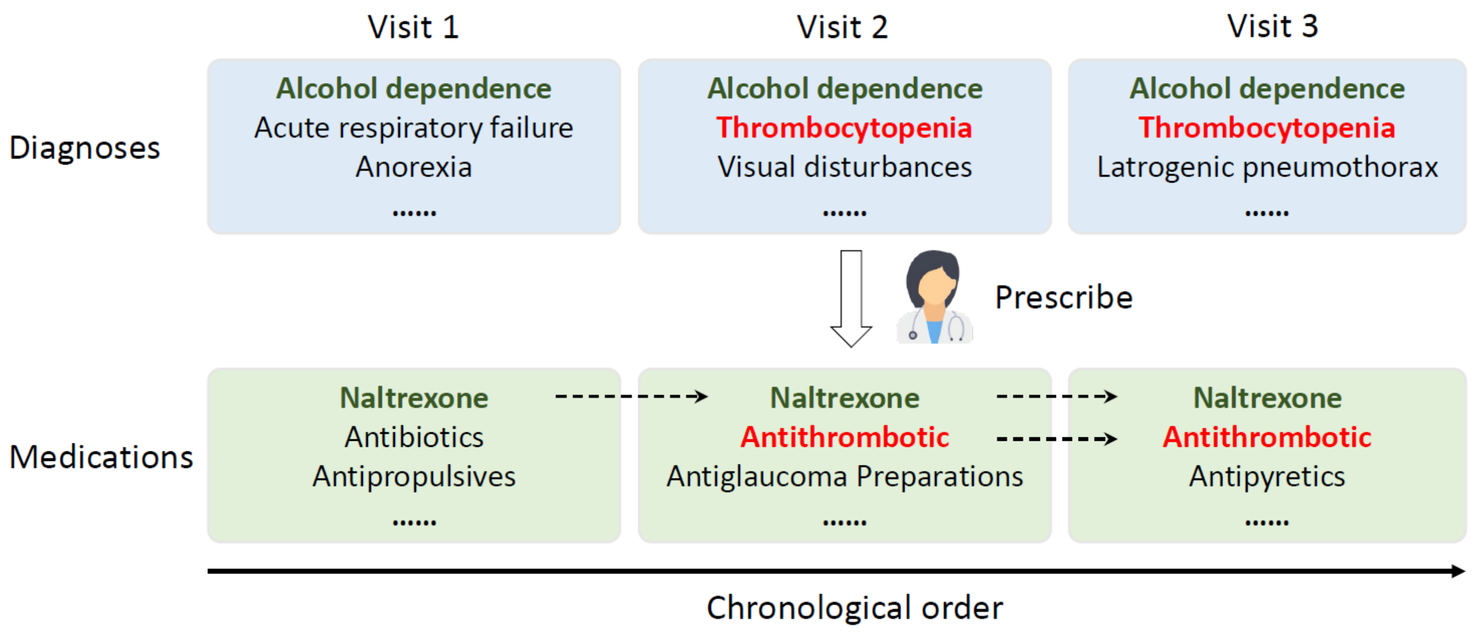
Conditional Generation Net for Medication Recommendation

The Web Conference 2022

Rui Wu[1], Zhaopeng Qiu[2], Jiacheng Jiang[2], Guilin Qi[1], Xian Wu[2]

[1]Southeast University, China, [2]Tencent

Background: Medication Recommendation

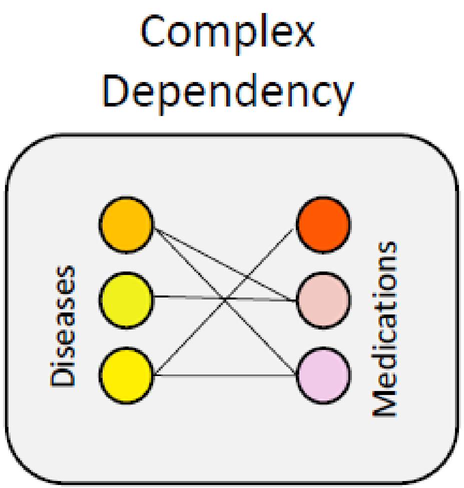


Medication recommendation aims to provide a set of medicines to treat the diagnosed diseases of a patient.

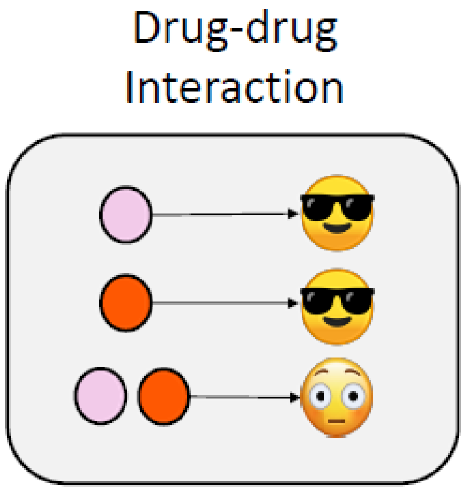
Challenge: Managing Multimorbidity

Patients are usually diagnosed with multiple diseases at one time.

- The doctor needs to select proper medicines for each disease;



- The doctor needs to avoid harmful drug-drug interactions among selected medicines.



The automatic medication recommendation that can assist doctors in decision making is urged.

Reference: https://sites.psu.edu/kdd2021tutorial/files/2021/08/KDD21_tutorial.pdf

Related Work

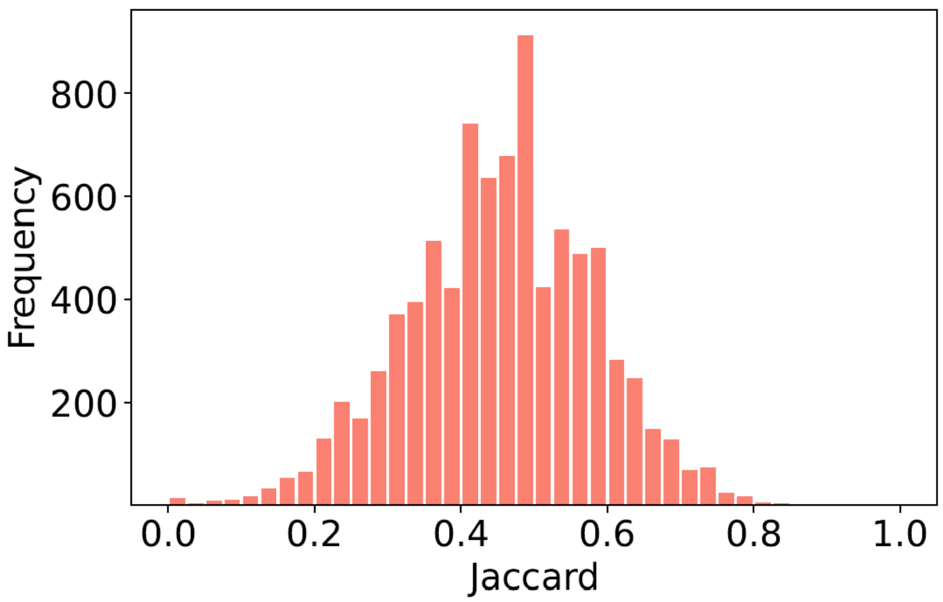
- Instance-based models
  - Use patient’s current diagnoses and procedures to conduct recommendations while ignoring the longitudinal patient history.
  - Fail to consider the historical disease development process.

- **Longitudinal models**
  - Designed to take use of the longitudinal patient history and capture the temporal dependencies.

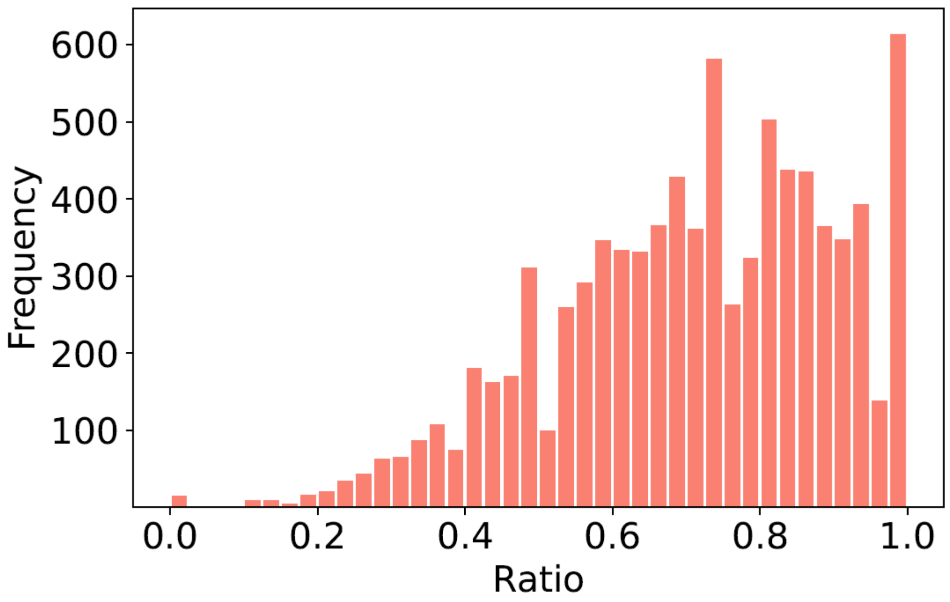
Existing longitudinal models usually consist of two stages:

- First aggregating the known information into a patient-level representation;
- Then conducting medication recommendation based on it.

Relationship between Medications



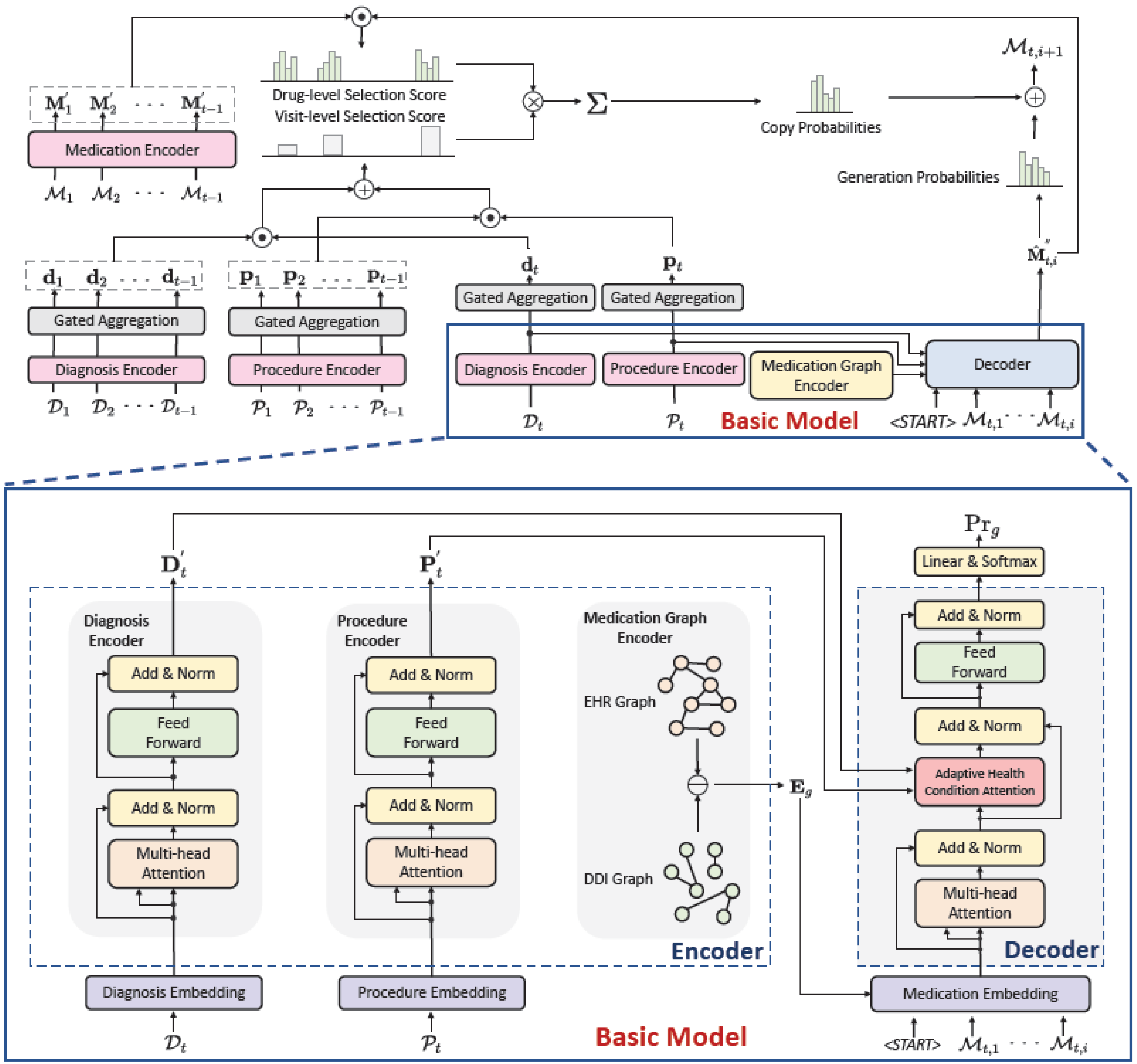
The histogram of Jaccard between current medications and historical medications



The histogram of the proportion of current medications that occur in past visits

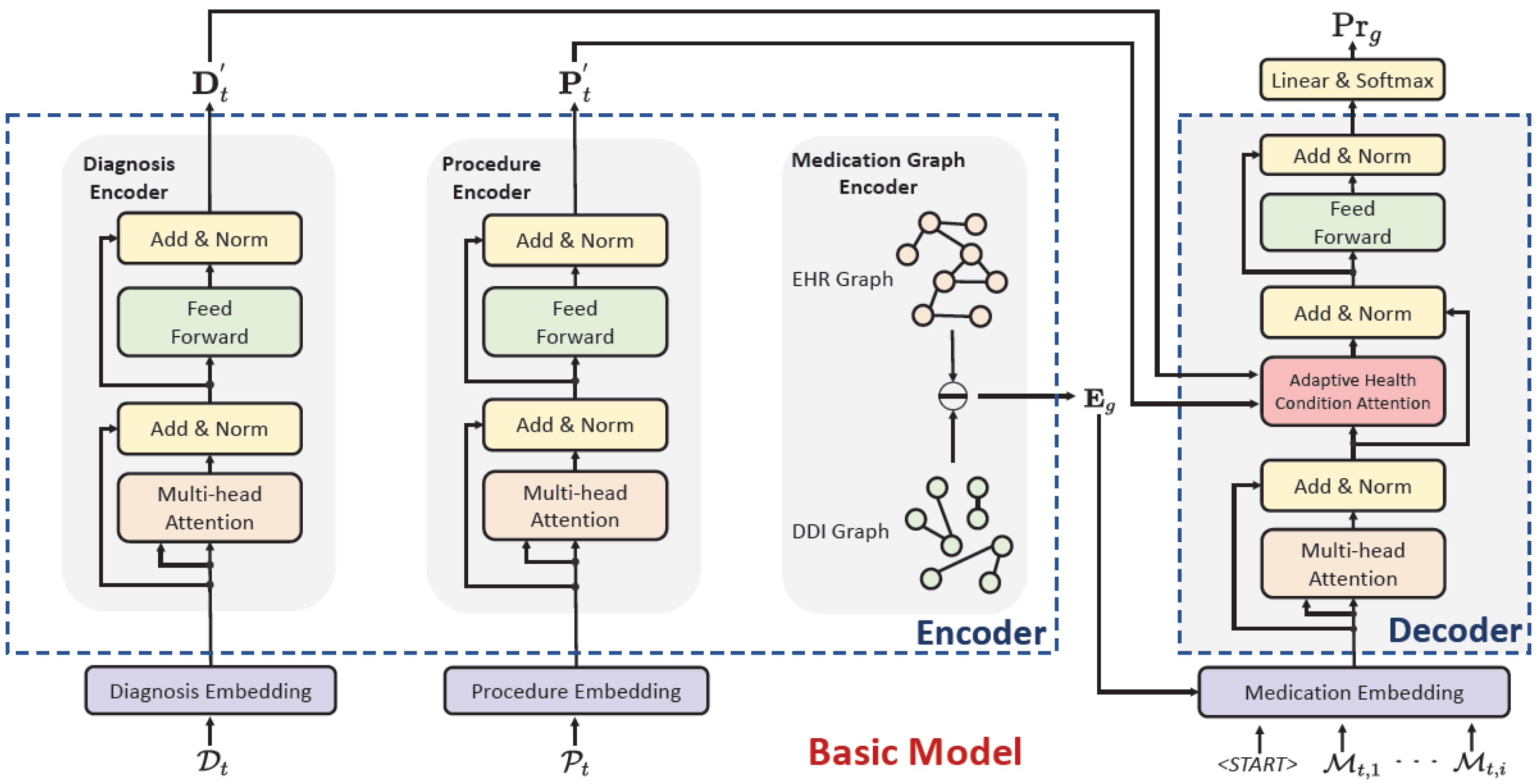
Taking use of historical information from a medication-level perspective could be useful.

Methodology: Conditional Generation Net (COGNet)



Basic Model

- Basic model recommends the medication combination only based on the patient's health condition in current visit.
- The basic model is an encoder-decoder generation model.

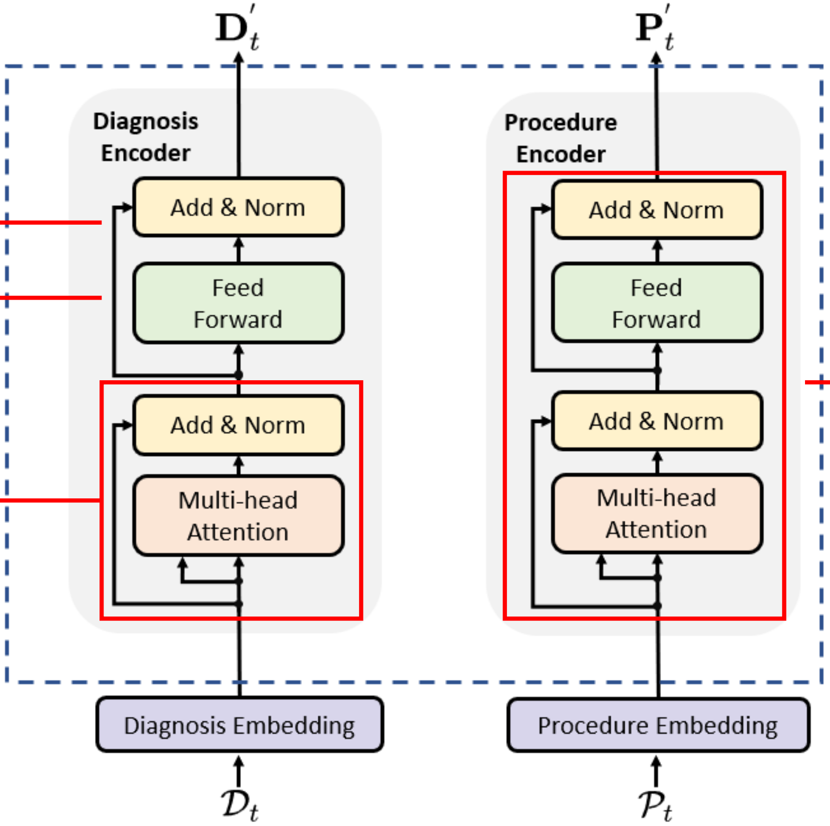


Diagnosis/Procedure Encoder

$$D'_t = \mathrm{Enc_d}(D_t) = \mathrm{LayerNorm}(H + \mathrm{FFN}(H))$$

$$\mathrm{FFN}(H) = \mathrm{ReLU}(HW_1^F + b_1^F)W_2^F + b_2^F$$

$$H = \mathrm{LayerNorm}(D_t + \mathrm{MH}(D_t, D_t, D_t))$$



$$P'_t = \mathrm{Enc_p}(P_t)$$

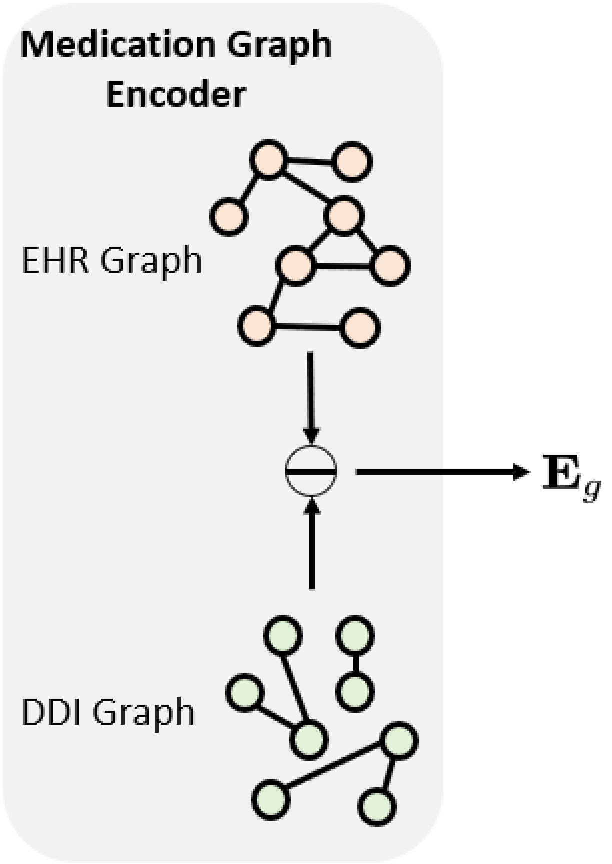
Medication Graph Encoder

co-occurrence relationship

$$G_e = GCN(ReLU(GCN(E_m, A_e))W_e^g, A_e)$$

DDI relationship

$$G_d = GCN(ReLU(GCN(E_m, A_d))W_d^g, A_d)$$



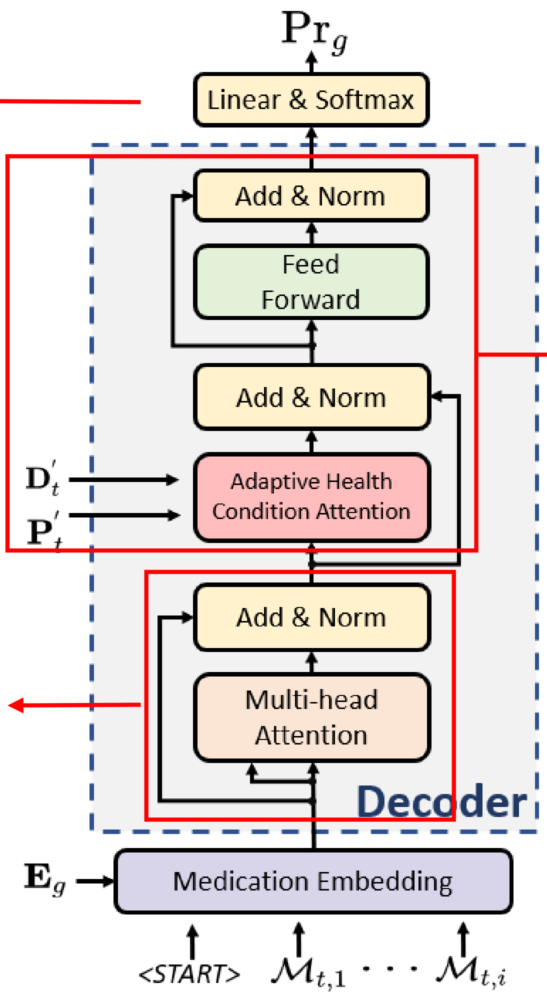
Relation-aware medication embedding

$$E_g = G_e - \lambda G_d$$

Medication Combination Decoder

3. Medication prediction

$$\Pr_g = \mathrm{Softmax}(\hat{\mathbf{M}}''_{t,i-1}\mathbf{W}_g + \mathbf{b}_g)$$



2. Adaptively model the uncovered diseases to guide the next medication recommendation

$$\hat{\mathbf{M}}''_t = \mathrm{LayerNorm}(\hat{\mathbf{M}}'_t + \mathrm{MH}(\hat{\mathbf{M}}'_t, \mathbf{D}'_t, \mathbf{D}'_t) + \mathrm{MH}(\hat{\mathbf{M}}'_t, \mathbf{P}'_t, \mathbf{P}'_t))$$

Query uncovered diseases

Query uncovered procedures

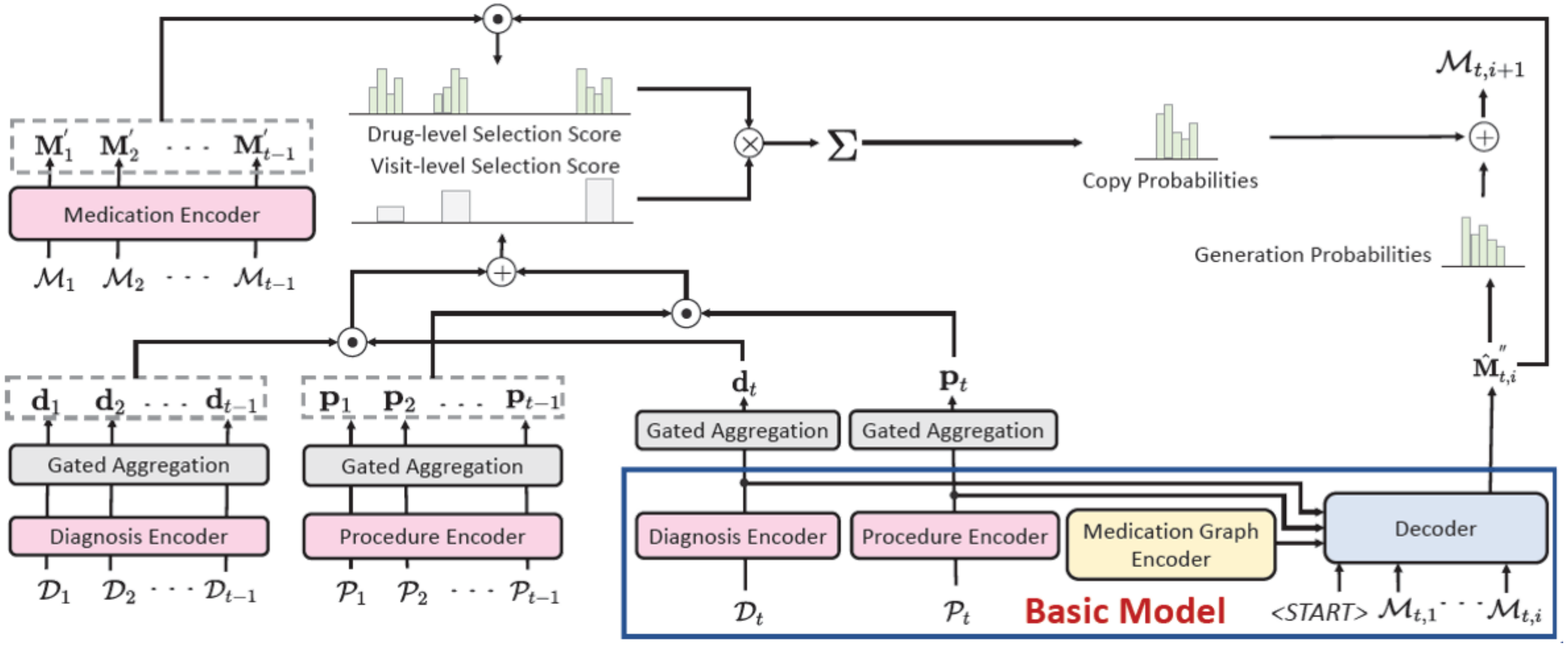
1. Relation-aware medication combination representation

$$\hat{\mathbf{M}}'_t = \mathrm{LayerNorm}(\hat{\mathbf{M}}_t + \mathrm{MH}(\hat{\mathbf{M}}_t, \hat{\mathbf{M}}_t, \hat{\mathbf{M}}_t))$$

$$\hat{\mathbf{M}}_t = \hat{\mathbf{M}}^m_t + \hat{\mathbf{M}}^g_t$$

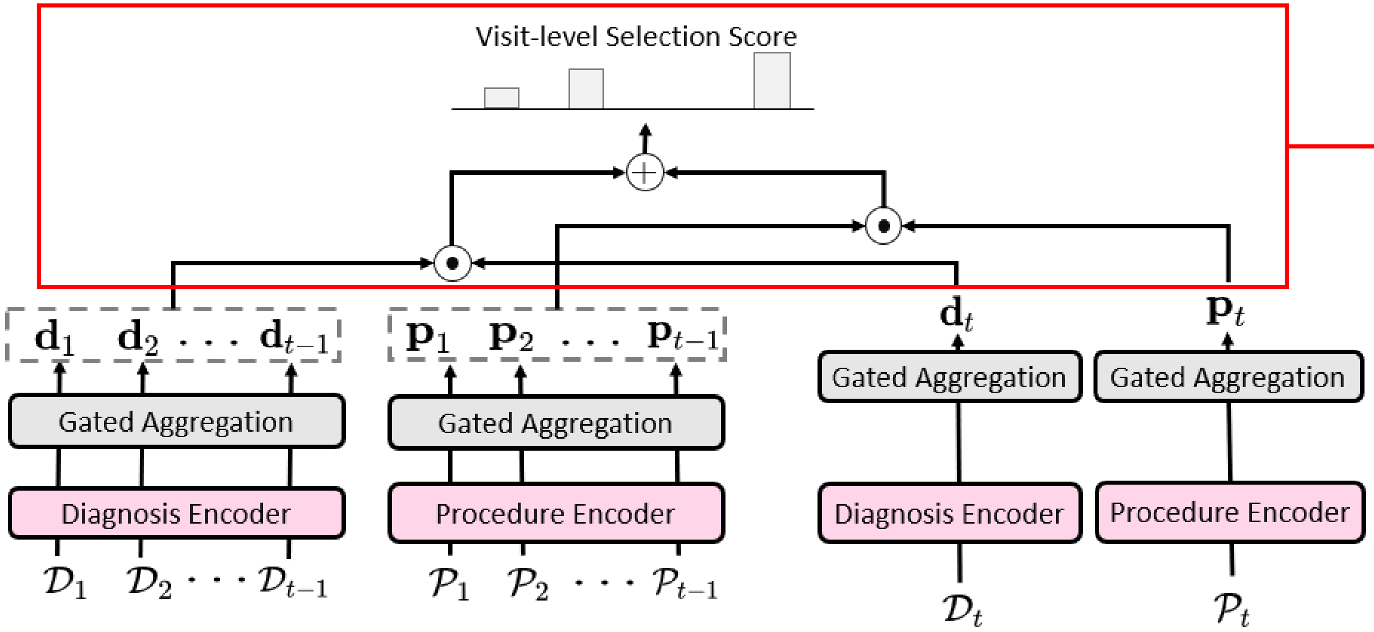
Copy Module

- Copy module compares the health conditions of current and historical visits and then copies the reusable medications to prescribe for current visit according to the condition changes.
- Copy module uses the **hierarchical selection mechanism** to conduct the copy process.



Visit-level Selection

Pick a similar visit by comparing health conditions



$$c_j = \mathrm{Softmax}\left(\frac{\mathrm{v}_j^d \cdot \mathrm{v}_t^d + \mathrm{v}_j^p \cdot \mathrm{v}_t^p}{\sqrt{s}}\right)$$

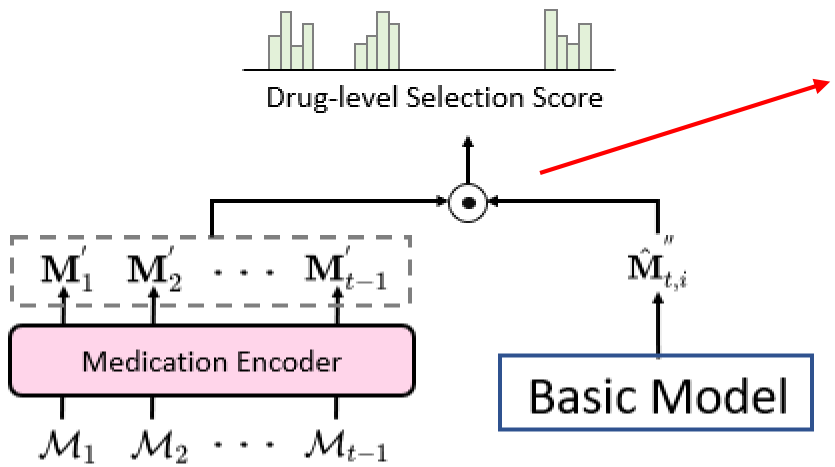
$$\mathrm{v}_j^d = \mathrm{Softmax}(\tanh(\mathrm{D}_j'\mathrm{W}_d^1 + \mathrm{b}_d^1)\mathrm{W}_d^2 + b_d^2)^\top \mathrm{D}_j'$$

$$\mathrm{v}_j^p = \mathrm{Softmax}(\tanh(\mathrm{P}_j'\mathrm{W}_p^1 + \mathrm{b}_p^1)\mathrm{W}_p^2 + b_p^2)^\top \mathrm{P}_j'$$

Drug-level Selection

Use the hidden state from the basic model to determine which historical medication is reusable in current situation

$$\mathbf{M}'_j = \mathrm{Enc_m}(\mathbf{M}_j)$$



$$q_{j,k} = \frac{\exp(\hat{q}_{j,k})}{\sum_{j=1}^{t-1}\sum_{k=1}^{|\mathcal{M}_j|}\exp(\hat{q}_{j,k})}$$

$$\hat{q}_{j,k} = \frac{(\hat{\mathbf{M}}''_{t,i-1}\mathbf{W}_c)\cdot\mathbf{M}'_{j,k}}{\sqrt{s}}$$

Copy Mechanism

1. Combine the visit-level and medication level scores to determine the copy probability

2. Combine the generation probabilities and copy probabilities to conduct the prediction

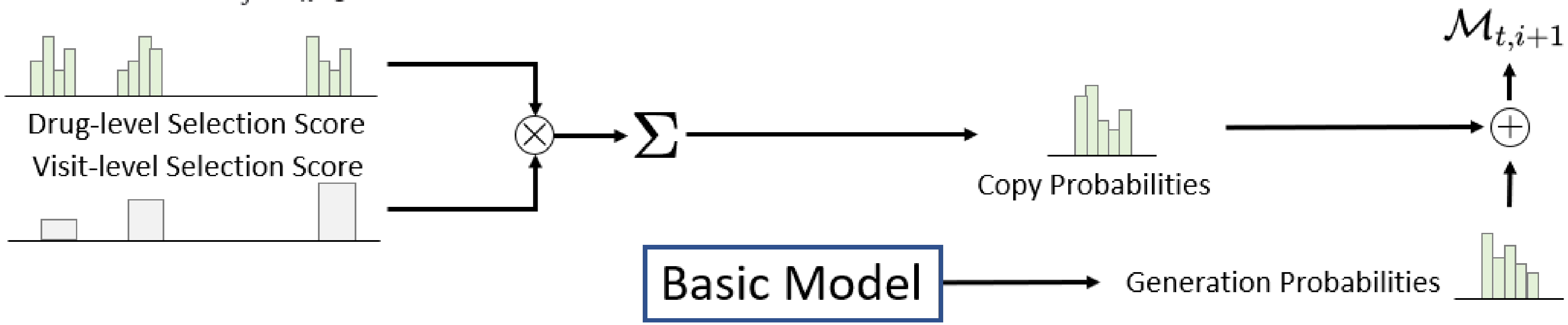
$$p_c^{(i)} = \frac{\hat{p}_c^{(i)}}{\sum_{i=1}^{|\mathcal{M}|} \hat{p}_c^{(i)}}$$

$$where\ \ \hat{p}_c^{(i)} = \sum_{j=1}^{t-1} \sum_{k=1}^{|\mathcal{M}_j|} q_{j,k} * c_j * 1_{\{\mathcal{M}_{j,k}=m_i\}}$$

$$\Pr_c = [p_c^{(1)}, p_c^{(2)}, ..., p_c^{(|\mathcal{M}|)}] \in \mathbb{R}^{|\mathcal{M}|}$$

$$\Pr = w_g * \Pr_g + (1 - w_g) * \Pr_c$$

$$w_g = \text{Sigmoid}(\hat{\mathbf{M}}''_{t,i-1}\mathbf{W}_f + b_f)$$



Results

Performance Comparison on MIMIC-III.

Model	Jaccard	F1	PRAUC	DDI	Avg. # of Drugs
LR	0.4865 ± 0.0021	0.6434 ± 0.0019	0.7509 ± 0.0018	0.0829 ± 0.0009	16.1773 ± 0.0942
ECC	0.4996 ± 0.0049	0.6569 ± 0.0044	0.6844 ± 0.0038	0.0846 ± 0.0018	18.0722 ± 0.1914
RETAIN	0.4887 ± 0.0028	0.6481 ± 0.0027	0.7556 ± 0.0033	0.0835 ± 0.0020	20.4051 ± 0.2832
LEAP	0.4521 ± 0.0024	0.6138 ± 0.0026	0.6549 ± 0.0033	0.0731 ± 0.0008	18.7138 ± 0.0666
DMNC	0.4864 ± 0.0025	0.6529 ± 0.0030	0.7580 ± 0.0039	0.0842 ± 0.0011	20.0000 ± 0.0000
GAMENet	0.5067 ± 0.0025	0.6626 ± 0.0025	0.7631 ± 0.0030	0.0864 ± 0.0006	27.2145 ± 0.1141
MICRON	0.5100 ± 0.0033	0.6654 ± 0.0031	0.7687 ± 0.0026	0.0641 ± 0.0007	17.9267 ± 0.2172
SafeDrug	0.5213 ± 0.0030	0.6768 ± 0.0027	0.7647 ± 0.0025	**0.0589 ± 0.0005**	19.9178 ± 0.1604
COGNet	**0.5336 ± 0.0011**	**0.6869 ± 0.0010**	**0.7739 ± 0.0009**	0.0852 ± 0.0005	28.0903 ± 0.0950

0.08379 in MIMIC-III.

Results

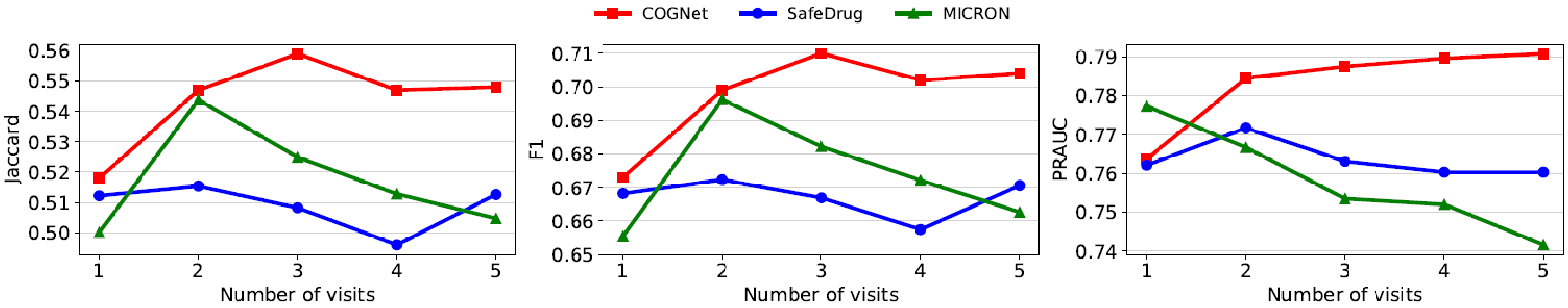
Ablation Study on MIMIC-III.

	Model	Jaccard	F1	PRAUC	DDI	Avg. # of Drugs
1	COGNet *w/o Copy*	0.5163 ± 0.0010	0.6713 ± 0.0009	0.7637 ± 0.0018	0.0842 ± 0.0005	28.3139 ± 0.0766
	COGNet *w/o* c_i	0.5119 ± 0.0016	0.6629 ± 0.0014	0.7588 ± 0.0014	0.0813 ± 0.0005	26.8944 ± 0.0953
	COGNet *w/o G*	0.5306 ± 0.0013	0.6836 ± 0.0012	0.7706 ± 0.0013	0.0840 ± 0.0002	29.1076 ± 0.0795
2	COGNet *w/o* $\mathcal{D}$	0.4937 ± 0.0011	0.6496 ± 0.0011	0.7443 ± 0.0014	0.0887 ± 0.0004	28.0519 ± 0.0995
	COGNet *w/o* $\mathcal{P}$	0.5117 ± 0.0010	0.6669 ± 0.0010	0.7625 ± 0.0016	0.0831 ± 0.0002	28.9554 ± 0.0885
	COGNet *w/o BS*	0.5266 ± 0.0021	0.6805 ± 0.0019	0.7729 ± 0.0013	0.0840 ± 0.0004	28.5592 ± 0.0701
	COGNet	**0.5336 ± 0.0011**	**0.6869 ± 0.0010**	**0.7739 ± 0.0009**	0.0852 ± 0.0005	28.0903 ± 0.0950

1. Copy mechanism with visit-level selection bring a significant improvement to the basic model;
2. Diagnosis and procedure information play a great role in medication recommendation;
3. Graphs and beam search also have contributions to the final result.

Analysis

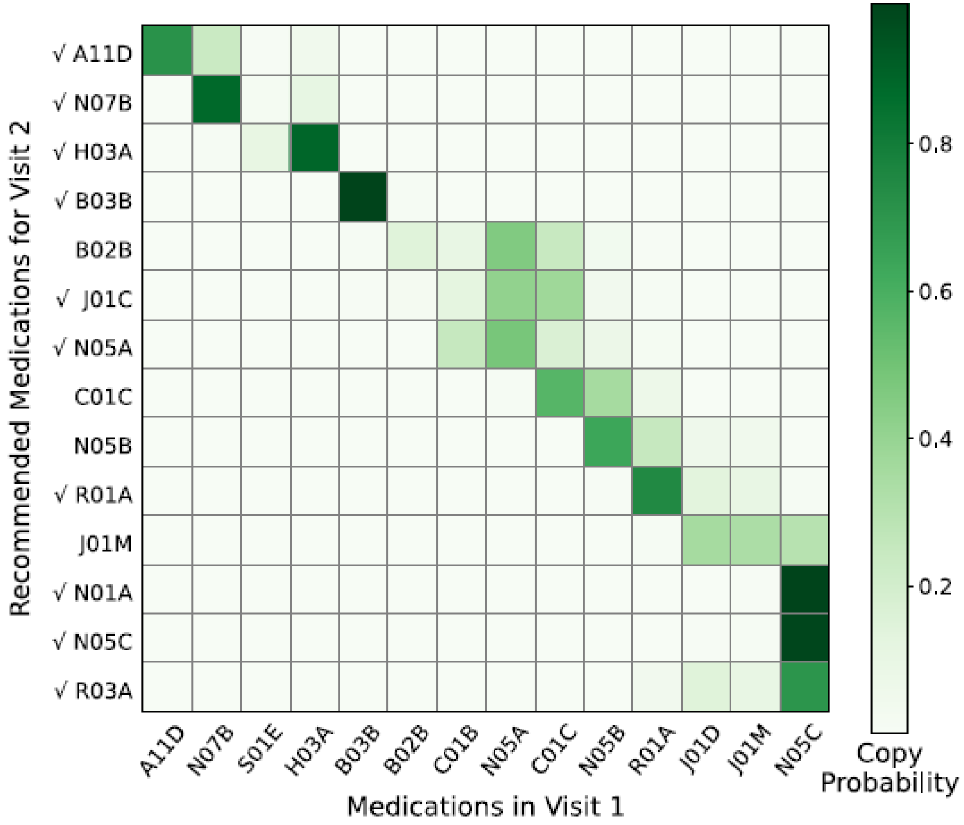
The effect of number of visits for various models.



- COGNet achieves relatively better performance with more visits.
  - The attention based hierarchical selection mechanism can more effectively incorporate the information of past visits;
  - Avoid the error accumulation problem.

Case Study

- Some reusable medications, like A11D, N07B and H03A, are correctly copied by assigning high probabilities to them in previous visit.
- Some new drugs, like J01C and R03A, can also be appropriately generated.



	Diagnoses	Medications
Visit 1	486, 0389, 5185, 78552, 5845, 5849, 99592, 34830, 2869, 2875, 2762, 5990	A11D, N07B, S01E, H03A, B03B, B02B, C01B, N05A, C01C, N05B, R01A, J01D, J01M, N05C
Visit 2	5307, 486, 0389, 99592, 78552, 5121, 5849, 53240, 53140, 2875, 28800	A11D, J01F, N07B, R05C, H03A, B03B, J01C, C01B, N05A, R01A, J01D, N01A, N05C, R03A

Conclusion

- We propose a medication recommendation model, COGNet, which can leverage historical medications to produce a more accurate recommendation;
- We develop a novel hierarchical selection mechanism, which chooses the reusable medicines to copy from both medication-level and visit-level perspectives;
- We conduct comprehensive experiments on a public dataset MIMIC-III to demonstrate the effectiveness of the proposed COGNet.
- Future work
  - Introduce more medical knowledge, like medical ontology and KG;
  - Introduce unstructured data in EHR, like clinical notes and medical images.

Thanks!