Generalizing Biomedical Relation Classification with Neural Adversarial Domain Adaptation

Anthony Rios

Ph.D. Candidate
Department of Computer Science
University of Kentucky

Email: anthonymrios@gmail.com

Biomedical Relations: Drug-Drug Interactions

Drug-Drug Interactions

- **1.5 million** adverse drug reactions (ADRs) each year (2009)
- **4 billion dollars** spent each year to prevent treatable ADRs.

Biomedical Relations

- Drug-Drug Interactions
- Protein-Protein Interactions
- Drug-Protein Interactions

Biomedical Relations

- Drug-Drug Interactions
- Protein-Protein Interactions
- Drug-Protein Interactions

Biomedical Relations

- Drug-Drug Interactions
- Protein-Protein Interactions
- Drug-Protein Interactions

SELECT DrugA **FROM** myTable **WHERE**
relation = 'kills_patient' **AND**
DrugB = 'warfarin'

DrugA	relation	DrugB
Diflunisal	kills_patient	warfarin
nevirapine	inhibits	warfarin

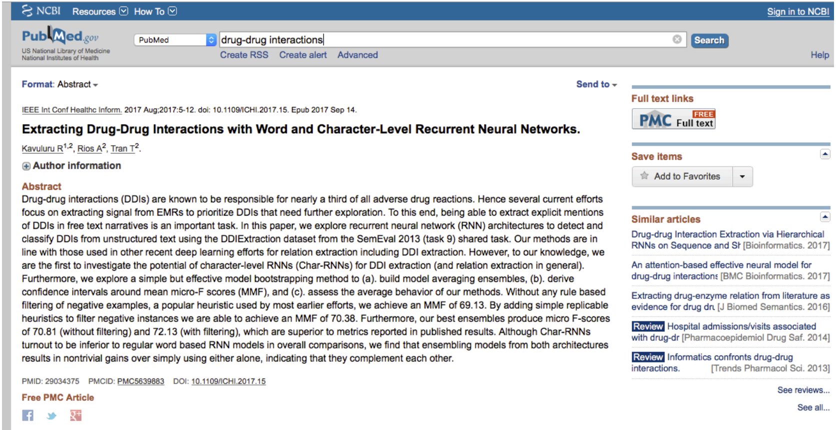
SELECT DrugA **FROM** myTable **WHERE**
relation = 'kills_patient' **AND**
DrugB = 'warfarin'

DrugA	relation	DrugB
Diflunisal	kills_patient	warfarin
nevirapine	inhibits	warfarin

Research Question: How to collect a comprehensive database (myTable) of biomedical relations?

Read PubMed: 23 Million Indexed Citations

Goal: Extract relations from articles indexed on PubMed.



NCBI Resources How To Sign in to NCBI

PubMed.gov US National Library of Medicine National Institutes of Health

PubMed | drug-drug interactions | Search

Create RSS Create alert Advanced Help

Format: Abstract Send to

IEEE Int Conf Healthc Inform. 2017 Aug;2017:5-12. doi: 10.1109/ICHI.2017.15. Epub 2017 Sep 14.

Extracting Drug-Drug Interactions with Word and Character-Level Recurrent Neural Networks.

Kavuluru R[1,2], Rios A[2], Tran T[2].

Author information

Abstract

Drug-drug interactions (DDIs) are known to be responsible for nearly a third of all adverse drug reactions. Hence several current efforts focus on extracting signal from EMRs to prioritize DDIs that need further exploration. To this end, being able to extract explicit mentions of DDIs in free text narratives is an important task. In this paper, we explore recurrent neural network (RNN) architectures to detect and classify DDIs from unstructured text using the DDIExtraction dataset from the SemEval 2013 (task 9) shared task. Our methods are in line with those used in other recent deep learning efforts for relation extraction including DDI extraction. However, to our knowledge, we are the first to investigate the potential of character-level RNNs (Char-RNNs) for DDI extraction (and relation extraction in general). Furthermore, we explore a simple but effective model bootstrapping method to (a). build model averaging ensembles, (b). derive confidence intervals around mean micro-F scores (MMF), and (c). assess the average behavior of our methods. Without any rule based filtering of negative examples, a popular heuristic used by most earlier efforts, we achieve an MMF of 69.13. By adding simple replicable heuristics to filter negative instances we are able to achieve an MMF of 70.38. Furthermore, our best ensembles produce micro F-scores of 70.81 (without filtering) and 72.13 (with filtering), which are superior to metrics reported in published results. Although Char-RNNs turnout to be inferior to regular word based RNN models in overall comparisons, we find that ensembling models from both architectures results in nontrivial gains over simply using either alone, indicating that they complement each other.

PMID: 29034375 PMCID: PMC5639883 DOI: 10.1109/ICHI.2017.15

Free PMC Article

Full text links

PMC FREE Full text

Save items

Add to Favorites

Similar articles

Drug-drug Interaction Extraction via Hierarchical RNNs on Sequence and Sł [Bioinformatics. 2017]

An attention-based effective neural model for drug-drug interactions [BMC Bioinformatics. 2017]

Extracting drug-enzyme relation from literature as evidence for drug dr [J Biomed Semantics. 2016]

Review Hospital admissions/visits associated with drug-dr [Pharmacoepidemiol Drug Saf. 2014]

Review Informatics confronts drug-drug interactions. [Trends Pharmacol Sci. 2013]

See reviews...

See all...

Biomedical Relation Extraction

Relation Extraction: Relation extraction is the task of automatically extracting **structured information** from **unstructured documents**.

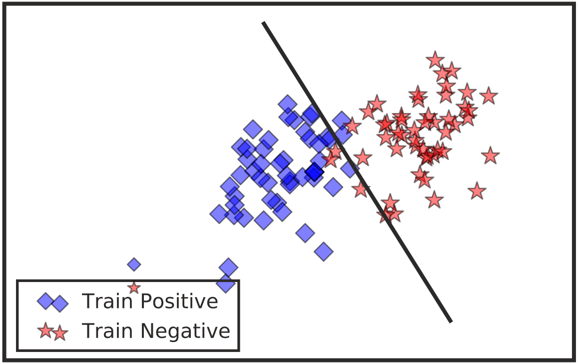
Protein-Protein Interaction Example: *Human cyclin E, a new cyclin that interacts with two members of the CDC2 gene family*

Drug-Drug Interaction Example: *The invitro interaction between nevirapine and the antithrombotic agent warfarin is complex.*

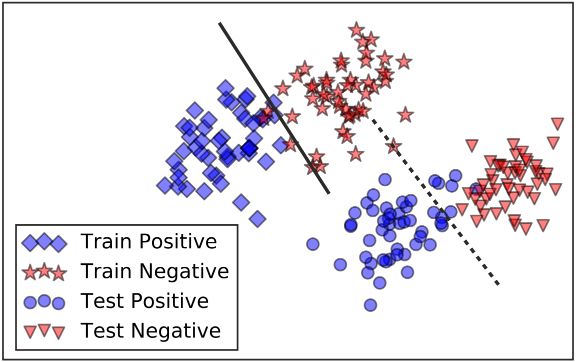
SUBJECT	RELATION	OBJECT
cyclin E	interacts_with	CDC2
nevirapine	interacts_with	warfarin

Challenges

Supervised Machine Learning



Covariate Shift

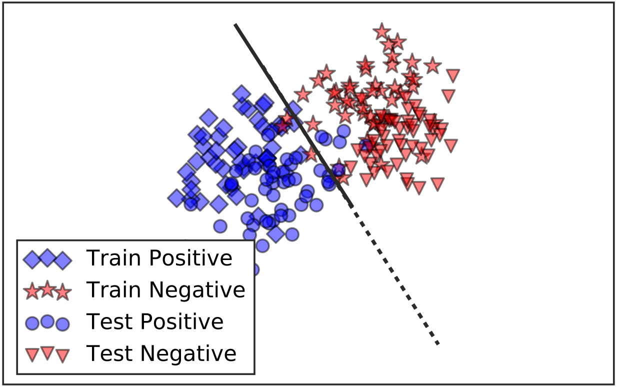


Domain Adaptation: Domain adaptation arises when the **source** data distribution is **different (but related)** to the **target** data distribution.

$$P(X) \neq P(X'); \; P(Y|X) \approx P(Y'|X')$$

Unsupervised Domain Adaptation: No labeled target data.

Covariate Shift



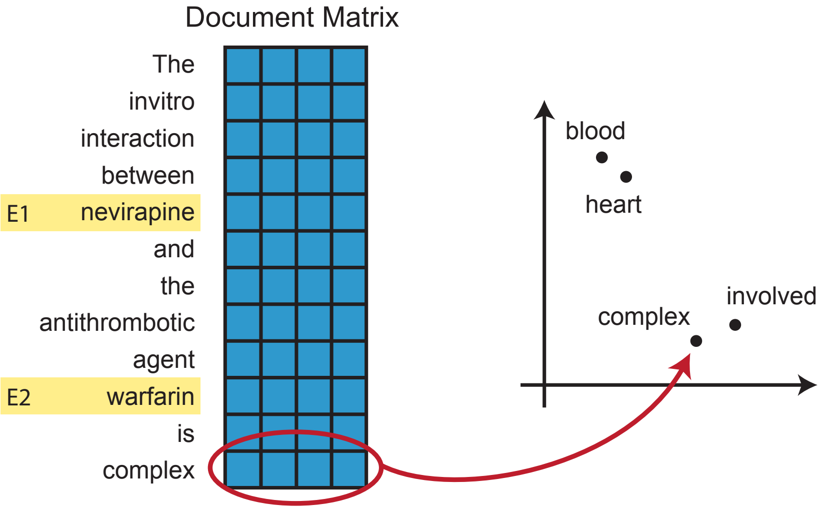
CNNs for Relation Extraction

CNNs for Relation Extraction

E1 E2

The invitro interaction between nevirapine and the antithrombotic agent warfarin is complex.

Relation Extraction CNN



CNNs for Relation Extraction

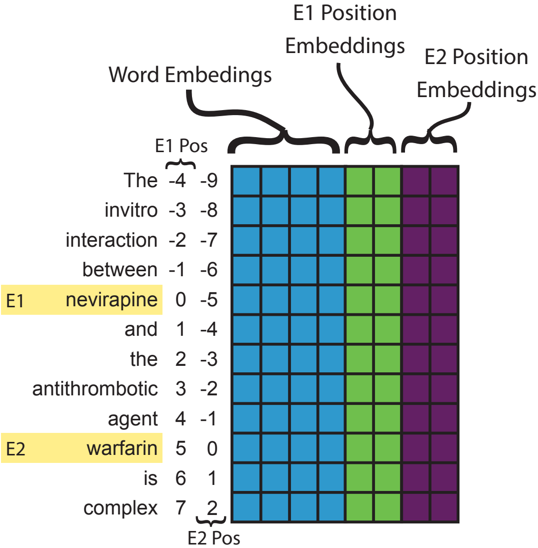
E1 E2 E3

The invitro interaction between nevirapine and the antithrombotic agent warfarin (Anasmol) is complex.

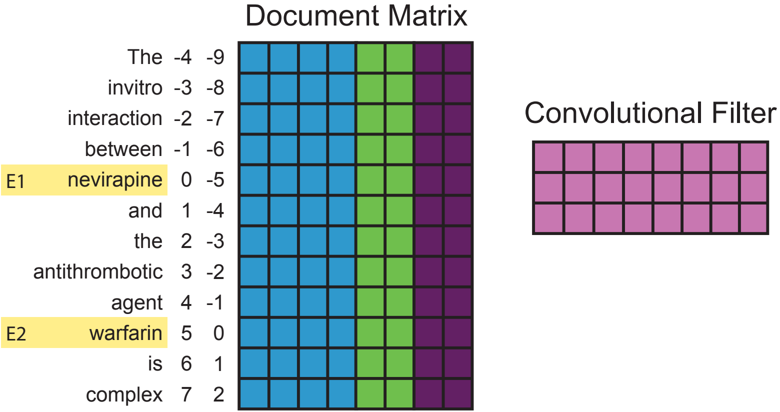
CNNs for Relation Extraction

					E1					E2		
	The	invitro	interaction	between	nevirapine	and	the	antithrombotic	agent	warfarin	is	complex.
Distance to E1	-4	-3	-2	-1	0	1	2	3	4	5	6	7
Distance to E2	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2

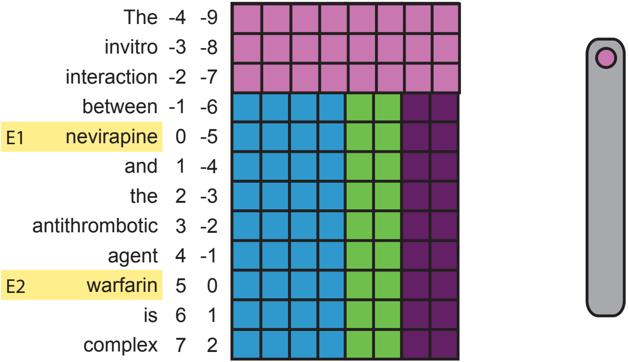
Relation Extraction CNN



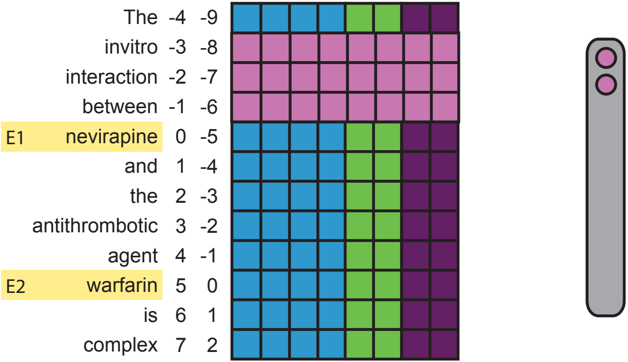
Convolutional Layer



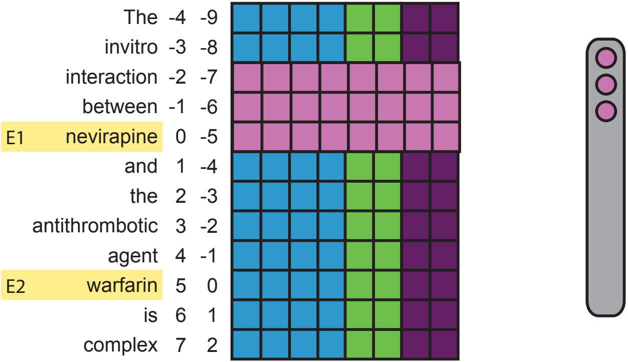
Convolutional Layer



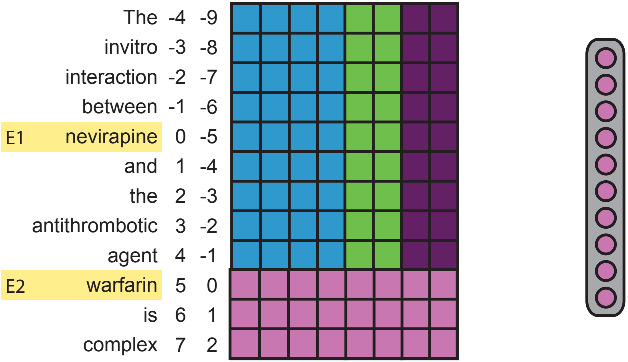
Convolutional Layer



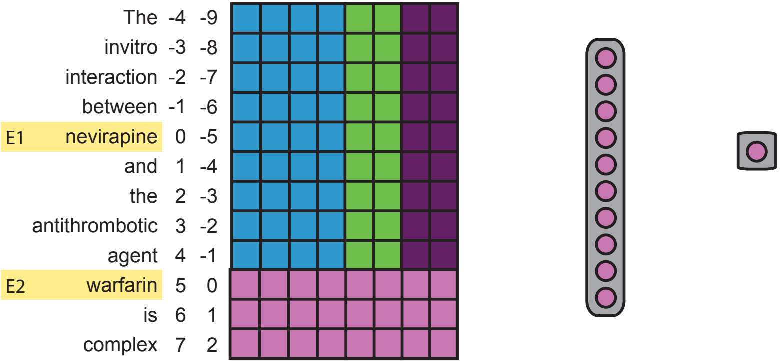
Convolutional Layer



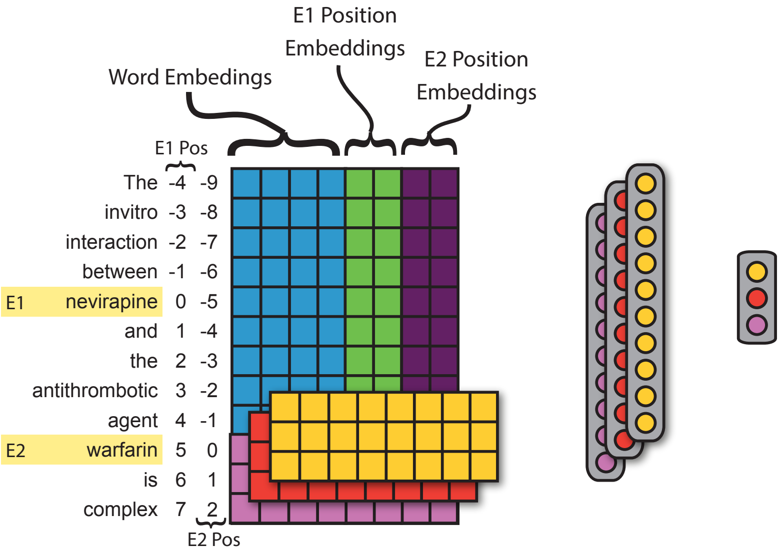
Convolutional Layer



Max-over-time Pooling

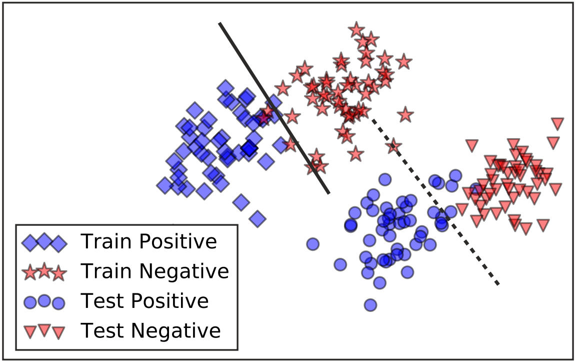


Relation Extraction CNN



Solution

Domain Adaptation

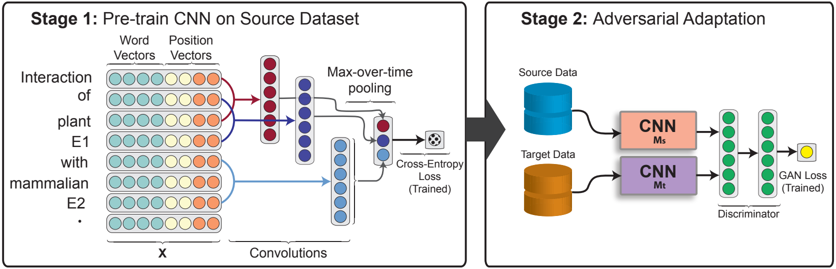


Domain Adaptation: Domain adaptation arises when the **source** data distribution is **different (but related)** to the **target** data distribution.

$$P(X) \neq P(X'); \; P(Y|X) \approx P(Y'|X')$$

Unsupervised Domain Adaptation: No labeled target data.

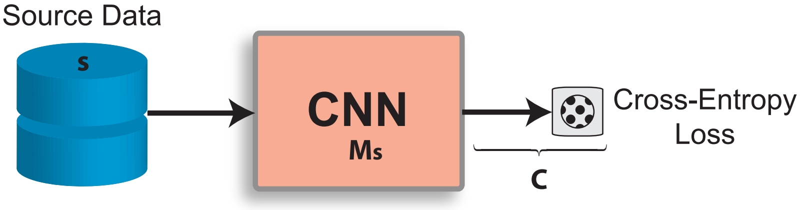
Train Model in 2 Stages



Our model uses **2 datasets** for training:

- **Source Dataset** – A labeled (maybe biased) dataset.
- **Target Dataset** – An unlabeled dataset that is more represented of the final test examples.

Stage 1: Classification Loss

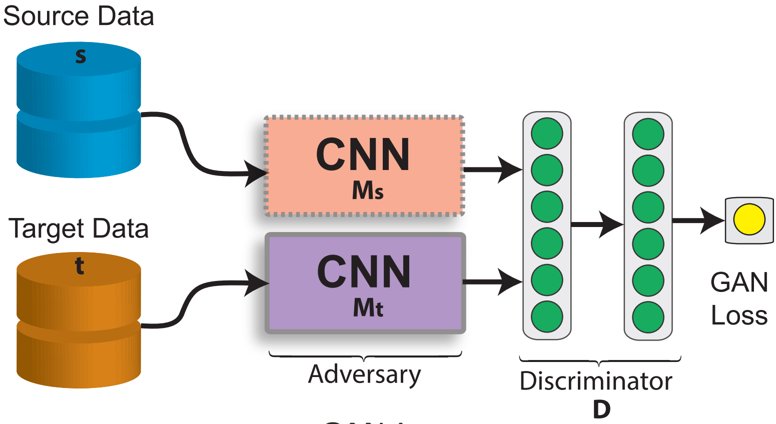


Binary Cross Entropy:

$$min_{\theta_C,\theta_M} \mathbb{E}_{(\mathbf{s},y)\sim\mathbf{S}}\left[-y\, log(C(\mathbf{s})) - (1-y)\, log(1-C(\mathbf{s}))\right],$$

Stage 2: Adversarial Domain Adaptation

Stage 2: GAN Loss

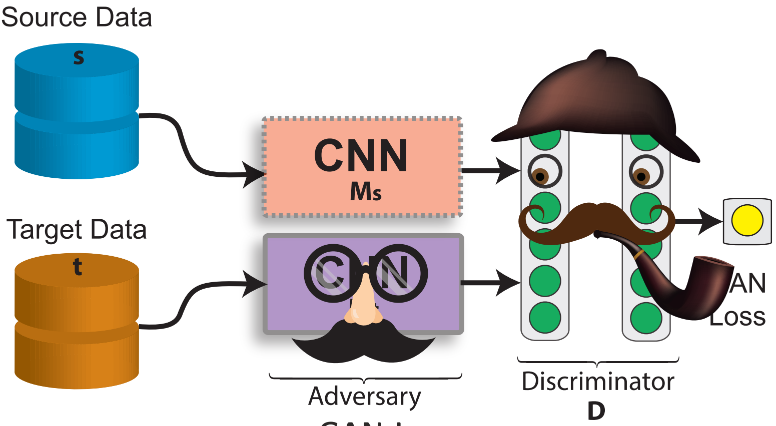


GAN Loss:

$$min_{\theta_D} \mathbb{E}_{\mathbf{s}\sim\mathbf{S}} \left[-\, log(D(M_s(\mathbf{s}))) \right] - \mathbb{E}_{\mathbf{t}\sim\mathbf{T}} \left[log(1 - D(M_t(\mathbf{t}))) \right],$$

$$min_{\theta_{M_t}} \mathbb{E}_{\mathbf{t}\sim\mathbf{T}} \left[-\, log(D(M_t(\mathbf{t}))) \right]$$

Stage 2: GAN Loss



GAN Loss:

$$min_{\theta_D} \mathbb{E}_{\mathbf{s} \sim \mathbf{S}} \left[-log(D(M_s(\mathbf{s}))) \right] - \mathbb{E}_{\mathbf{t} \sim \mathbf{T}} \left[log(1 - D(M_t(\mathbf{t}))) \right],$$

$$min_{\theta_{M_t}} \mathbb{E}_{\mathbf{t} \sim \mathbf{T}} \left[-log(D(M_t(\mathbf{t}))) \right]$$

Results

Results

Datasets:

- A – Protein-Protein Interactions
- B – Protein-Protein Interactions
- C – Drug-Drug Interactions

	B ⇒ A	A ⇒ B	B ⇒ C	C ⇒ B	A ⇒ C	C ⇒ A	AVG
CNN	0.4522	0.3672	0.3975	0.2213	0.1583	0.2793	0.3126
Bi-LSTM	0.4688	0.2959	0.4087	0.1721	0.1858	0.2580	0.2982
CNN RevGrad	0.4731	0.4255	0.4196	0.3611	0.3131	0.3072	0.3833
Bi-LSTM RevGrad	0.4641	0.4011	0.3941	0.3720	0.2772	0.3529	0.3769
Adv-CNN (Ours)	**0.4879**	0.5413	0.4419	**0.4853**	0.4596	**0.4471**	**0.4772**
Adv-Bi-LSTM (Ours)	0.4851	**0.5654**	**0.4447**	0.449	**0.4657**	0.4344	0.4746

Table: F1-score for all pair wise combinations (source ⇒ target) of the three datasets

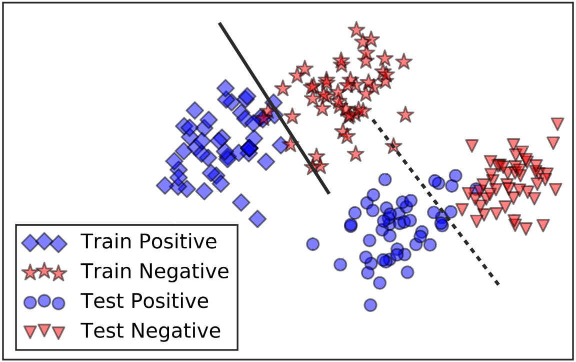
Biomedical Relation Extraction

Relation Extraction: Relation extraction is the task of automatically extracting **structured information** from **unstructured documents**.

> **Protein-Protein Interaction Example:** *Human cyclin E, a new cyclin that interacts with two members of the CDC2 gene family*
>
> **Drug-Drug Interaction Example:** *The invitro interaction between nevirapine and the antithrombotic agent warfarin is complex.*

E1	E1 Type	Relation	E2	E2 Type
cyclin E	Protein	interacts_with	CDC2	Protein
nevirapine	Drug	interacts_with	warfarin	Drug

Covariate Shift



Domain Adaptation: Domain adaptation arises when the **source** data distribution is **different (but related)** to the **target** data distribution.

$$P(X) \neq P(X'); \; P(Y|X) \approx P(Y'|X')$$

Unsupervised Domain Adaptation: No labeled target data.

Future Work: Domain Adaptation

- Simulated data to real data.
- Shift between hospital A and hospital B.
- Multi-source domain adaptation.

Future Work: Domain Adaptation

- Simulated data to real data.
- Shift between hospital A and hospital B.
- Multi-source domain adaptation.

Future Work: Domain Adaptation

- Simulated data to real data.
- Shift between hospital A and hospital B.
- Multi-source domain adaptation.

Other Areas of Research

- Extreme Multi-label Classification (10k+ classes).
- Social Medical Monitoring.

Thank You!

Email: anthonymrios@gmail.com
Code: https://github.com/AnthonyMRios/relation-extraction-rnn

Acknowledgements

University of Kentucky

NIH NLM

NCBI