Extracting Drug Information with Apache cTAKES and ClearTK

Timothy A Miller[1] and Dmitriy Dligach[2]

[1]Computational Health Informatics Program, Boston Children's Hospital, Harvard Medical School

[2]Department of Computer Science, Loyola University Chicago

Background

Adverse drug events cause harm to patients and are not well-understood, especially in pediatric populations. We are interested in designing machine learning algorithms for extracting adverse drug events from texts in the electronic health record.

This dataset largely annotates *physician-inferred adverse drug events*. These represent cases where ADEs were believed by clinicians to have occurred, and then documented.

Data

Example ADEs (ADEs in Red, Drugs in Blue)

Elevated LFTs/coagulopathy: Secondary to shock liver versus **tylenol** toxicity.

During the admission the pt was thought to have had a **ceftriaxone** allergic reaction (**morbilliform drug rash**) and ceftriaxone was added to allergy list.

She did not experience significant nausea during hospitalization, but continued to have **diarrhea** related to her **chemotherapy** which was treated with Lomotil.

Your **rash** appeared to improve after discontinuation of the **antibiotics piperacillin** and **tazobactam**, but the exact cause was unknown.

Hypotension - The patient was initially hypotensive in the [**Hospital Unit Name 153**]. This was after receiving **Propofol** and then other **sedating medications**.

Methods

We participated in all 3 tracks:

- Entity extraction
- Relation identification
- End-to-end relation extraction

Our focus was on adapting general-purpose clinical NLP tools to this new task. We used:

- ClearTK [1] entity extraction API to train a BIO tagger for entity extraction (above right)

increase to 50mg po BID for seven days then

Duration	O	O	O	O	O	B	I	I	O
Frequency	O	O	O	O	B	O	O	O	O
Strength	O	O	B	O	O	O	O	O	O
Route	O	O	O	B	O	O	O	O	O
Drug	O	O	O	O	O	O	O	O	O

- Features based on characters, word, and part-of-speech of focus word and context words (right)

increase to 50mg po BID for seven days then

ID	increase	to	50mg	po	BID	for	seven	days	then
POS	VB	TO	N	ADJ	PN	P	CD	NS	P
Char	L*	L*	N*L*	L*	U*	L*	L*	L*	L*

Apache cTAKES [2] Relation Extractor module for relation classification [3]:

- She did not experience significant **nausea** during hospitalization, but continued to have **diarrhea** related to her **chemotherapy** which was treated with **Lomotil**
- Specify allowed types for each argument:
  - Arg1 : Medication Mention
  - Arg2 : Sign/Symptom/Disease/Disorder
- Compare all pairs within covering type (paragraph)
  - (Chemotherapy, nausea) -> False
  - (Chemotherapy, diarrhea) -> True
  - (Lomotil, nausea) -> False
  - (Lomotil, diarrhea) -> False

Results

Entities	F1
Drug	0.9375
ADE	**0.3092**
Dosage	0.9082
Duration	**0.7106**
Form	0.9461
Frequency	0.9520
Reason	**0.5077**
Route	0.9369
Strength	0.9655
Average	**0.9052**
Task Median	0.9052
Task Average	0.8467

Relations	F1
ADE	**0.7635**
Dosage	0.9503
Duration	**0.8223**
Form	0.9660
Frequency	0.9136
Reason	**0.7497**
Route	0.9416
Strength	0.9428
Average	**0.9067**
Task Median	0.8934
Task Average	0.810

End to end Relations	F1
ADE	0.2758
Dosage	0.8780
Duration	0.5992
Form	0.9219
Frequency	0.8784
Reason	0.4195
Route	0.8894
Strength	0.9139
Average	**0.8249**
Task Median	0.8037
Task Average	0.7485

The classifier for all tasks is a linear SVM using the Liblinear library [4].

- With minimal domain-specific modeling, we were able to obtain results > 0.9 in tasks 1 and 2, and 0.82 for the end-to-end task.
- Novel challenges of the task – ADE and Reason – were some of the worst performing
- Scoring metric incentivized work on most prevalent categories

Blue indicates final scores, **bold** indicates trouble areas

Error Analysis

Reason/ADE

- Using Sign/Symptom/Disease/Disorder as Argument type is still too fine-grained (Recall errors)
- Drug categories are sometimes annotated as well as drugs, not always in ontologies (Recall errors)
- Reason and ADE relations are typically much farther apart than other relations, harder to memorize context between arguments (Recall errors)

Duration

- Confusion with time expressions (Precision errors)

References

[1] S. Bethard, P. Ogren, L. Becker, ClearTK 2.0: Design Patterns for Machine Learning in UIMA., LREC (2014).

[2] G. Savova, J. Masanz, P. Ogren, Mayo clinical Text Analysis and Knowl- edge Extraction System (cTAKES): architecture, component evaluation and applications, Journal of the American Medical Informatics Associa- tion (2010).

[3] D. Dligach, S. Bethard, L. Becker, T. Miller, G. K. Savova, Discovering body site and severity modifiers in clinical texts., Journal of the American Medical Informatics Association : JAMIA 21 (????) 448–54.

[4] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, C.-J. Lin, LIBLINEAR: A Library for Large Linear Classification, Journal of Machine Learning Research 9 (2008) 1871–1874.

Acknowledgements and Contact

Research reported in this poster was supported by the National Library Of Medicine of the National Institute of Health under Award Number R01LM012918 and by the National Institute of General Medical Sciences under Award Number R01GM114355.

Contact:

Email: timothy.miller@childrens.harvard.edu, ddligach@luc.edu

Web: http://scholar.harvard.edu/tim-miller