Master in Artificial Intelligence

The Task: SemEval 2013

Dataset

Course Plan

Advanced Human Language Technologies



UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH





Outline

The Task: SemEval 2013

Dataset

Course Plan

1 The Task: SemEval 2013

2 Dataset

SemEval-2013 Task 9: DDI Extraction

The Task: SemEval 2013

Dataset
Course Plan

- SemEval International Conference on Semantic Evaluation.
 Several tasks or challenges are posed every edition
- SemEval 2013 Task 9: DDIExtraction Detect drug names and interactions among them described in text.
- Documents Documents extracted from DrugBank (Drug description leaflets database) and MedLine (abstracts of medical papers)
- Participants 11 research teams from around the world (1 Cuba, 1 Italy, 1 Finland, 2 Germany, 1 Portugal, 2 Spain, 3 USA)
- Subtasks
 - Drug name recognition and classification (NERC)
 - Drug-Drug interaction recognition and classification (DDI)

Outline

The Task: SemEval 2013

Dataset Course Plan 1 The Task: SemEval 2013

2 Dataset

DDI Extraction Dataset

The Task: SemEval 2013

Dataset

Course Plan

DrugBank document example

```
<document id="DDI-DrugBank.d483">
   <sentence id="DDI-DrugBank.d483.s0" text="Increases in serum creatinine have occurred when CEFOTAN was given alone.">
      <entity id="DDI-DrugBank.d483.s0.e0" charOffset="49-55" type="brand" text="CEFOTAN"/>
   </sentence>
   <sentence id="DDI-DrugBank.d483.s1"</pre>
           text="If CEFOTAN and an aminoglycoside are used concomitantly, renal function should be carefully monitored,
                because nephrotoxicity may be potentiated.">
      <entity id="DDI-DrugBank.d483.s1.e0" charOffset="3-9" type="brand" text="CEFOTAN"/>
      <entity id="DDI-DrugBank.d483.sl.el" charOffset="18-31" type="group" text="aminoglycoside"/>
      </sentence>
   (...)
   <sentence id="DDI-DrugBank.d483.s4"</pre>
           text="As with other cephalosporins, high concentrations of cefotetan may interfere with measurement of
                serum and urine creatinine levels by Jaffe reaction and produce false increases in the levels
                of creatinine reported.">
      <entity id="DDI-DrugBank,d483.s4.e0" charOffset="14-27" type="group" text="cephalosporins"/>
      <entity id="DDI-DrugBank.d483.s4.e1" charOffset="53-61" type="drug" text="cefotetan"/>
      </sentence>
</document>
```

DDI Extraction Dataset

MedLine document example

The Task: SemEval 2013

Dataset

```
<document id="DDI-MedLine.d16">
  <sentence id="DDI-MedLine.d16.s0"</pre>
        text="Interaction of ketamine and halothane in rats.">
     <entity id="DDI-MedLine.d16.s0.e0" charOffset="15-22" type="drug" text="ketamine"/>
     <entity id="DDI-MedLine.d16.s0.e1" charOffset="28-36" type="drug" text="halothane"/>
     </sentence>
  <sentence id="DDI-MedLine.d16.s1"</pre>
         text="The interaction of intramuscularly injected ketamine and its N-demethylated metabolite (metabolite I)
             with halothane was evaluated in rats. ">
     <entity id="DDI-MedLine.d16.s1.e0" charOffset="44-51" type="drug" text="ketamine"/>
     <entity id="DDI-MedLine.d16.s1.e1" charOffset="107-115" type="drug" text="halothane"/>
     </sentence>
  <sentence id="DDI-MedLine.d16.s5"</pre>
         text="The half-life of ketamine in plasma and brain was longer in the presence of halothane than when
             ketamine was given alone. ">
     <entity id="DDI-MedLine.dl6.s5.e0" charOffset="17-24" type="drug" text="ketamine"/>
     <entity id="DDI-MedLine.d16.s5.e1" charOffset="76-84" type="drug" text="halothane"/>
     <entity id="DDI-MedLine.d16.s5.e2" charOffset="96-103" type="drug" text="ketamine"/>
     </sentence>
</document>
```

Outline

The Task: SemEval 2013

Dataset

Course Plan

1 The Task: SemEval 2013

2 Dataset

Course plan for lab sessions

The Task: SemEval 2013

Dataset

Course Plan

- Sessions 1-2: Develop a ML system for Task 9.1: Drug NERC.
 - The B-I-O approach. Feature extraction.
- Sessions 3-4: Develop a ML system for Task 9.2: DDI detection
 - Sentence classification. Feature extraction.
- Sessions 5-7: Develop DNN alternatives for Tasks 9.1 and 9.2
 - NERC Sequence label approach with NN
 - DDI Sentence classification with NN

Lab project must be developed in teams of 2 people

Course plan for lab sessions

The Task: SemEval 2013 Dataset

Course Plan

■ Warning 1: Lab sessions may not be enough. You'll probably need to work on the assignments at home.

> ■ Warning 2: Even if you can complete the work in class, sessions are every each other week, so you may forget what you were doing... It is good that you do work a bit at home to get things freshened-up