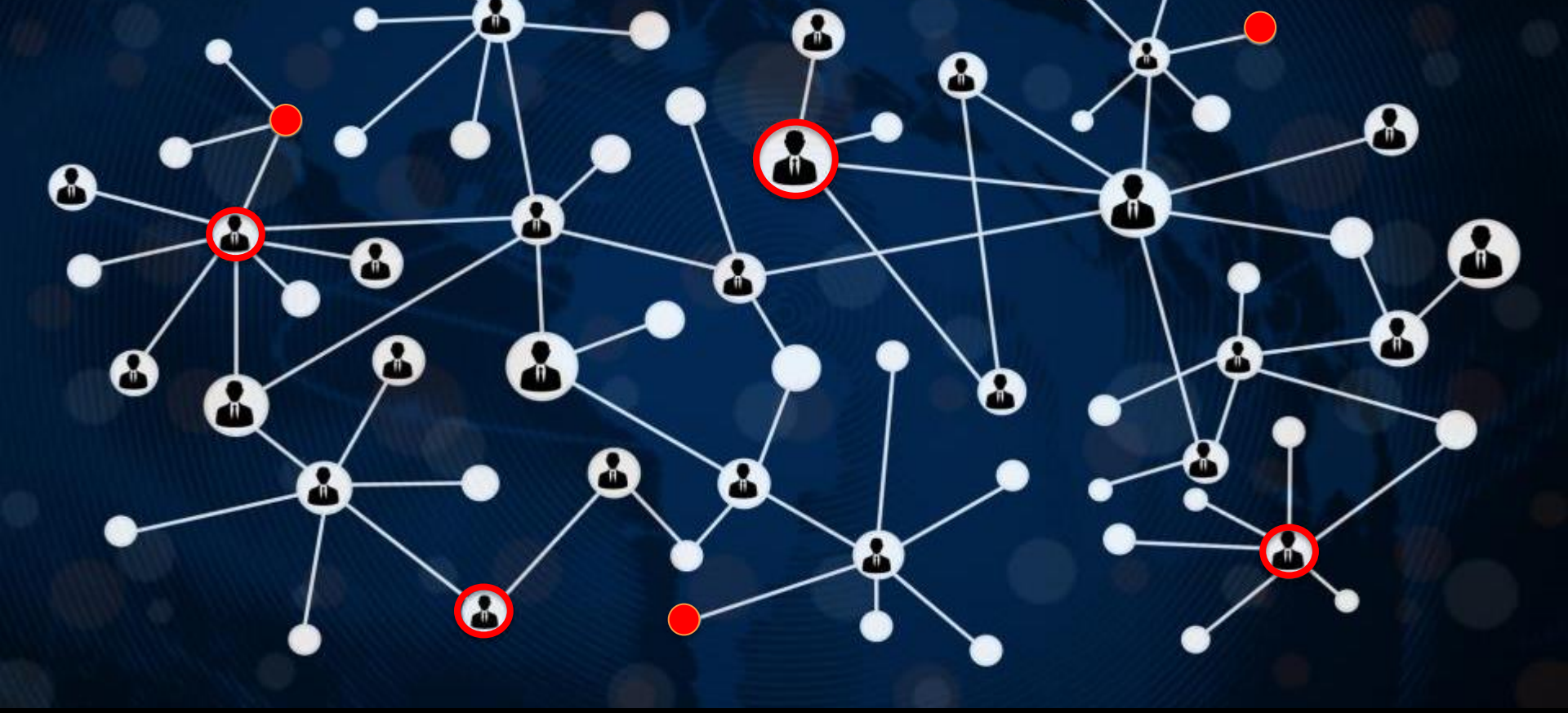


Adopting MaxEnt to Identification of Bullying Incidents in Social Networks

Maral Dadvar
Web-based Information Systems and Services, Stuttgart Media University
University Library Johann Christian Senckenberg

Outline

- What is cyberbullying?
- Challenges and shortages
- Existing approaches
- Dataset
- Maximum Entropy
- What is next?

What is Cyberbullying?

Cyberbullying is

- "An aggressive, intentional act
- carried out by a group or individual
- using electronic forms of contact
- repeatedly or over time
- against a victim who cannot easily defend herself."

Espelage et al. 2003

What is Cyberbullying?

- **A global problem**
  - USA 50% (10-20 % repeatedly)
  - European Union 18%
  - Japan: 10%

- **Depression, low self-esteem, low functionality, cases of suicide.**

- **Cyberbullying vs. physical bullying:**
  - Spread fast
  - Wider range of audience
  - Persistence and durable
  - It happens everywhere

Challenges

- Lack of sufficient and appropriate dataset
- Privacy issues
- Demographic information
- Labeling process

Dadvar et al. 2014

Existing Approaches

- **Statistical Regressions and Machine learning Classificatio**
  - Supervised learning approaches
  - Sentiment and content based features

    Example: profanity , personal pronouns

  Shortcomings
  - Identifying the bullying comments and not the bully users
  - Foul words are used among teenagers as a sign of friendship and close relationships.
  - Being bullied and becoming a victim of cyberbullying is also dependen on the personality of each person.

Our Earlier Contributions

- **Hybrid Approach**
  - Integrating social studies' findings into technical solutions
  - Expert systems
  - Users' features and Activity features
  - Identification of potential bully users

- **Cyberbullying Dataset**

Dadvar et al. 2012, 2013, 2014

DOI: 10.13140/RG.2.1.3083.4649

YouTube NL

Upload

maral dadvar
Watch Later
Watch History
Playlists

What to watch
My subscriptions
Social

SUBSCRIPTIONS
Red Hot Chili Peppers

M

Subscribe 110

Videos Discussion About

2 others liked 6 months ago

A Disney Land Birthday Video
by **Markus Aurelius David DeMott** • 6 months ago • 1,157 views
A Disney Land Birthday Video for Jessica ^_^

110 subscribers
433,574 views

subscribed to a channel 11 months ago

leokimvideo
344 videos
I'm just a guy making kids happy by presenting them with my views about today's toys. I cover Thomas And Friends, The Trash Pack, Disney Cars, Surprise Eggs, Chuggington, Angry Birds, Pez Candy Toys, Club P…
CHANNEL Subscribe 134,079

Joined Jan 20, 2011

s commented and liked 1 year ago
Incredible!!!

Sleepy Man Banjo Boys then with Rob Dyrdek Kickflips Chevy Sonic ove World's Largest Skateboard
by **Markus Aurelius David DeMott** • 1 year ago • 771 views
This is Truly Awesome...(-:AWE INSPIRING:-) with a dash of (-:How the Heck Did They Do That:-)

YouTube
Broadcast Yourself™

Dataset

- Top 3 videos of the YouTube video categories
- 4 months of users' activities (April – June 2012)
  - Network Activities:
    - Posting comments, Subscribing, Uploading, …
  - Profile information
    - Age, Signup date, …

- 54,050 comments from 3825 users (on average 15 comments per user)
- Manually labelled as bullying and non-bullying
- Inter-annotator agreement = 93%
- 12% Bullying comments, 88% Non- bullying comments

Problem Not Solved

YouTube

Maximum Entropy

- Statistical learning method
- Developed and used in other fields
- Mainly used for classification with imbalances datasets
- Estimates the multivariate distribution of incidents (here the bully users) in feature-space according to the principle of maximum entropy.
  - The best approximation of an unknown distribution is the one with maximum entropy (the most spread out) subject to known constraints. The constraints are defined by the expected value of the distribution, which is estimated from a set of incidents.

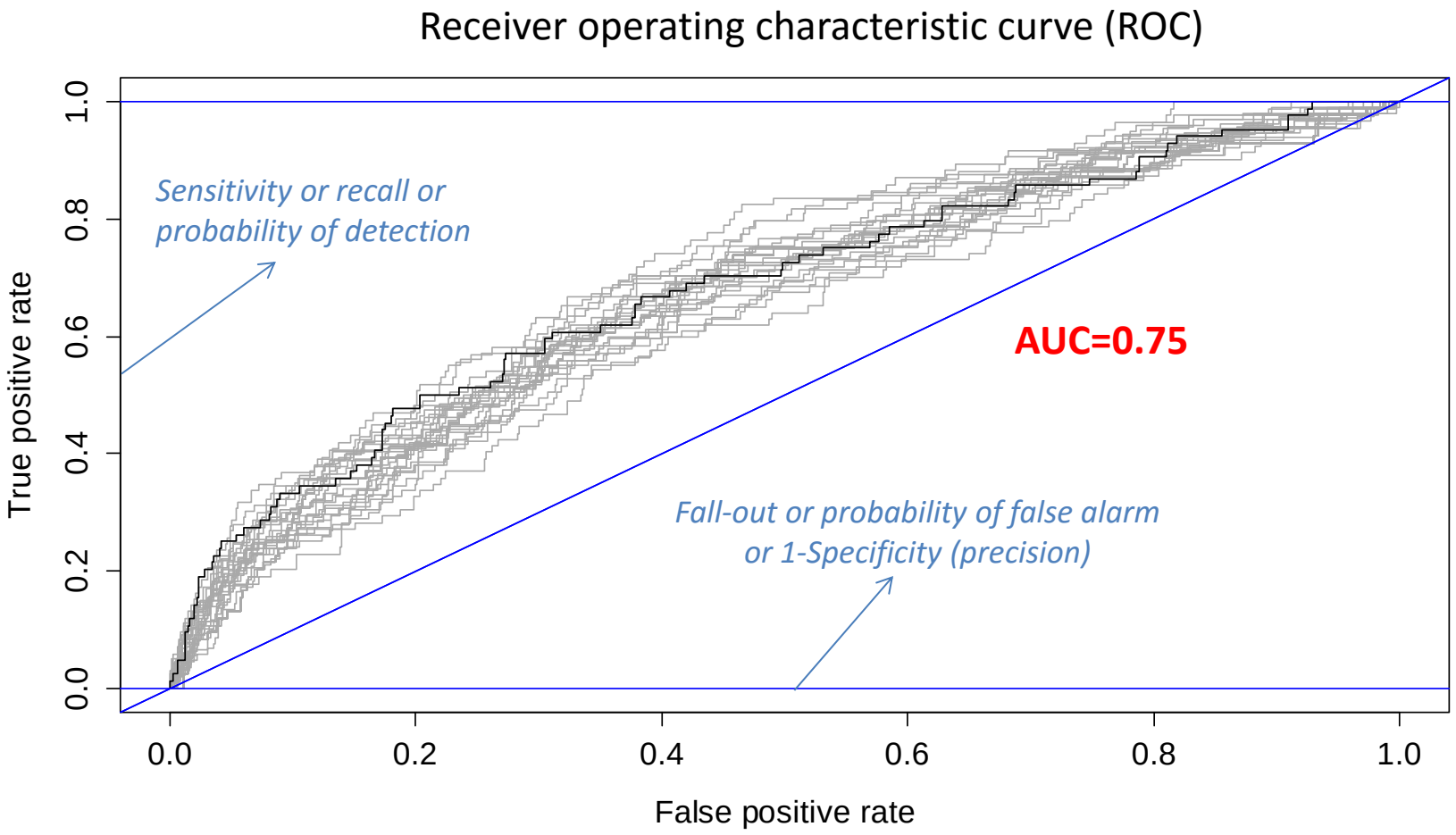
Maximum Entropy

- Maxent software package (version 3.3.3) for species distribution and environmental niche modelling
  - requires incident data, often called incident-only data plus features for the whole datasets.
  - It is very robust to limited amount of training data (i.e. small sample size)
  - results are amenable to interpretation of the form of the feature response functions.

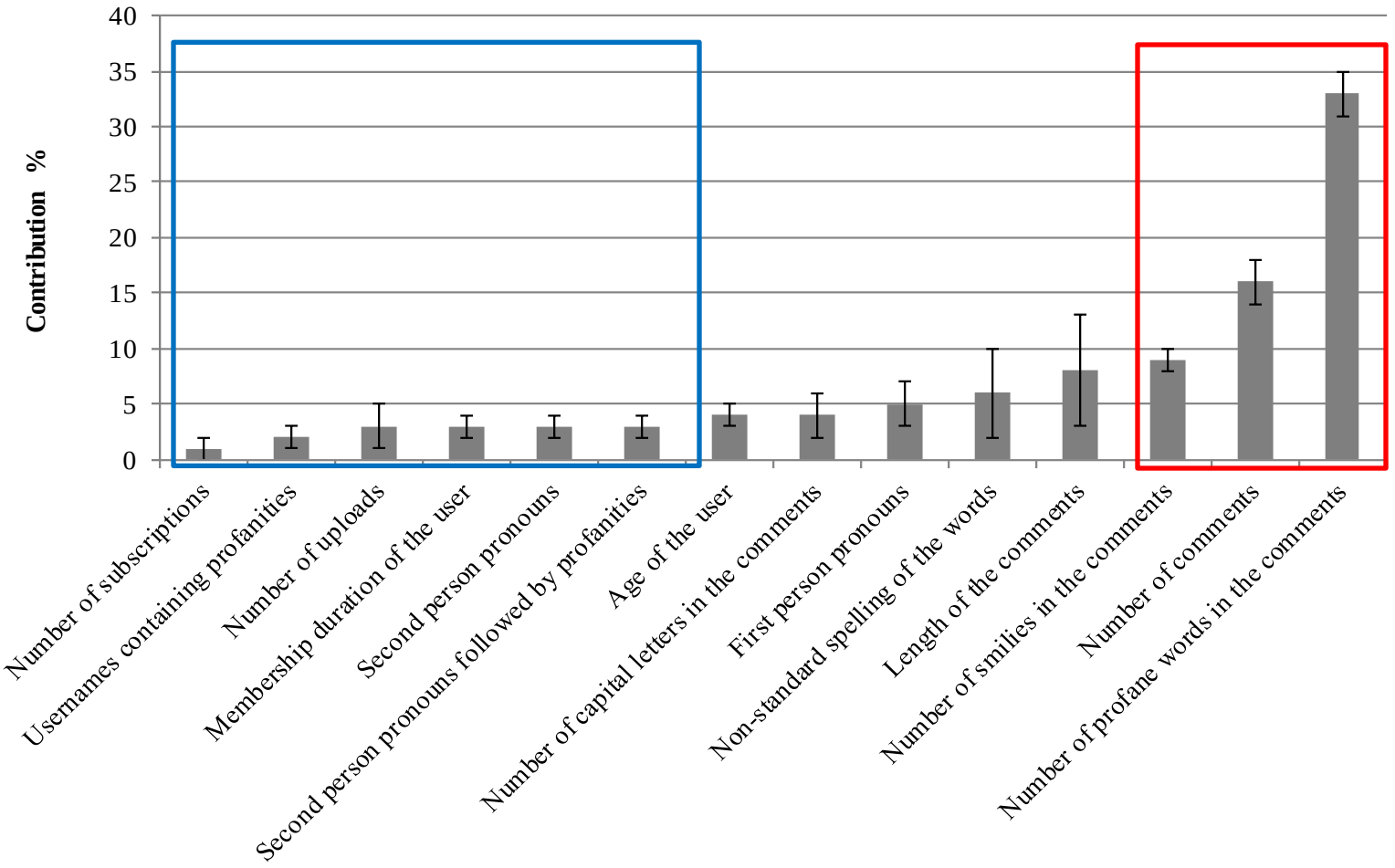
Features Space

Feature Set	Feature Name
User features	Age
	Membership Duration
Content Features	Length of the comments
	Profane words in the username
	Profanities and bullying sensitive topics
	Second person pronouns
	First person pronouns
	Non-standard spellings
Activity Features	Number of uploads
	Number of subscriptions
	Number of posted comments

Results


Receiver operating characteristic curve (ROC)
True positive rate
False positive rate
0.0
0.2
0.4
0.6
0.8
1.0
Sensitivity or recall or
probability of detection
AUC=0.75
Fall-out or probability of false alarm
or 1-Specificity (precision)

Results


Contribution %
40
35
30
25
20
15
10
5
0
Number of subscriptions
Usernames containing profanities
Number of uploads
Membership duration of the user
Second person pronouns
Second person pronouns followed by profanities
Age of the user
Number of capital letters in the comments
First person pronouns
Non-standard spelling of the words
Length of the comments
Number of smilies in the comments
Number of comments
Number of profane words in the comments

Results

Model	AUC	AUC.sd
Maximum Entropy (MaxEnt)	0.75	0.032
Generalized Linear Model (GLM)	0.64	0.034
Random Forests (RF)	0.69	0.032
Support Vector Machine (SVM)	0.59	0.032

Conclusion

- MaxEnt can be a recommended for cases with imbalanced datasets or rare number of target incidents.

  For example, cyberterrorism detection, Identification of online sexual predators

- Language independent and adaptable to other social networks

Next

- **Spatial Features**

  Geographical location of the users

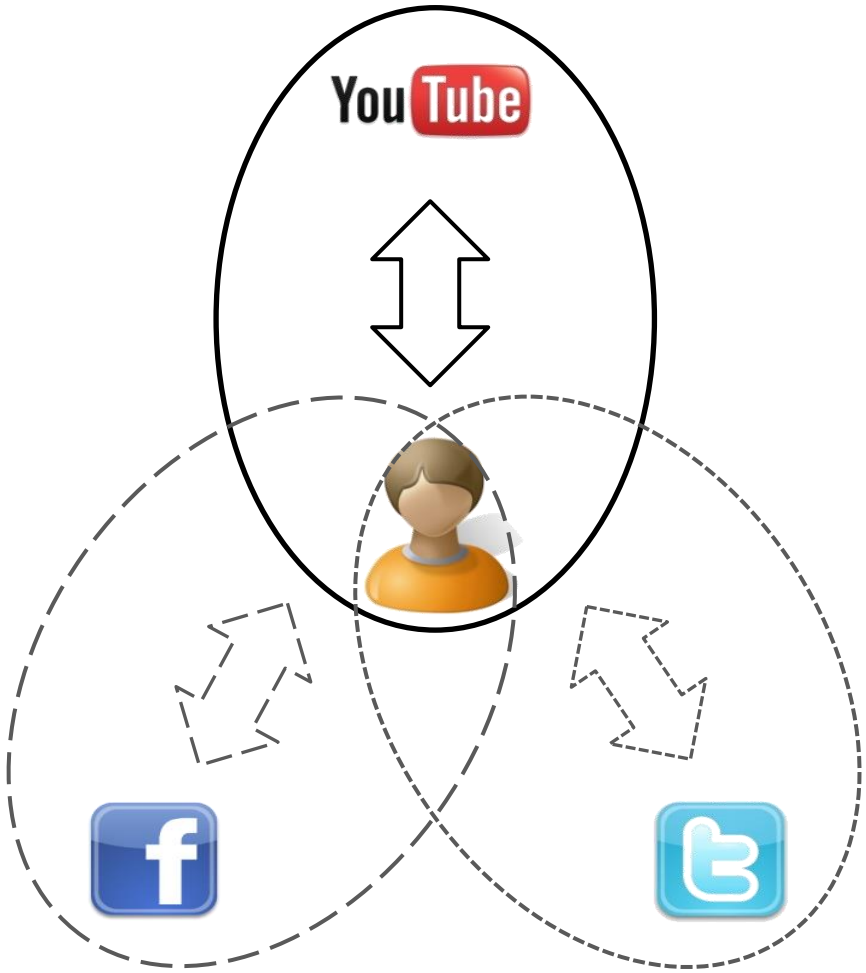
- **Temporal features**

  Time trend of users online activities

- **Victim identification**
- **Public effect**
- **Post-bullying behaviour**

Next

- **Cross-systems analysis**
- 6.2 % link to all
- 42.8% link to at least one of



Thank you

m.dadvar@hdtm.stuttgart.de
aidin.niamir@senckenberg.de

HOCHSCHULE DER MEDIEN

JOHANN WOLFGANG GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN