A KERNEL-BASED APPROACH FOR IRONY AND SARCASM DETECTION IN ITALIAN

UNITOR @ evalita2018

Andrea Santilli, Danilo Croce, Roberto Basili

Università di Roma Tor Vergata

Turin 13 December 2018

INTRODUCTION

UNITOR approach for Irony and Sarcasm detection

- Task is modeled as a cascade of kernel-based SVM classifier

AIM: derive an effective task modeling

- Each kernel function is based on a specific set of irony/ sarcasm specific features

The adopted modeling allows to rank **first** in the sarcasm detection task

SOME ISSUES IS IRONY/SARCASM DETECTION AND (POSSIBLE) SOLUTIONS

📚 Lexical Sparseness (especially in short messages and reduced datasets)

😔 Negative evaluation towards a target

😐 Irony is often not explicit in (short) messages

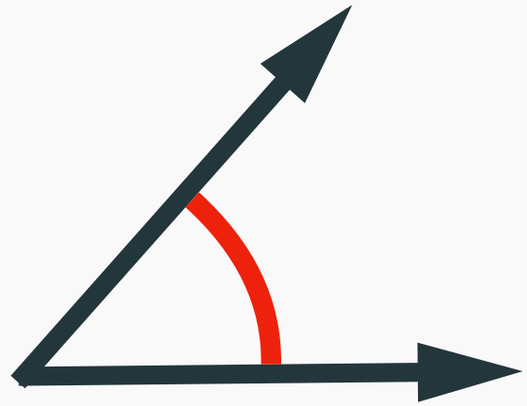
IRONY/SARCASM FEATURES

IRONY SPECIFIC FEATURES

Goal: define capture irony features

- Irony Specific BOW (*ISBOW*)

Cosine similarity between the word embedding vector of the considered word and mean topic centroid



“La mia macchina è veloce e scattante come un bradipo”

POLARITY BASED FEATURES

Goal: exploit negative evaluation towards a target

DPL: large scale Polarity Lexicon (~150K)

Buono: **0.7** 0.3 **0.0**

- Distributional Polarity Lexicon Sum (*DSUM*)

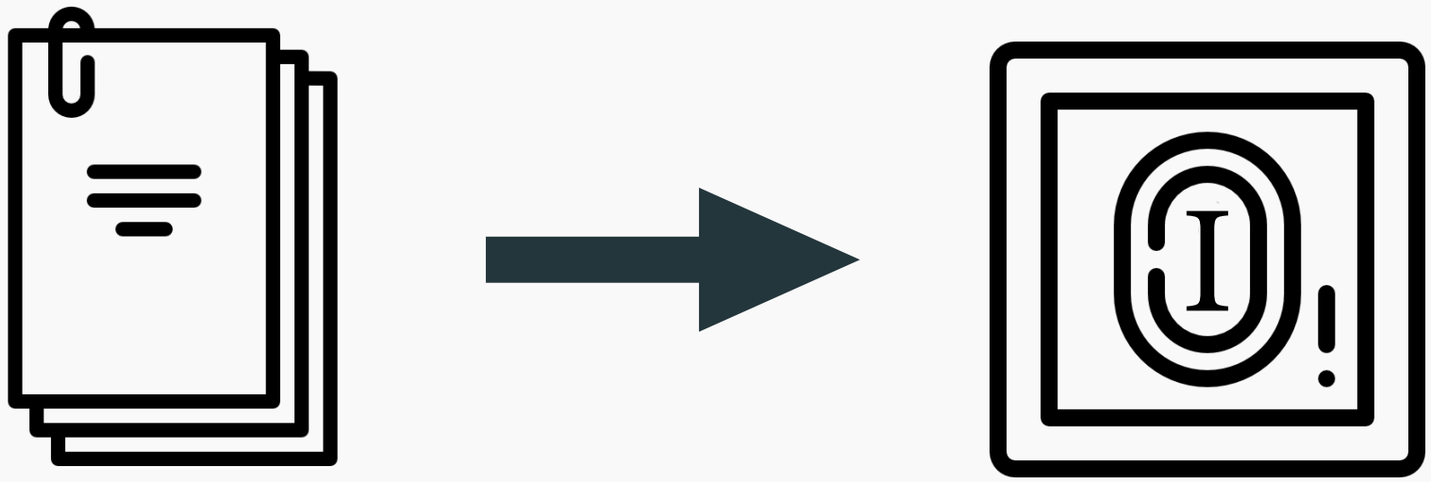
$$\sum_{w \in Noun} \frac{w^p}{|Noun|} \oplus \sum_{w \in Verb} \frac{w^p}{|Verb|} \oplus \sum_{w \in Adj} \frac{w^p}{|Adj|} \oplus \sum_{w \in Adv} \frac{w^p}{|Adv|} \oplus \sum_{w \in \tau} \frac{w^p}{|\tau|}$$

- Distributional Polarity Lexicon BoW (*DBOW*)

buono_pos**:** **0.7** buono_neu**:** **0.3** buono_neg**:** **0.0**

IRONY CORPUS FEATURES

Goal: define an Irony specific Lexicon and Irony patterns



- Irony Corpus BOW (ICBOW)
- Irony Corpus weighted BOW (ICwBOW)
- Irony Corpus weighted Mean (ICM)

EXPERIMENTS AND RESULTS

TRAINING

Training Set		
Ironic	Not-Ironic	Total
2023	1954	3977

10-fold cross validation

Train effettuato sullo split di:

- **Irony:** All dataset (3977)
- **Sarcasm:** Ironic tweets (2023)

TRAINING

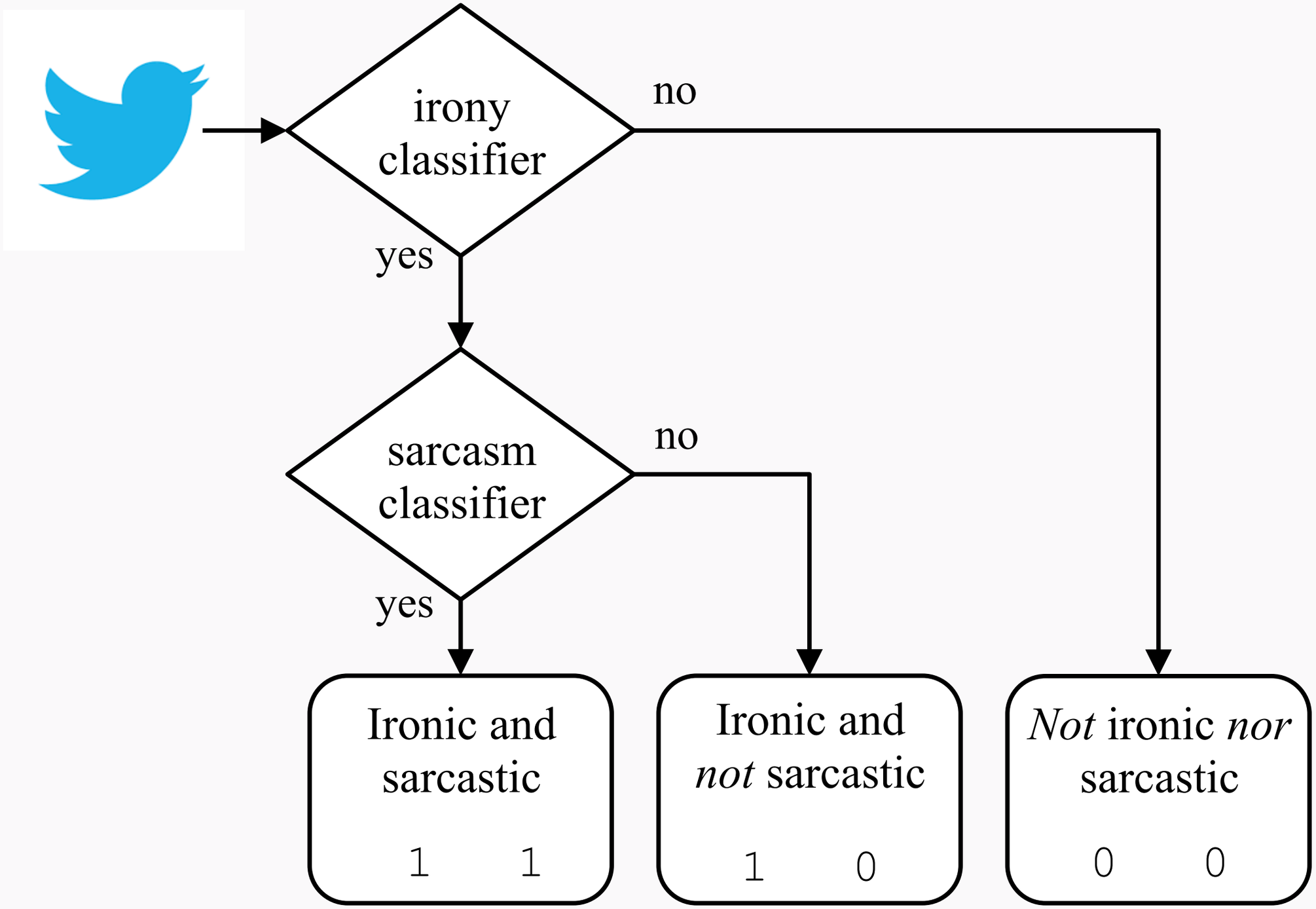
Each feature modeling generate a separate feature space

We also added some “standard” feature modeling

- Bag-of-Words
- Bag-of-Word Ngrams
- Bag-of-Character Ngrams
- Some syntethic feature (capitalized letters, elongations, …)

The resulting system is implemented as two binary classifiers in cascade

CASCADE SVM


irony classifier
no
yes
sarcasm classifier
no
yes
Ironic and sarcastic
1 1
Ironic and *not* sarcastic
1 0
Not ironic *nor* sarcastic
0 0

Irony Detection

	Not Ironic			Ironic			Mean
	P	R	F1	P	R	F1	F1
1st	.785	.643	.707	.696	.823	.754	**.731**
2nd*	.771	.617	.686	.680	.816	.741	.714
6th(u)	.778	.577	.662	.662	.834	.739	.700
7th(c)	.764	.593	.668	.666	.816	.733	.700
BL	.501	1.00	.668	1.00	.000	.000	.334

Sarcasm Detection

	Not Sarcastic			Sarcastic			Mean
	P	R	F1	P	R	F1	F1
1st(c)	.362	.584	.447	.492	.407	.446	**.520**
2nd(u)	.355	.553	.432	.469	.449	.459	.518
4th*	.344	.566	.428	.344	.566	.428	.508
BL	.296	.132	.183	1.00	.000	.000	.199

CONCLUSION

- 😎 UNITOR ranked **first** in the sarcasm detection task
- 😎 The proposed irony and sarcasm specific features, especially **Polarity Based Features**, were beneficial to detect irony and sarcasm
- 🤨 Non trivial for human reader, lot of information *out of context*

Thanks for your attention

Code available: https://github.com/andry9454/ironySarcasmDetection