Learning Electronic Health Records through Hyperbolic Embedding of Medical Ontologies

Presenter: Qiuhao Lu

University of Oregon

Medical Ontologies

- Advantages:
  - Knowledge management
    - support the indispensable integration of medical knowledge and data
    - help build more powerful and more interoperable information systems in healthcare
  - Interoperability
    - support the need of the healthcare process to transmit, re-use and share patient data
  - Accessibility
    - provide semantic-based criteria to support different statistical aggregations for different purposes
- Limitations:
  - Some remain skeptical about the impact that ontologies may have on the design and maintenance of real-world healthcare information systems.

Well-used Medical Ontologies

- Concepts in medical ontologies are widely used for research, representation and reorganization of medical terms, health management, and clinical purposes.
  - **ICD-10**: ICD-10 is the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD), a medical classification list by the World Health Organization (WHO). ICD codes are widely used by physicians and coders to record patient symptoms and diagnoses for the purpose of submitting claims to payers for reimbursement and for clinical research.
  - **UMLS**: The Unified Medical Language System is a compendium of many controlled vocabularies in the biomedical sciences. UMLS is widely used as a means of standardizing the documentation of clinical concepts found in a patient's electronic health record.
  - **LOINC**: Logical Observation Identifiers Names and Codes is one of the standards used in the US for identifying medical laboratory observations.
  - **NDC**: National Drug Code is a universal product identifier for human drugs in the US.
  - **SNOMED CT**: SNOMED Clinical Terms is a systematically organized computer processable collection of medical terms providing codes, terms, synonyms and definitions used in clinical documentation and reporting.
  - **MeSH**: Medical Subject Headings is a comprehensive controlled vocabulary for the purpose of indexing journal articles and books in the life sciences; it serves as a thesaurus that facilitates searching.

Applying Medical Ontologies

- How these Ontologies are used in downstream applications
  - Used for semantic annotation of text
    - identifying cancer related causes of death from death certificates [1][5]
    - measuring semantic similarity between medical concepts, which is central to a number of techniques in medical informatics and information retrieval [2]
  - Used for representation learning, the embeddings of which are further explored
    - representation learning of medical concepts [3]
    - predicting 30-day ICU readmissions using disease embeddings as features [4]
  - Used as an external knowledge source
    - multi-class classification of cause of death from plaintext autopsy reports [6]
    - automatic new concepts insertion [7]

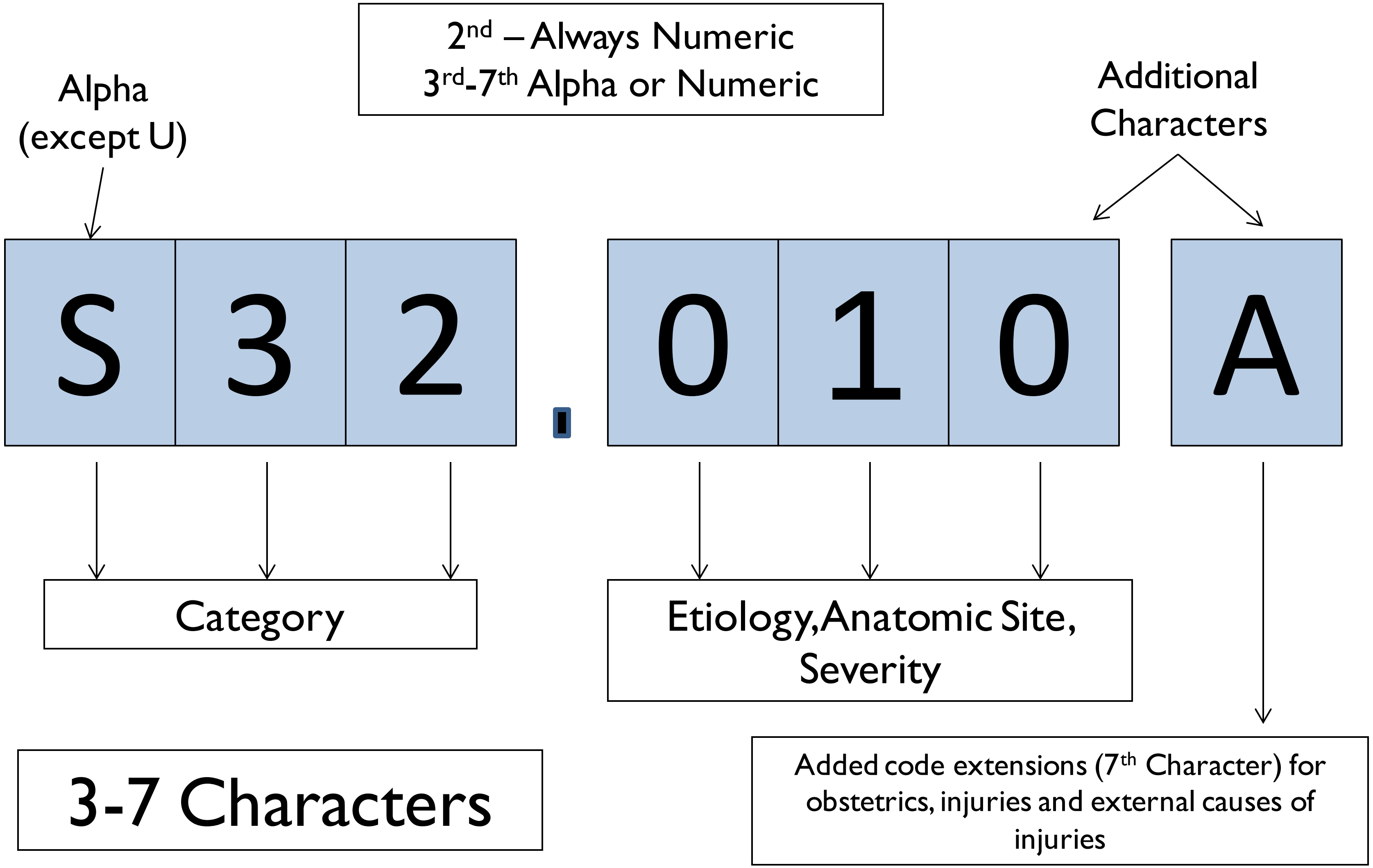
Embedding Medical Ontologies

- One way to utilize medical ontologies is embedding, previous approaches:
  - Cui2vec: learning medical concept embeddings from multimodal medical data [27]
    - Embed insurance claims, clinical notes, full text biomedical journal articles into a common space.
  - Word2vec: learning medical concepts from insurance claims or clinical notes [3]
    - Skip-gram
    - Embed insurance claims or clinical notes to medical concepts
  - Neural language models [2]
    - Skip-gram but based on pre-defined UMLS concepts.
    - Data from PMC abstracts

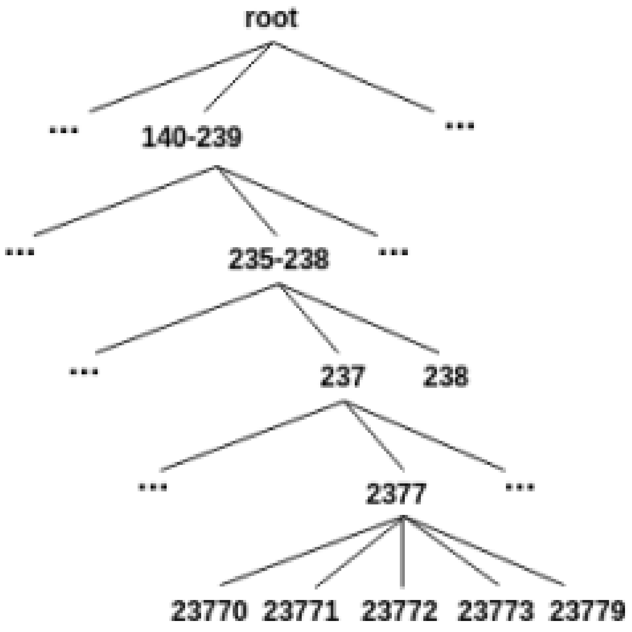
Hyperbolic Embedding

- Ontologies are Graphs.
  - Hyperbolic graph (embedding) methods efficiently learn embeddings of large datasets with tree-like structure, scale-free networks, etc. [16]
    - Medical Ontologies mainly have tree-like structures.
  - Other state-of-the-art graph embedding methods typically learn embeddings in Euclidean vector spaces, and they share a fundamental limitation: their ability to model complex patterns is inherently bounded by the dimensionality of the embedding space.
  - Examples include [16][17][18]

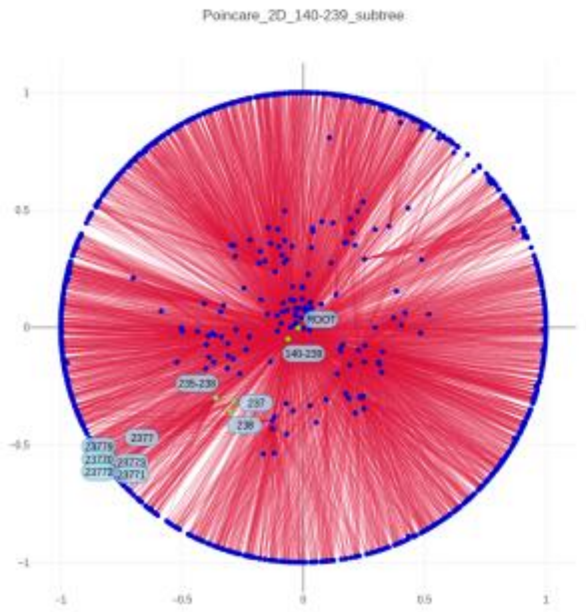
Example: ICD-10-CM (diagnosis) Code Format



Tree Structure of ICD codes



(a) Hierarchical Structure of the Subtree



(b) Visualization of the 2-D Hyperbolic Embeddings of the Subtree

Case Study: Unplanned 30-day Readmission Prediction

- Most of the existing works on predicting unplanned 30-day readmissions are developed with certain components from EHR such as vital signs and lab tests. Those events are typically aggregated over an observation window with feature transformation and then fed into a predictor such as logistic regression or random forest for the prediction task [19][20].
- Deep learning-based methods on this task (and similar tasks in the medical domain) are under-explored. Generally they use CNN or RNN to capture the time information among the medical events. For example, [21] predicts hospital readmissions via learning interpretable patient representations by capturing both local and global contexts from EHRs through a RNN model. [22] proposes a LSTM model to capture the progression patterns for Parkinson's disease.
- There is also some solid work that is from the angle of natural language processing which focuses on predicting readmissions based on medical textual notes from EHRs [23].

Unplanned 30-day Readmission Prediction

- Lin et al.'s work [24] gives the state-of-the-art performance in unplanned 30-day all-cause ICU readmission prediction on MIMIC-III dataset [25].
  - MIMIC-III Dataset
    - a large, freely-available database comprising deidentified health-related data associated with over 40,000 ICU patients, which includes information such as demographics, vital sign measurements made at the bedside (~1 data point per hour), laboratory test results, procedures, medications, caregiver notes, imaging reports, and mortality
  - Method
    - Features used include 17 types of selected chart events (diastolic blood pressure, capillary refill rate, etc.) over a 48-hour time window, ICD-9 embeddings of diagnosis, and demographic information of patients
    - Input features are fed into a LSTM-CNN model as shown in Figure 1
  - Metric
    - The Area under the Receiver Operating Characteristic (AUROC) as primary metric. It measures the overall performance of the recall with respect to different false positive rate.
    - AUROC is a commonly used metric for the readmission prediction task, since the dataset is usually highly imbalanced.
    - Along with some other metrics, including positive case recall rate (sensitivity), accuracy, etc.

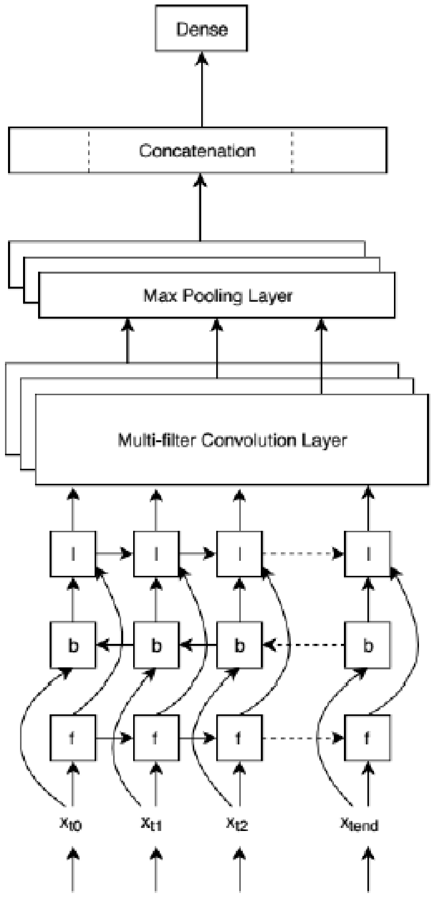
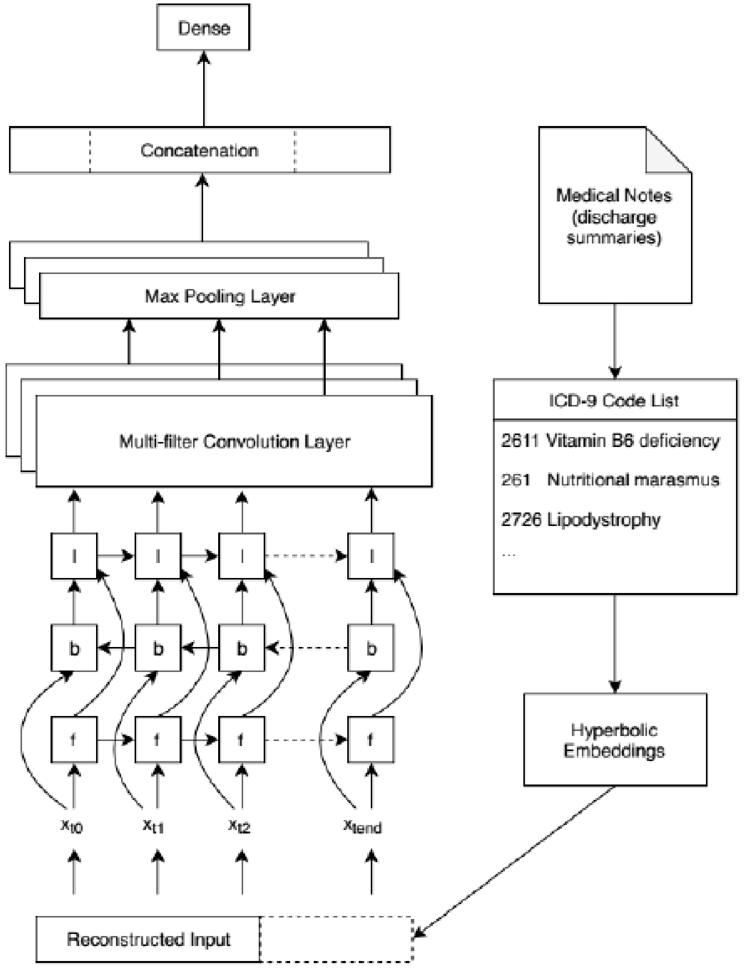


Figure 1: LSTM-CNN model

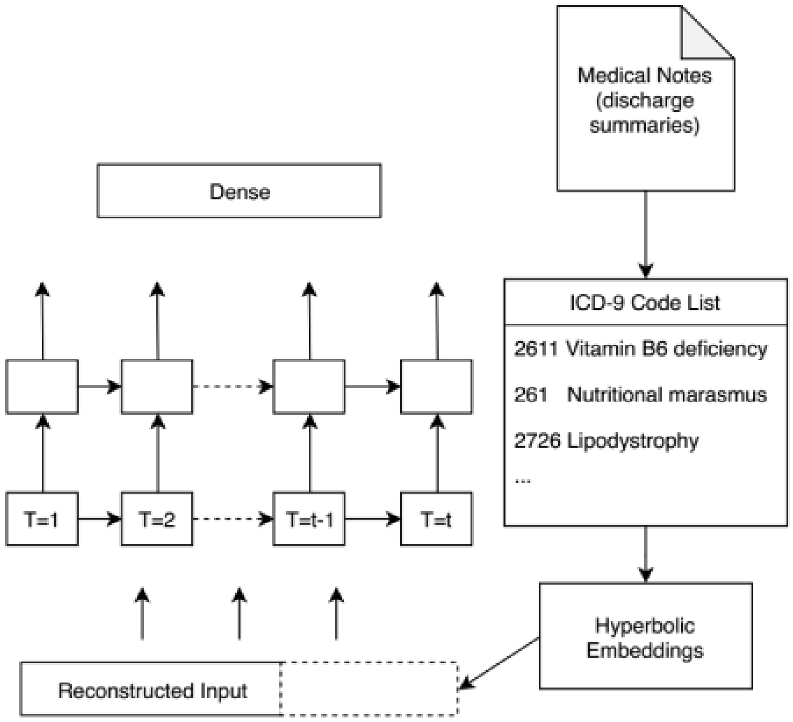
Incorporating hyperbolic ICD-9 embeddings for ICU readmission prediction

- Our Novel Method
  - Motivated by the structural property of hierarchical data (medical ontology) that can be efficiently embedded into hyperbolic space, we learn the **Poincare** embeddings of ICD-9 ontology following the work of [16].
  - Extract ICD-9 codes from the discharge summaries of ICU patients' EHRs, the Poincare embeddings of which are used to reconstruct the input for the LSTM-CNN model



Case Study: In-hospital Mortality Prediction

- Harutyunyan et al.'s work [11] is a widely accepted benchmark in in-hospital mortality prediction on the MIMIC-III dataset [7].
  - MIMIC-III Dataset
    - Same as mentioned
  - Method
    - Features used include 17 types of selected chart events (Glascow coma scale, Heart Rate, etc.) over a 48-hour time window.
    - Input features are fed into a LSTM-based prediction model.
  - Metric
    - AUROC as mentioned
- We simply reconstruct the input features by concatenating the embeddings of diagnosis (in the form of ICD-9 codes).



Evaluations

- Intrinsic Evaluation
  - Compare and demonstrate how the similarities of medical concepts from ICD-9 are retained in the embedding spaces
  - We randomly pick 20,000 concept pairs from the ICD-9 ontology and compute the mentioned 4 kinds of ontology-based similarities between them. Then we compute the distance-based similarities over these pairs for the comparing embeddings. Finally we calculate the Pearson Correlation Coefficients between the above two kinds of sequences as shown in Table 1.

- Extrinsic Evaluation
  - See whether any performance improvement on 30-day unplanned ICU readmission prediction and mortality prediction can be gained. Results are shown in the following tables.

Intrinsic Evaluation

Method	Dim	Measurement			
		WUP	LCH	RESNIK	RADA
Poincare	10	**0.5720**	**0.6797**	**0.5784**	**0.7278**
	100	**0.5866**	**0.6902**	**0.5977**	**0.7351**
	300	**0.6042**	**0.7046**	**0.6007**	**0.7491**
ComplEx	10	0.4279	0.3169	0.4320	0.3036
	100	0.2265	0.2018	0.2094	0.1774
	300	0.1307	0.1432	0.1134	0.1141
DistMult	10	0.4297	0.3621	0.4174	0.3410
	100	0.1941	0.1827	0.1964	0.1521
	300	0.1204	0.1223	0.1161	0.0922
transE	10	0.0483	0.0269	0.0709	0.0130
	100	0.4159	0.3682	0.3494	0.3658
	300	0.4355	0.3958	0.3862	0.3912
Rescal	10	0.4108	0.2884	0.4364	0.2952
	100	0.2522	0.2756	0.1986	0.2523
	300	0.1243	0.1355	0.1166	0.1039

Table 1: Pearson Correlation Coefficients for Different Embeddings of ICD-9

Extrinsic Evaluation 1: 30-day Unplanned ICU Readmission Prediction

Embedding	Acc	Pre-0	Pre-1	Re-0	Re-1	**A.R**	A.P
Poincaré	0.7223	0.9035	0.3740	0.7361	0.6655	0.7786	0.4827
ComplEx	0.6621	0.9141	0.3306	0.6423	0.7454	0.7591	0.4236
Distmult	0.6426	0.9126	0.3172	0.6169	0.7508	0.7534	0.4243
TransE	0.7254	0.9062	0.3789	0.7366	0.6782	**0.7876**	0.4875
Rescal	0.6544	0.9160	0.3264	0.6303	0.7562	0.7661	0.4456

*Acc: Accuracy, Pre: Precision, Re: Recall, A.R: AUC under ROC, A.P: AUC under PRC

Table 2: Performance on ICU Readmission Prediction Without Discharge Summaries

Embedding	Acc	Pre-0	Pre-1	Re-0	Re-1	**A.R**	A.P
Poincaré	0.7481	0.8993	0.4005	0.7766	0.6310	0.7851	0.4819
ComplEx	0.6705	0.9101	0.3342	0.6565	0.7263	0.7602	0.4341
Distmult	0.6678	0.9067	0.3303	0.6565	0.7151	0.7606	0.4327
TransE	0.7536	0.9039	0.4100	0.7779	0.6511	**0.7882**	0.4957
Rescal	0.6399	0.9209	0.3197	0.6067	0.7801	0.7684	0.4454

*Acc: Accuracy, Pre: Precision, Re: Recall, A.R: AUC under ROC, A.P: AUC under PRC

Table 3: Performance on ICU Readmission Prediction With Discharge Summaries

Extrinsic Evaluation 2: In-Hospital Mortality Prediction

Embedding	Acc	Pre-0	Pre-1	Re-0	Re-1	**A.R**	A.P
Poincaré	0.8814	0.8977	0.6350	0.9737	0.2912	0.8722	0.5543
ComplEx	0.8947	0.9165	0.6701	0.9662	0.4380	0.8915	0.6104
Distmult	0.8988	0.9152	0.7115	0.9730	0.4243	**0.8956**	0.6247
TransE	0.8888	0.9019	0.6930	0.9777	0.3211	0.8854	0.5717
Rescal	0.8913	0.9019	0.7239	0.9809	0.3188	0.8954	0.6025

*Acc: Accuracy, Pre: Precision, Re: Recall, A.R: AUC under ROC, A.P: AUC under PRC

Table 4: Performance on Mortality Prediction Without Discharge Summaries

Embedding	Acc	Pre-0	Pre-1	Re-0	Re-1	**A.R**	A.P
Poincaré	0.8882	0.9128	0.6338	0.9626	0.4128	0.8756	0.5760
ComplEx	0.8972	0.9012	0.8220	0.9895	0.3073	0.8958	0.6312
Distmult	0.8941	0.9267	0.6330	0.9529	0.5183	0.8959	0.6218
TransE	0.8882	0.8995	0.7043	0.9802	0.3004	0.8852	0.5771
Rescal	0.8929	0.9141	0.6654	0.9669	0.4197	**0.8979**	0.6042

*Acc: Accuracy, Pre: Precision, Re: Recall, A.R: AUC under ROC, A.P: AUC under PRC

Table 5: Performance on Mortality Prediction With Discharge Summaries

Conclusion and Future Work

- Conclusion
  - In this study, we present a new method to incorporate the medical notes in patients' EHR data to improve the state-of-the-art model on ICU readmission prediction, and the widely accepted benchmark on in-hospital mortality prediction. We specially leverage the hyperbolic embeddings of the ICD-9 ontology in our proposed method. To the best of our knowledge, we are the first to do so and achieve promising results.
- Future Work
  - Though the hyperbolic embeddings (Poincare) of ICD-9 perform well, they are learned from very limited information (i.e., the hierarchical structure). We will try to train a better and joint embedding with the hierarchy of medical ontologies, the textual descriptions of concepts, and even patients' EHR data in the hyperbolic space.
  - In the mortality prediction task, the hyperbolic embeddings are not as good as all the other embeddings. We will try to incorporate task-specific supervision to derive the embeddings of medical concepts.
  - Readmission and mortality predictions are two important tasks for the healthcare system. We will explore more sophisticated and effective models to solve these tasks. Based on our research, we have found that useful and valuable information exists in the medical notes (e.g., discharge summaries) of patients' EHRs, which supports prediction; so we will try new ways to leverage the textual information.

Thank you!

References

[1] Koopman, Bevan, et al. "Automatic ICD-10 classification of cancers from free-text death certificates." *International journal of medical informatics* 84.11 (2015): 956-965.

[2] De Vine, Lance, et al. "Medical semantic similarity with a neural language model." *Proceedings of the 23rd ACM international conference on conference on information and knowledge management*. ACM, 2014.

[3] Choi, Youngduck, Chill Yi-I. Chiu, and David Sontag. "Learning low-dimensional representations of medical concepts." *AMIA Summits on Translational Science Proceedings* 2016 (2016): 41.

[4] Lin, Yu-Wei, et al. "Analysis and Prediction of Unplanned Intensive Care Unit Readmission using Recurrent Neural Networks with Long Short-Term Memory." *bioRxiv* (2018): 385518.

[5] Boyd, Andrew D., et al. "Physician nurse care: A new use of UMLS to measure professional contribution: Are we talking about the same patient a new graph matching algorithm?." *International journal of medical informatics* 113 (2018): 63-71.

[6] Mujtaba, Ghulam, et al. "Automatic ICD-10 multi-class classification of cause of death from plaintext autopsy reports through expert-driven feature selection." *PloS one* 12.2 (2017): e0170242.

[7] Liu, Hao, et al. "Using Convolutional Neural Networks to Support Insertion of New Concepts into SNOMED CT." *AMIA Annual Symposium Proceedings*. Vol. 2018. American Medical Informatics Association, 2018.

[8] Hamilton, William L., Rex Ying, and Jure Leskovec. "Representation learning on graphs: Methods and applications." *arXiv preprint arXiv:1709.05584* (2017).

[9] Ahmed, Amr, et al. "Distributed large-scale natural graph factorization." *Proceedings of the 22nd international conference on World Wide Web*. ACM, 2013.

[10] Cao, Shaosheng, Wei Lu, and Qiongkai Xu. "Grarep: Learning graph representations with global structural information." *Proceedings of the 24th ACM international on conference on information and knowledge management*. ACM, 2015.

[11] Perozzi, Bryan, Rami Al-Rfou, and Steven Skiena. "Deepwalk: Online learning of social representations." *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014.

[12] Grover, Aditya, and Jure Leskovec. "node2vec: Scalable feature learning for networks." *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2016.

[13] Kipf, Thomas N. and Max Welling. “Semi-Supervised Classification with Graph Convolutional Networks.” ICLR. 2017

[14] Hamilton, Will, Zhitao Ying, and Jure Leskovec. "Inductive representation learning on large graphs." *Advances in Neural Information Processing Systems*. 2017.

[15] Li et al., “Gated Graph Sequence Neural Networks.” ICLR. 2016

[16] Nickel, Maximillian, and Douwe Kiela. "Poincaré embeddings for learning hierarchical representations." *Advances in neural information processing systems*. 2017.

[17] Sala, Frederic, et al. "Representation tradeoffs for hyperbolic embeddings." *International Conference on Machine Learning*. 2018.

[18] Chamberlain, Benjamin Paul, James Clough, and Marc Peter Deisenroth. "Neural embeddings of graphs in hyperbolic space." *arXiv preprint arXiv:1705.10359* (2017).

[19] Basu Roy, Senjuti, et al. "Dynamic hierarchical classification for patient risk-of-readmission." *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*. ACM, 2015.

[20] Hammill, Bradley G., et al. "Incremental value of clinical data beyond claims data in predicting 30-day outcomes after heart failure hospitalization." *Circulation: Cardiovascular Quality and Outcomes* 4.1 (2011): 60-67.

[21] Xiao, Cao, et al. "Readmission prediction via deep contextual embedding of clinical concepts." *PloS one* 13.4 (2018): e0195024.

[22] Baytas, Inci M., et al. "Patient subtyping via time-aware LSTM networks." *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. ACM, 2017.

[23] Rumshisky, A., et al. "Predicting early psychiatric readmission with natural language processing of narrative discharge summaries." *Translational psychiatry* 6.10 (2016): e921.

[24] Lin, Yu-Wei, et al. "Analysis and Prediction of Unplanned Intensive Care Unit Readmission using Recurrent Neural Networks with Long Short-Term Memory." *bioRxiv* (2018): 385518.

[25] Johnson, Alistair EW, et al. "MIMIC-III, a freely accessible critical care database." Scientific data 3 (2016): 160035.

[26] Finlayson, Samuel G., Paea LePendu, and Nigam H. Shah. "Building the graph of medicine from millions of clinical narratives." Scientific data 1 (2014): 140032.

[27] Beam, Andrew L, et al. “Clinical Concept Embeddings Learned from Massive Sources of Multimodal Medical Data.” *arXiv preprint arXiv:1804.01486* (2018).