

Master in Artificial Intelligence

SemEval 2013

Dataset

Course Plan

Advanced Human Language Technologies



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Facultat d'Informàtica de Barcelona



Outline

SemEval 2013

Dataset

Course Plan

1 SemEval 2013

2 Dataset

3 Course Plan

SemEval-2013 Task 9: DDI Extraction

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

SemEval 2013

Dataset

Course Plan

1 SemEval 2013

2 Dataset

3 Course Plan

DDI Extraction Dataset

SemEval 2013

Dataset

Course Plan

DrugBank document example

```
- <document id="DDI-DrugBank.d193">
  - <sentence id="DDI-DrugBank.d193.s0" text="A drug interaction study was performed in which ERBITUX was administered in combination with irinotecan.">
    <entity id="DDI-DrugBank.d193.s0.e0" charOffset="48-54" type="brand" text="ERBITUX"/>
    <entity id="DDI-DrugBank.d193.s0.e1" charOffset="93-102" type="drug" text="irinotecan"/>
    <pair id="DDI-DrugBank.d193.s0.p0" e1="DDI-DrugBank.d193.s0.e0" e2="DDI-DrugBank.d193.s0.e1" ddi="false"/>
  </sentence>
  - <sentence id="DDI-DrugBank.d193.s1" text="There was no evidence of any pharmacokinetic interactions between ERBITUX and irinotecan.">
    <entity id="DDI-DrugBank.d193.s1.e0" charOffset="66-72" type="brand" text="ERBITUX"/>
    <entity id="DDI-DrugBank.d193.s1.e1" charOffset="78-87" type="drug" text="irinotecan"/>
    <pair id="DDI-DrugBank.d193.s1.p0" e1="DDI-DrugBank.d193.s1.e0" e2="DDI-DrugBank.d193.s1.e1" ddi="false"/>
  </sentence>
</document>
```

DDI Extraction Dataset

MedLine document example

SemEval 2013

Dataset

Course Plan

```
- <document id="DDI-MedLine.d19">
  <sentence id="DDI-MedLine.d19.s0" text="Anaesthesia and the epileptic pateint. "/>
  <sentence id="DDI-MedLine.d19.s1" text="A review. "/>
  <sentence id="DDI-MedLine.d19.s2" text="A review is presented of some of the problems that may arise in association
  with anaesthesia for epileptic patients. "/>
  <sentence id="DDI-MedLine.d19.s3" text="There is the possibility of precipitating anticonvulsant drug toxicity. "/>
  - <sentence id="DDI-MedLine.d19.s4" text="Numerous drug interactions are possible with some anticonvulsant agents,
  such as phenobarbitone and phenytoin, which affect hepatic microsomal enzyme systems. ">
    <entity id="DDI-MedLine.d19.s4.e0" charOffset="50-70" type="group" text="anticonvulsant agents"/>
    <entity id="DDI-MedLine.d19.s4.e1" charOffset="81-94" type="drug" text="phenobarbitone"/>
    <entity id="DDI-MedLine.d19.s4.e2" charOffset="100-108" type="drug" text="phenytoin"/>
    <pair id="DDI-MedLine.d19.s4.p0" e1="DDI-MedLine.d19.s4.e0" e2="DDI-MedLine.d19.s4.e1" ddi="false"/>
    <pair id="DDI-MedLine.d19.s4.p1" e1="DDI-MedLine.d19.s4.e0" e2="DDI-MedLine.d19.s4.e2" ddi="false"/>
    <pair id="DDI-MedLine.d19.s4.p2" e1="DDI-MedLine.d19.s4.e1" e2="DDI-MedLine.d19.s4.e2" ddi="false"/>
  </sentence>
  - <sentence id="DDI-MedLine.d19.s5" text="There is the risk of convulsions occurring in susceptible patients following
  the use of the new anaesthetic agents which are capable of inducing CNS excitability.">
    <entity id="DDI-MedLine.d19.s5.e0" charOffset="96-113" type="group" text="anaesthetic agents"/>
  </sentence>
</document>
```

Outline

SemEval 2013

Dataset

Course Plan

1 SemEval 2013

2 Dataset

3 Course Plan

Course plan for lab sessions

SemEval 2013
Dataset
Course Plan

- **Session 1:** Develop a rule-based baseline for Task 9.1 - Drug NERC
- **Sessions 1-2:** Develop a ML-based system for Task 9.1 - Drug NERC
- **Delivery of NERC systems**
- **Session 3:** Develop a rule-based baseline for Task 9.2 - DDI detection
- **Sessions 3-4:** Develop a ML-based system for Task 9.2 - DDI detection
- **Delivery of DDI systems**
- **Sessions 5-7:** Develop DNN alternatives for Tasks 9.1 and 9.2
- **Delivery of DNN alternatives**

Lab project must be developed in teams of 2 people

Course plan for lab sessions

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 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