

Sending AI Algorithms to Medical School: Transfer Learning in Medicine Towards Enhanced Prediction of Opioid Overdose

Sarah Mullin¹, Brianne MacKenzie¹, Kenneth Leonard^{1, 4}, Shageenth Sandrakumar¹, Frank LeHouillier¹, and Peter L. Elkin^{1,2,3,4}

¹Department of Biomedical Informatics, University at Buffalo, The State University of New York, USA, ²Department of Veterans Affairs, WNY VA, USA,

³Faculty of Engineering, University of Southern Denmark, ⁴Department of Psychiatry, University at Buffalo

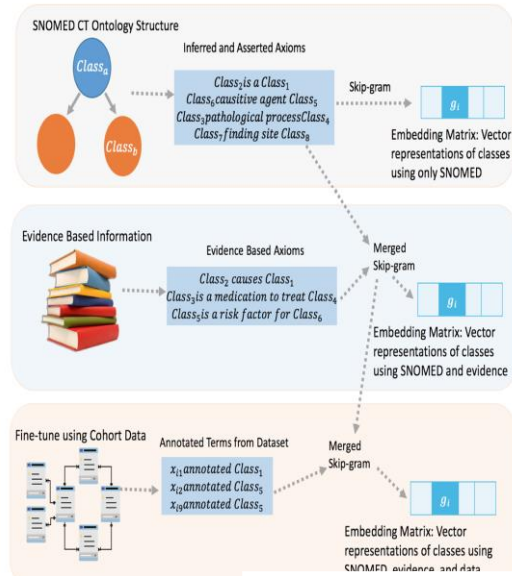
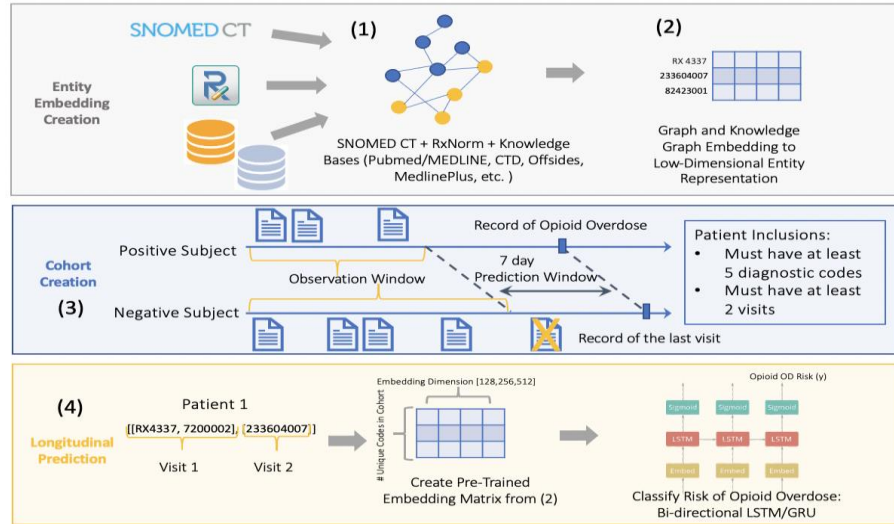
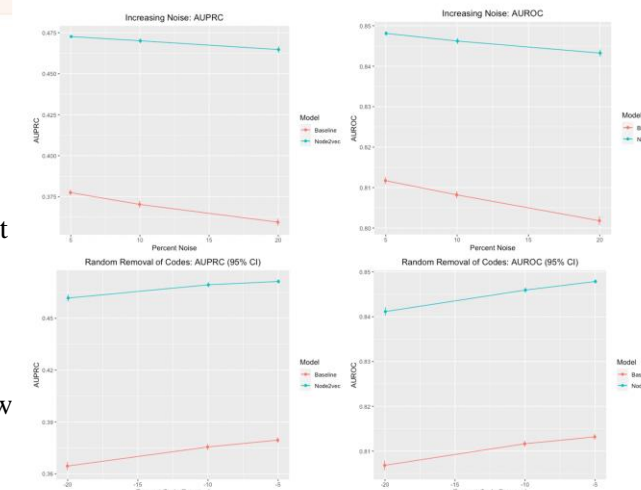
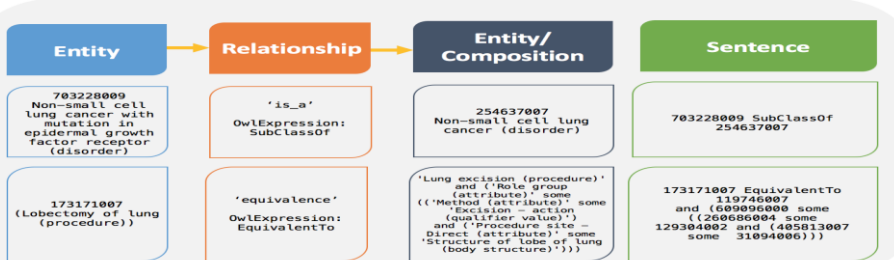


Table 1. Gated-Recurrent Unit RNN and Bi-Directional LSTM Accuracy Measures

		AURO C	AUPR C	Precision	Recall	F1
GRU	TransE	0.789	0.392	0.685	0.139	0.231
	RotatE	0.833	0.444	0.599	0.282	0.384
	DeepWalk	0.821	0.427	0.615	0.261	0.367
	Node2vec	0.844	0.472	0.466	0.509	0.487
	Cui2vec	0.791	0.391	0.501	0.354	0.415
Bi-Directional LSTM	Baseline	0.783	0.356	0.468	0.315	0.377
	TransE	0.834	0.412	0.616	0.238	0.344
	RotatE	0.839	0.437	0.629	0.251	0.359
	DeepWalk	0.794	0.384	0.546	0.261	0.353
	Node2vec	0.841	0.476	0.659	0.268	0.381
	Cui2vec	0.832	0.442	0.587	0.258	0.358
Baseline	0.809	0.401	0.563	0.263	0.358	



Conclusions:

- Using Ontology Terms to build Machine Learning Applications is more accurate
- Pre-trained Embeddings with Medical Knowledge Improves this Accuracy and provides some insulation against missing or bad training data

Method: Pipeline for using graph-aware embeddings in predictive algorithms. (1) First, we extract SNOMED CT and RxNorm terminological information. We then add external knowledge base [subject-relation-object] triples. Finally, we extract opioid specific triples from guidelines. (2) Graph embeddings (Node2vec and DeepWalk) and Knowledge graph embeddings (TransE and RotatE) are trained on these triples. (3) We build our alcohol use disorder cohort by selecting positive opioid overdose subjects and controls, negative subjects with up to an observation window of 2 years and a prediction window of 7 days. (4) For longitudinal predictive algorithms, the embeddings are used as prior information which the models fine-tune on the cohort data.