FOR CLINICAL EVENTS

PRIVACY-PRESERVING HETEROGENEOUS NETWORK EMBEDDING Gustavo L. de Oliveira¹, Ricardo M. Marcacini¹, Maria da Graça C. Pimentel¹

Introduction

Electronic Health Records (EHR) databases have become increasingly popular in data mining and machine learning tasks because can be used to support medical decision-making. One of the challenges is to obtain an appropriate representation from EHR datasets that considers both structured and unstructured patient data as well their relationships.

Heterogeneous Information Networks (HIN's) are a promising alternative for datasets that contain interrelated multi-typed data such as patients and clinical events. Furthermore, HIN's can be mapped to a latent space (i.e., embedding space) to preserve the structure in a low dimensional vector space.

Many projects make the network embeddings publicly for machine learning tasks which demand omitting the original information network to preserve sensitive information. Also, laws of data protection (e.g. GPDR and LGPD) obligate data holders to conceal users' private information. Studies show that inference attacks allow reconstruction of the network structure to predict sensitive and private features.

Techniques to sanitize network data through noise or graph pruning have not been successful in mitigating attacks. Moreover, existing privacy-preserving network embedding methods consider only homogeneous networks formed by a single type of node and relationship. Thus, we raise the following question: how to learn privacy-preserving embeddings from clinical events heterogeneous networks?

We propose the Privacy-Preserving Heterogeneous Information Network Embeddings (PHINE) approach for clinical events. Our approach explores (GAE) Graph Autoencoder to learn how to preserve important HIN information (utility features) and minimize inference attacks and omit sensitive information (private features). We evaluated PHINE considering the embedding usefulness for classification tasks and its ability to preserve private information from inference attacks. The experimental results indicate that our approach presents a competitive tradeoff between privacy-preserving and utility feature prediction.

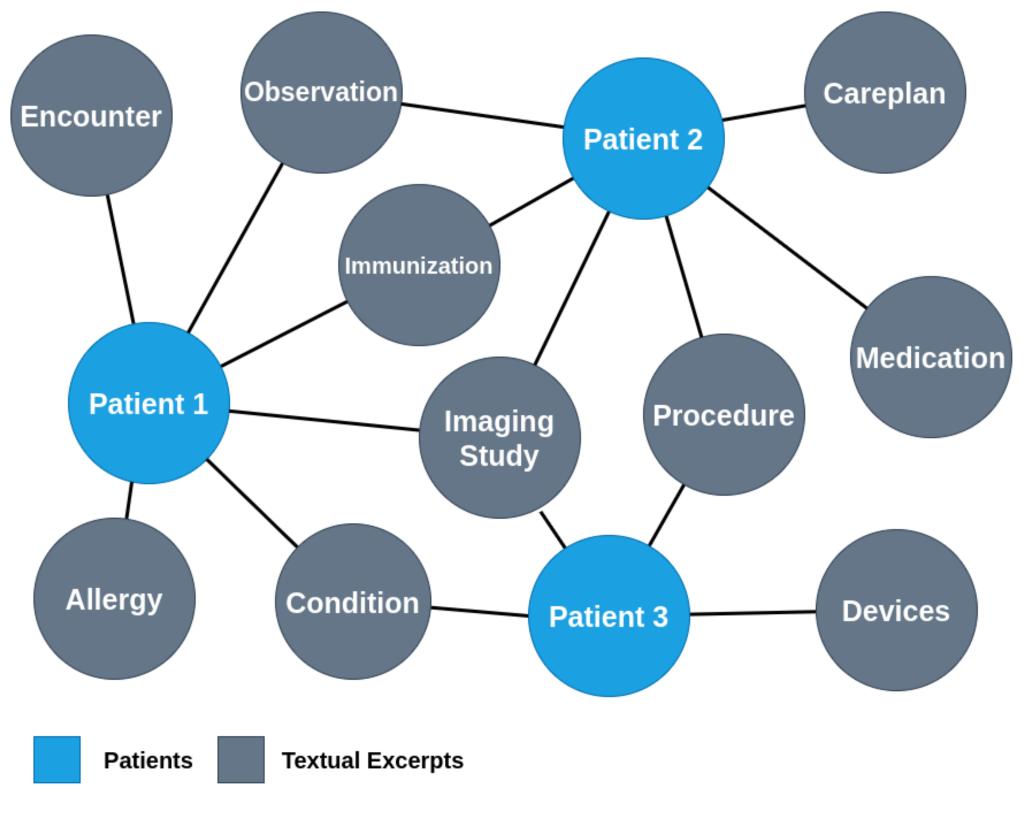


Fig. 1: Example of heterogeneous network of clinical events.

¹Institute of Mathematical and Computer Science, University of São Paulo, Brazil

Proposed Approach

Heterogeneous Network for Clinical Events

Clinical events are mapped on a heterogeneous network using two types of nodes: (i) patients and (ii) textual excerpts extracted from the EHR. The patient-type nodes have their respective private vector features. The links indicate when a patient is related to some textual information extracted from the EHR [1]. Moreover, the network topology indicates when two or more patients share the same utility features extracted from the texts as shown in Figure 3.

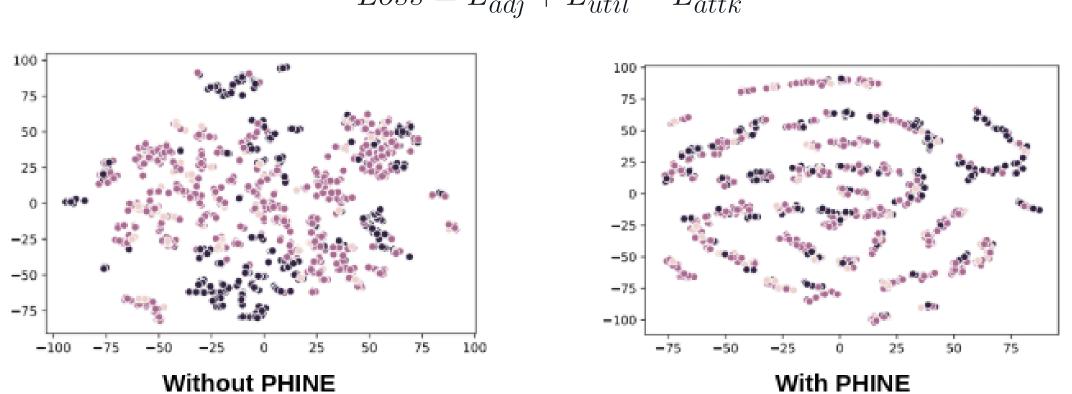
To obtain the utility features for patients, we use a regularization framework for graphbased learning (Equation 1), to propagate the utility features information to the patient nodes considering the network topology. We extend the regularization framework for heterogeneous networks proposed by [2] to deal with specific characteristics of clinical events. After the regularization, the Feature Matrix is a learned utility features matrix for all nodes in the heterogeneous network, including patient type nodes.

$$Q(\mathbf{X}) = \frac{1}{2} \sum_{o_i, o_j \in \mathcal{O}} w_{o_i, o_j} (x_{o_i} - x_{o_j})^2 + \mu \sum_{o_k \in \mathcal{O}_T} (x_{o_k} - u_{o_k})^2 + \mu \sum_{o_k \in \mathcal{O}_T} (x_{o_k} - u_{o$$

Privacy-Preserving Heterogeneous Network Embedding

If the matrix generated on the previous step will be public for machine learning tasks, then an attacker who obtains a partial set of private features can train a classifier (or regressor) to infer the other users' private features. PHINE employs Graph Convolutional Networks (GCN) to conceal private information and keep the utility information no embeddings data. The decoder step of GCN was extended to consider the reconstruction of utility features and private features via link prediction. In this sense, we adapted the privacy-preserving embedding technique for homogeneous networks proposed by [2] to deal with heterogeneous networks. An important aspect that we consider during training is that this step is related to patient-type nodes. Thus, we use a trick that generates a fake private label for all nodes that do not belong to the patient type, thereby allowing the model training for heterogeneous networks.

The GCN final loss function is described in Equation 2. Note that while the terms L_{adj} (result from loss function to optimize the link prediction) and L_{util} (result from loss function to infer utility features) must be minimized during training to reduce the reconstruction error of the links and the utility features prediction, the term L_{attk} (result from loss function to infer private features) receives a negative sign to penalize the loss function when embeddings can infer private features. Thus, during training, we want a trade-off between preserving utility features and avoiding inference attacks.



 $Loss = L_{adj} + L_{util} - L_{attk}$



Experimental Evaluation and Results

(1)



(2)

We used a synthetic EHR generator called Synthea for the experimental evaluation. Patient nodes have eight attributes: node type, ethnicity, birth, death, gender, age category, marital and city. Textual information extracted from medical diagnoses, treatments, symptoms, and observations were added as nodes in the heterogeneous network and connected to the respective patients. The network's final structure was composed of 1834 nodes and 78055 edges. We compared PHINE to the following baselines: (1) kNN-HN: Utility and private label prediction without privacy-preserving mechanisms; (2) Data Sanitization + GCN; (3) Inferences by class distribution.

We defined two experimental evaluation scenarios. In the first, we selected as (binary) labels for utility features the occurrence of "Prediabetes", and the "Gender" was selected as a private feature. In the second, the occurrence of "Otitis media" was selected as a label (binary) for the utility features and the "Marital" status (Married, Single and Others) was selected as private feature (Table 1). The goal is to maximize the Accuracy (ACC) and F-Measure (F1) metrics for the utility prediction task and minimize these metrics for private features prediction task (which simulates an inference attack).

PHINE presented promising results for privacy-preserving, especially for the second scenario (Figure 2). Inference attacks for the private attribute Marital were reduced to values close to Inference by class distribution for the F1 metric.

	Utility Feature (Otitis media)		Private Feature (Marital)	
	ACC	F1	ACC	F1
kNN-HN	0.92 ± 0.02	0.80 ± 0.06	0.74 ± 0.01	0.62 ± 0.01
Data Sanitization + GCN	0.92 ± 0.02	0.80 ± 0.04	0.75 ± 0.03	0.59 ± 0.02
PHINE	0.90 ± 0.01	0.73 ± 0.02	0.60 ± 0.03	0.49 ± 0.01
Inferences by class distribution	0.79 ± 0.01	0.50 ± 0.02	0.33 ± 0.02	0.44 ± 0.03

Table. 1: Utility feature (Otitis media) and Private feature (Marital) evaluation.

Learn More About This Work

We release our code on GitHub with a Coolab Notebook. Scan the following QR Code.



References

- [1] Anahita Hosseini et al. "HeteroMed: Heterogeneous Information Network for Medical Diagnosis". In: (Apr. 2018).
- [2] Kaiyang Li et al. Adversarial Privacy Preserving Graph Embedding against Inference Attack. 2020. arXiv: 2008.13072 [cs.LG].