Computational Phenotyping from EHR data and Medical Ontologies for Predictive Analytics

William K. Cheung
Kejing Yin, Dong Qian, Lihong Song, Ken Cheong

Dept of Computer Science
Hong Kong Baptist University

Benjamin C.M. Fung

School of Information Studies
McGill University
Canada

Jonathan Poon

Hospital Authority
Hong Kong

Supported by
RGC GRF Grant 12202117

How to get started?

- Critical Care Units
- 2001 - 2012
- 38,597 adult patients
- 53,423 distinct hospital admissions
- Age (med) = 65.8
- In-hospital mortality = 11.5%
- LOS @ICU (med) = 2.1d
- LOS @HOS (med) = 6.9d
- …

EHR Data Analytics: Plug-and-Play?

DEPARTMENT OF
COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Electronic Health Records (EHR):

- Patient demographics
- Medication prescriptions (ATC)
- Diagnoses (ICD-10)
- Laboratory tests (LOINC)
- ...

☺ Providing opportunities for predictive analytics
(mortality, next diagnosis, length of stay, …)

☹ Heterogeneous data types

Complex (different sources, different codes, …)

Missing, noisy, biased (collection process, reimbursement process, …)

Hripcsak, George, and David J. Albers. "Next-generation phenotyping of electronic health records." *Journal of the American Medical Informatics Association* 20.1 (2012): 117-121.

Computational Phenotyping

Suppose you want to identify diabetes patients.

- Searching by diagnosis codes is not good enough.

Toy examples:

Diabetes Diagnoses?	✓	✓	✗	✓	✗
Diabetes Medications?	✓	✗	✓	✗	✗
High blood glucose?	✓	✗	✓	✓	✗
Case patient?	Yes	Probably Not	Yes	Yes	No

Instead, use *<u>the combination of diagnoses, medications, procedures, laboratory tests, etc.</u>* to identify patients with certain conditions.

Phenotypes
(observable properties)

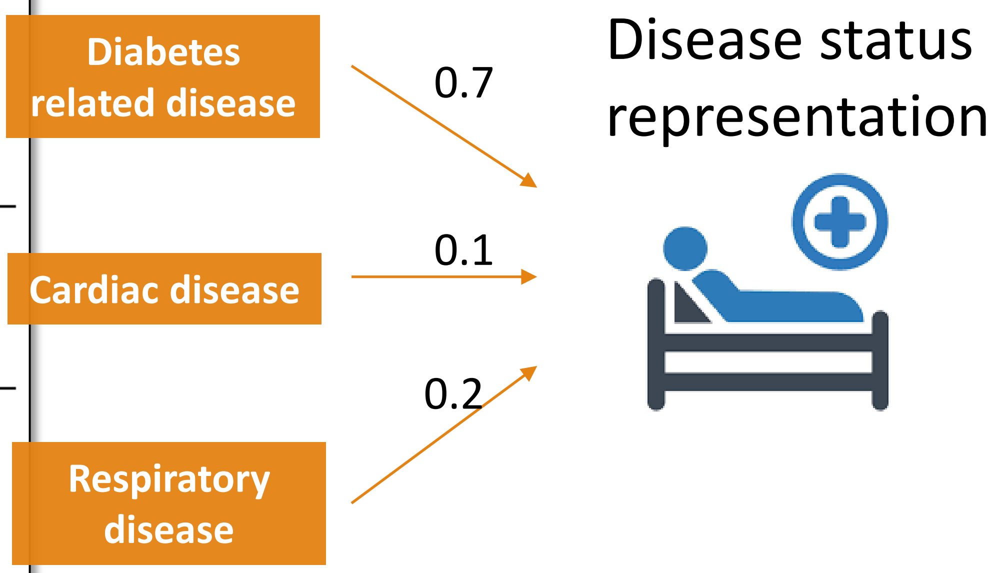
Computational Phenotyping

DEPARTMENT OF
COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Phenotypes

Table 2: Three Examples of Phenotypes

Diagnoses	Medication
Diabetes mellitus Other diseases of lung Acute kidney failure Essential hypertension ...	Insulin Insulin Human Regular
Cardiac dysrhythmias Heart failure Other diseases of lung	Amiodarone HCl Metoprolol Furosemide ...
Other diseases of lung Cardiac dysrhythmias Heart failure Chronic airway obstruction, not elsewhere classified	Albuterol Diltiazem Ipratropium Bromide MDI Fluticasone Propionate ...



Hripcsak, George, and David J. Albers. "Next-generation phenotyping of electronic health records." *Journal of the American Medical Informatics Association* 20.1 (2012): 117-121.

Computational Phenotyping

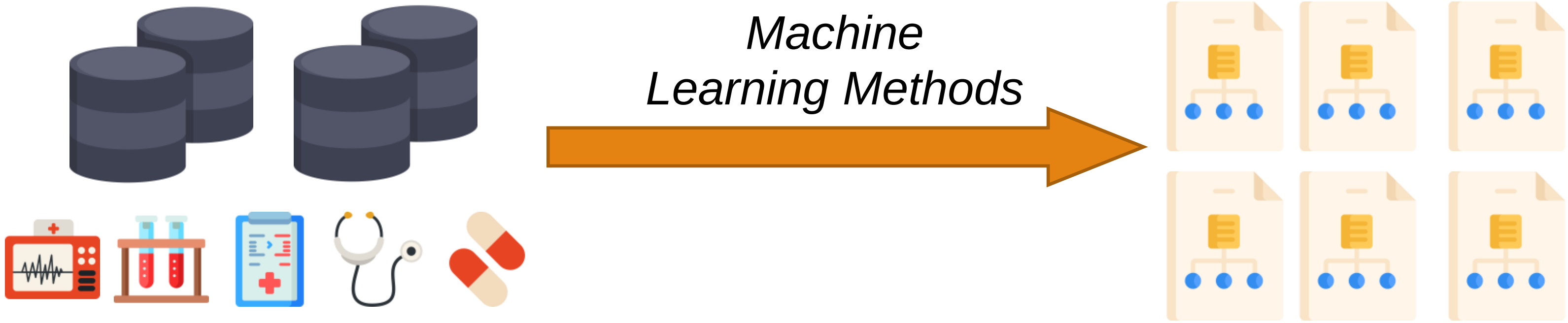
DEPARTMENT OF COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Phenotypes:

The combination of clinically meaningful items (e.g. diagnoses and medications) that reveals the true disease status.

Computational Phenotyping:

The process of automatically discovering meaningful phenotypes from the raw EHR data.



Machine Learning Methods
Natural Language Processing (NLP)
Deep Learning
Matrix Factorization
Tensor Factorization

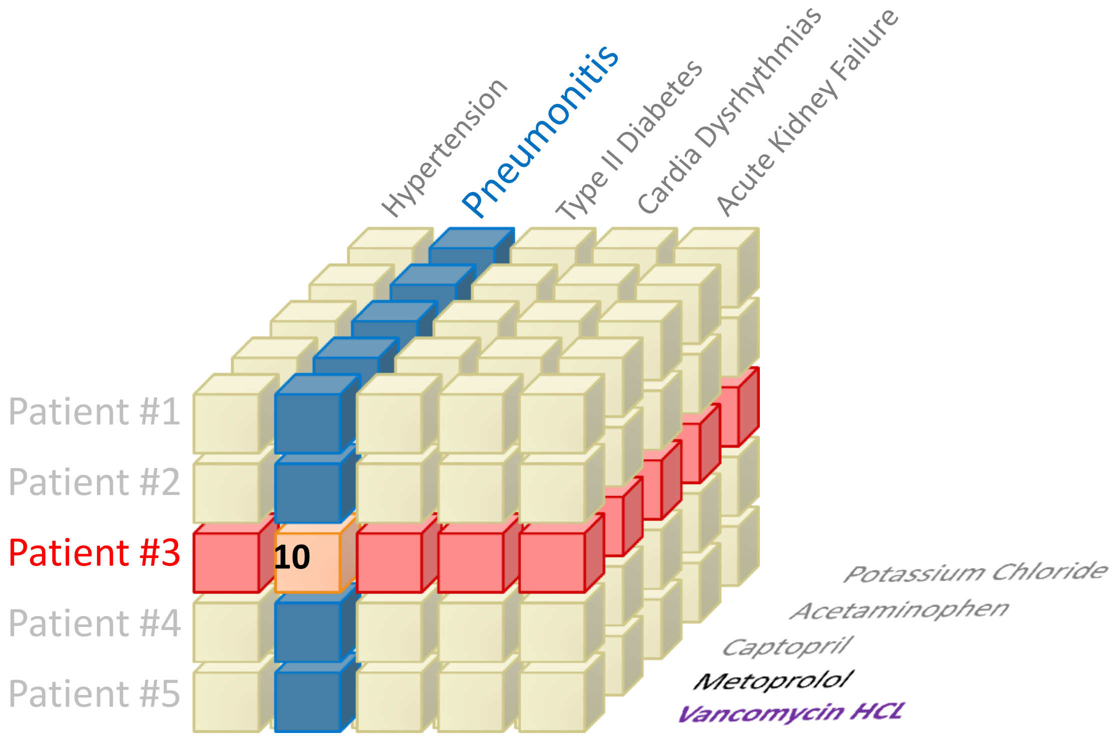
[1] Kirby, Jacqueline C., et al. "PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability." *Journal of the American Medical Informatics Association* 23.6 (2016): 1046-1052.
[2] Ho, Joyce C., et al. "Limestone: High-throughput candidate phenotype generation via tensor factorization." *Journal of biomedical informatics* 52 (2014): 199-211.
[3] Yang, Kai, et al. "TaGiTeD: Predictive Task Guided Tensor Decomposition for Representation Learning from Electronic Health Records." *AAAI*. 2017.

Hidden Interaction Tensor Factorization [IJCAI-18]

for Joint Learning of Phenotypes and Diagnosis-Medication Correspondence

Yin, Kejing, et al. "Joint Learning of Phenotypes and Diagnosis-Medication Correspondence via Hidden Interaction Tensor Factorization." *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*. 2018.

Tensor Factorization for Phenotyping

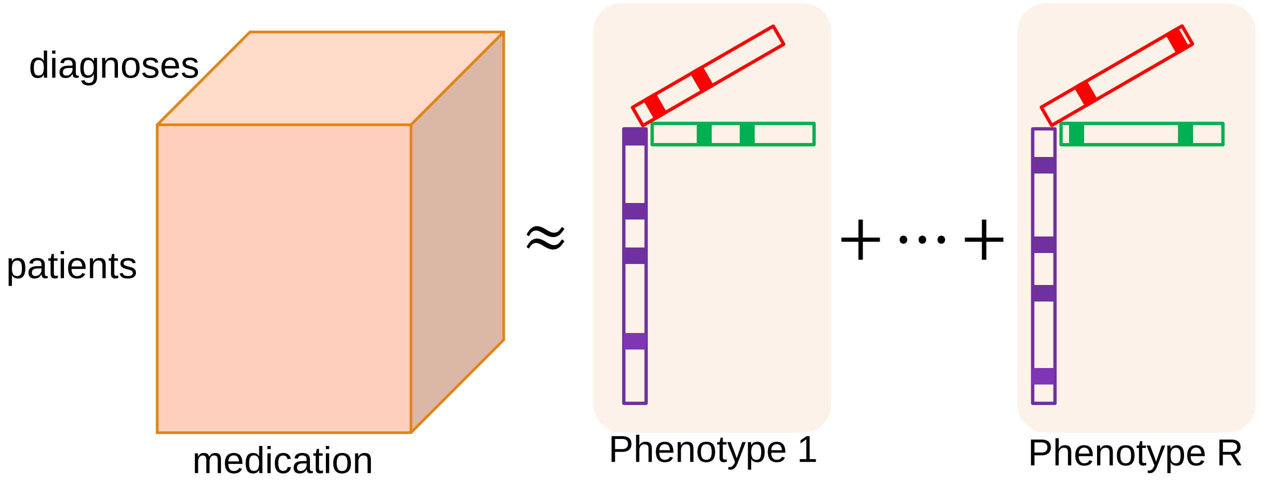


Patient #3 is prescribed with Vancomycin HCL for ten times in response to Pneumonitis.

[1] Ho, Joyce C., et al. "Limestone: High-throughput candidate phenotype generation via tensor factorization." *Journal of biomedical informatics* 52 (2014): 199-211.
[2] Ho, Joyce C., Joydeep Ghosh, and Jimeng Sun. "Marble: high-throughput phenotyping from electronic health records via sparse nonnegative tensor factorization." *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 2014.
[3] Wang, Yichen, et al. "Rubik: Knowledge guided tensor factorization and completion for health data analytics." *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* ACM, 2015.
[4] Kim, Yejin, et al. "Discriminative and distinct phenotyping by constrained tensor factorization." *Scientific reports* 7.1 (2017): 1114.
[5] Yang, Kai, et al. "TaGiTeD: Predictive Task Guided Tensor Decomposition for Representation Learning from Electronic Health Records." *AAAI.* 2017.
[6] Henderson, Jette, et al. "Granite: Diversified, Sparse Tensor Factorization for Electronic Health Record-Based Phenotyping." *2017 IEEE International Conference on Healthcare Informatics (ICHI)*, 2017.

Tensor Factorization for Phenotyping

- Non-negative CP factorization for computational phenotyping:



Approximation with sum of R rank-one tensors: $\mathcal{X} \approx \hat{\mathcal{X}} = \sum_{r=1}^{R} a_r \circ b_r \circ c_r$

Minimize the reconstruction error: $\min \ \mathrm{Error}(\mathcal{X}, \hat{\mathcal{X}})$

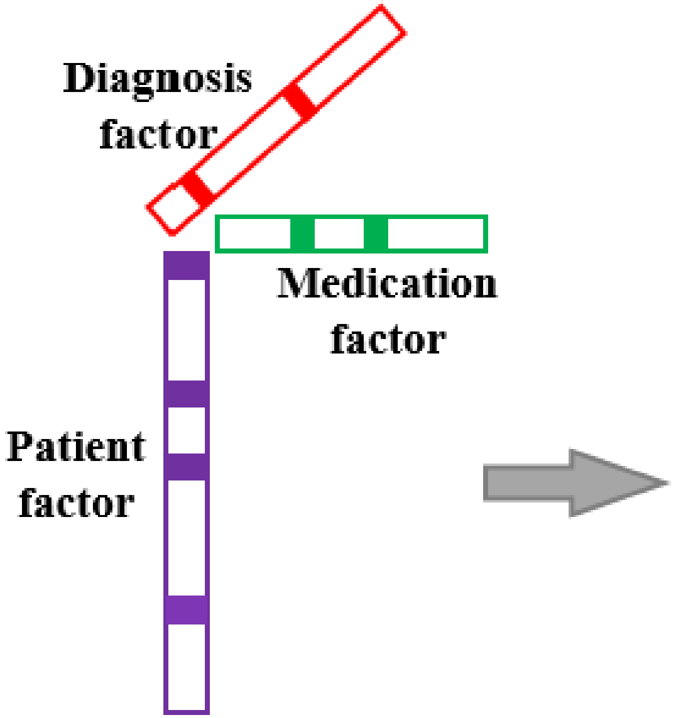
Interaction patterns are captured by the rank-one tensors.

[1] Kolda, T. G., & Bader, B. W. (2008). Tensor Decompositions and Applications. *SIAM Review*, 51(3)
[2] Chi, Eric C., and Tamara G. Kolda. On tensors, sparsity, and nonnegative factorizations. *SIAM Journal on Matrix Analysis and Applications* 33.4 (2012): 1272-1299.

Tensor Factorization for Phenotyping

DEPARTMENT OF
COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

- Phenotype extraction from rank-one tensor:



Diagnosis factor

Diagnosis entry	value
Cardiac dysrhythmias	**0.7**
Diabetes mellitus	0
Heart failure	**0.3**
Asthma	0
Pneumonitis	0

Medication factor

Medication entry	value
Insulin	0
Metoprolol	**0.55**
Furosemide	**0.45**
Albuterol	0
Diltiazem	0

[1] Kolda, T. G., & Bader, B. W. (2008). Tensor Decompositions and Applications. *SIAM Review*, 51(3)
[2] Chi, Eric C., and Tamara G. Kolda. On tensors, sparsity, and nonnegative factorizations. *SIAM Journal on Matrix Analysis and Applications* 33.4 (2012): 1272-1299.

Research Challenge

DEPARTMENT OF
COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

- Interaction information is often **missing** in the records.

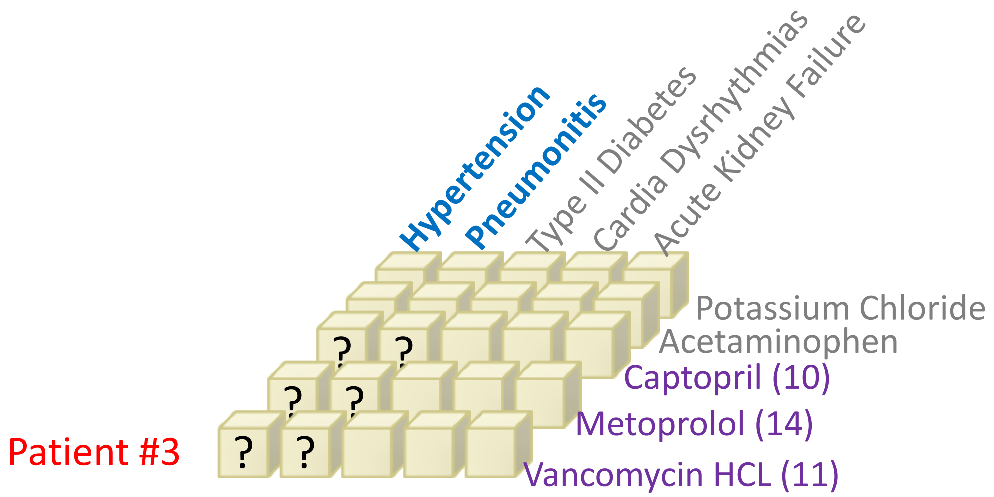
List of medications

Vancomycin HCL	11
Metoprolol	14
Captopril	10
…	…

Correspondence?
Unknown!

List of diagnoses

Essential Hypertension
Pneumonitis
Type II Diabetes
…



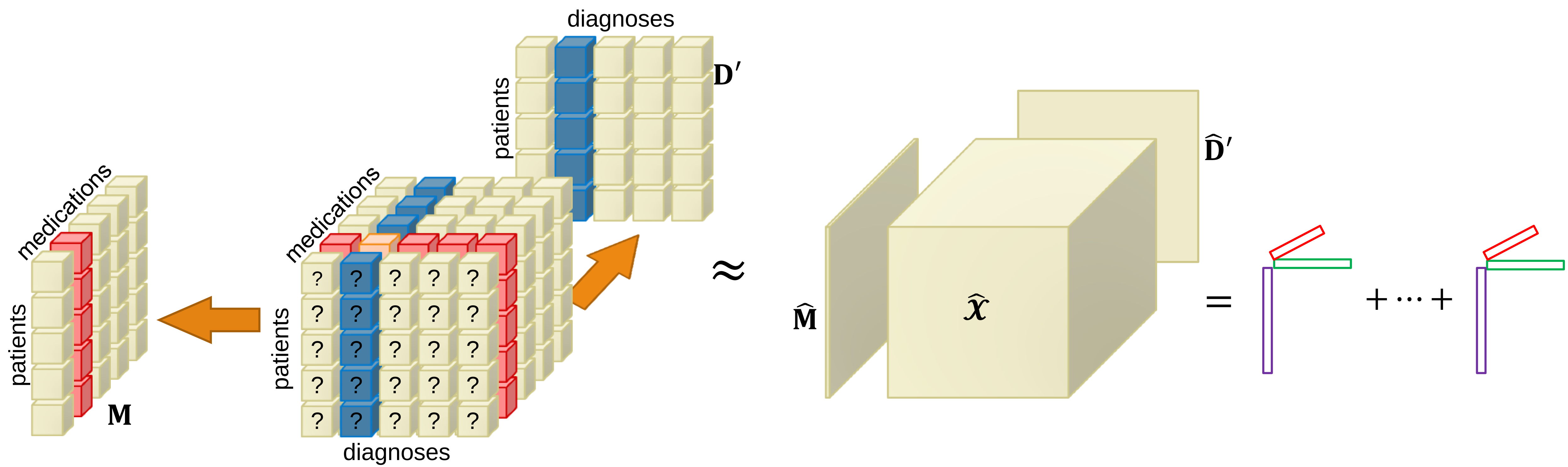
How to fill in the entries?

How to factorize the tensor when we do not observe it?

Hidden Interaction Tensor Factorization

DEPARTMENT OF COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Key Idea



Interaction tensor $\mathcal{X}$:
NOT observed

Experimental Results

DEPARTMENT OF COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Diagnosis-Medication Correspondence

Table 1: Top Five Corresponding Medications for Three Diagnoses Inferred by HITF and Rubik

Cardiac dysrhythmias(39.0%)		Diabetes mellitus(25.3%)		Asthma(5.5%)	
HITF	**Rubik**	**HITF**	**Rubik**	**HITF**	**Rubik**
Furosemide(0.08)	Potassium Chloride(0.08)	Insulin(0.64)	Insulin(0.09)	Albuterol 0.083% Neb Soln(0.46)	Potassium Chloride(0.08)
Potassium Chloride(0.07)	Insulin(0.06)	Insulin Human Regular(0.05)	Potassium Chloride(0.07)	Ipratropium Bromide Neb(0.39)	Insulin(0.06)
Metoprolol(0.06)	Furosemide(0.06)	Aspirin(0.05)	Furosemide(0.06)	Furosemide(0.08)	Furosemide(0.05)
Amiodarone HCl(0.05)	Magnesium Sulfate(0.04)	Furosemide(0.03)	Magnesium Sulfate(0.03)	Heparin(0.06)	Magnesium Sulfate(0.04)
Heparin Sodium(0.04)	Acetaminophen(0.03)	Atorvastatin(0.03)	Acetaminophen(0.03)	[illegible]	Acetaminophen(0.03)

unrelated

Relevant drug identified by HITF gets much higher weight

Relevant drugs inferred only by HITF

Evaluated by a clinician:

"There is qualitative superiority of HITF method over the Rubik method."

Experimental Results

Clinical relevance of the Phenotypes

Table 2: Three Examples of Phenotypes

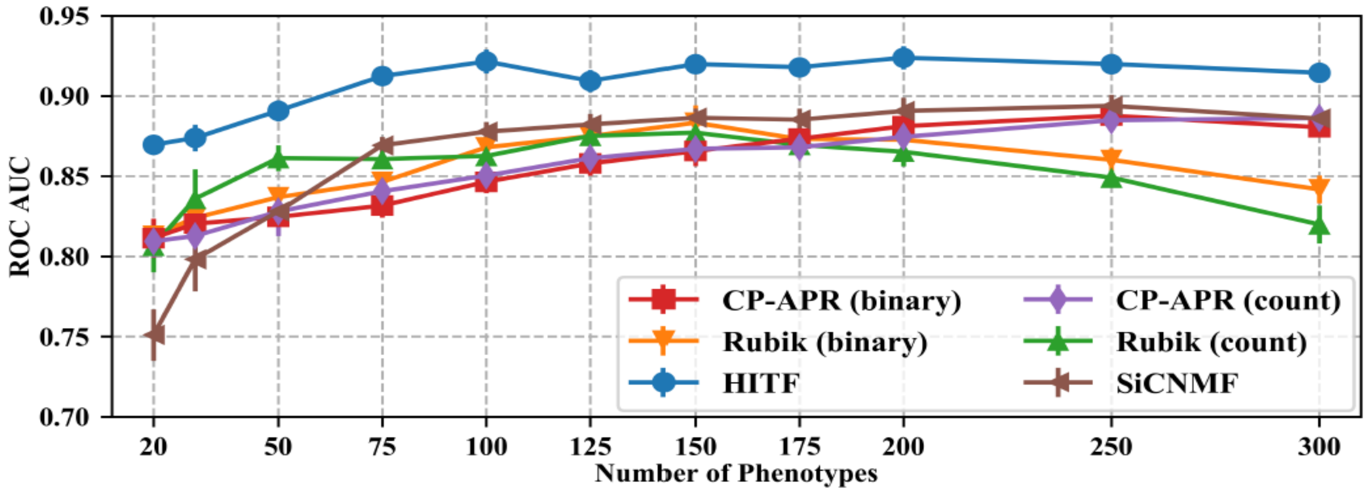
	Diagnoses	Medication
Diabetes related disease	Diabetes mellitus Other diseases of lung Acute kidney failure Essential hypertension ...	Insulin Insulin Human Regular
Cardiac disease	Cardiac dysrhythmias Heart failure Other diseases of lung	Amiodarone HCl Metoprolol Furosemide ...
Respiratory disease	Other diseases of lung Cardiac dysrhythmias Heart failure Chronic airway obstruction, not elsewhere classified	Albuterol Diltiazem Ipratropium Bromide MDI Fluticasone Propionate ...

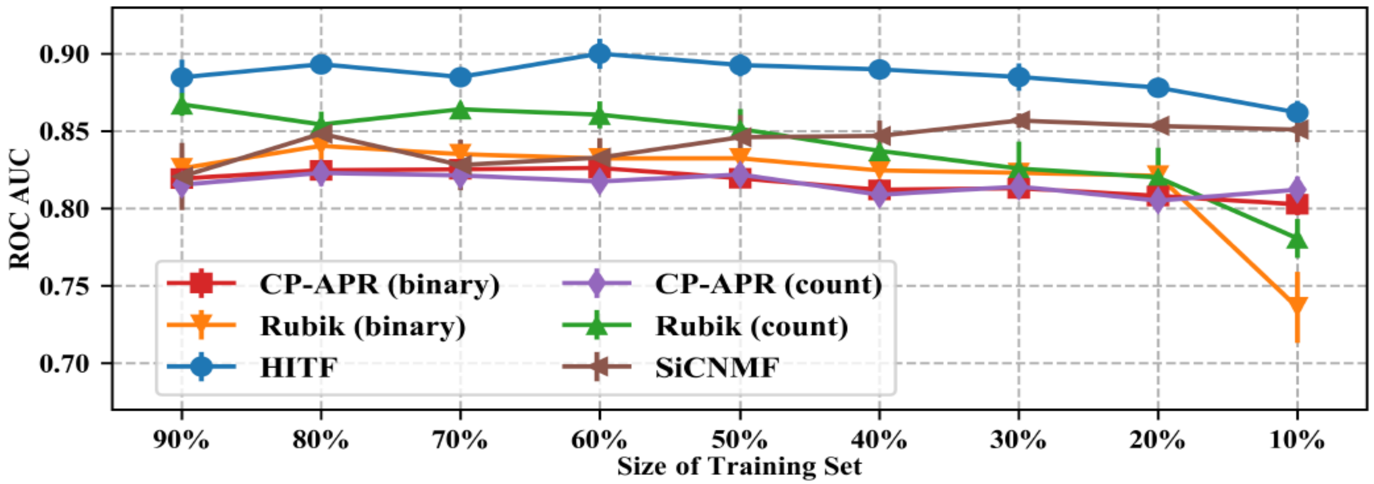
According to the clinician, phenotypes inferred by HITF are clinically relevant.

Experimental Results

DEPARTMENT OF
COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Mortality prediction





- HITF outperforms all baselines consistently in terms of mortality prediction task.
- More robust against small size of training set.

Patients can be effectively represented by phenotypes derived using HITF.

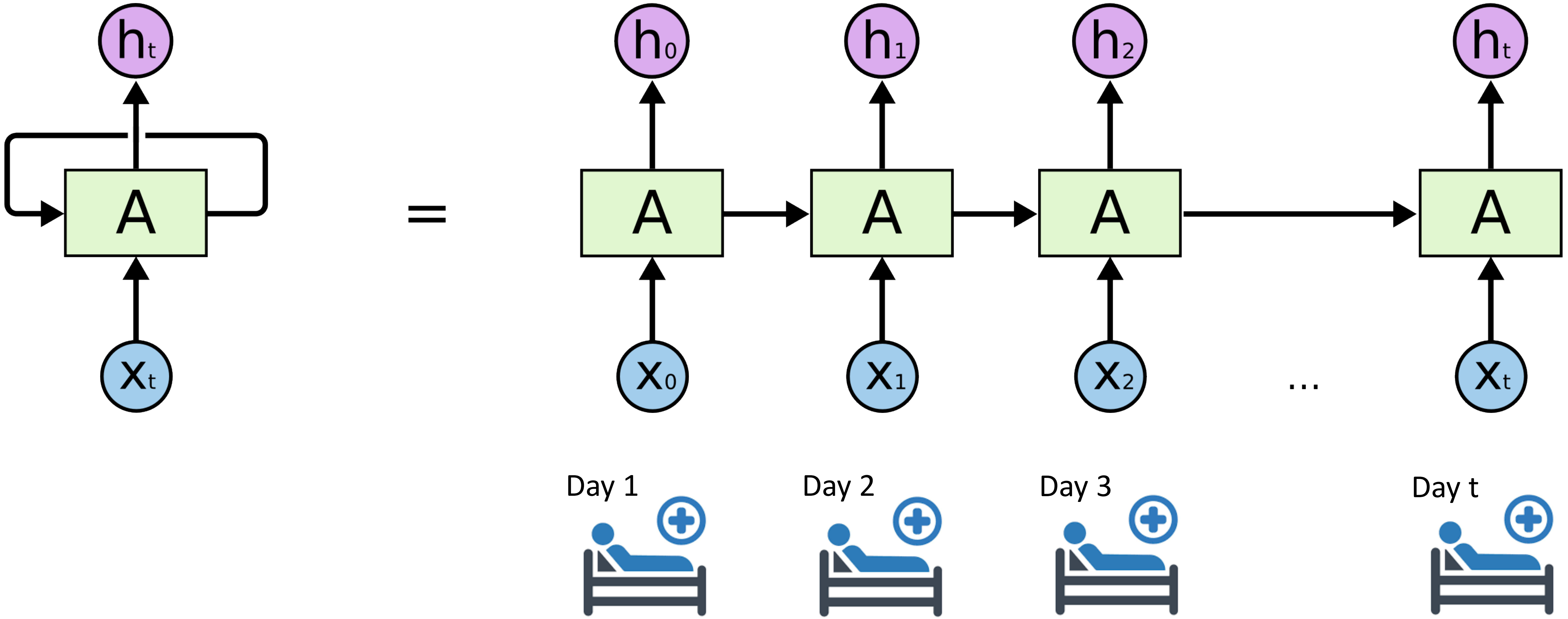
Collective Non-negative Tensor Factorization [AAAI-19]

with RNN regularization for Joint Learning of Static Phenotypes and Dynamic Patient Representation

Yin, Kejing, et al. "Learning Phenotypes and Dynamic Patient Representations via RNN Regularized Collective Non-negative Tensor Factorization.“ *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*. 2019.

Recurrent Neural Network

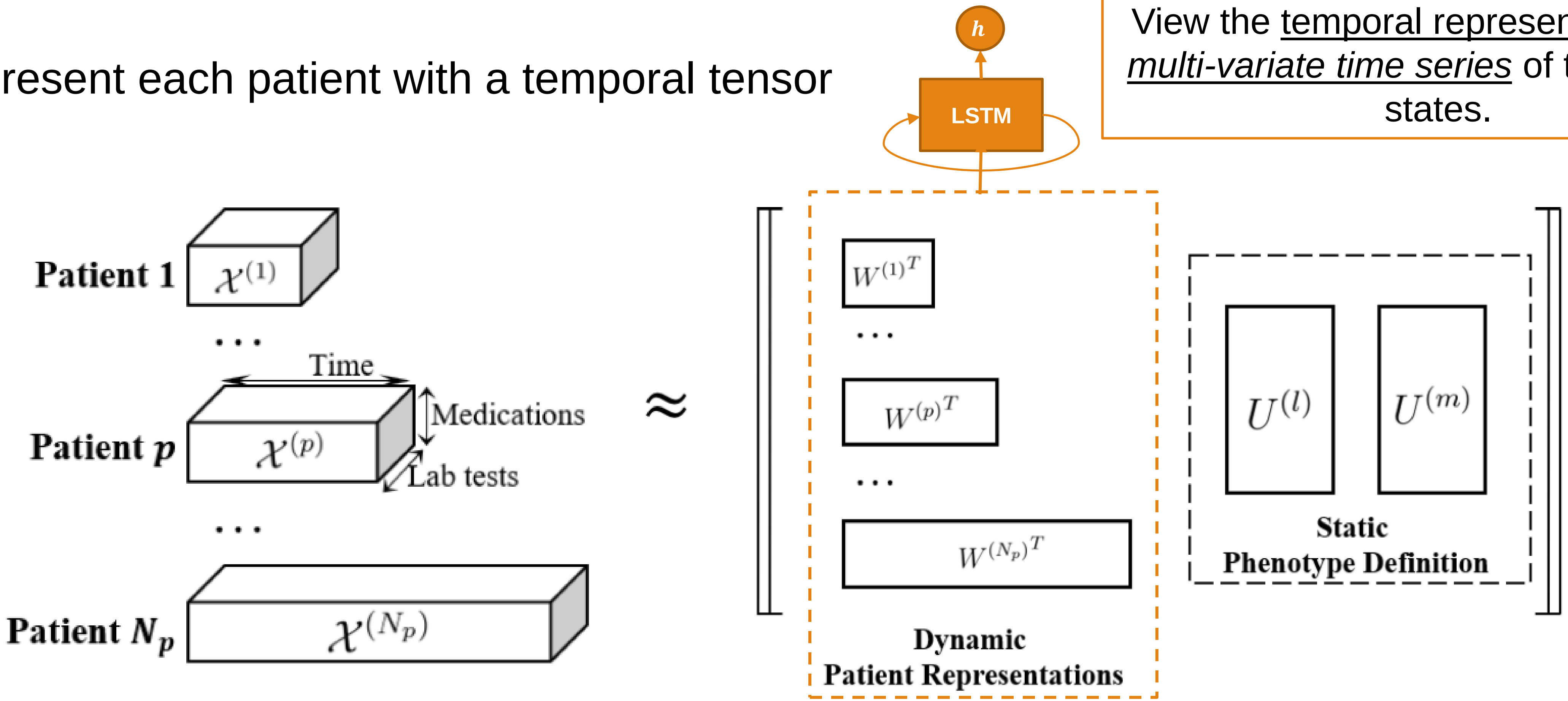
DEPARTMENT OF
COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系



Collective Non-negative Tensor Factorization

Represent each patient with a temporal tensor

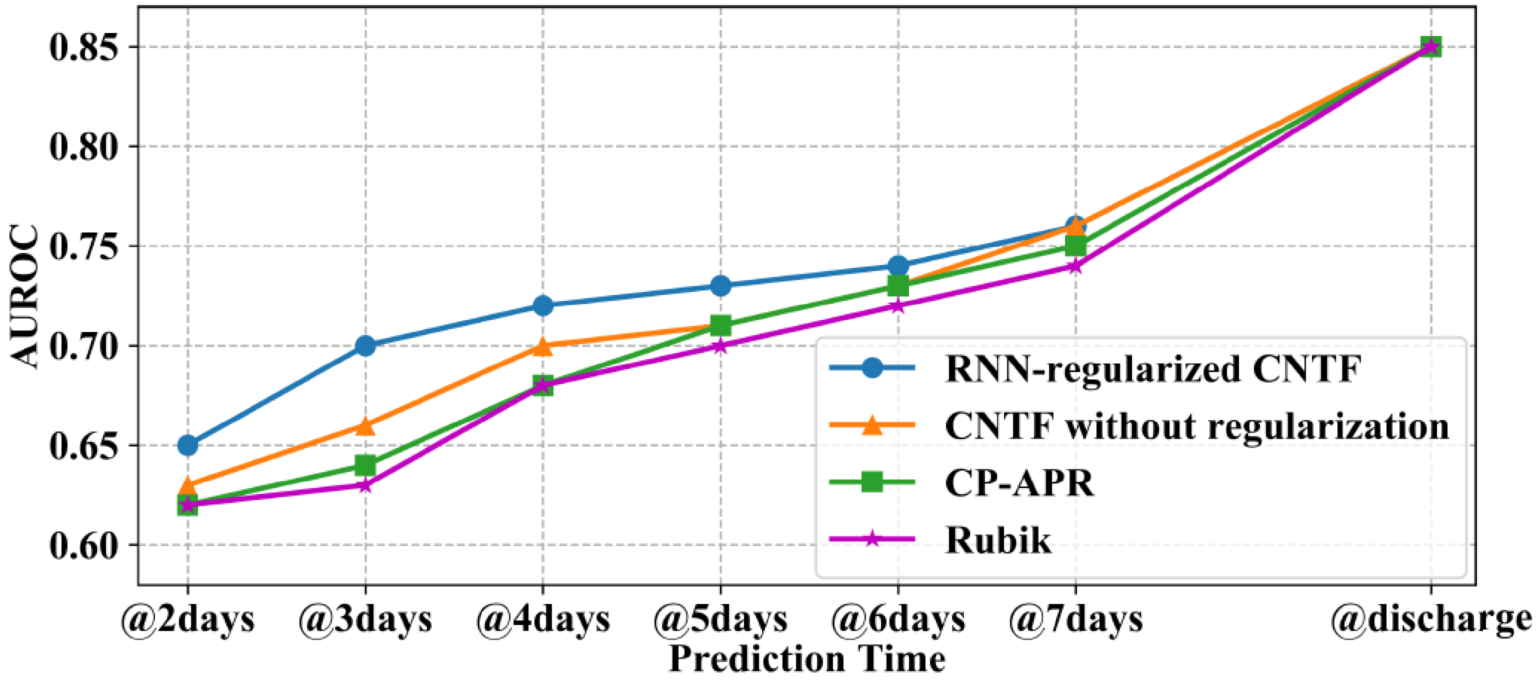
View the temporal representation as a *multi-variate time series* of the disease states.



RNN Regularized CNTF

DEPARTMENT OF
COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

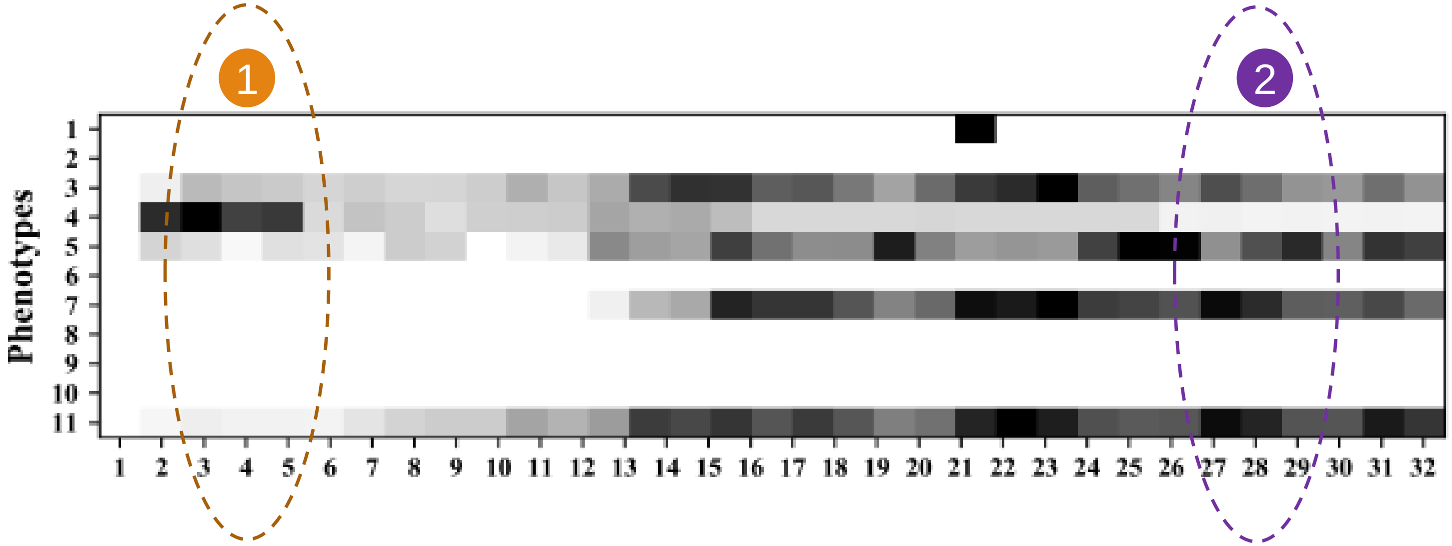
Results: Mortality Prediction



Higher prediction rate is resulted

Dynamic Patient Representation

DEPARTMENT OF COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系



1 High value for phenotype 4 (Chronic Heart Disease)

2 High value for
phenotype 3 (Other Disease of the Lung),
phenotype 5 (Cardiac Dysrhythmias),
phenotype 7 (Acute Kidney Failure),
phenotype 11 (Cardiac Dysrhythmias with Heart Failure)

"Patient admitted with existing condition, chronic heart disease, which is treated unsuccessfully, and eventually developed multiple organ failure." *(Supported by reviewing the clinical notes.)*

RNN Regularized CNTF

DEPARTMENT OF COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Results: Phenotypes

Our proposed model

Clinically much more meaningful, evaluated by a medical expert.

Phenotype 1	Phenotype 4	Phenotype 9
Chronic kidney disease (CKD) (0.536)	Other forms of chronic ischemic heart disease (0.507) Cardiac dysrhythmias (0.372) Essential hypertension (0.024)	Other diseases of lung (0.876)
RBC (Urine) (0.200) Osmolality, Measured (Blood) (0.117) Protein/Creatinine Ratio (Urine) (0.069)	Hematocrit (Blood) (0.072) Red Blood Cells (Blood) (0.071) Hemoglobin (Blood) (0.070)	pO2 (Blood Gas) (0.253) pCO2 (Blood Gas) (0.237) pH (Blood Gas) (0.215)
Hydromorphone (0.336) (0.038) 33)	Acetaminophen (0.188)	Acetaminophen (0.113)

"The disease state CKD is indeed associated with elevated RBC in urine due to renal tubular necrosis, elevated blood osmolality due to electrolyte retention in the vascular system, and elevated protein loss in the urine leading to an abnormal protein/creatinine ratio."

"Phenotype 9 corresponds to the diagnosis Other Disease of the Lung and abnormal laboratory tests pO2, pCO2, pH of the arterial blood gas. Again, this correlates well with the clinical context, where reduced oxygen levels and pH, and elevated carbon dioxide levels all indicate the presence of acute respiratory failure (which is classified under the "other disease of lung" in the ICD-9 coding system)."

Baseline: Rubik

Phenotype 1	Phenotype 2	Phenotype 3
Other diseases of lung (0.045) Septicemia (0.040) Certain adverse effects not elsewhere classified (0.039)	Other diseases of lung (0.040) Acute kidney failure (0.036) Certain adverse effects not elsewhere classified (0.032)	Acute kidney failure (0.039) Other diseases of lung (0.037) Cardiac dysrhythmias (0.033)
Glucose(Blood) (0.019) Red Blood Cells(Blood) (0.019) Hematocrit(Blood) (0.019)	Hematocrit(Blood) (0.017) Red Blood Cells(Blood) (0.017) Glucose(Blood) (0.017)	Glucose(Blood) (0.018) Hematocrit(Blood) (0.018) Red Blood Cells(Blood) (0.018)
Vancomycin (0.017) Insulin (0.015) Potassium Chloride (0.015)	Vancomycin (0.013) Potassium Chloride (0.013) Pantoprazole Sodium (0.012)	Vancomycin (0.015) Potassium Chloride (0.014) Heparin (0.014)

Multiple Ontological Representations (MMORE) [IJCAI-19]

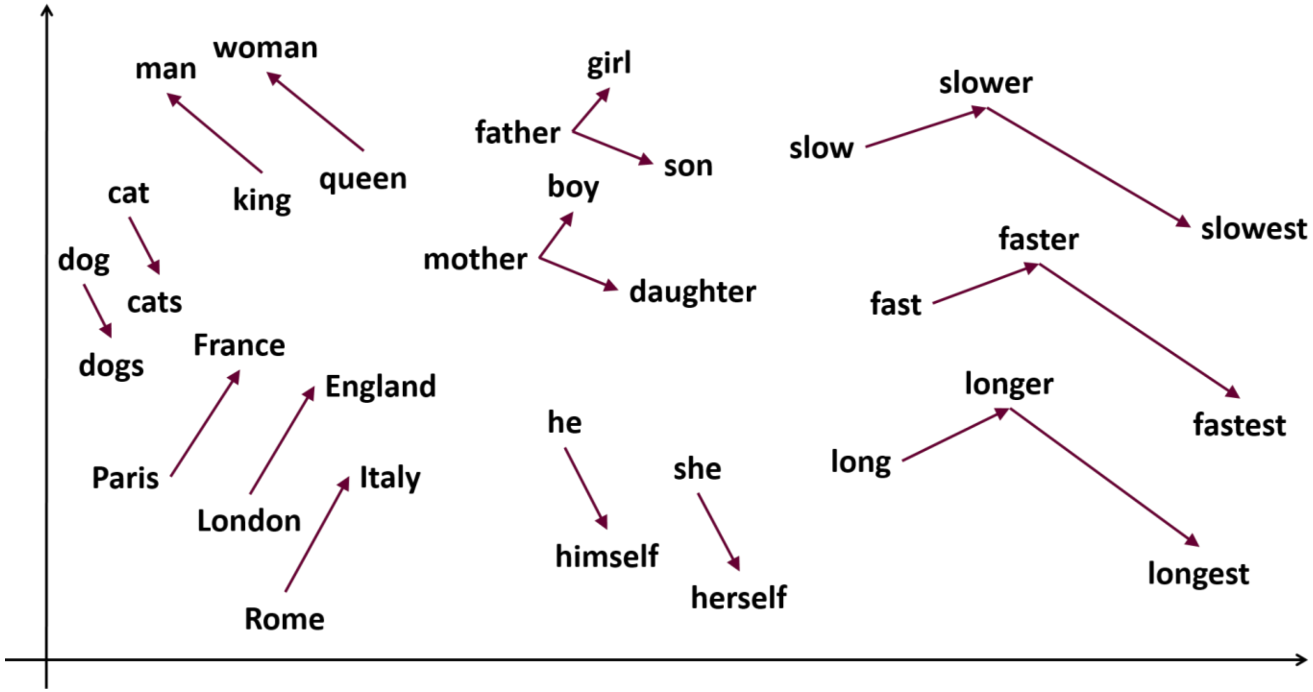
for learning medical concept representations from medical ontologies and EHR

Song, Lihong, et al. “Medical concept embedding with multiple ontological representations ." *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*. 2019.

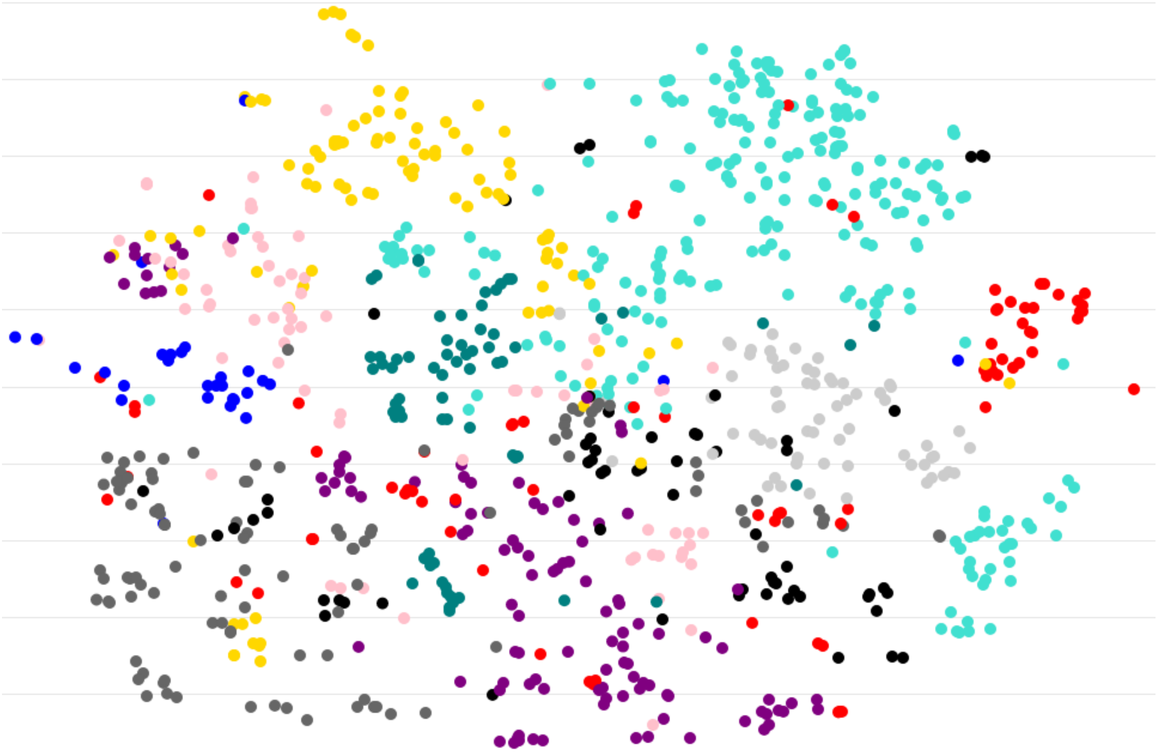
Representation Learning for Medical Concepts

DEPARTMENT OF COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Word2Vec



Med2Vec



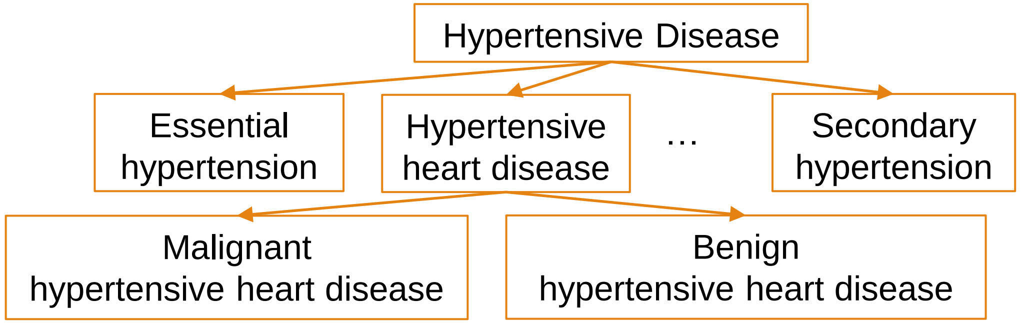
colors: categories　　dots: medical concepts

Choi, Edward, et al. "Multi-layer representation learning for medical concepts." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016.

Research Challenge

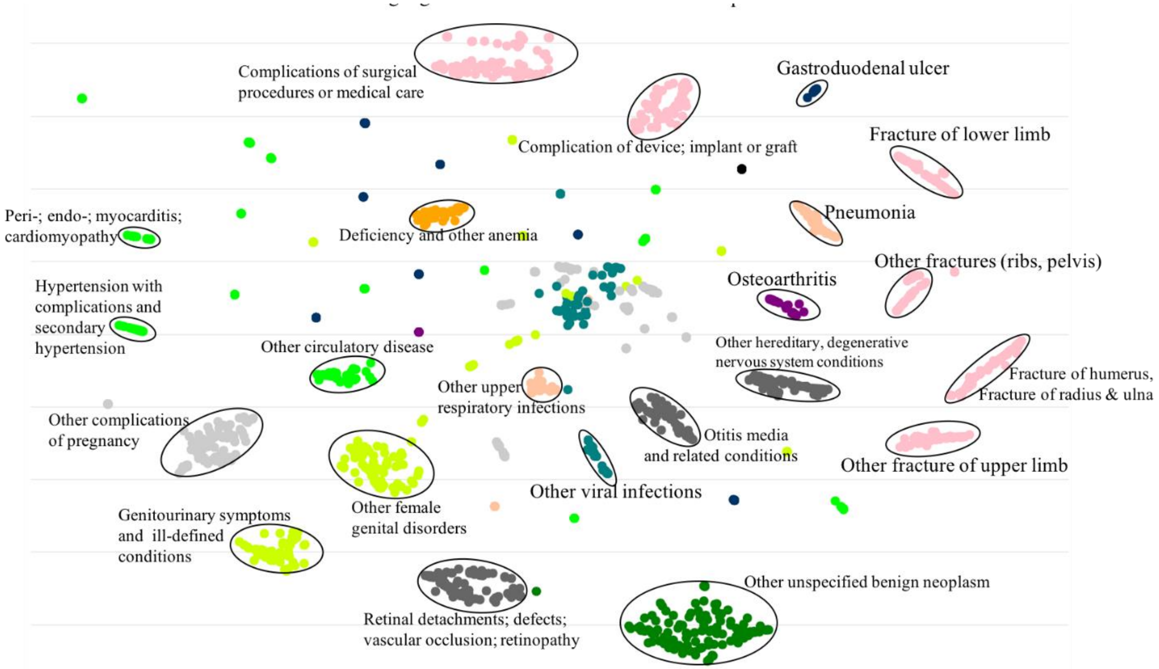
DEPARTMENT OF COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

- **Inconsistency between medical ontologies and EHR**



Example: ICD-9 ontology

GRAM model (KDD ’17)



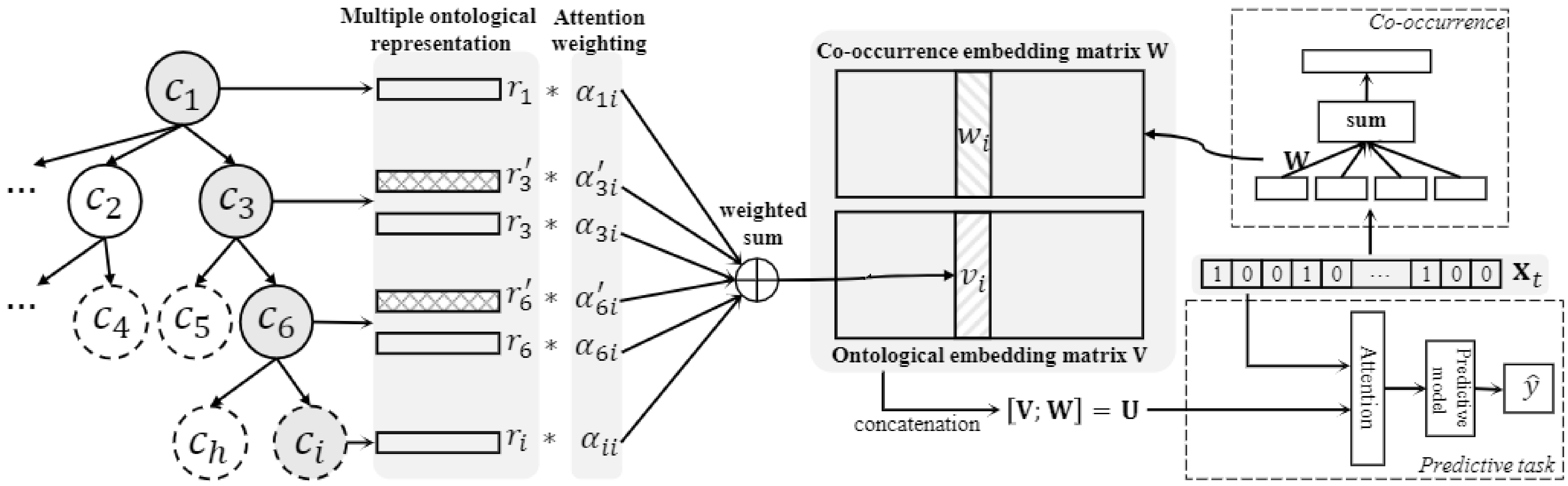
- Good enough?
- Medical concepts under the same category should co-occur with other concepts in EHR in a similar manner. Correct? *E.g.*, essential hypertension & secondary hypertension.

Choi, Edward, et al. "GRAM: Graph-based attention model for healthcare representation learning." *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017.

MMORE

DEPARTMENT OF
COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

- Key Idea: Multiple representations for each ontological category



Experimental Results

DEPARTMENT OF COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Next-admission Diagnosis Prediction

Measure the predictive performance by $Accuracy@k = \frac{\text{\# of true positives in the top } k \text{ predictions}}{\text{\# of positives}}$

Utilize only the EHR data

Mainly focus on medical ontologies

Consider both the ontologies and the EHR co-occurrence

- Less sensitive to the medications
- Ontologies could serve the role to "regularize" the learned representations

Size of training data are varied to train models

Data	Model	20%	40%	60%	80%
Dx	RETAIN	0.4422	0.4447	0.4449	0.4545
	Med2Vec	0.5064	0.5187	0.5200	0.5290
	GRAM	0.4980	0.5218	0.5409	0.5498
	MMORE	**0.5205**	**0.5426**	**0.5548**	**0.5618**
Dx & Rx	RETAIN	0.4422	0.4447	0.4449	0.4547
	Med2Vec	0.4920	0.4967	0.4979	0.5110
	GRAM	0.5057	0.5285	0.5426	0.5548
	MMORE	**0.5243**	**0.5498**	**0.5619**	**0.5689**

Dx for the diagnosis, Rx for the medication

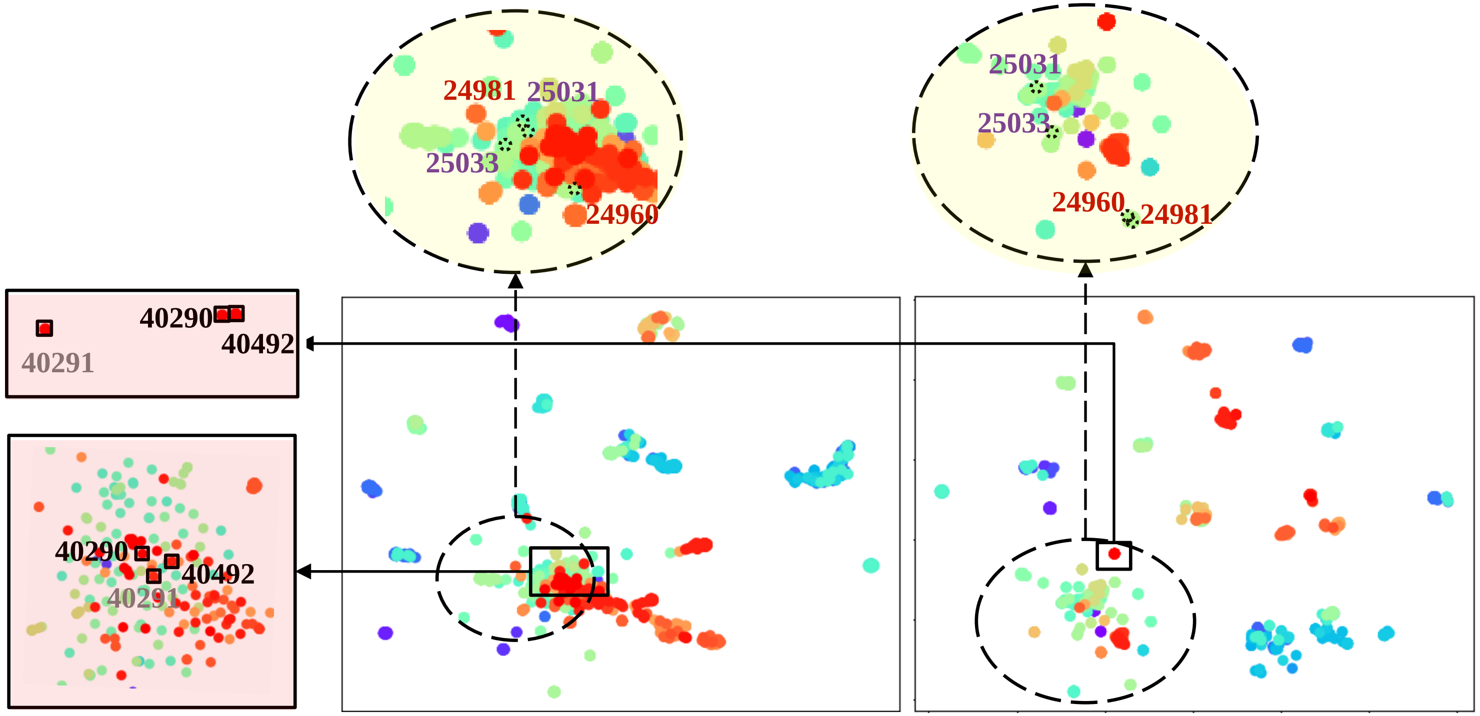
Experimental Results

DEPARTMENT OF COMPUTER SCIENCE
HONG KONG BAPTIST UNIVERSITY
香港浸會大學計算機科學系

Case study

Diabetes with neurological manifestations & Diabetes with other manifestations

Hypertensive heart disease with or without heart failure



(a) MMORE w/o MORE (b) MMORE

24960	**Secondary diabetes mellitus** with neurological manifestations, not stated as uncontrolled, or unspecified
24981	**Secondary diabetes mellitus** with other specified manifestations, uncontrolled
25031	**Diabetes with other coma,** type I [juvenile type], not stated as uncontrolled
25033	**Diabetes with other coma,** type I [juvenile type], uncontrolled

40291	Unspecified hypertensive heart disease **with heart failure**
40290	Unspecified hypertensive heart disease **without heart failure**
40492	Hypertensive heart and chronic kidney disease, unspecified, **without heart failure** and with chronic kidney disease stage V or end stage renal disease

Learned representations align with both EHR and medical ontologies

Experimental Results: Phenotyping

- Applying Non-negative Matrix Factorization to Attention Matrix
- Basis factors try to group related concepts together (phenotypes)

Phenotype	
Phenotype 1 Dx: Atrial fibrillation; Congestive heart failure, NOS; ... *Rx: Warfarin; Heparin; ...*	Heart diseases
Phenotype 2 Dx: Cirrhosis of liver w/o mention of alcohol; Dx: Alcoholic cirrhosis of liver; ... *Rx: Lactulose; Folic acid; ...*	Liver diseases
Phenotype 3 Dx: Chronic airway obstruction, NEC; Dx: Obstructive chronic bronc w/ (acute) exacerbation; ... *Rx: Ipratropium bromide; Albuterol sulfate; ...*	Respiratory diseases

Closing Remarks

- Three ML methods proposed for EHR Data Analytics.
  - Tensor Factorization -> HITF model
  - Tensor Factorization + RNN -> CNTF model
  - Representation Learning + Ontology -> MMORE model

- Future Research Directions:
  - More data modalities (e.g., vital signs)
  - Going beyond categorical ontology (e.g., SNOMED-CT)
  - Continuous-time modelling (from ICU to primary care data)

Thank you!

Q&A

References

[1] Adoption of Electronic Health Record Systems among U.S. Non-Federal Acute Care Hospitals: 2008-2015.
[2] Johnson AEW, Pollard TJ, Shen L, Lehman L, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG. MIMIC-III, a freely accessible critical care database. *Scientific Data*, 2016.
[3] Hripcsak, George, and David J. Albers. "Next-generation phenotyping of electronic health records." *Journal of the American Medical Informatics Association* 20.1 (2013): 117-121.
[4] Wang, Yichen, et al. "Rubik: Knowledge guided tensor factorization and completion for health data analytics." *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2015.
[5] Kirby, Jacqueline C., et al. "PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability." *Journal of the American Medical Informatics Association* 23.6 (2016): 1046-1052.
[6] Ho, Joyce C., et al. "Limestone: High-throughput candidate phenotype generation via tensor factorization." *Journal of biomedical informatics* 52 (2014): 199-211.
[7] Yang, Kai, et al. "TaGiTeD: Predictive Task Guided Tensor Decomposition for Representation Learning from Electronic Health Records." *AAAI*. 2017.
[8] Jennifer Pacheco and Will Thompson. Northwestern University. Type 2 Diabetes Mellitus. PheKB; 2012 Available from: https://phekb.org/phenotype/18
[9] Ho, Joyce C., Joydeep Ghosh, and Jimeng Sun. "Marble: high-throughput phenotyping from electronic health records via sparse nonnegative tensor factorization." *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014.
[10] Kim, Yejin, et al. "Discriminative and distinct phenotyping by constrained tensor factorization." *Scientific reports* 7.1 (2017): 1114.
[11] Yang, Kai, et al. "TaGiTeD: Predictive Task Guided Tensor Decomposition for Representation Learning from Electronic Health Records." *AAAI*. 2017.
[12] Henderson, Jette, et al. "Granite: Diversified, Sparse Tensor Factorization for Electronic Health Record-Based Phenotyping." *2017 IEEE International Conference on Healthcare Informatics (ICHI)*, 2017.
[13] Kolda, T. G., & Bader, B. W. (2008). Tensor Decompositions and Applications. *SIAM Review*, 51(3)
[14] Chi, Eric C., and Tamara G. Kolda. On tensors, sparsity, and nonnegative factorizations. *SIAM Journal on Matrix Analysis and Applications* 33.4 (2012): 1272-1299.
[15] Choi, E., Bahadori, M. T., Song, L., Stewart, W. F., & Sun, J. (2017, August). GRAM: graph-based attention model for healthcare representation learning. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 787-795). ACM.
[16] Choi, E., Bahadori, M. T., Searles, E., Coffey, C., Thompson, M., Bost, J. & Sun, J. (2016, August). Multi-layer representation learning for medical concepts. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1495-1504). ACM.