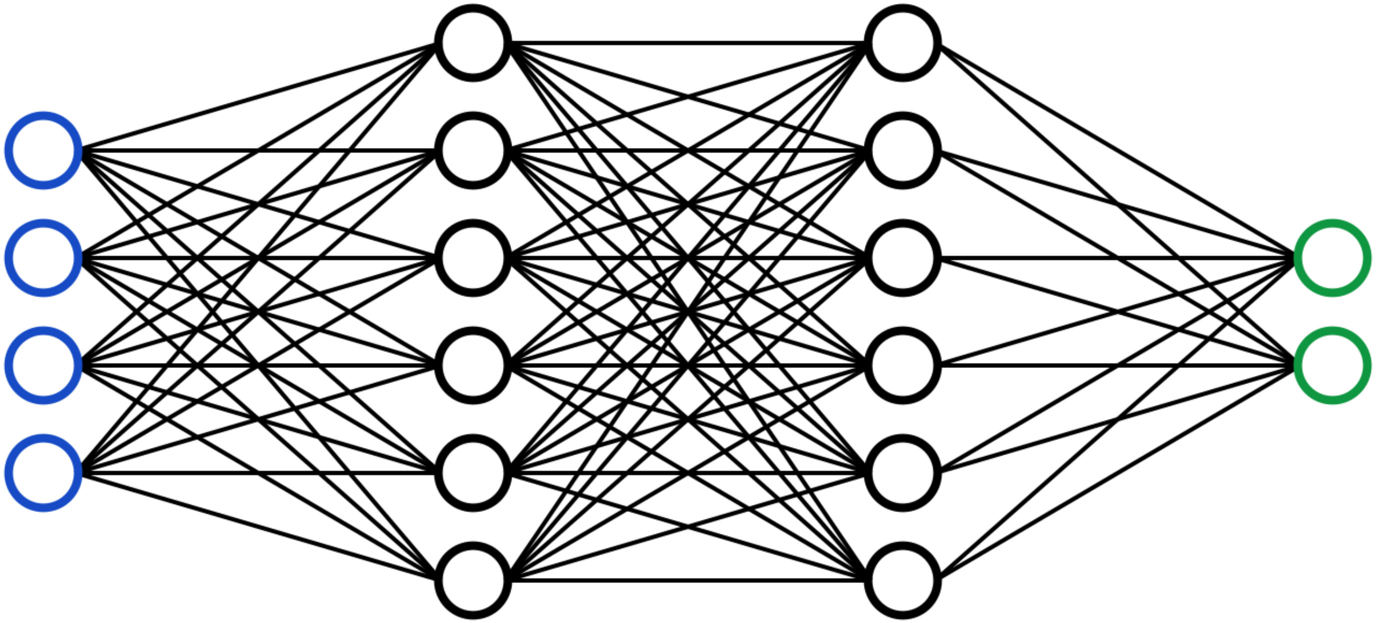


Detecting Sarcasm and Irony in Texts: An Overview

Denis Reibel | 07.03.2022

Table of Content

Research Questions

(1) What are the main problems when detecting sarcasm or irony in texts?

(2) Which groups of methods can be identified from previous works?

(3) What are good baseline algorithms for sarcasm detection? And how do more complex models perform in comparison to those baselines?

Language Theory – Difference between irony & sarcasm

Many different definitions, Scientific dispute

➡ No clear borders

➡ Both mean opposite of what they describe

➡ Humans do not share mutual understanding of the concept

Treated as the same concepts in this work.

How Humans comprehend sarcasm?

- Context Information
- Gesture / Emphasis / Mimic
- Textual Markers: #sarcasm, /s

Problems in sarcasm detection

1) **Context Information**
   - Example: Knowing a person is a comedian, background knowledge
   - Algorithms missing that Information

2) **Interrater Agreement**
   - No common understanding about sarcasm/irony
   - Humans label data very differently

3) **Growing language / fast-paced changes [1]**
   - Language changes fast-paced (especially on social media)
   - Many new (slang-) words created regularly

[1] T. Jain, N. Agrawal, G. Goyal and N. Aggrawal, "Sarcasm detection of tweets: A comparative study“, 2017

Methods

6 Groups of Methods identified

baselines

- Simple Algorithms

- Good for comparison

[2]: Decision Tree, Naive Bayes on linguistic features

[3]: semantic markers (heavy punctuation, special emoticons, 'lol')

[2] Reyes, A., Rosso, P. & Veale, T., "A multidimensional approach for detecting irony in Twitter", 2013

[3] Paula Carvalho, Luís Sarmento, Mário J. Silva, and Eugénio de Oliveira, „Clues for detecting irony in user-generated contents: oh...!! it‘s „so easy" ;-)“, 2009

Methods

Pattern-based

Key Idea:

Detecting typical sarcastic word-patterns in the data
(often depending of POS-Tagging)

"I love it when I have to stand up early."

"I love cats when they scratch me."

↓

Sarcastic Pattern: "… I love NOUN when …"
(based on [4])

[2] Mondher Bouazizi and Tomoaki Otsuki Ohtsuki „A Pattern-Based Approach for Sarcasm Detection on Twitter", 2016

Methods

Rule-based

Key Idea:

Researchers try to define logical rules that describe sarcasm

rule	**tweet**	**final sentiment**
If there is a single hashtag denoting sarcasm, we flip the sentiment to negative.	It’s not like I wanted to eat breakfast anyway. #sarcasm	negative
If there is more than one hashtag, we look at any sentiment contained in those hashtags.	The best feeling in the world is being ignored. #bestthingever #no	negative

(based on [5])

[5] Diana Maynard & M.A. Greenwood, “Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis”, 2014

Methods

Learning-based (also: Statistic-based)

Key Idea:

Algorithms *learn* special features or distributions from the given input data.

Typical Examples:

Naïve Bayes, Support Vector Machines (SVM) [6], k-Nearest Neighbours (kNN) [7]

[6] Debanjan Ghosh, Weiwei Guo, and Smaranda Muresan, "Sarcastic or Not: Word Embeddings to Predict the Literal or Sarcastic Meaning of Words", 2015

[7] Oren Tsur, Dmitry Davidov, and Ari Rappoport, „A Great Catchy Name: Semi-Supervised Recognition of Sarcastic Sentences in Online Product Reviews", 2010

Methods

Deep learning methods

Key Idea:

Neural networks automatically detect sarcastic features on training data. Trained network can be transferred on other (evaluation) data.

Examples:

[8] CNN + Author Information

[9] RoBERTa-Architecture

[8] Silvio Amir, Byron C. Wallace, Hao Lyu, Paula Carvalho, and Mário J. Silva, “Modelling Context with User Embeddings for Sarcasm Detection in Social Media”, 2016

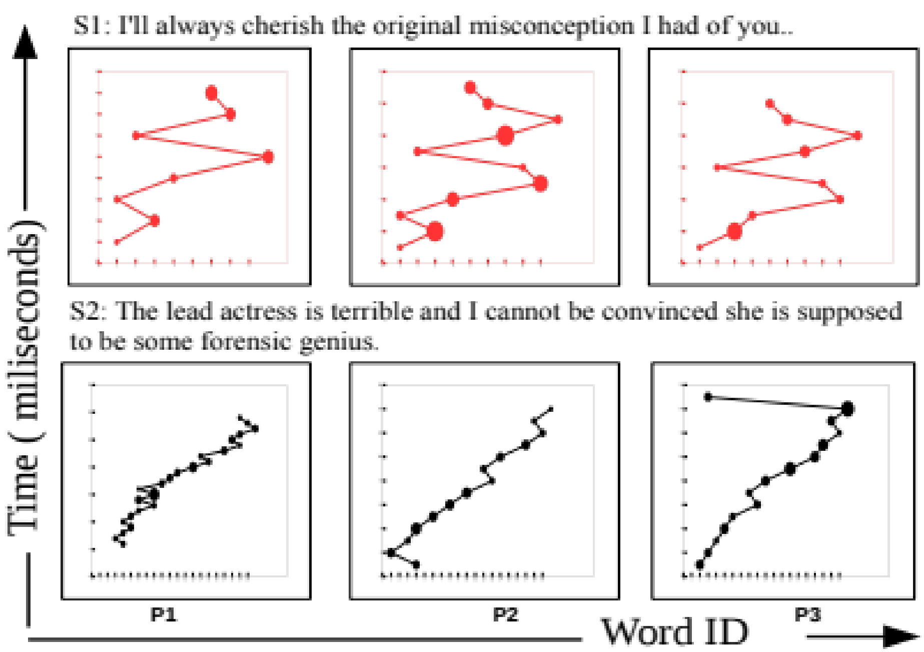
[9] Saurav Pradha, Malka N. Halgamuge, and Nguyen Tran Quoc Vinh, „Effective Text Data Preprocessing Technique for Sentiment Analysis in Social Media Data. 2019

Methods

Out of the box methods

Methods that don’t follow the usual approaches or a used in slightly different use-cases but are still worth mentioning.

Example: Eye-Patterns [10] as additional information



sarcastic sentence

non-sarcastic sentence

[10] Abhijit Mishra, Diptesh Kanojia, Seema Nagar, Kuntal Dey, „Harnessing Cognitive Features for Sarcasm Detection“ 2016

Comparison

		Accuracy	Precision	Recall	F1-Measure	Data
baseline methods	[26]	-	0,780	0,740	0,760	Tweets (annotated by hashtags)
	[6]	-	0,611	-	-	user-comments in online-newspapers
pattern-based methods	[5]	0,831	0,911	0,734	0,811	Tweets (annotated by hashtags)
rule-based methods	[20]	-	0,910	0,910	0,910	Tweets (annotated manually)
	[1]	-	0,970	0,920	0,950	Tweets (annotated manually)
learning-based methods	[29]	-	0,766	0,813	0,788	Amazon Reviews (semi-supervised, annotated manually)
	[10]	-	0,870	0,856	0,863	Tweets (annotated by hashtags)
deep learning methods	[3]	0,872	-	-	-	Tweets (pre-existent datasets)
	[22]	0,910	0,900	0,900	0,900	4 pre-existent datasets

Table 3: Reported performances of previous explained methods

Values can not be directly compared because they are based on different data !

Discussion

1) Human labelling vs. Hashtag labelling
   - Human labelling is error prone and costly
   - Hashtag labelling also has many errors but is time-efficient

2) Role of Context
   - Context is not always available
   - Difficult to implement into the model
   - Can boost performance

Conclusion

- Difficult field with lots of non-trivial problems
    - No mutual understanding of concepts, labelling problems, …
- Many different methods to solve the Problem
    - No right or wrong, it depends on the use-case
- Current trend goes towards state-of-the-art deep learning techniques
    - Neural Networks, BERT-architectures or Transformers

THANK YOU!

Any further questions?

denis.reibel@uni-ulm.de