated adenocarcinoma: (mass near the chest wall of the left upper lung).

Treatment:

After the relevant examinations were completed after admission, the respiratory department was consulted to exclude surgical contraindications. On April 27, 2012, the right frontal lobe (functional area) lesion was resected in the supine position under general anesthesia. The postoperative recovery was good. Consistent with lung cancer metastasis; The head MRI was reexamined, and the postoperative blood biochemical test showed γ-glutamyltransferase 99.3U/L↑. The gastroenterology department was asked to consult and recommend the simultaneous liver protection treatment (Ademetionine 1,4-Butanedisulfonate Enteric Coated Tablets) during the prevention of epilepsy outside the hospital, and the patient was discharged after being advised to take precautions.

Discharge diagnosis: 1. secondary epilepsy, 2. right central region lesion, 3. postoperative status of right parieto-occipital lobe metastasis tumor, 4. chronic hepatitis B, 5. postoperative status of lung cancer

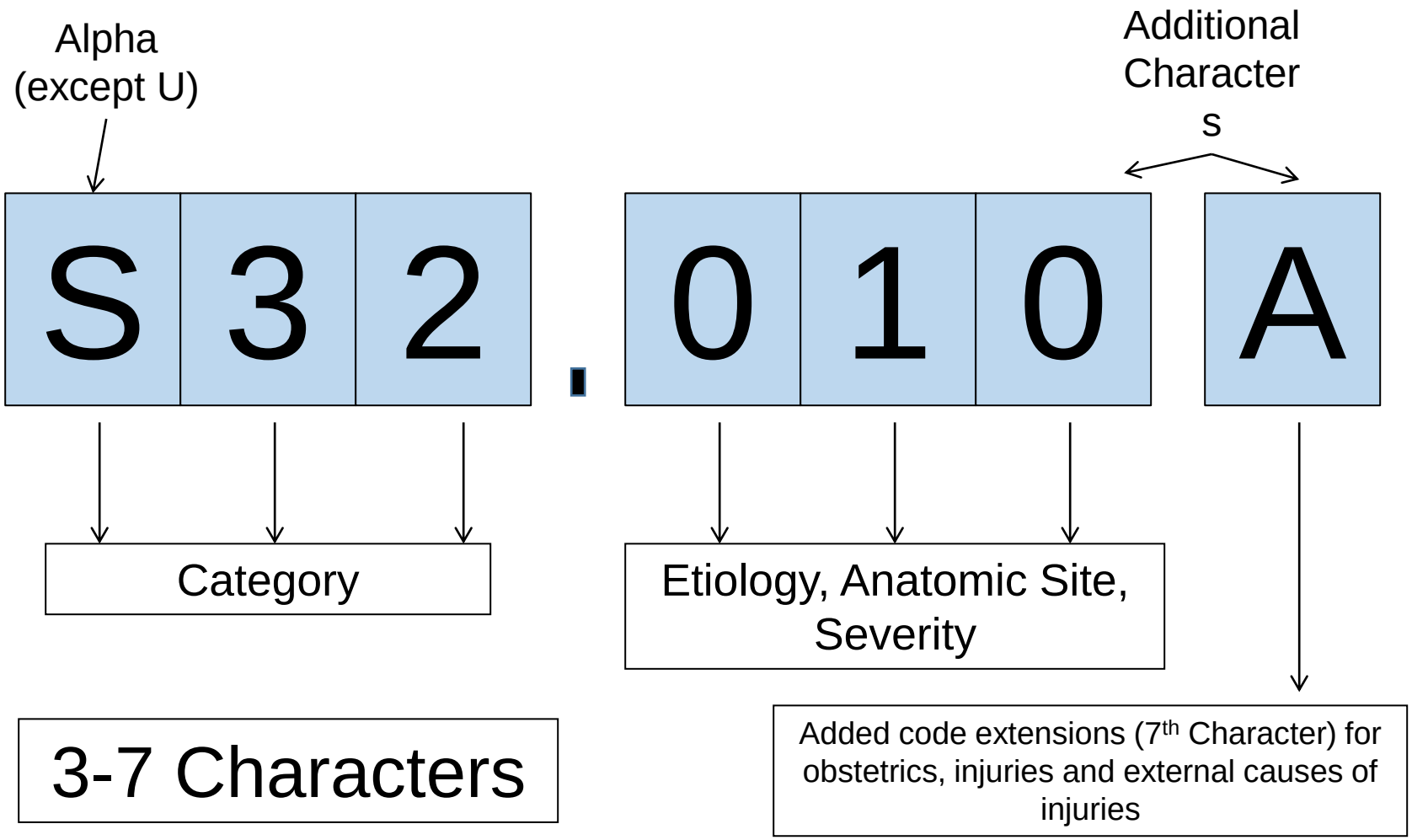
Discharge condition: .....

Discharge instructions: .....

Discharge diagnosis:

1. **secondary epilepsy**
2. right central region lesion
3. postoperative status of right parieto-occipital lobe metastasis tumor
4. chronic hepatitis B
5. postoperative status of lung cancer

The ICD-10 Coding System



I63.40	Cerebral infarction due to embolism of unspecified cerebral artery
I63.49	Of other cerebral artery
I63.411	Of right middle cerebral artery
I63.412	Of left middle cerebral artery
I63.419	Of unspecified middle cerebral artery
I63.421	Of right anterior cerebral artery
I63.422	Of left anterior cerebral artery
I63.429	Of unspecified anterior cerebral artery
I63.431	Of left posterior cerebral artery
I63.432	Of right posterior cerebral artery
I63.439	Of unspecified posterior cerebral artery
I63.441	Of right cerebellar artery
I63.442	Of left cerebellar artery
I63.449	Of unspecified cerebellar artery

The Challenge of Find the Right Principle Diagnosis

出院小结

入院情况：	患者因“发现左肺占位及左侧胸腔积液22天”入院。患者于入院一月前出现左侧胁部疼痛，无头痛、恶心、呕吐……
入院诊断：	左肺占位胸腔积液
住院诊治：	入院后完善各项检查，予以氨酚羟考酮止痛、尿毒清颗粒保护肾功能，注射用头孢曲松钠抗感染治疗……
出院诊断：	1左肺癌胸腔积液2慢性肾功能不全

Various free text discharge summary

Comprehensive Medicine Knowledge

Complex Coding System

Principal Diagnosis	Lung Cancer
ICD-10	C34.900x001

Coding for one patient takes 30min in average;
Lacks experienced coder

Example

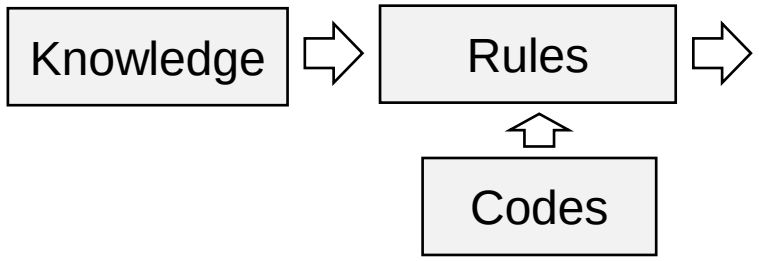
Discharge Summary	
Admission Diagnosis：	1. Secondary epilepsy 2. Right central zone lesion 3. Postoperative right parietal occipital metastasis 4. Chronic hepatitis B 5. Postoperative lung cancer.
Discharge Diagnosis：	1. Secondary epilepsy 2. Right central zone lesion 3. Postoperative right parietal occipital metastasis 4. Chronic hepatitis B 5. Postoperative lung cancer.
Admission Conditions：	The patient was admitted to the hospital because of "lung cancer brain metastasis surgery for more than 1 year, epilepsy for more than 5 months". Case characteristics at admission: ****. Physical examination: ****. Auxiliary examination: ****. Brain CT: ****. Lung CT: ****
Treatment History：	…After consulting with the respiratory department to exclude surgical contraindications, the patient underwent supine right frontal lobe (functional area) lesion resection under general anesthesia on April 27, 2012. The patient recovered well after the operation... The postoperative pathological report showed (right frontal lobe lesion) metastatic moderately differentiated adenocarcinoma, which was consistent with lung cancer metastasis in combination with the clinical history…

Secondary epilepsy 

Secondary malignant brain tumor 

NLP Solutions: Rule-based

Rule-based Approaches



Tuberculous exadative pleurisy 结核性渗出性胸膜炎
右侧结核性渗出性胸膜炎，左侧结核性胸膜炎，结核性渗出性胸膜炎，双侧结核性渗出性胸膜炎，右侧渗出性结核性胸膜炎
^(.){0,3}(左|右|双)?(侧)?结核(性)?渗出(性型)胸膜炎|^(.){0,3}(左|右|双)?(侧)?渗出(性型)结核(性)?胸膜炎

Regular expressions or IF THEN rules

- Relies on experts
- Site-specific
- Limited performance

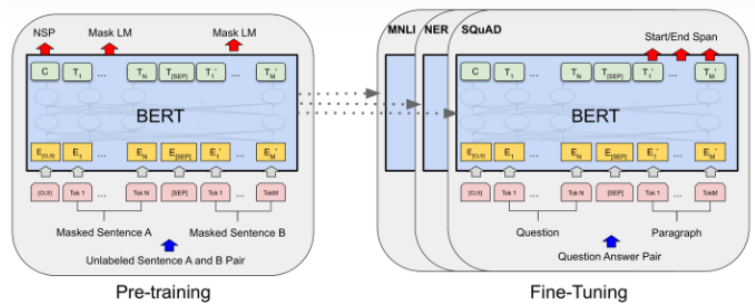
NLP Solutions: Machine Learning

Machine Learning Approaches

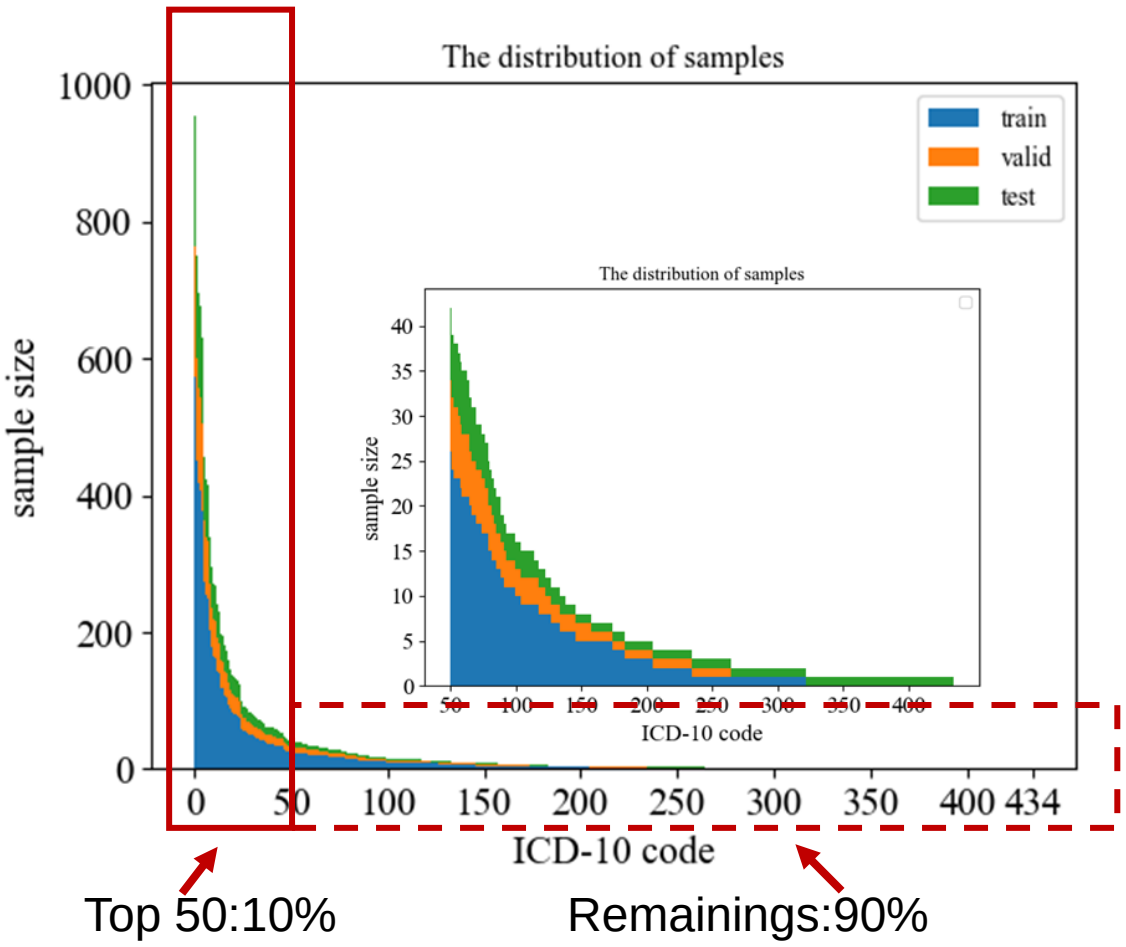
Feature Extraction(based on Statistical Language Models) ⇨ ML Models

$d_{i,1}$	$d_{i,2}$	$d_{i,3}$	$d_{i,4}$	$d_{i,5}$	⋯	$d_{i,51}$
Valvular heart disease	Mild-to-moderate mitral valve insufficiency	Coronary atherosclerosis	Grade 3 hypertension (to be investigated)	-	-	-

Sematic Extraction (based on pre-trained model) ⇨ DL Models (with fine-tuning)

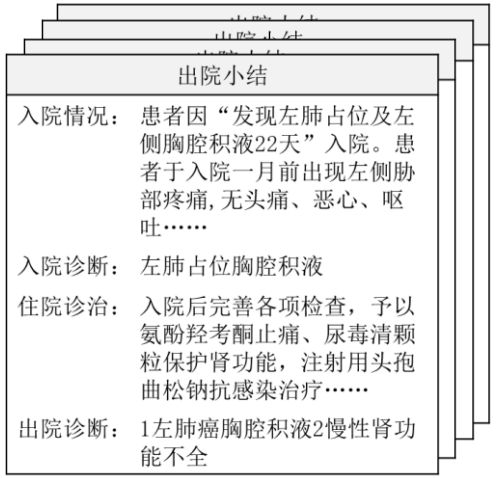


- Length limit of the input text
- Scale limit of the code



How does the LLM Approach perform?

An Experiment



Randomly select 100 discharge summaries from EMR

1 **Directly ask ChatGPT 3.5 Turbo without prompt design**

2

Prompt	
角色：	你现在是一位大型三甲医院的疾病主诊断编码员；
目标：	请根据我提供的病人信息，筛选出一个疾病主诊断的ICD-10细目编码；
约束条件：	你可以参考其他模型给出的答案<****>
约束条件：	直接回答，不需要解释；
病人信息：	*****

Prompt Engineering

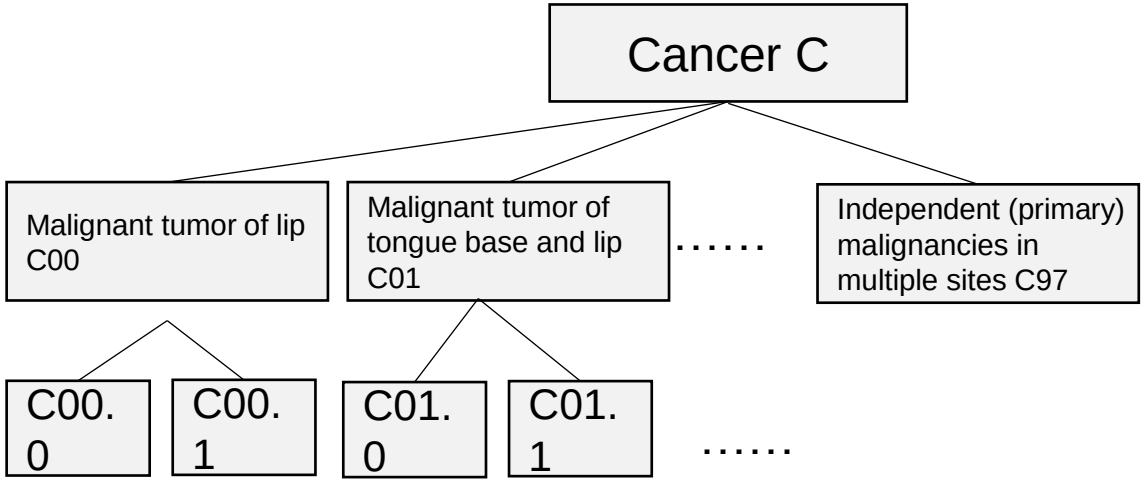
	ChatGPT	ChatGPT-constraint
HR@5	0.39	0.54
HR@1	0.20	0.28

- **Not as good as expected**
- **Retrain and/or instruction tuning need excessive computing power**

How does an experienced human coder do the job?

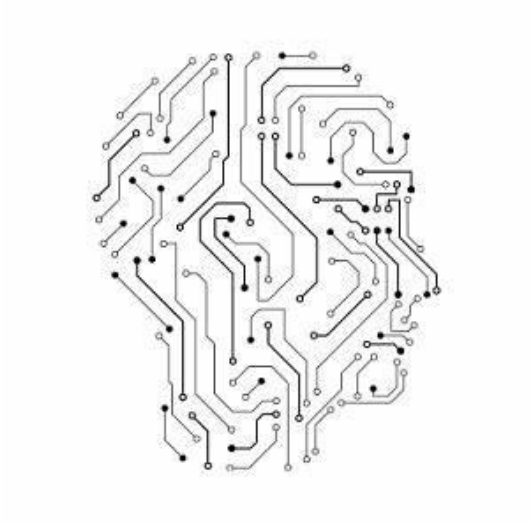
Discharge Summary	
Admission Diagnosis：	1. Secondary epilepsy 2. Right central zone lesion 3. Postoperative right parietal occipital metastasis 4. Chronic hepatitis B 5. Postoperative lung cancer.
Discharge Diagnosis：	1. Secondary epilepsy 2. Right central zone lesion 3. Postoperative right parietal occipital metastasis 4. Chronic hepatitis B 5. Postoperative lung cancer.
Admission Conditions：	The patient was admitted to the hospital because of "lung cancer brain metastasis surgery for more than 1 year, epilepsy for more than 5 months". Case characteristics at admission: ****. Physical examination: ****. Auxiliary examination: ****. Brain CT: ****. Lung CT: ****
Treatment History：	…After consulting with the respiratory department to exclude surgical contraindications, the patient underwent supine right frontal lobe (functional area) lesion resection under general anesthesia on April 27, 2012. The patient recovered well after the operation... The postoperative pathological report showed (right frontal lobe lesion) metastatic moderately differentiated adenocarcinoma, which was consistent with lung cancer metastasis in combination with the clinical history…

ICD-10 Codes



1.Find the important part (word/character, sentence)
2.Find possible code groups
3.Compare against a group of possible codes
4.Decide which ones are most appropriate to choose

Can we teach AI working as a human coder?

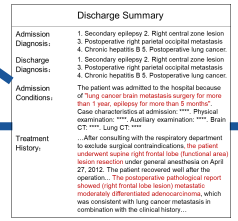


Discharge Summary	
Admission Diagnosis:	1. Secondary epilepsy 2. Right central zone lesion 3. Postoperative right parietal occipital metastasis 4. Chronic hepatitis B 5. Postoperative lung cancer.
Discharge Diagnosis:	1. Secondary epilepsy 2. Right central zone lesion 3. Postoperative right parietal occipital metastasis 4. Chronic hepatitis B 5. Postoperative lung cancer.
Admission Conditions:	The patient was admitted to the hospital because of "lung cancer brain metastasis surgery for more than 1 year, epilepsy for more than 5 months". Case characteristics at admission: ****. Physical examination: ****. Auxiliary examination: ****. Brain CT: ****. Lung CT: ****
Treatment History:	...After consulting with the respiratory department to exclude surgical contraindications, the patient underwent supine right frontal lobe (functional area) lesion resection under general anesthesia on April 27, 2012. The patient recovered well after the operation... The postoperative pathological report showed (right frontal lobe lesion) metastatic moderately differentiated adenocarcinoma, which was consistent with lung cancer metastasis in combination with the clinical history...

2.Find possible code groups

4.Decide which ones are most appropriate to choose

3.Compare against a group of possible codes



1.Find the important part (word/character, sentence)

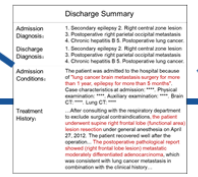
OLR-Net: Object Label Retrieval Network for principal diagnosis extraction

The idea

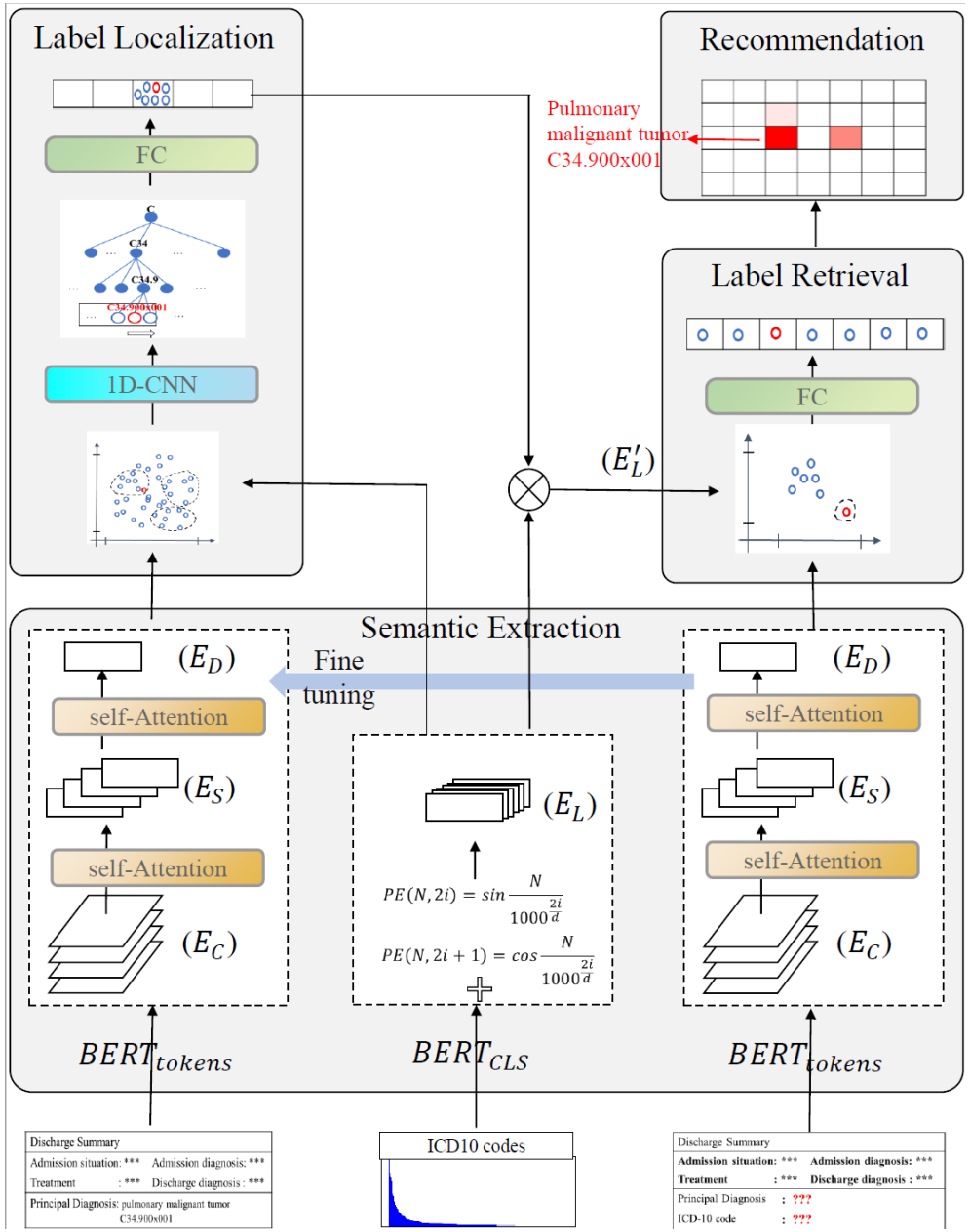
2.Find possible code groups

4.Decide which ones are most appropriate to choose

3.Compare against a group of possible codes



1.Find the important part (word/character, sentence)

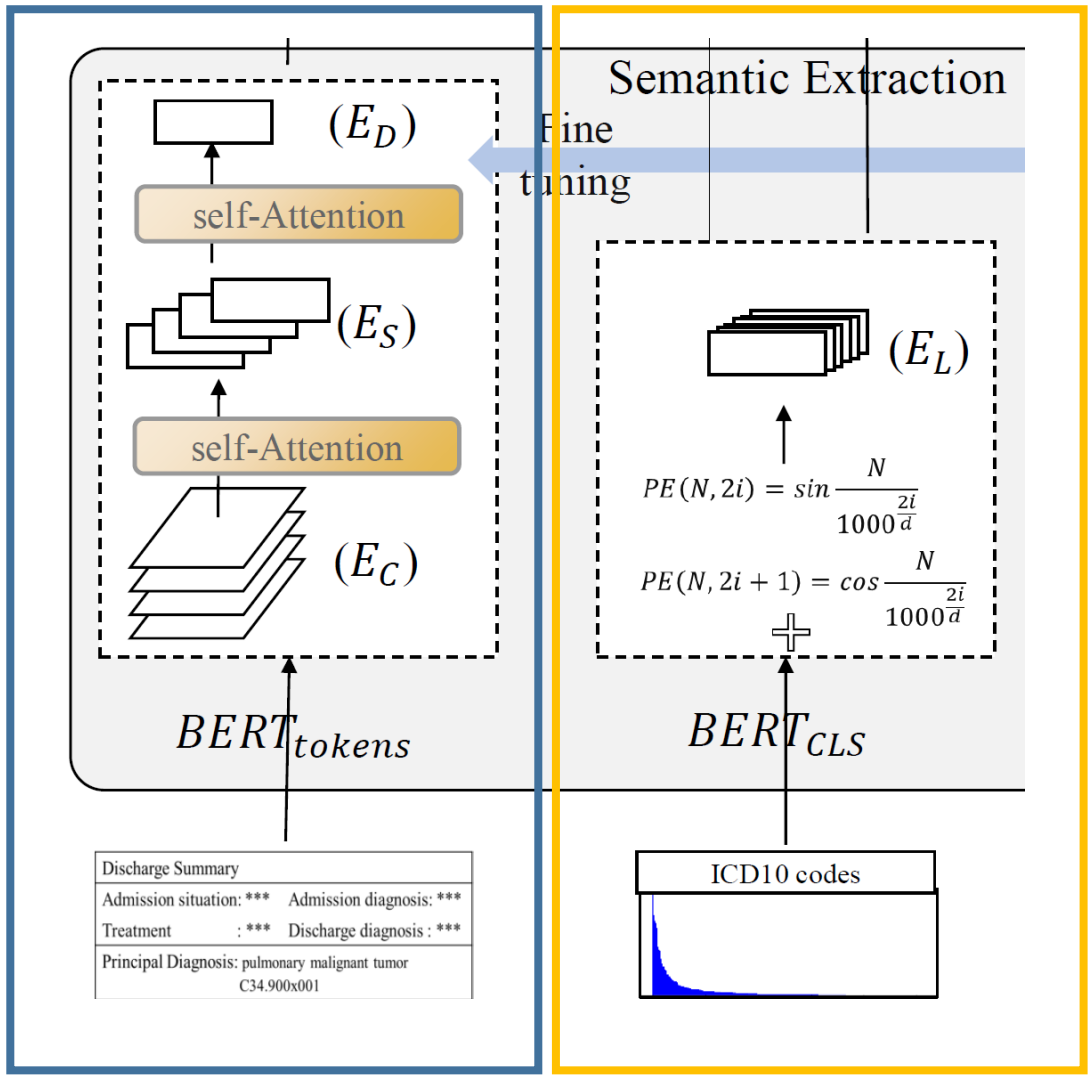


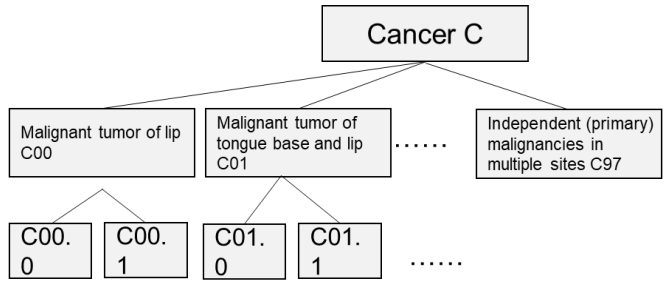
Sematic Extraction

Discharge Summary	
Admission Diagnosis:	1. Secondary epilepsy 2. Right central zone lesion 3. Postoperative right parietal occipital metastasis 4. Chronic hepatitis B 5. Postoperative lung cancer.
Discharge Diagnosis:	1. Secondary epilepsy 2. Right central zone lesion 3. Postoperative right parietal occipital metastasis 4. Chronic hepatitis B 5. Postoperative lung cancer.
Admission Conditions:	The patient was admitted to the hospital because of "lung cancer brain metastasis surgery for more than 1 year, epilepsy for more than 5 months". Case characteristics at admission: ****. Physical examination: ****. Auxiliary examination: ****. Brain CT: ****. Lung CT: ****
Treatment History:	…After consulting with the respiratory department to exclude surgical contraindications, the patient underwent supine right frontal lobe (functional area) lesion resection under general anesthesia on April 27, 2012. The patient recovered well after the operation... The postoperative pathological report showed (right frontal lobe lesion) metastatic moderately differentiated adenocarcinoma, which was consistent with lung cancer metastasis in combination with the clinical history…

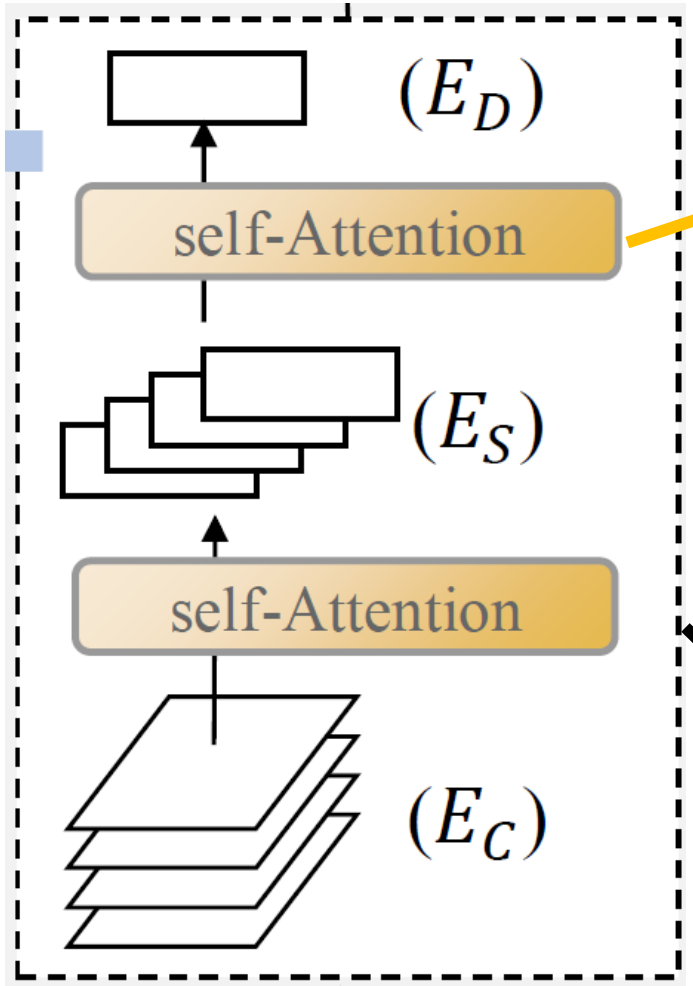
Extract features of the discharge summary

Extract features of the ICD-10 code lists





Document Embedding



$$Q = GELU(W_Q \cdot X + b_Q)$$
$$K = GELU(W_K \cdot X + b_K)$$
$$V = GELU(W_V \cdot X + b_V)$$
$$X_{attention} = softmax(\frac{Q \cdot K^T}{\sqrt{d}}) \cdot V$$
$$X_{residual} = X_{attention} + X$$
$$X_{out} = GVP(X_{residual})$$

$$E_C = BERT_C$$
$$E_S = f_C(E_C)$$
$$E_{Con.} = Concatenate(E_S^1, E_S^2, ..., E_S^{d_s})$$
$$E_D = f_S(E_{Con.})$$

Attention Is All You Need

Ashish Vaswani[*] Google Brain avaswani@google.com
Noam Shazeer[*] Google Brain noam@google.com
Niki Parmar[*] Google Research nikip@google.com
Jakob Uszkoreit[*] Google Research usz@google.com
Llion Jones[*] Google Research llion@google.com
Aidan N. Gomez[*] [†] University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser[*] Google Brain lukaszkaiser@google.com
Illia Polosukhin[*] [‡] illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

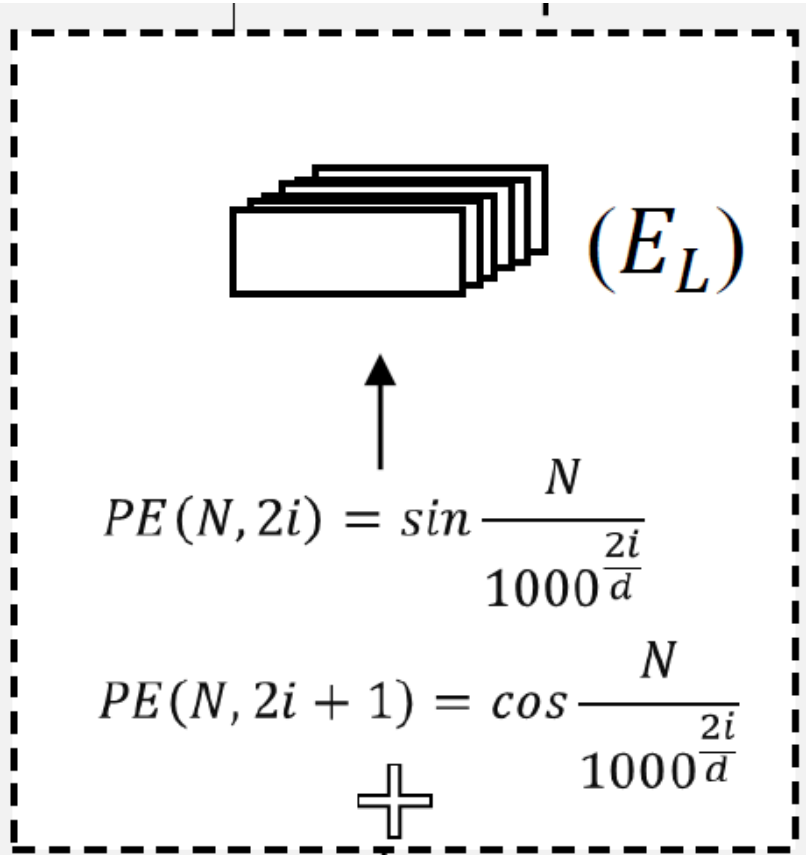
[*]Equal contribution. Listing order is random. Jakob proposed replacing RNNs with self-attention and started the effort to evaluate this idea. Ashish, with Illia, designed and implemented the first Transformer models and has been crucially involved in every aspect of this work. Noam proposed scaled dot-product attention, multi-head attention and the parameter-free position representation and became the other person involved in nearly every detail. Niki designed, implemented, tuned and evaluated countless model variants in our original codebase and tensor2tensor. Llion also experimented with novel model variants, was responsible for our initial codebase, and efficient inference and visualizations. Lukasz and Aidan spent countless long days designing various parts of and implementing tensor2tensor, replacing our earlier codebase, greatly improving results and massively accelerating our research.

[†]Work performed while at Google Brain.

[‡]Work performed while at Google Research.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

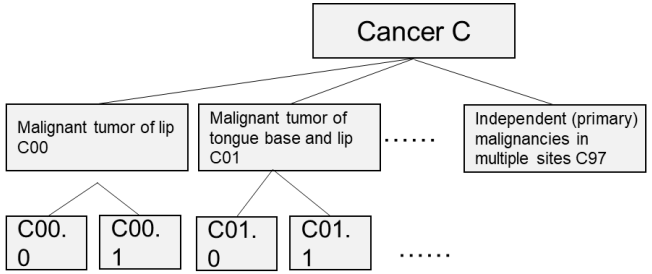
Label Embedding



$$E_L = BERT_L + PE$$

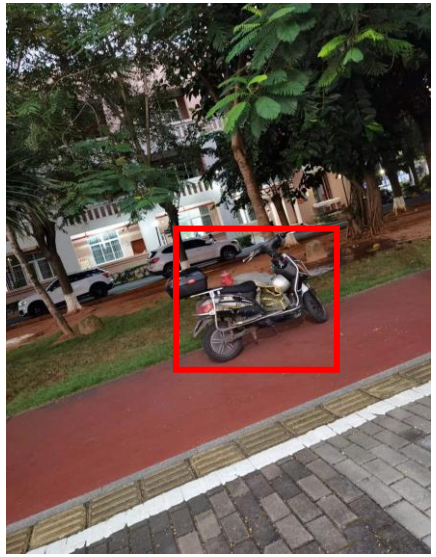
$$PE(N,2i) = \sin \frac{N}{10000^{\frac{2i}{d}}}$$

$$PE(N,2i+1) = \cos \frac{N}{10000^{\frac{2i}{d}}}$$



Label Localization

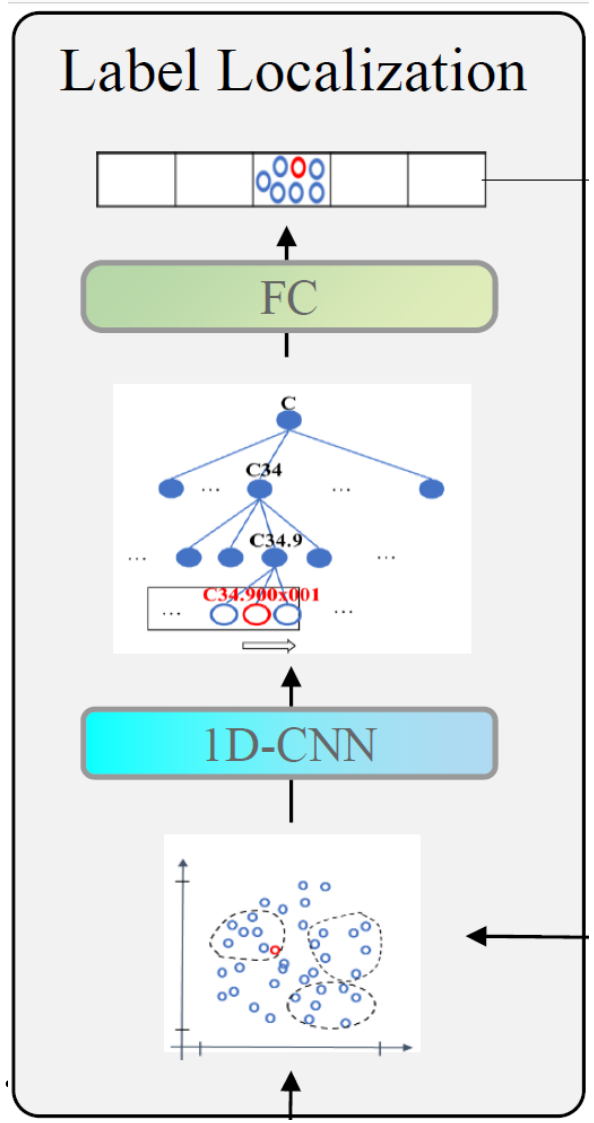
Region proposals



出院小结	
入院情况：	患者因“发现左肺占位及左侧胸腔积液22天”入院。患者于入院一月前出现左侧胸部疼痛，无头痛、恶心、呕吐……
入院诊断：	左肺占位胸腔积液
住院诊治：	入院后完善各项检查，予以氨酚羟考酮止痛、尿毒清颗粒保护肾功能，注射用头孢曲松钠抗感染治疗……
出院诊断：	1左肺癌胸腔积液2慢性肾功能不全

⇒ Code

Code1-Code1i	Code2-Code2i	Code3-Code3i	Code4-Code4i	Code5-Code5i



$$X_{attention} = \frac{E_L \cdot E_D}{\sqrt{d}}$$

$$X_{CNN} = f_{CNN}(W_{CNN}, X_{attention}, \text{“}Same\text{”})$$

$$X_{FC} = f_{FC}(W_{FC}, X_{CNN})$$

$$Y_{localization} = softmax(X_{Dense})$$

Label Retrieval and Code Recommendation

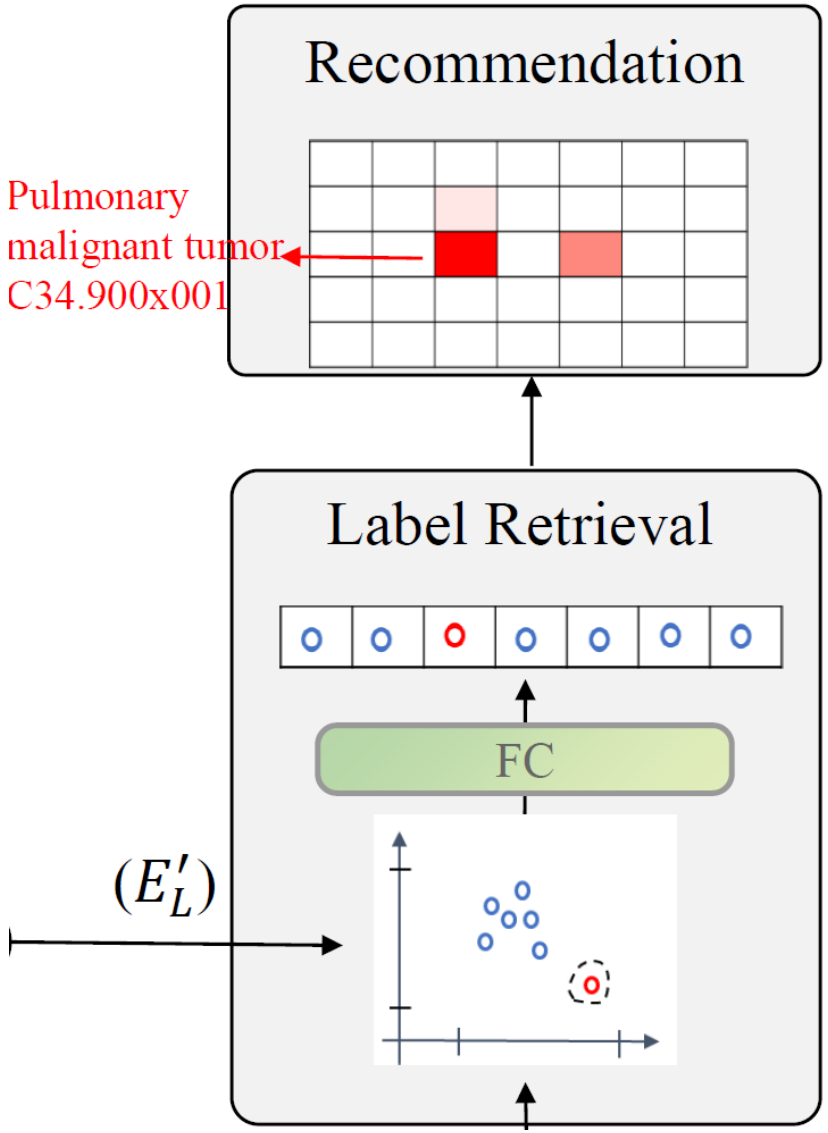
$$P(B_i^n) = P(A_i B_i^n) = P(B_i^n | A_i) \cdot P(A_i)$$

Top n diagnosis ranked by their probabilities

$$E_L' = E_L \cdot Y_{localization}$$

$$X_{attention}' = \frac{E_L' \cdot E_D}{\sqrt{d}}$$

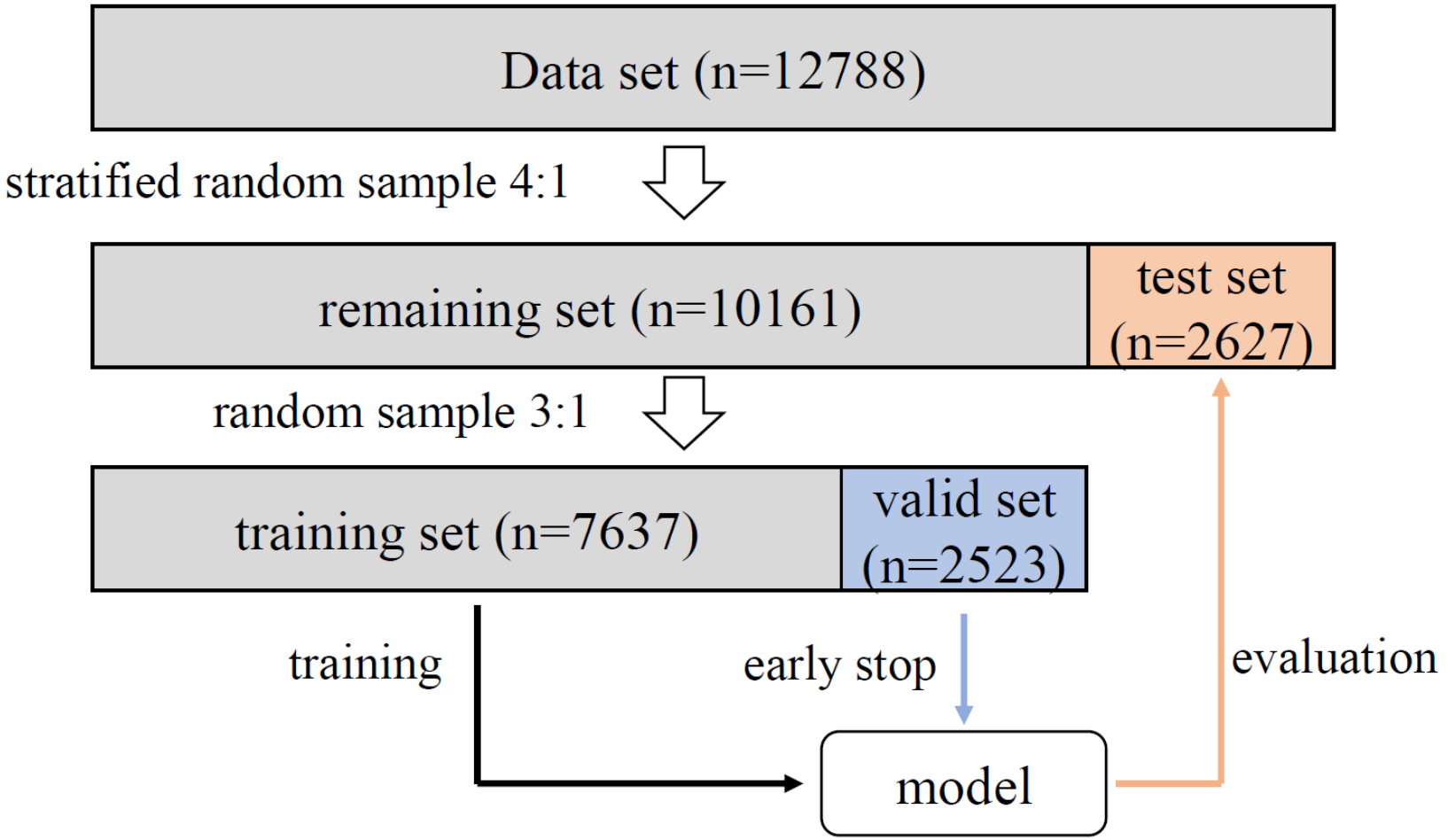
$$Y_{sim} = softmax(X_{attention}')$$



Experiments

The data set

All patients diagnosed as cancer in Hainan PLAGH



Distribution of diagnosis

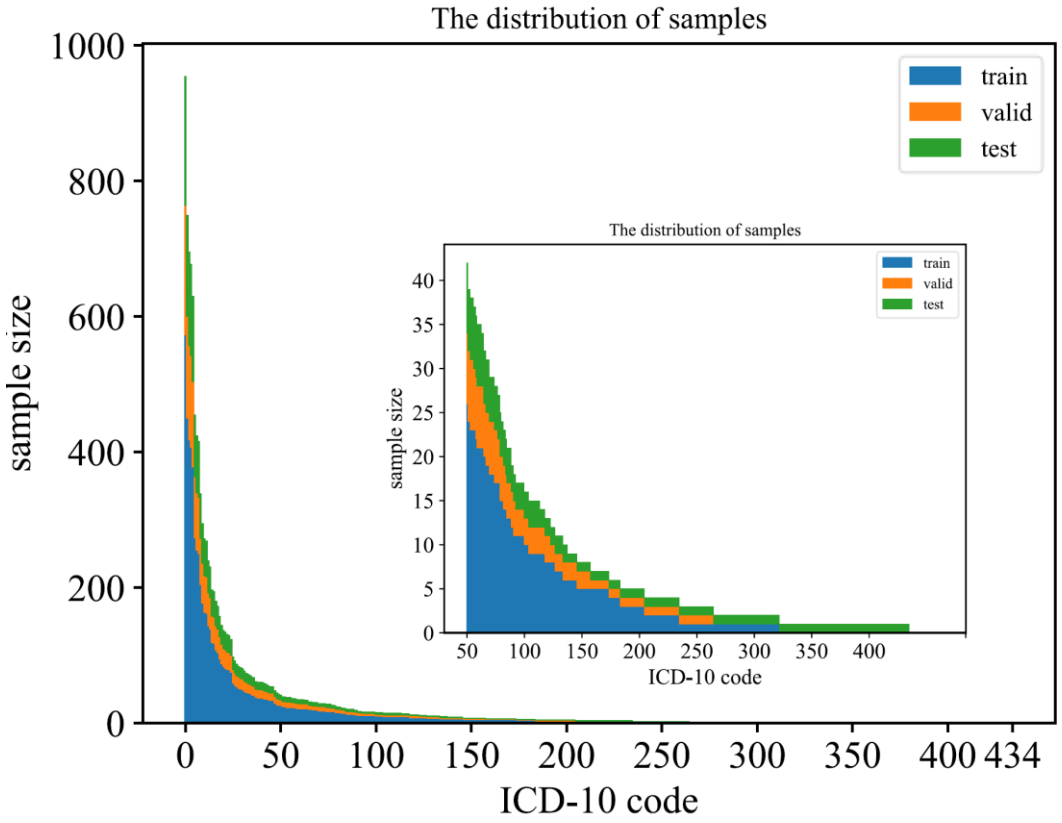


Table 1. The distribution of samples on test set and training set.

	n *	training set	test set
full	$1 \leq n \leq 573$	7637(100.00%)	2627(100.00%)
top50	$26 \leq n \leq 573$	6035(79.02%)	1987(75.64%)
few-shot	$1 < n \leq 10$	662(8.67%)	187(7.12%)
	$10 < n \leq 20$	516(6.76%)	149(5.67%)
	$20 < n \leq 30$	419(5.49%)	130(4.95%)
	$30 < n \leq 40$	354(4.64%)	113(4.30%)
	$40 < n \leq 50$	360(4.71%)	115(4.38%)
	$1 < n \leq 50$	2311(30.26%)	694(26.42%)
one-shot	$n = 1$	87(1.14%)	87(3.31%)
zero-shot	$n = 0$	0	113(4.30%)

* the sample size for the ICD-10 code on training set.

Performance of our approach

(a) Full codes

	HR@5	HR@1	MRR@5	microAUC	macroAUC
CNN	0.7480	0.4636	0.5726	0.9616	0.9433
BiGRU	0.7609	0.5059	0.6061	0.9536	0.9298
ML-Net	0.7678	0.4831	0.5958	0.9685	0.9397
ML-NET+HS-Att	0.8238	0.5634	0.6665	0.9832	0.9021
JointLAAT	0.7693	0.5325	0.6274	0.9619	0.9359
JointLAAT+HS-Att	0.8188	0.5660	0.6632	0.9648	0.9476
OLR-Net	**0.8778**	**0.6483**	**0.7423**	**0.9898**	**0.9851**

(b) Few-shot codes

	HR@5	HR@1	MRR@5	microAUC	macroAUC
CNN	0.4914	0.1715	0.2796	0.9286	0.9285
BiGRU	0.4654	0.1340	0.2463	0.9202	0.9197
ML-Net	0.4755	0.0836	0.2102	0.9238	0.9262
ML-NET+HS-Att	0.6571	0.1888	0.3511	0.9678	0.9698
JointLAAT	0.5216	0.1369	0.2708	0.9272	0.9297
JointLAAT+HS-Att	0.6311	0.2435	0.3822	0.9565	0.9648
OLR-Net	**0.8213**	**0.7426**	**0.6012**	**0.9824**	**0.9833**

$$HR@K = \frac{1}{N} \cdot \sum_{i=1}^{N} hit(i)$$

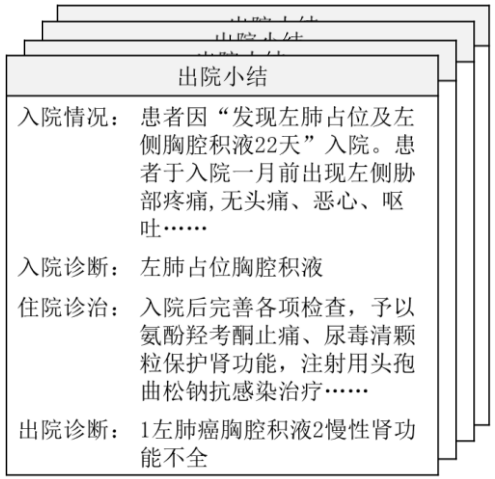
$$MRR@K = \frac{1}{N} \cdot \sum_{i=1}^{N} \frac{1}{p_i}$$

Being completely honest…

	HR@5	HR@1	MRR@5	microAUC	macroAUC
CNN	0.7480	0.4636	0.5726	0.9616	0.9433
BiGRU	0.7609	0.5059	0.6061	0.9536	0.9298
ML-Net	0.7678	0.4831	0.5958	0.9685	0.9397
ML-NET+HS-Att	0.8238	0.5634	0.6665	0.9832	0.9021
JointLAAT	0.7693	0.5325	0.6274	0.9619	0.9359
JointLAAT+HS-Att	0.8188	0.5660	0.6632	0.9648	0.9476
OLR-Net	**0.8778**	**0.6483**	**0.7423**	**0.9898**	**0.9851**

This is not yet satisfying…

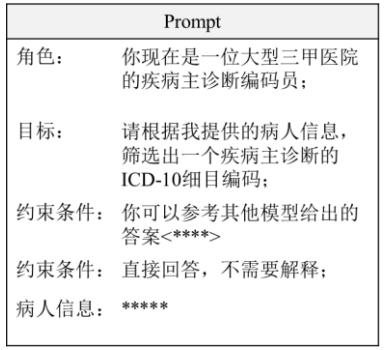
Let’s see if LLMs can help us



Randomly select 100 discharge summaries from EMR

1 **Directly ask ChatGPT 3.5 Turbo without prompt design**

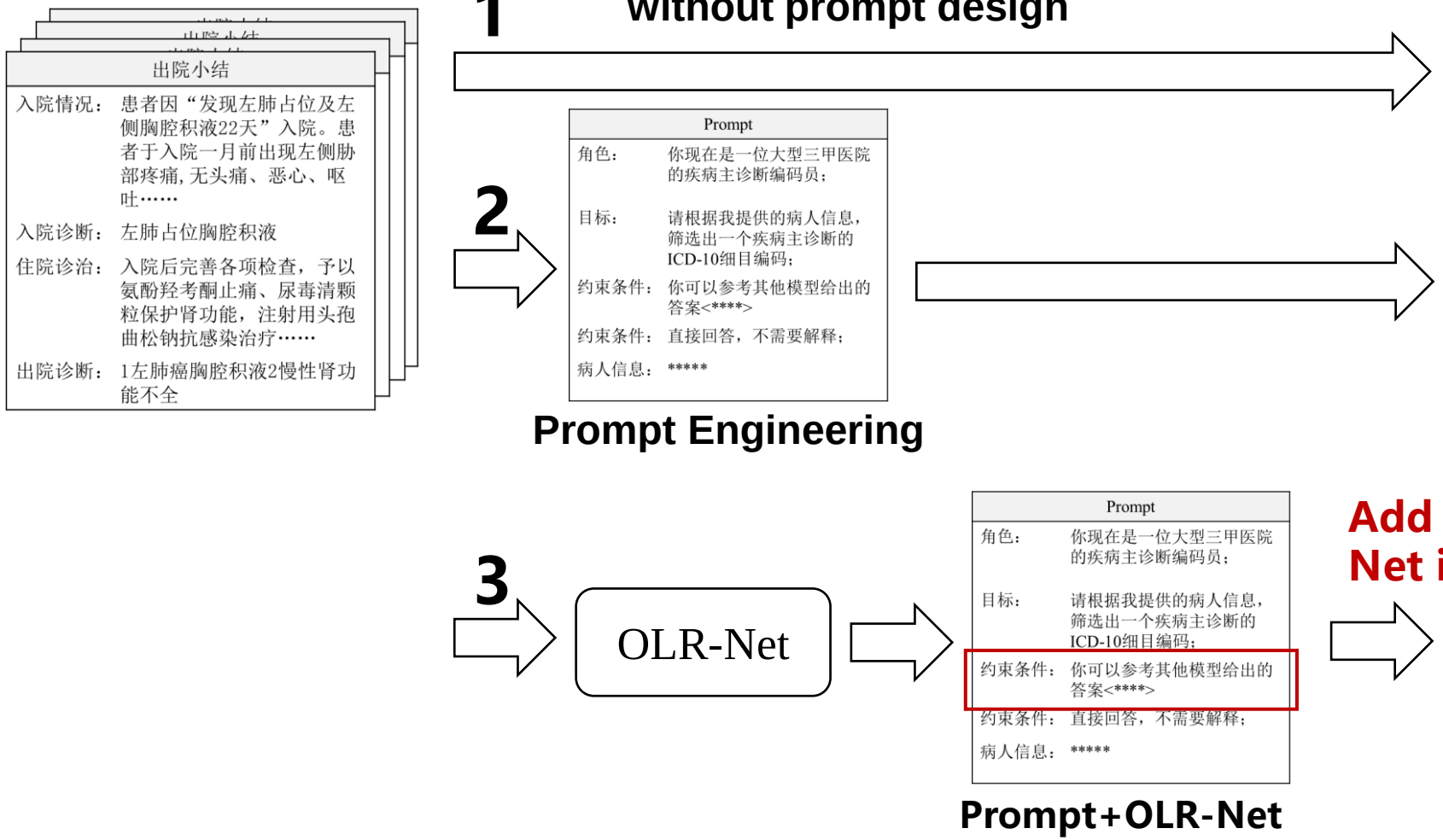
2



Prompt Engineering

	ChatGPT	ChatGPT-constraint
HR@5	0.39	0.54
HR@1	0.20	0.28

Using OLR-Net Jointly with LLMs



Using OLR-Net Jointly with LLMs

Prompt 1 ChatGPT Jointly with OLR-Net

Application Programming Interface: GPT-3.5 Turbo
Input: discharge summary record, Top-5 recommendation of OLR-Net
Output: principal diagnosis
You are the medical record manager. I will provide you with the possible Chinese primary diagnoses and ICD-10 codes. The possible Chinese primary diagnoses: <****>; the possible ICD-10 codes: <****>.
You need give one most possible Chinese primary diagnosis and ICD-10 code through provided information.
Discharge Summary Record:
admission situation: <****>
admission diagnoses: <****>
discharge diagnoses: <****>
treatment : <****>
Please directly answer without explanation

Table 7. The HR@K for experiments associated with ChatGPT.

	OLR-Net	ChatGPT	ChatGPT jointly with OLR-Net
HR@5	**0.85**	0.54	-
HR@1	0.65	0.28	**0.72**

Publication

ELSEVIER

Computers in Biology and Medicine

Volume 182, November 2024, 109130

OLR-Net: Object Label Retrieval Network for principal diagnosis extraction

Kai Wang [a b], Xin Tan [d a], Shan Nan [a], Lei Sang [c], Han Chen [c], Huilong Duan [a d]

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.compbiomed.2024.109130 Get rights and content

Highlights

- OLR-Net has been introduced to extract principle diagnosis from discharge summaries.
- Region proposals can locate diagnoses of interest from long text of medical record.
- Hierarchical self-attention can make use of hidden relationships in medical text.
- The proposed OLR-Net can extract less frequent principal diagnosis.
- The prediction of OLR-Net can serve as information enhancement for ChatGPT.


海南大学生物医学工程学院
Hainan University School of Biomedical Engineering
Thank you