cTAKES 3.0 - Side Effect

cTAKES 3.0 Links

Apache cTAKES web site
Documentation:

- cTAKES 3.0
- cTAKES 3.0 User Install Guide
- cTAKES 3.0 Developer Install Guide
- cTAKES 3.0 Component Use Guide
- Documentation Conventions

Overview of Side Effect

The side effect extraction system extracts physician-asserted drug side effects from clinical notes. Alternatively the system can extract sentences that possibly contain both side effects and causative drugs, which cover higher number of side effect occurrences but needs human validation to extract individual causative drugs and side effects.

Analysis engines (Annotator) Descriptors

SideEffectAggregateTAE.xml

An aggregated engine defines a pipeline used by side effect extraction on plain text with Mayo format section tags.

SideEffectAggregate_CDA_TAE.xml

An aggregated engine defines a pipeline used by side effect extraction on CDA documents.

SideEffectSentenceAggregateTAE.xml

An aggregated engine defines a pipeline used by side effect sentence extraction on plain text with Mayo format section tags.

SideEffectSentenceAggregate_CDA_TAE.xml

An aggregated engine defines a pipeline used by side effect sentence extraction on CDA documents.

SideEffectAnnotator.xml

This annotator uses pattern matching rules to identify side effect and causative drug.
Parameters

- sectionToIgnore
  sections to ignore for side effect extraction (a default one uses Mayo Clinic section convention)
  The following parameters are side effect indication keywords in pattern matching rules. Users may use their own keywords:
- hasPatternOfDrugCauseVerbPse
- hasPatternOfPseDueToDrug
- hasPatternOfDrugDueToPse
- hasPatternOfDiscontDrugBecausePse
- sideEffectWord
- hasPatternOfNotePseWithDrug
- hasPatternOfDrugMadePse
- hasPatternOfPseAfterDrug

Resources
sideEffectDic: - A:: dictionary file that contains the drug and known side effect used for the dictionary lookup rule. The default file includes sample cases

SESentenceClassifierAnnotator.xml

An annotator to extract sentences that contains side effect and causative drugs.
Parameters
PathOfModel - A machine learning model of side effect sentence classification

PSESentenceAnnotator.xml

This annotator extracts spans of sentences that contain potential side effects (i.e., signs/symptoms and diseases/disorders) and drugs and add them to the type PSESentence. Note that certain sections (in clinical notes) defined in parameter, SectionsToIgnore are ignored.
Parameters
SectionToIgnore: sections not to be considered for this annotation (a default one uses Mayo Clinic section convention)

SESentenceFeatureAnnotator.xml

This annotator extracts features for side effect sentence classification and add them to the type PSESentenceFeature.
Parameters
MetaKeywordsFile:
A file of keywords used for side effect sentence classification

LookupWindowAnnotator_sideEffect.xml

A customized lookup window annotator for side effect extraction to properly catch potential side effect candidates.

DictionaryLookupAnnotator_sideEffect.xml

A customized dictionary lookup annotator for side effect extraction.
Resources
LookupDescriptorFile: a lookup descriptor file used for this project
SnomedIndexReader: custom resource specifier
RxnormIndexReader: custom resource specifier
OrangeBookIndexReader: custom resource specifier
CsvSEFile: a file for supplemental named entities. A user may use this file to add extra named entities that are not in a dictionary (the format should be first_word|full_named_entity|typeID(e.g., drug=1, disease/disorder=2, signs/symptoms=3, etc.). The default file includes sample cases.

CAS Consumers

- SideEffectCasConsumer
  Write side effects and causative drugs to the file in the format of:

  ```
  documentName|sideEffect|beginOffset|endOffset|drug|beginOffset|endOffset|sentence
  ```

- SideEffectSentenceCasConsumer
  Write side effects sentences to the file in the format of:

```
document_name|side_effect_sentence|beginOffset|endOffset
```

Collection Processing Engine (CPE)

SideEffect_CDA_CPE.xml

A CPE for both side effect and side effect sentence extraction on CDA documents.

SideEffectSentenceCPE.xml

A CPE for both side effect and side effect sentence extraction on plain text with Mayo format section tags.
Resources
libsvm-2.91.jar: The support vector machine (SVM) classification tool used to side effect sentence classification.

Type System - SideEffectTypeSystem.xml

Types used in this project. Click JCasGen to generate necessary types.

How to

- Run side effect extraction
  Use SideEffectCPE.xml for plain text with Mayo section tags or SideEffect_CDA_CPE.xml for CDA documents
- Run both side effect and side effect sentence extraction
  Use SideEffectSentenceCPE.xml for plain text with Mayo section tags or SideEffectSentence_CDA_CPE.xml for CDA documents

  > Input file can be either plain text with Mayo format section tags or CDA documents (if you use a CDA format, take out SimpleSegmentWithTagsAnnotator in the TAE workflow)

- Customize side effect extraction rules
  The method, getSideEffectsWithPrioritizedRule() in SideEffectAnnotator.java contains the list of rules used for side effect extraction. A user can modify existing rules or create new rules.