Extracting food-drug interactions from scientific literature

Tsanta Randriatsitohaina
tsanta@limsi.fr

LIMSI - CNRS, Université Paris-Saclay, France

Supervisor: Thierry Hamon

Context

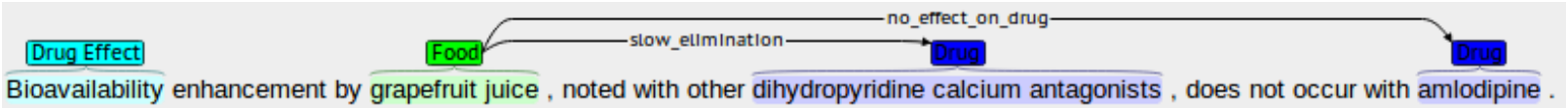
- Food-drug interaction $\Longrightarrow$ Adverse effects
- Less known and sparse in unstructured data
  E.g. : food-drug interactions
  Grapefruit juice *increases effect of other* **dihydropyridine calcium antagonists**.

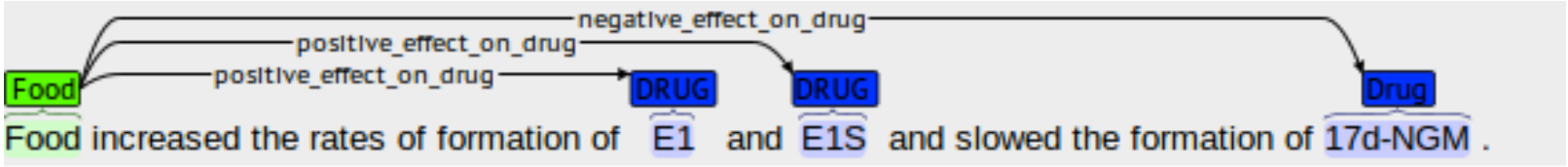
unlike for drug-drug interaction or drug adverse effect (DrugBank [1] or Theriaque [2])

Goal: Automatic identification of interaction statements between drug and food in abstracts of scientific articles issued from the Medline database.

Approach: Use of NLP methods for scientific abstracts mining

Problematic





- Variable mention of drugs and foods in abstracts
- Fine description of interactions
- Unbalanced learning set

$\Longrightarrow$ 14 types of relation, 831 sentences

Corpora

- 639 Medline abstracts with the query

```
(FOOD DRUG INTERACTIONS"[MH] OR "FOOD DRUG INTERACTIONS*" )
                                AND ("adverse effects*")
```

- Brat annotation by an intern in pharmacy [Hamon et al.17]

Relation	#	%	Relation	#	%
unspecified relation	530	58,8%	no effect on drug	109	12,1%
decrease absorption	53	5,9%	improve drug effect	6	0,7%
positive effect on drug	21	2,3%	without food	13	1,4%
negative effect on drug	88	9,8%	speed up absorption	1	0,1%
increase absorption	39	4,3%	worsen drug effect	8	0,9%
slow elimination	15	1,7%	new side effect	4	0,4%
slow absorption	15	1,7%			
Total	902			100%	

Grouping relation

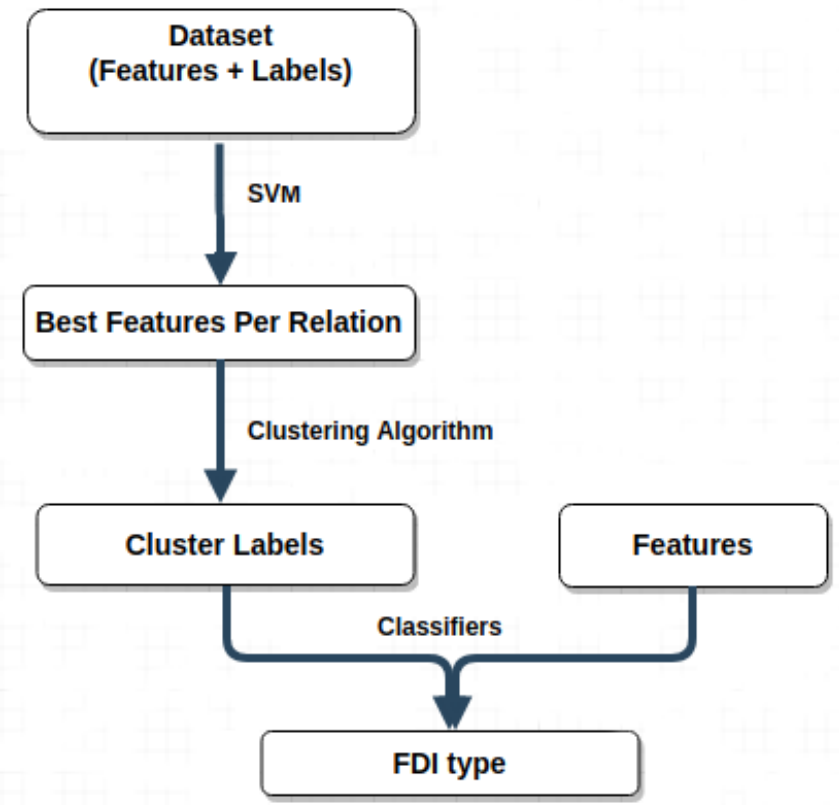
- Intuitive grouping
- Unsupervised clustering
- Drug-drug interaction Domain adaptation

Intuitive grouping (ARNP)

1. **Non-precised relation**
2. **No effect**
3. **Reduction**
   decrease absorption, *slow absorption*, *slow elimination*
4. **Augmentation**
   increase absorption, *speed up absorption*
5. **Negative**
   new side effect, *negative effect on drug*, *worsen drug effect*, *without food*, *negative effect on drug*, *worsen drug effect*, *side effect*, *without food*
6. **Positive**
   positive effect on drug, *improve drug effect*

Relation Clustering

- Relation representation method
- Clustering method on types of relation



Results

- New grouping scheme: (1) *decrease absorption*, *increase absorption*, (2) *improve drug effect*, *new side effect*, *worsen drug effect*, which refer to effect of drug, *speed up absorption*, *slow absorption*, *without food*, *positive effect on drug*, (3) *negative effect on drug*, (4) *no effect on drug*, (5) *slow elimination*
- Improvement on F1-score with 200 features: from 0.41 with ARNP and non-clustered data to 0.58
- Reduction of the impact of the imbalance of data: Difference of macro and micro F1 from 0.23 to 0.09

Domain adaptation - Drug-drug interaction

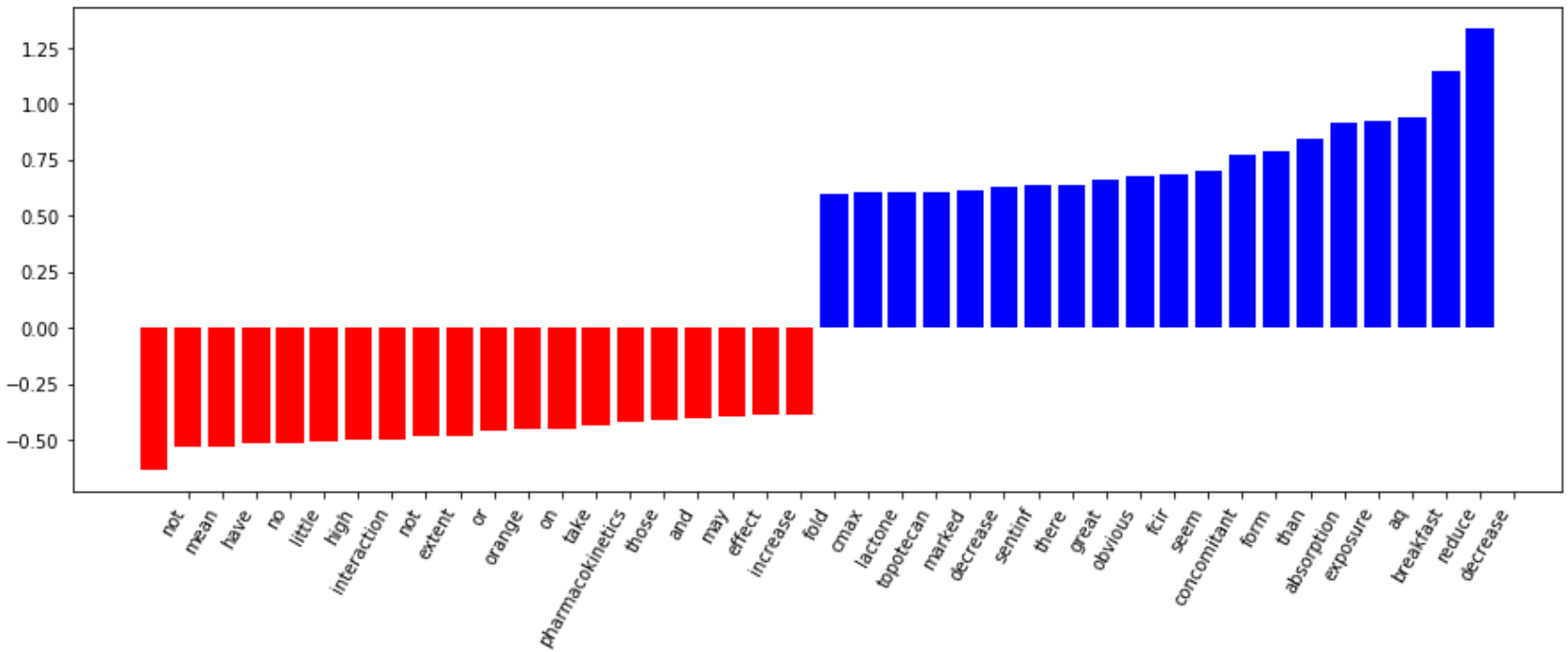
Correspondence of type DDI-FDI (Line 1) et percentage of FDI instances affected to type DDI (Lines 2-5)

Relation (Rel), *Decrease absorption (Dec)*, *No effect on drug (No)*, *Increase absorption (Inc)*, *Negative effect on drug (Neg)*,*Positive effect on drug (Pos)*, *New side effect (New)*, *Without food (Wout)*, *Improve drug effect (Imp)*, *Slow elimination (Sl-e)*, *Slow absorption (Sl-a)*, *Worsen drug effect (Wors)*, *Speed up absorption(Speed)*, *Advice (A)*, *Mecanism (M)*, *Effect (E)*, *Interaction (Int)*

FDI	Rel	Dec	No	Inc	Neg	Pos	New	Wout	Imp	Sl-e	Wors	Sl-a	Speed
DDI	M	M	M	M	E	E	E	A	E	M	E	M	M
Advice	7	4	10	8	12	24	0	**54**	17	0	0	0	0
Effect	**50**	7	31	13	**69**	**48**	**100**	23	**83**	13	**75**	0	0
Int	15	0	0	0	7	0	0	0	0	0	0	0	0
Mecha	28	**89**	**59**	**79**	11	29	0	23	0	**87**	25	**100**	**100**

$\Longrightarrow$ F1-score from 0.41 on the initial labels to 0.78 on the new labels

20 best and worst SVM features coefficient for Decrease absorption relation



Hamon (Thierry), Tabanou (Vincent), Mougin (Fleur), Grabar (Natalia) et Thiessard (Frantz). – POMELO: Medline corpus with manually annotated food-drug interactions. *In : Proceedings of Biomedical NLP Workshop associated with RANLP 2017*, pp. 73–80. – Varna, Bulgaria, September 2017.