Opinion Mining and Sentiment Analysis Part 2

Isabelle Augenstein
augenstein@di.ku.dk
@IAugenstein
http://isabelleaugenstein.github.io/

Web Science Lecture
19 March 2019

UNIVERSITY OF COPENHAGEN

Recap: Tasks

Sentiment Analysis

- Assess the sentiment, i.e. feeling expressed in a text automatically
- Detect words / phrases in texts expressing such feelings
- *"I am concerned the imposed austerity measures will hurt the economy in the long term"*

Opinion Mining

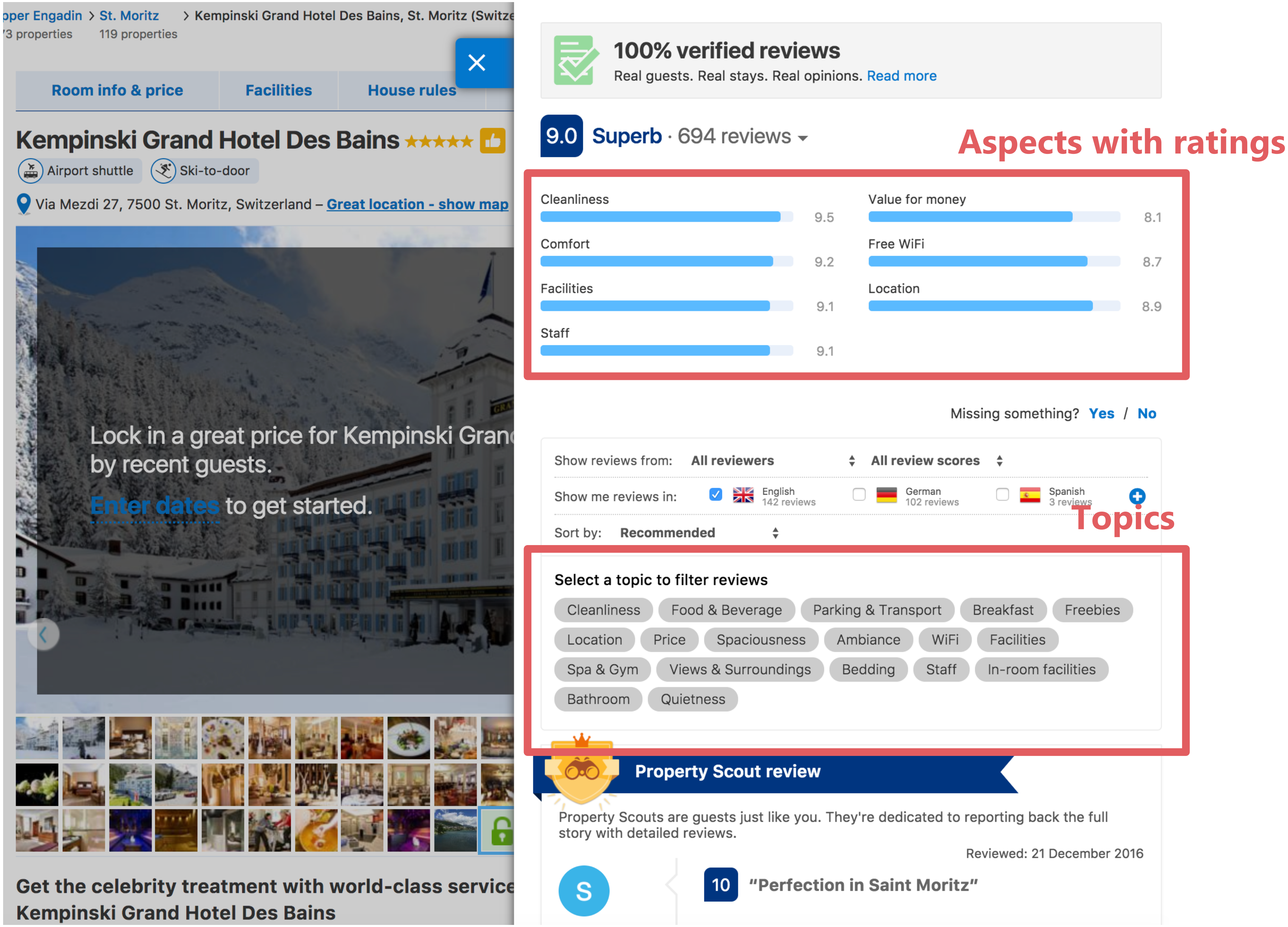
- Extract opinions about entities from texts
- *"I wasn't satisfied with the location of the hotel"*, *"A one-night stay was surprisingly cheap"* -> location: negative; price: positive

- In practice, the two terms are used interchangeably

Lecture Overview

- Tasks
  - Target-Based Sentiment Analysis
  - Aspect-Based Sentiment Analysis
  - Stance Detection
  - Rumour Detection
  - Affect Extraction
  - Emojis and Emotion
  - Sentiment and Personality
- Approaches

Applications



Room info & price
Facilities
House rules
Kempinski Grand Hotel Des Bains
Airport shuttle
Ski-to-door
Via Mezdi 27, 7500 St. Moritz, Switzerland – Great location - show map
Get the celebrity treatment with world-class service
Kempinski Grand Hotel Des Bains
100% verified reviews
Real guests. Real stays. Real opinions. Read more
9.0 Superb · 694 reviews
Aspects with ratings
Cleanliness 9.5
Comfort 9.2
Facilities 9.1
Staff 9.1
Value for money 8.1
Free WiFi 8.7
Location 8.9
Missing something? Yes / No
Show reviews from: All reviewers
All review scores
Show me reviews in: English 142 reviews German 102 reviews Spanish 3 reviews
Sort by: Recommended
Topics
Select a topic to filter reviews
Cleanliness
Food & Beverage
Parking & Transport
Breakfast
Freebies
Location
Price
Spaciousness
Ambiance
WiFi
Facilities
Spa & Gym
Views & Surroundings
Bedding
Staff
In-room facilities
Bathroom
Quietness
Property Scout review
Property Scouts are guests just like you. They're dedicated to reporting back the full story with detailed reviews.
Reviewed: 21 December 2016
10 "Perfection in Saint Moritz"

Select a topic to filter reviews

Cleanliness | Food & Beverage | Parking & Transport | Breakfast | Freebies

Location | Price | Spaciousness | Ambiance | WiFi | Facilities

Spa & Gym | Views & Surroundings | Bedding | Staff | In-room facilities

Bathroom | Quietness

Reviewed: 18 December 2017

S

Sandhya75

Switzerland

Age group: 35 – 44

7 reviews

1 helpful vote

10 **"Perfect stay"**

The rooms are rather small and I was missing a tea/coffee maker . But that is all I could mention

The location is perfect for skiing activities. Cross country im front of the door and downhill across the street. And you can rent the skis inhouse!

Breakfast is very broad and the included a la carte choices just makes breakfast a feast.

The staff is attentive and very friendly.

Stayed in December 2017

Helpful

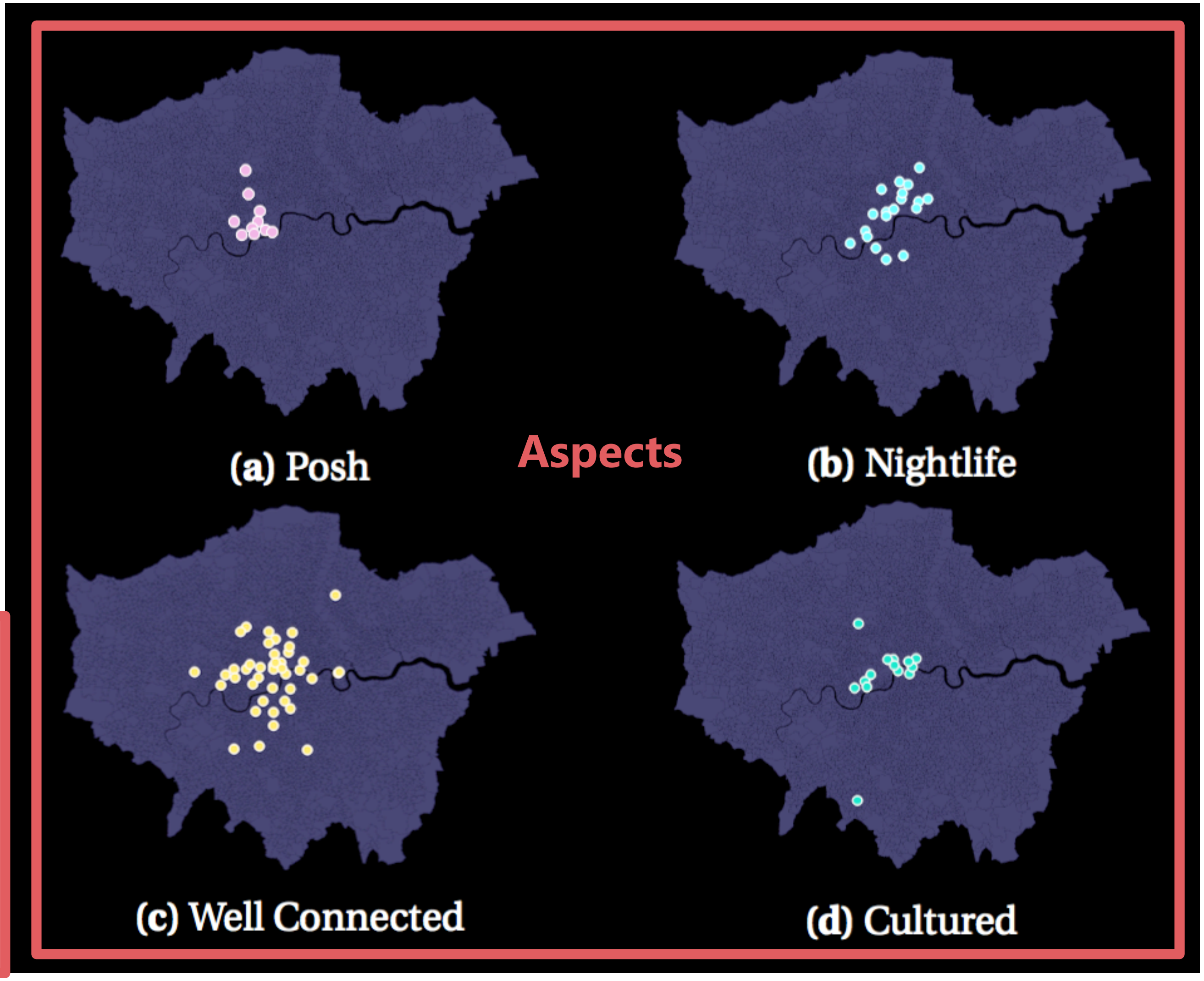
Neighborhoods

Saeidi et al., 2017

Q: What area of London should i live in?

A: Cool areas to live in at the moment are: / Clapham / Balham / Battersea / Hoxton / Camden

Targets



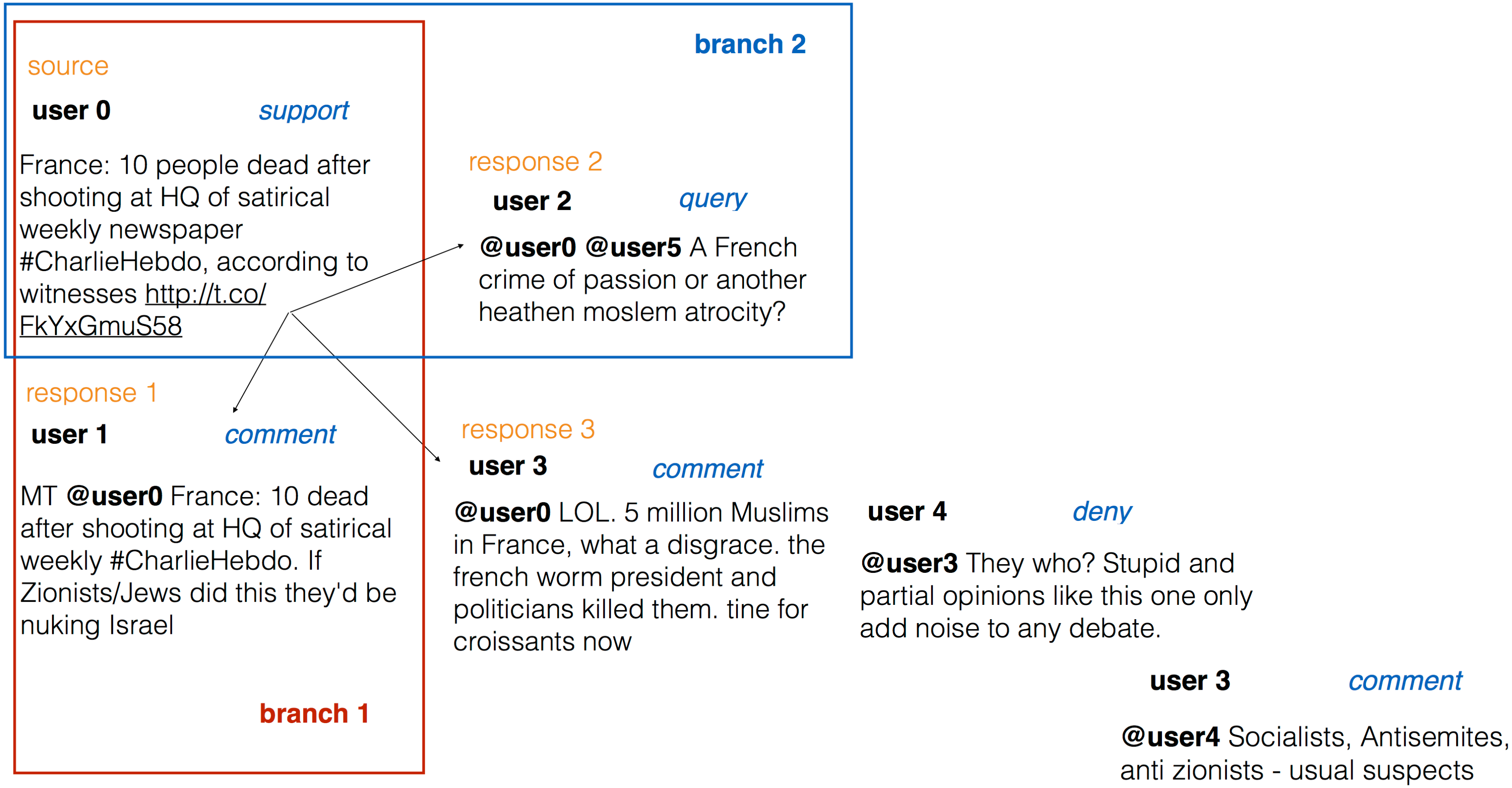
Understanding stance towards political entities

No more #NastyWomen or #BadHombres

Task: Is tweet positive, negative or neutral towards a given target (Donald Trump)?

Augenstein et al. (2017)

Tracking Rumours on Social Media



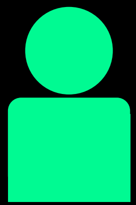
Kochkina, Liakata, **Augenstein** (2017)

Zubiaga, Kochkina, Liakata, Procter, Lukasik, Bontcheva, Cohn, **Augenstein** (2017)

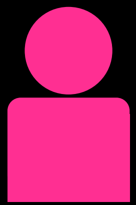
Approaches & Tasks

Perspective

The new **iPhone** is way better than the **Samsung Galaxy**!



Apple



Samsung

TARGET-DEPENDENT SA

This is a great area for **living**, but a bit boring to **go out**.

ASPECT-BASED SA

Aspect-based Sentiment

This is a great area for **living**, but a bit boring to **go out**.

Aspect detection: → LIVE → NIGHTLIFE

Sentiment classification: positive negative

SentiHood: Target-Based Aspect-Based Sentiment Analysis

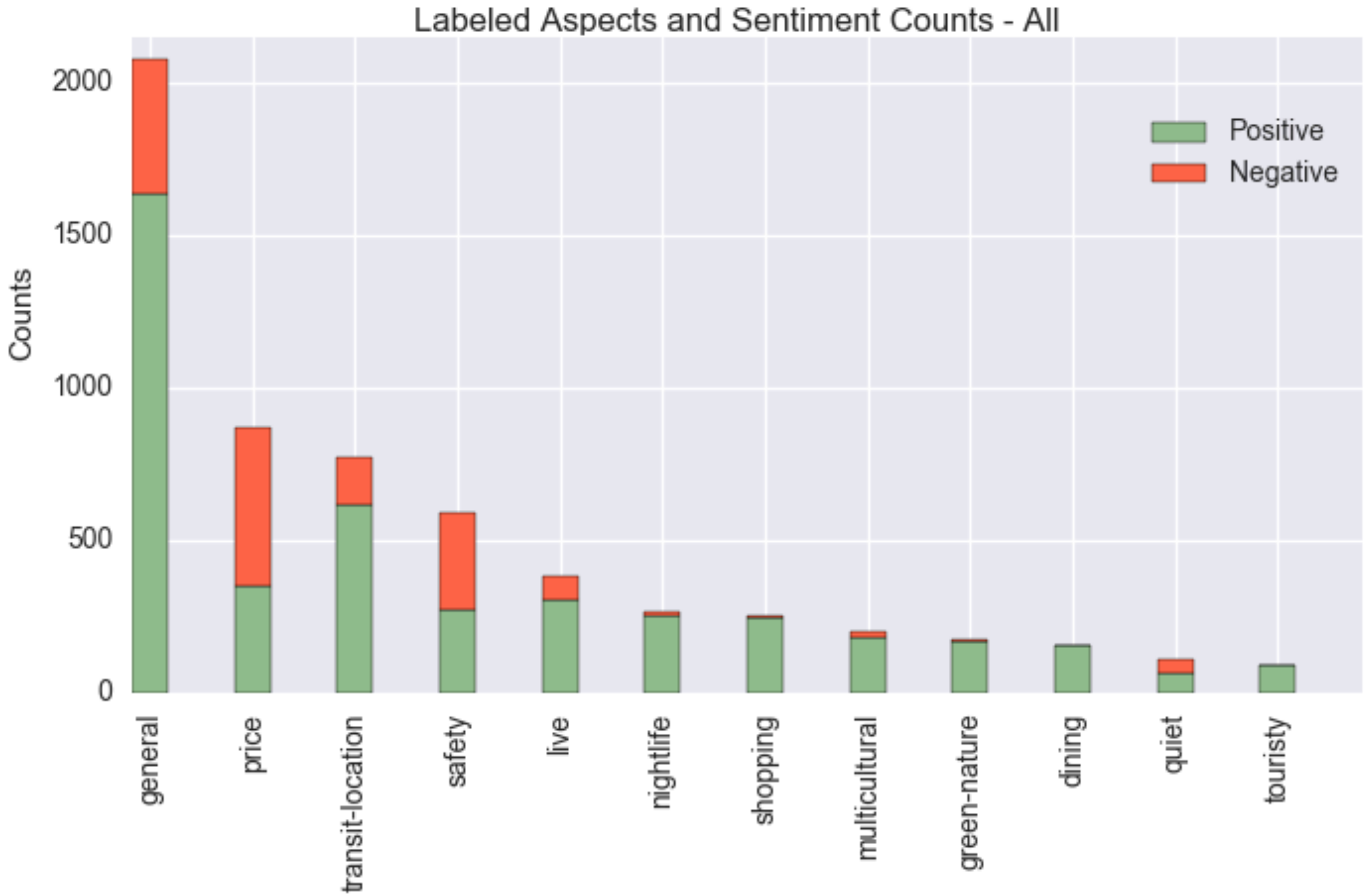
Sentence	Labels
The cheap parts of London are **Edmonton** and **Tottenham** and they are all poor, crime ridden and crowded with immigrants	(Edmonton,price,Positive) (Tottenham,price,Positive) (Edmonton,safety,Negative) (Tottenham,safety,Negative)
Hampstead area, more expensive but a better quality of living than in **Tufnell Park**	(Hampstead,price,Negative) (Hampstead,live,Positive)

Named Entity Masking

Sentence	Labels
location1 is very safe and **location2** is too far	(location1,safety,Positive) (location1,transit-location,None) (location2,safety,None) (location2,transit-location,Negative)

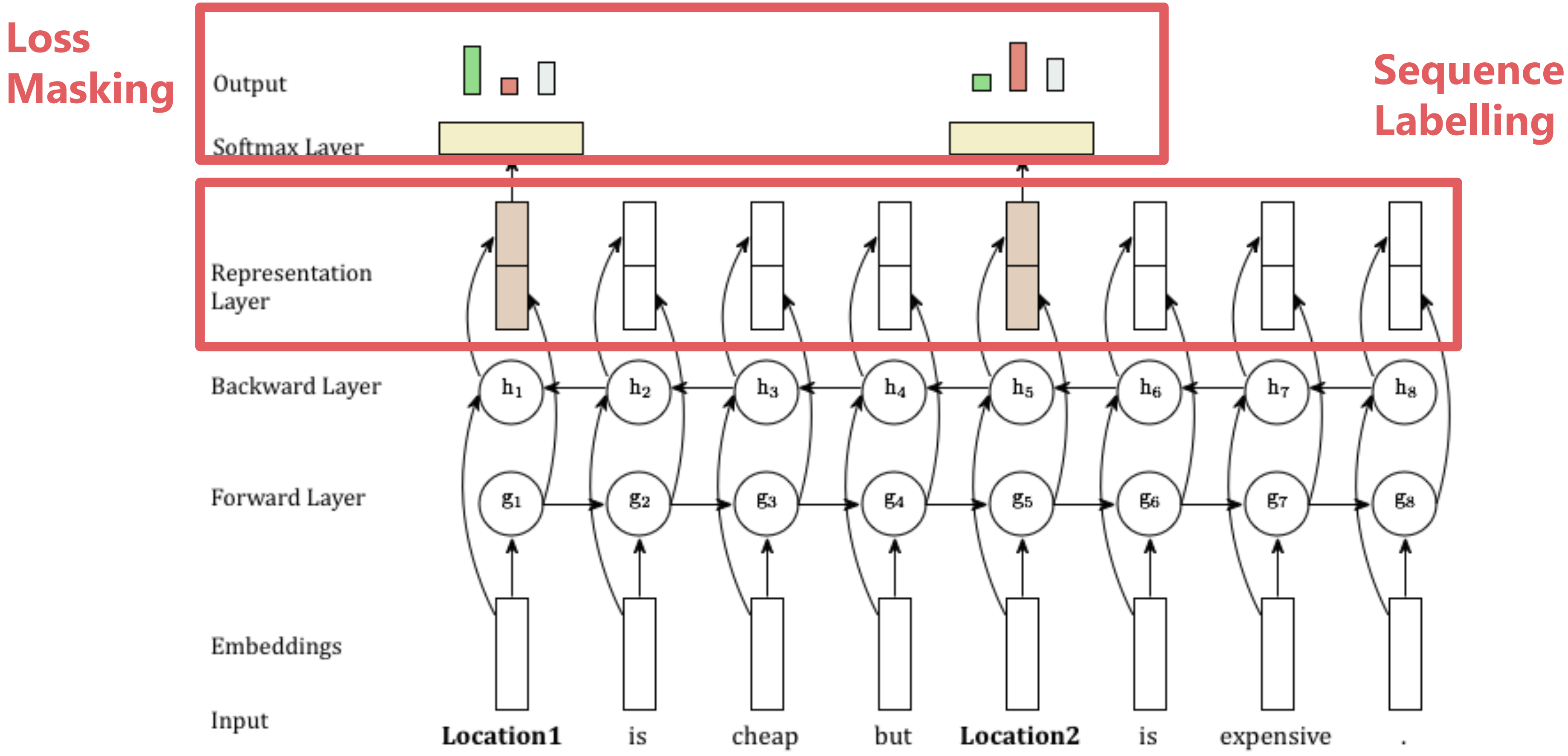
Saeidi, Marzieh; Bouchard, Guillaume; Liakata, Maria; Riedel, Sebastian (2016). SentiHood: Targeted Aspect Based Sentiment Analysis Dataset for Urban Neighbourhoods. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics.

SentiHood: Target-Based Aspect-Based Sentiment Analysis



Saeidi, Marzieh; Bouchard, Guillaume; Liakata, Maria; Riedel, Sebastian (2016). SentiHood: Targeted Aspect Based Sentiment Analysis Dataset for Urban Neighbourhoods. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics.

SentiHood: Target-Based Aspect-Based Sentiment Analysis



Saeidi, Marzieh; Bouchard, Guillaume; Liakata, Maria; Riedel, Sebastian (2016). SentiHood: Targeted Aspect Based Sentiment Analysis Dataset for Urban Neighbourhoods. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics.

SentiHood: Target-Based Aspect-Based Sentiment Analysis

Model	Aspect (F_1)	Sentiment (Accuracy)	Aspect (AUC)	Sentiment (AUC)
LR-Left-Right	0.683	0.847	0.903	0.875
LR-Mask(ngram)	**0.697**	0.853	0.918	0.885
LR-Mask(ngram+POS)	*0.393*	**0.875**	**0.924**	**0.905**
LSTM-Final	0.689	0.820	0.898	0.854
LSTM-Location	**0.693**	0.819	0.897	0.839

Model	Price	Safety	Transit	General
LR - Mask (n-gram + POS)	**0.940**	**0.960**	**0.879**	0.864
LSTM - Final	0.875	0.932	0.836	**0.869**

Saeidi, Marzieh; Bouchard, Guillaume; Liakata, Maria; Riedel, Sebastian (2016). SentiHood: Targeted Aspect Based Sentiment Analysis Dataset for Urban Neighbourhoods. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics.

SentiHood: Target-Based Aspect-Based Sentiment Analysis

Sentence	Aspect	Predicted	Label
location1 is not a nice cheap residential area to live trust me i was born and raised there	Price	Positive	Negative
I think you'd find it tough to find something affordable in **location1**	Price	Positive	Negative
I can't recommend **location1** for affordability	Price	Negative	Negative
I only know about **location1**, most people prefer location2	General	None	None
I only know about location1, most people prefer **location2**	General	Positive	Positive

Saeidi, Marzieh; Bouchard, Guillaume; Liakata, Maria; Riedel, Sebastian (2016). SentiHood: Targeted Aspect Based Sentiment Analysis Dataset for Urban Neighbourhoods. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics.

Target-Based vs. Aspect-Based Sentiment Analysis vs. Stance Detection

- Target-Based:
  - Sentiment towards a concrete target, e.g. a product
- Aspect-Based:
  - Sentiment towards a concrete sub-aspect of a target (price, cleanliness, size, $\cdots$)
- Stance Detection:
  - Like target-based, but target not always mentioned / can be implicit or user-defined

Stance Detection

- Determine attitude expressed in document/paragraph/sentence towards a topic/statement/target
- Different classification schemes:
  - positive, negative, neutral (SemEval 2016 Task 6, RTE, SNLI)
  - support, deny, query, comment (SemEval 2017 Task 8 RumourEval)
  - agree, disagree, discuss, unrelated (Fake News Challenge)

Stance Detection with Conditional Encoding

No more #NastyWomen or #BadHombres

Task: Is tweet positive, negative or neutral towards a given target (Donald Trump)?

Problems:
- Interpretation depends on target
- Target not always mentioned in tweet
- No training data for test target

Augenstein et al., SemEval 2016, EMNLP 2016

Stance Detection with Conditional Encoding

- Challenges
  - **Model**:
    Learn a model that interprets the tweet stance towards a target that might not be mentioned in the tweet itself
  - **Training Data**:
    Learn model without labelled training data for the target with respect to which we are predicting the stance

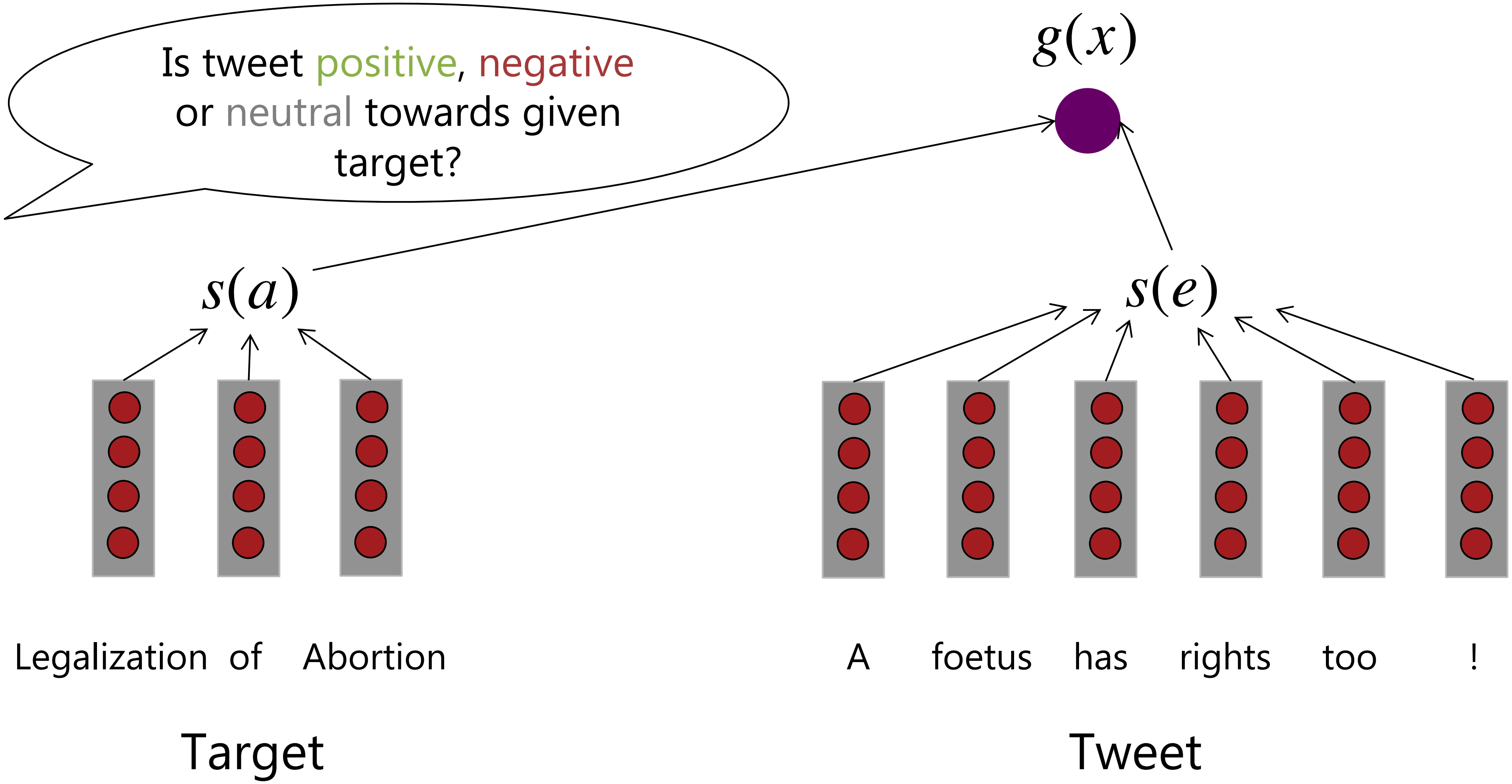
Stance Detection with Conditional Encoding

- Weakly Supervised Setting
  - Weakly label *Donald Trump* tweets using hashtags / expressions, evaluate on *Donald Trump* tweets

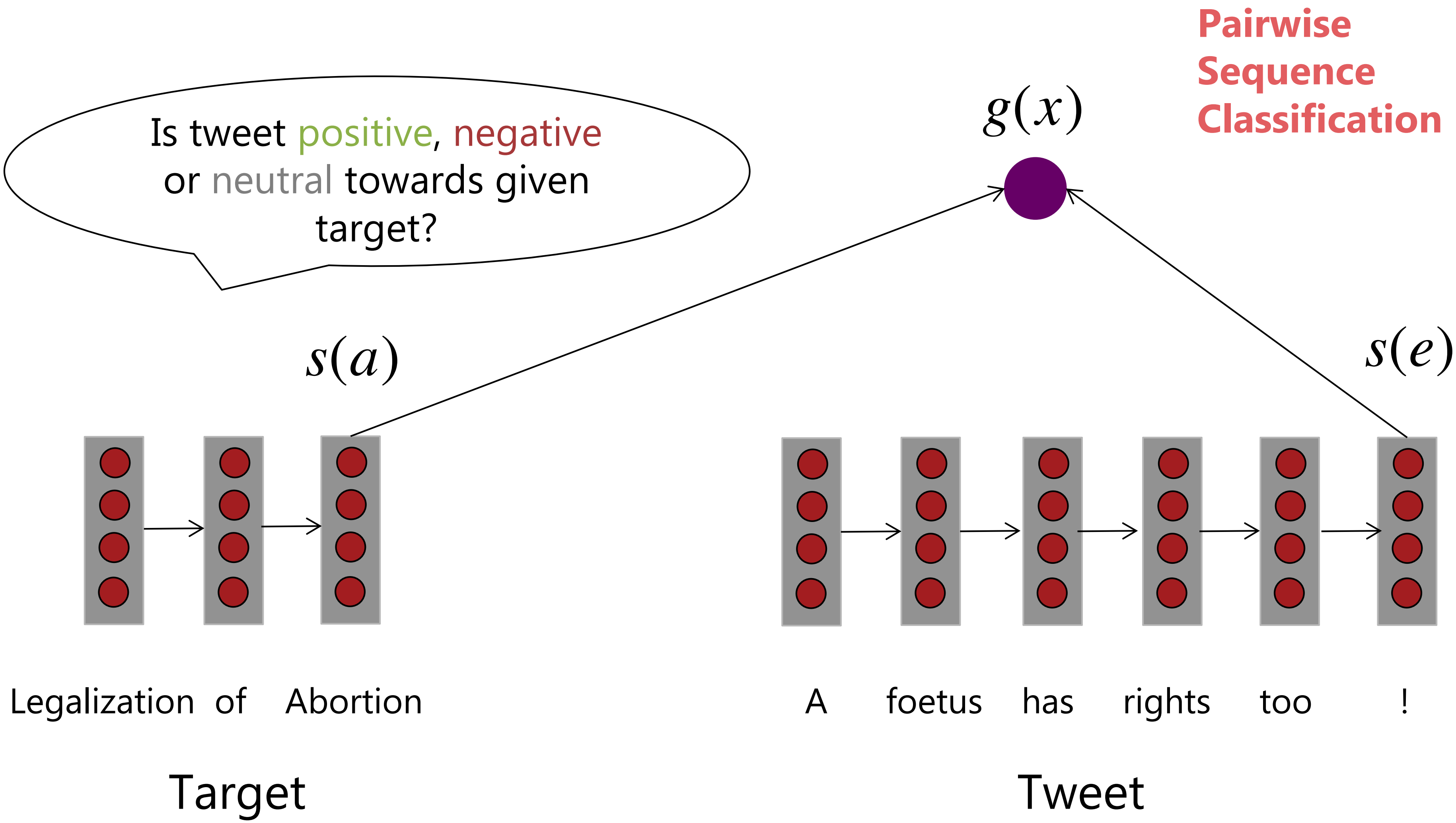
    positive:
    make(?)america(?)great(?)again
    trump(?)(for|4)(?)president

    negative:
    #dumptrump
    #notrump

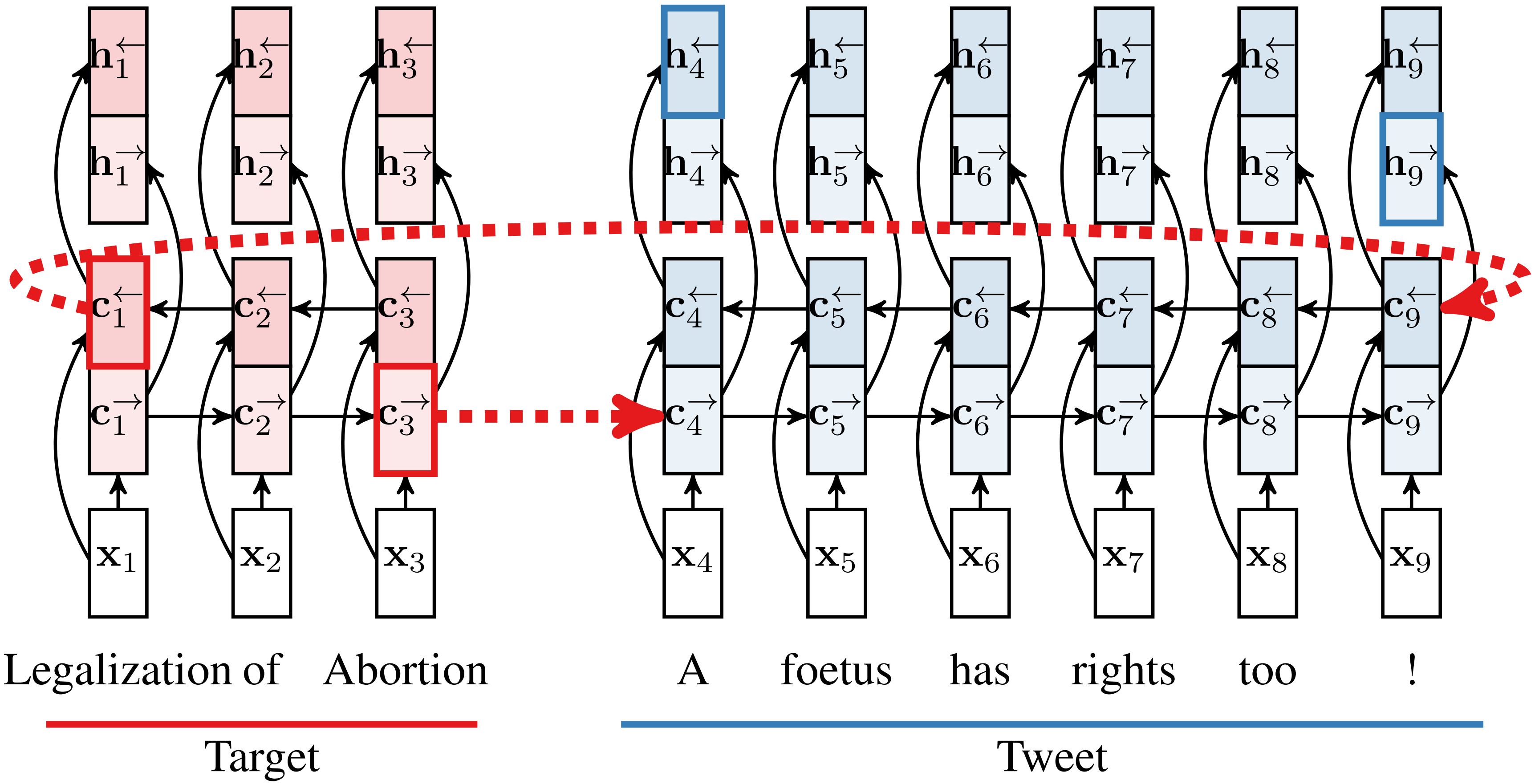
Stance Detection Model: Sum of Word Embeddings



Stance Detection Model: Concatenated Sequence Representations



Stance Detection Model: Bidirectional Conditional Encoding



Stance Detection

Python + Tensorflow exercises

Absalon -> Web Science -> Files -> EXERCISES -> stance.zip

Tracking Rumours on Social Media

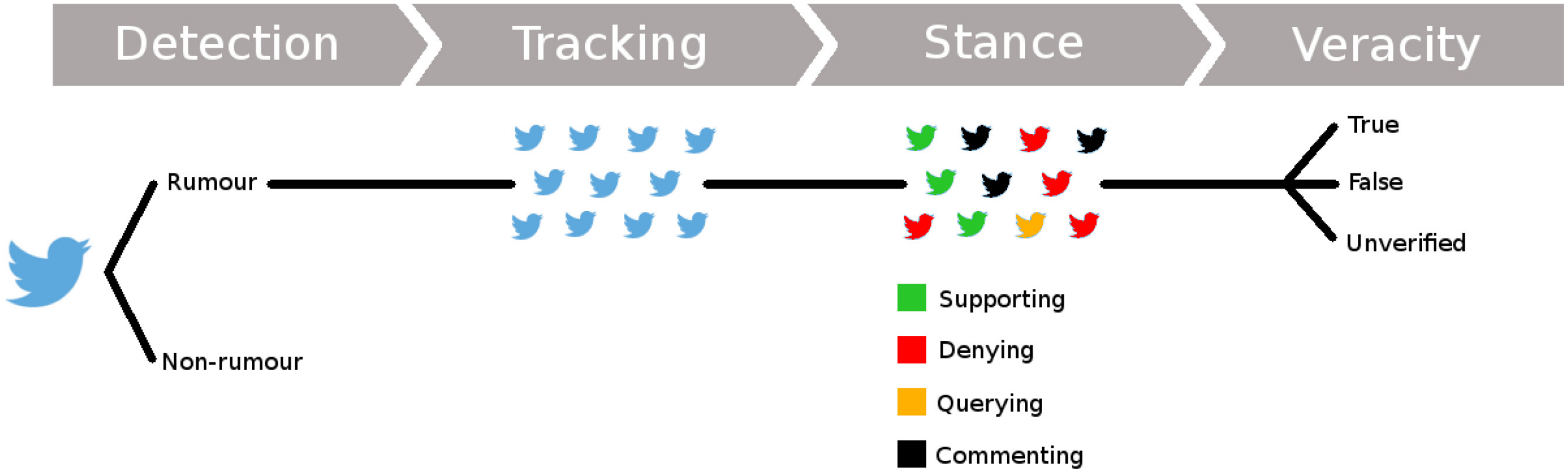
Example Rumours (10 in total, 2 of those only in test)

- **Putin missing**: from March 2015 - Russian president Vladimir Putin did not appear in public for 10 days. Rumours emerged he had been ill or killed. *Denied* by Putin himself on 11th day.
- **Gurlitt collection**: from November 2014 - Bern Museum of Fine Arts to accept a collection of modernist masterpieces kept by the son of a Nazi-era art dealer. *Confirmed*.

Kochkina, Liakata, **Augenstein** (2017)

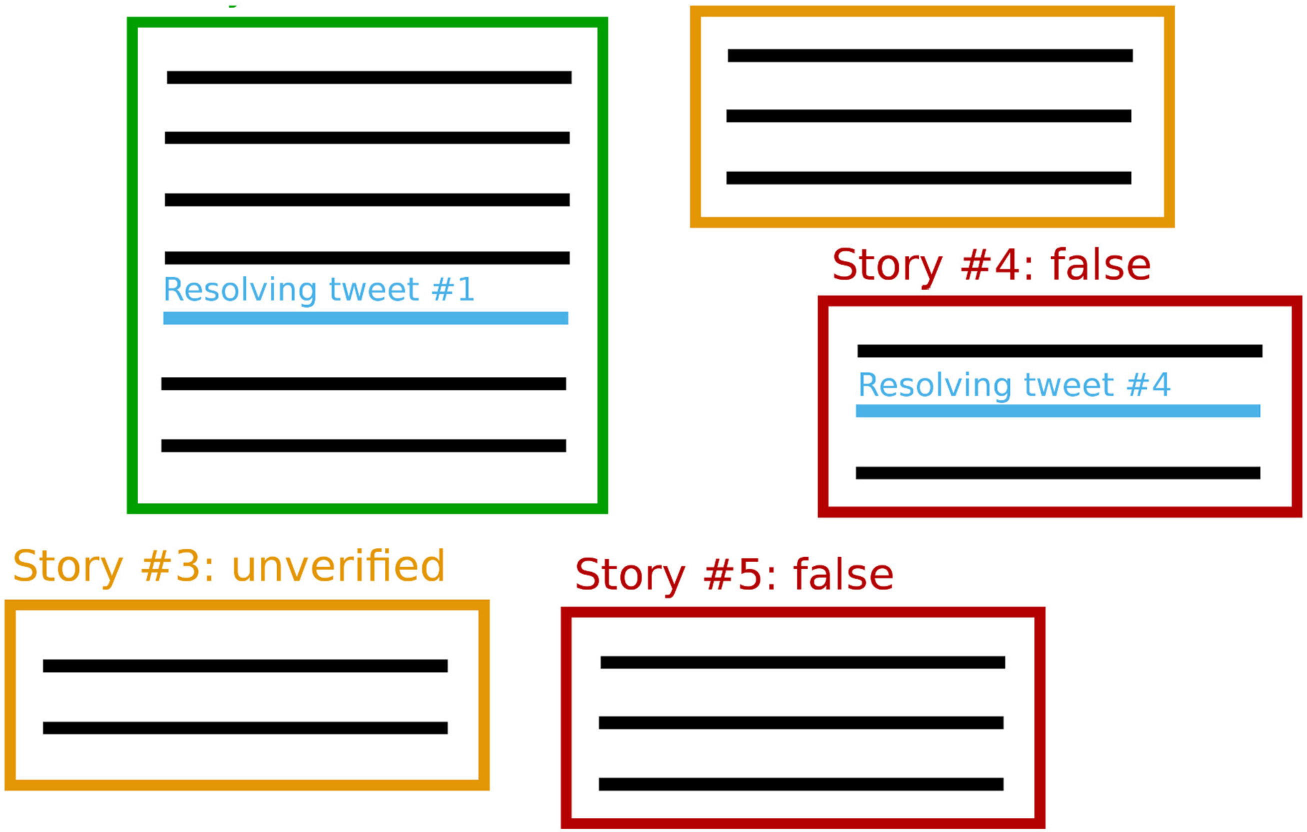
Zubiaga, Kochkina, Liakata, Procter, Lukasik, Bontcheva, Cohn, **Augenstein** (2017)

From Rumour Detection to Rumour Veracity Prediction



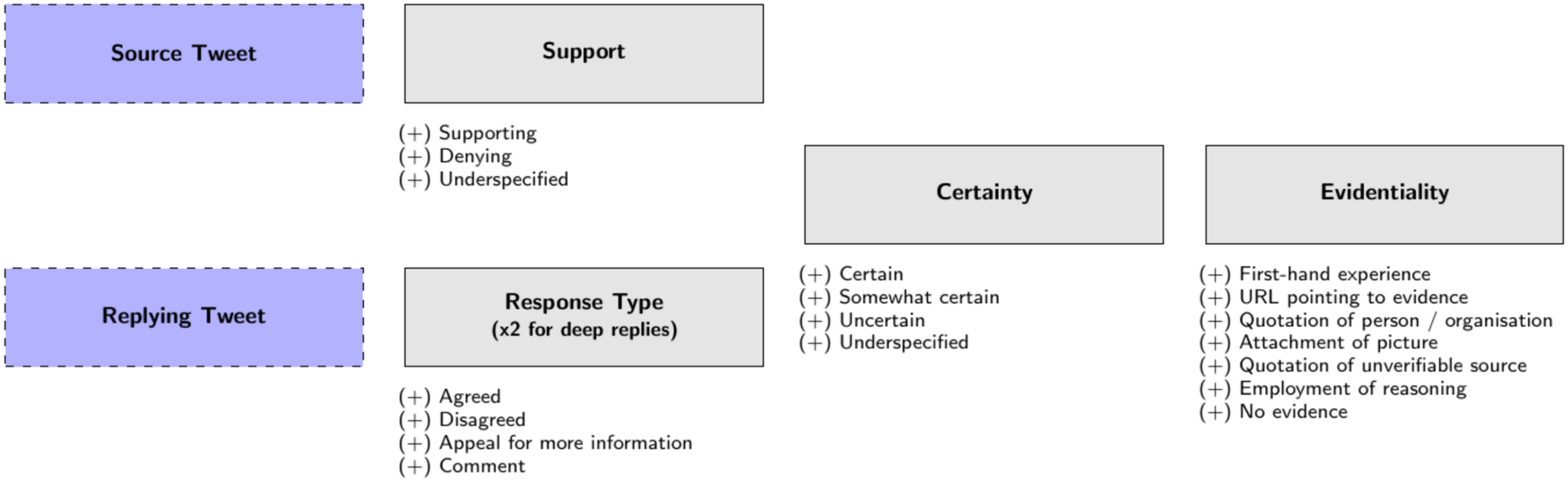
Arkaitz Zubiaga, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter. 2018. Detection and Resolution of Rumours in Social Media: A Survey. ACM Comput. Surv. 51, 2, Article 32 (February 2018), 36 pages. DOI: https://doi.org/10.1145/3161603

From Rumour Detection to Rumour Veracity Prediction



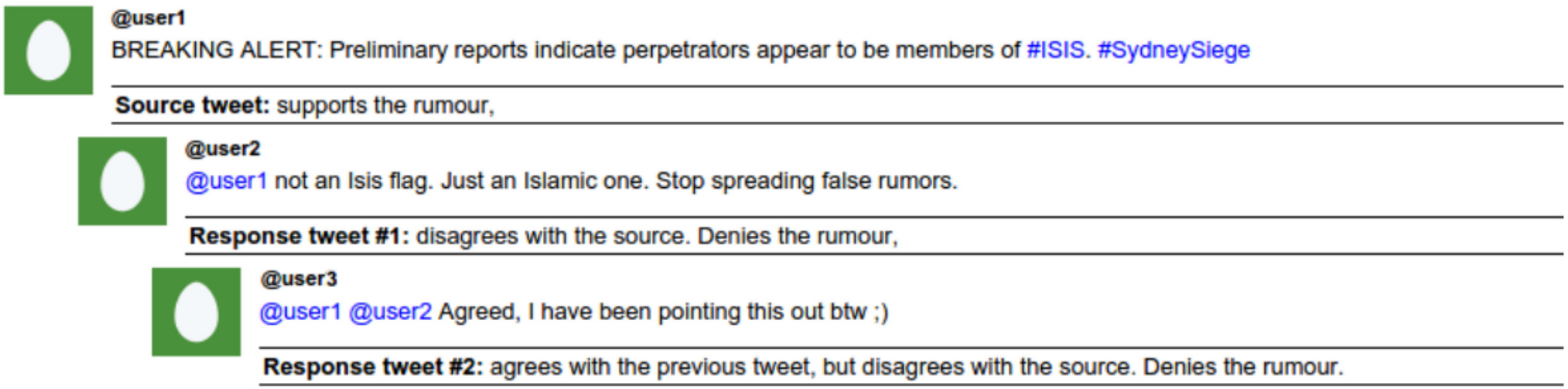
Arkaitz Zubiaga, Maria Liakata, Rob Procter, Geraldine Wong Sak Hoi, Peter Tolmie. Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads. PLOS ONE. 2016. [IF 2014: 3.234]

From Rumour Detection to Rumour Veracity Prediction

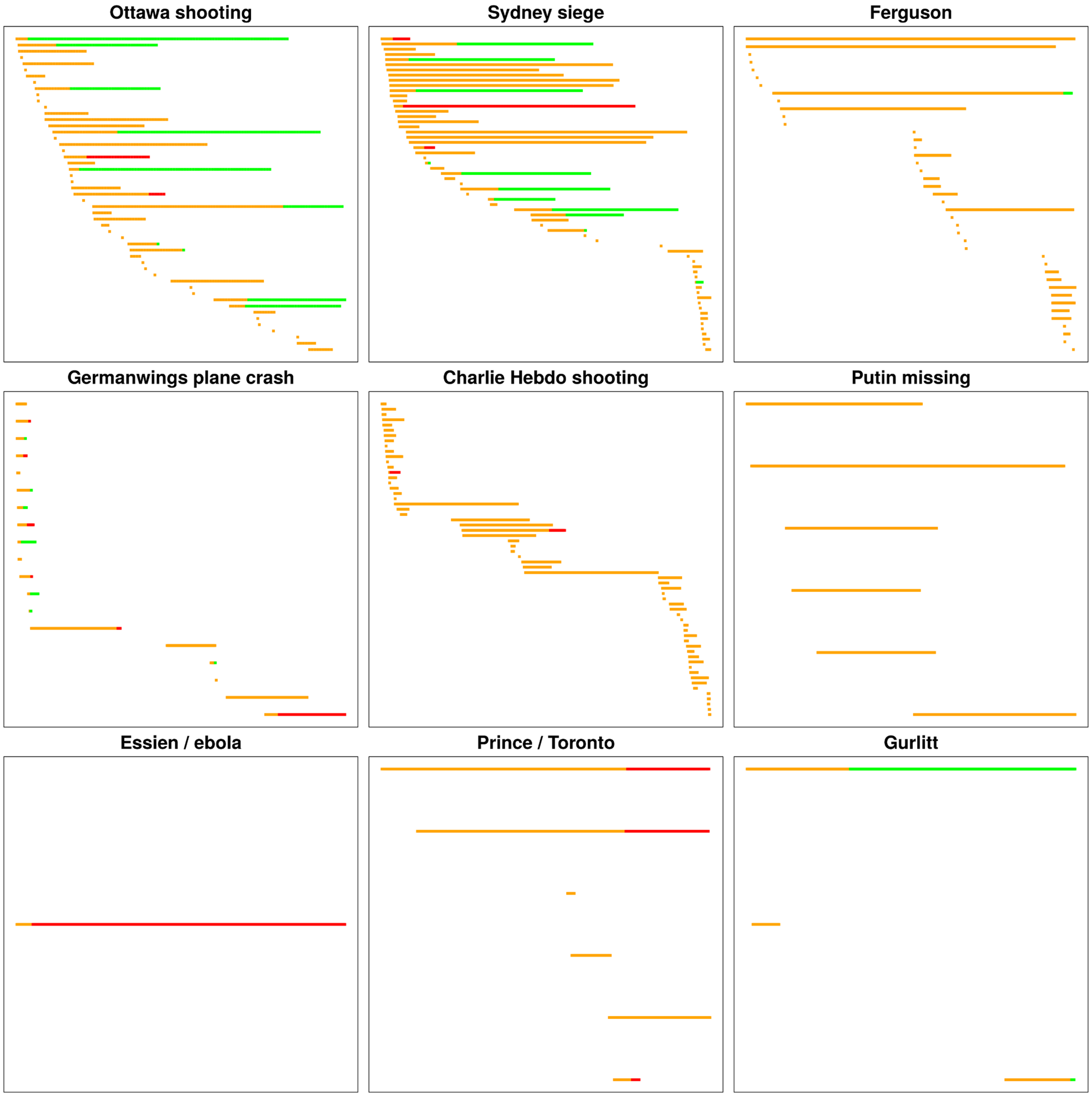


Arkaitz Zubiaga, Maria Liakata, Rob Procter, Geraldine Wong Sak Hoi, Peter Tolmie. Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads. PLOS ONE. 2016. [IF 2014: 3.234]

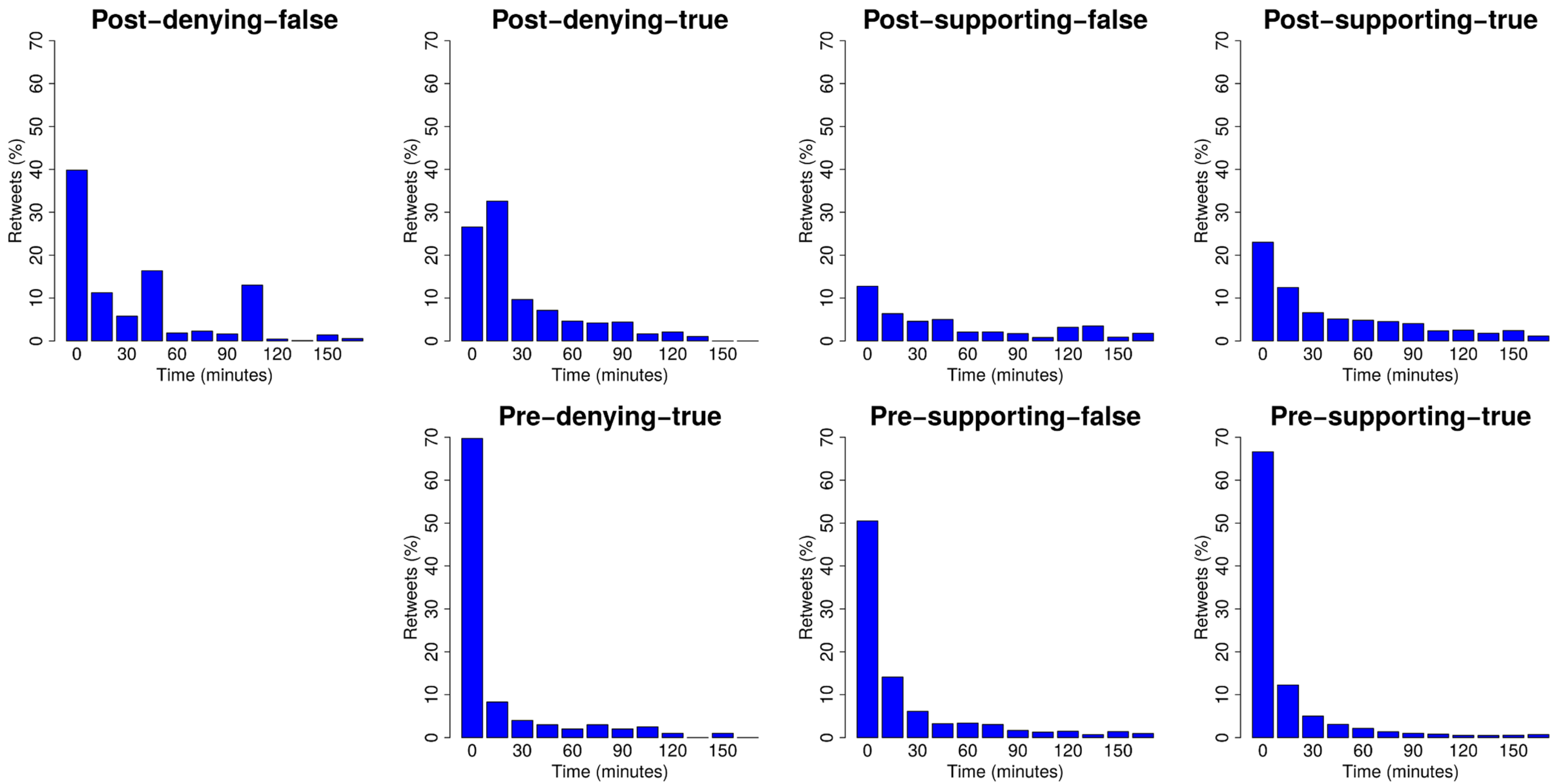
From Rumour Detection to Rumour Veracity Prediction



Arkaitz Zubiaga, Maria Liakata, Rob Procter, Geraldine Wong Sak Hoi, Peter Tolmie. Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads. PLOS ONE. 2016. [IF 2014: 3.234]

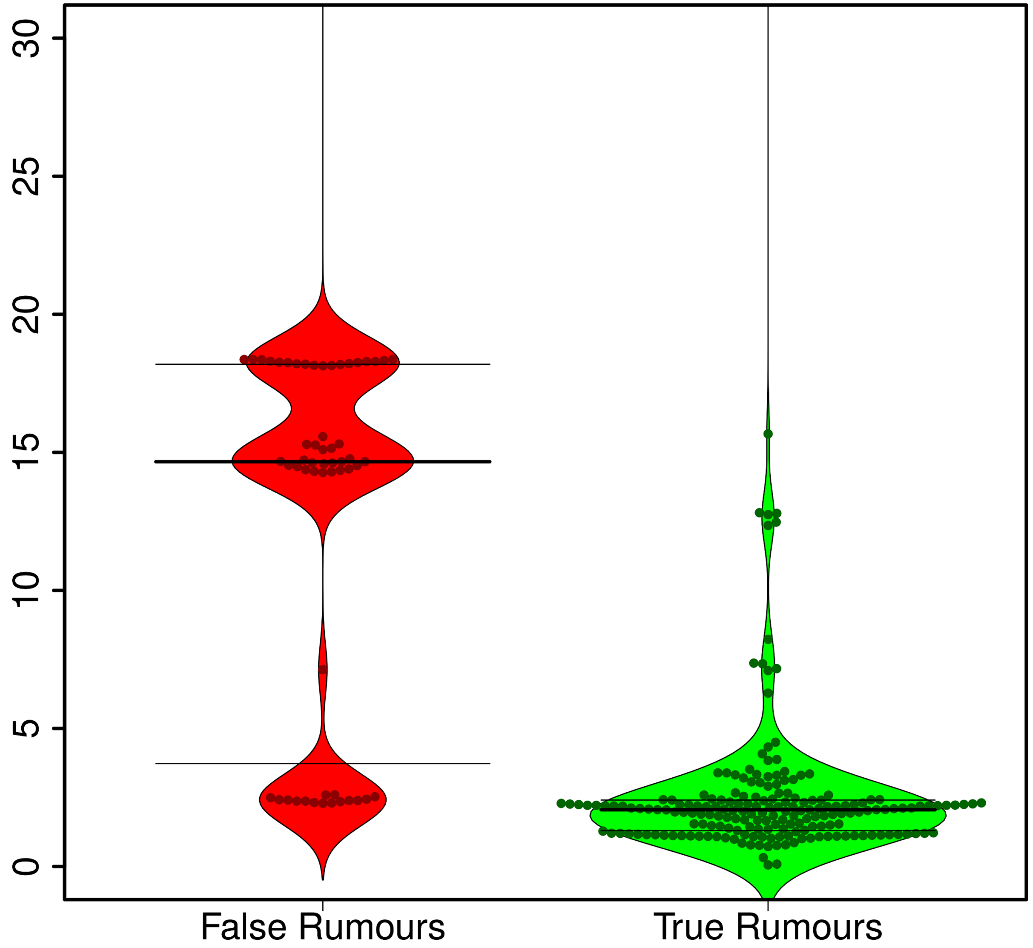

Ottawa shooting
Sydney siege
Ferguson
Germanwings plane crash
Charlie Hebdo shooting
Putin missing
Essien / ebola
Prince / Torontto
Gurlitt

From Rumour Detection to Rumour Veracity Prediction



Retweet timelines showing the percentage of retweets that each type of tweet gets in 15 minute steps. Higher retweet percentages at the beginning represent a high **interest in spreading the tweet** in the very first minutes.

From Rumour Detection to Rumour Veracity Prediction



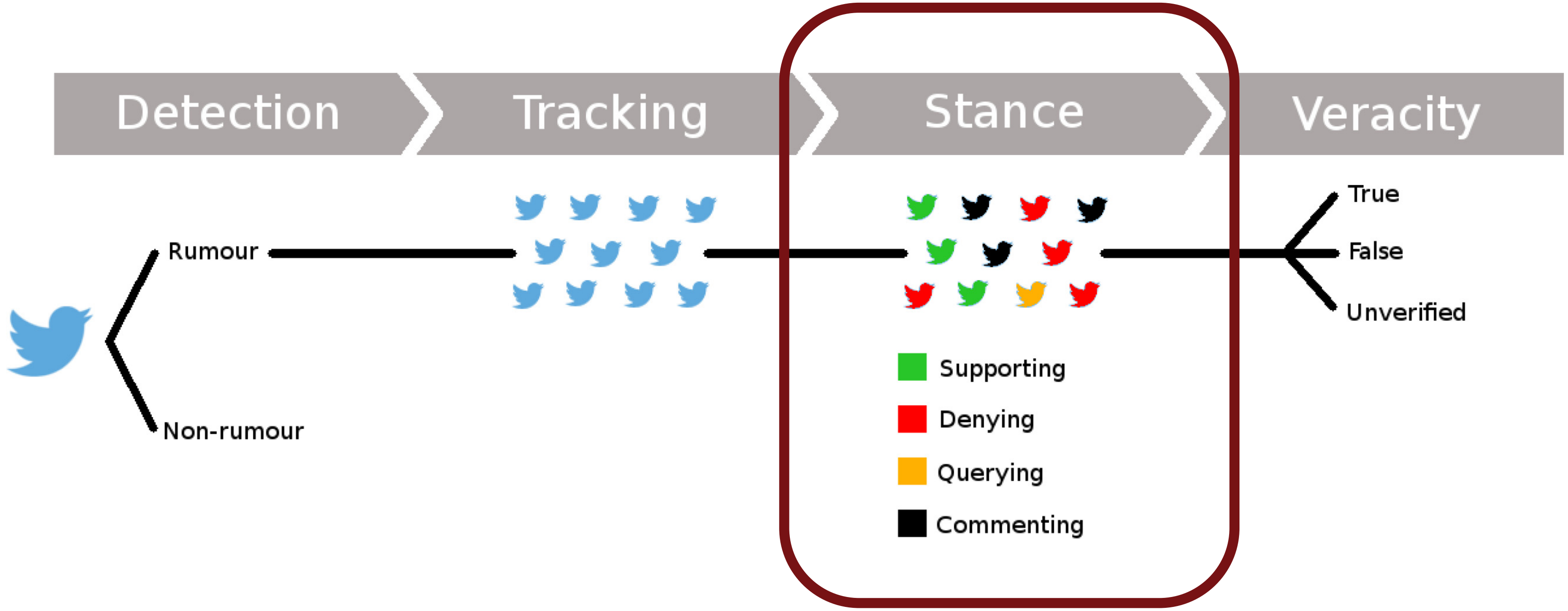
Distribution of delays (in hours) in resolving false (red) and true (green) rumours. Horizontal lines represent 25, 50, and 75 percentiles.

Conclusions: From Rumour Detection to Rumour Veracity Prediction

- True rumours tend to be **resolved faster** than false rumours.
- Rumours in their unverified stage produce a distinctive **burst in the number of retweets** within the first few minutes, substantially more than rumours proven true or false.
- The prevalent tendency of users is to **support every unverified rumour**.
- Level of **certainty** does not change over the course of the rumour lifecycles.
- Users **provide evidence** in their tweets when rumours are yet to be resolved, but less so after resolution.
- Highly **reputable users** such as news organisations tend to support rumours, irrespective of them being eventually confirmed or debunked, tweet with certainty and provide evidence within their tweets.

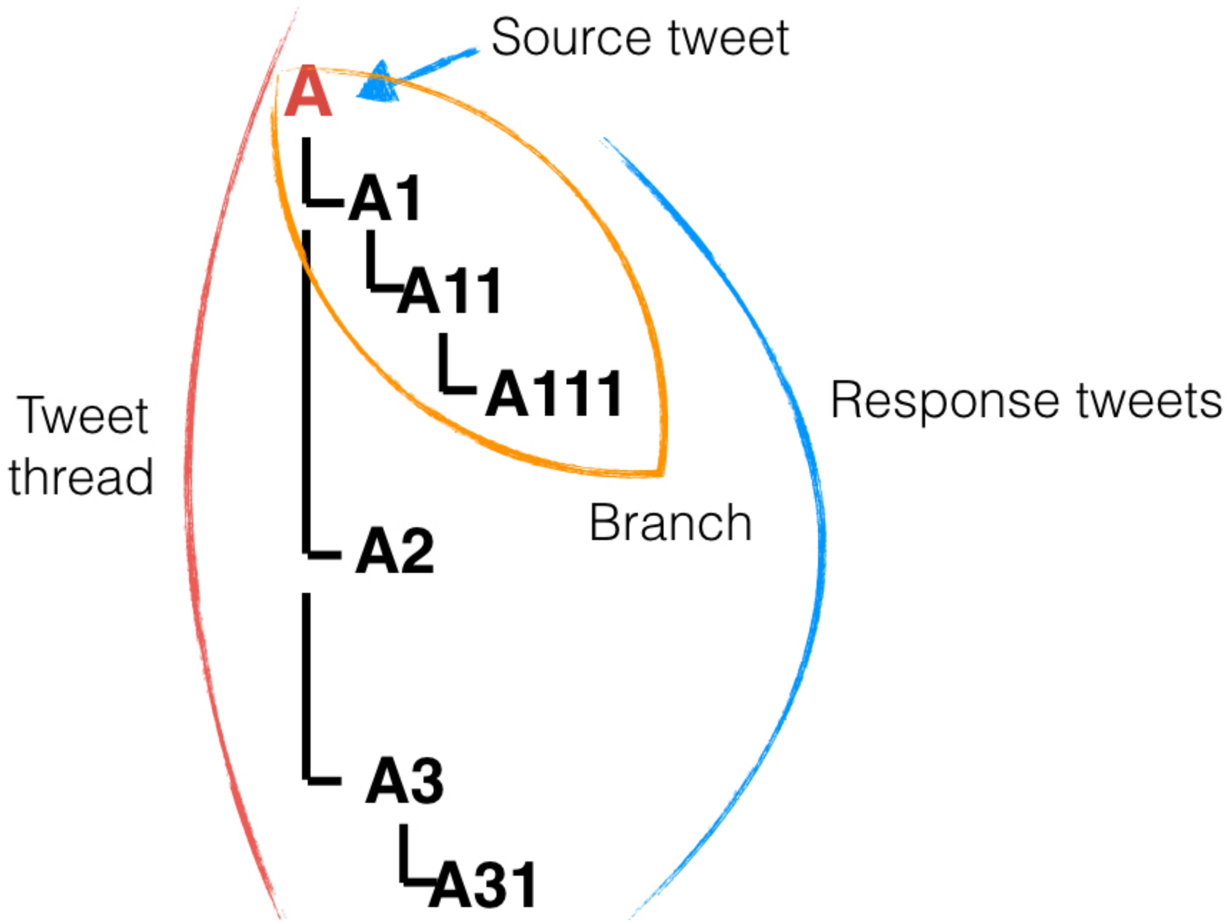
Arkaitz Zubiaga, Maria Liakata, Rob Procter, Geraldine Wong Sak Hoi, Peter Tolmie. Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads. PLOS ONE. 2016. [IF 2014: 3.234]

From Rumour Detection to Rumour Veracity Prediction



Arkaitz Zubiaga, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter. 2018. Detection and Resolution of Rumours in Social Media: A Survey. ACM Comput. Surv. 51, 2, Article 32 (February 2018), 36 pages. DOI: https://doi.org/10.1145/3161603

Stance Detection for Conversational Structures



SemEval 2017 RumourEval Task A

Stance Detection for Conversational Structures

branch 2

source

user 0 *support*

France: 10 people dead after shooting at HQ of satirical weekly newspaper #CharlieHebdo, according to witnesses http://t.co/FkYxGmuS58

response 2

user 2 *query*

@user0 @user5 A French crime of passion or another heathen moslem atrocity?

response 1

user 1 *comment*

MT **@user0** France: 10 dead after shooting at HQ of satirical weekly #CharlieHebdo. If Zionists/Jews did this they'd be nuking Israel

branch 1

response 3

user 3 *comment*

@user0 LOL. 5 million Muslims in France, what a disgrace. the french worm president and politicians killed them. tine for croissants now

user 4 *deny*

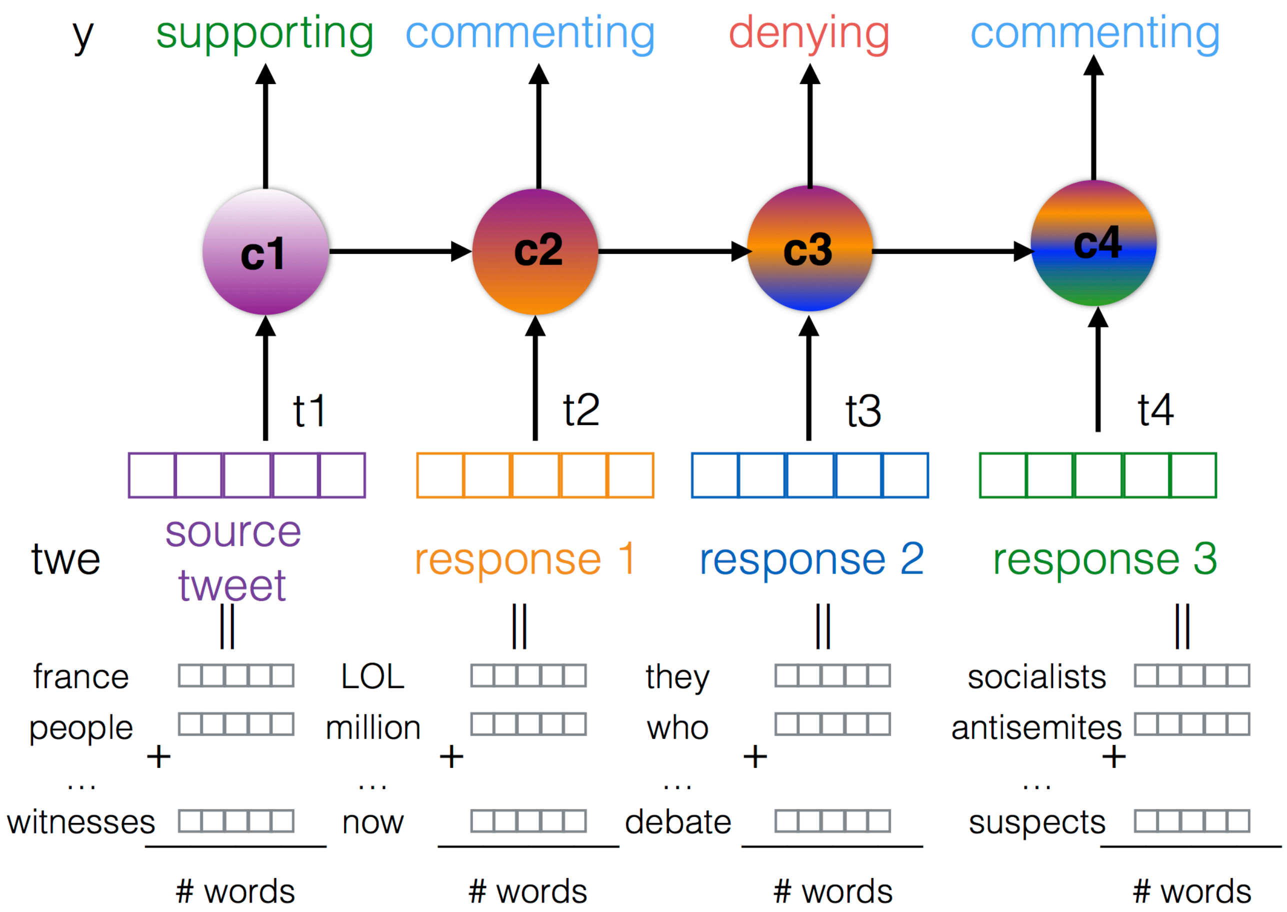
@user3 They who? Stupid and partial opinions like this one only add noise to any debate.

user 3 *comment*

@user4 Socialists, Antisemites, anti zionists - usual suspects

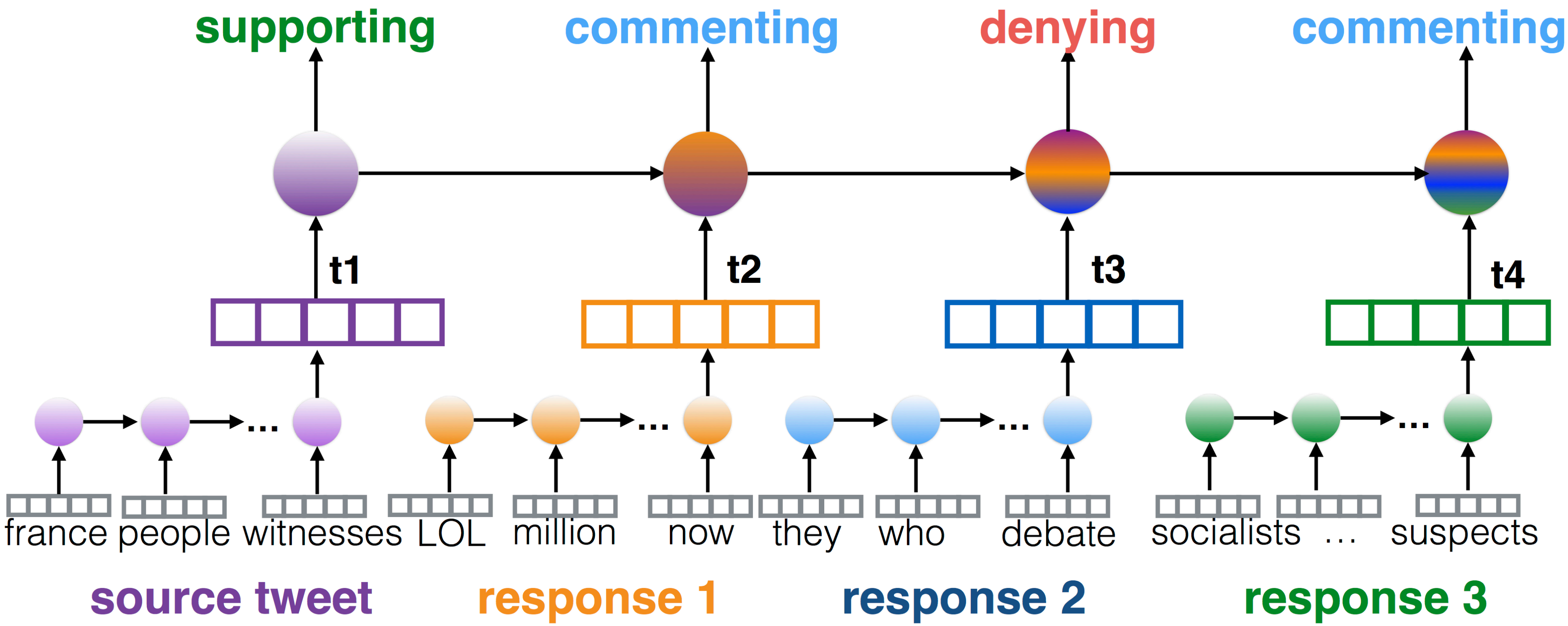
SemEval 2017 RumourEval Task A

Stance Detection for Conversational Structures



SemEval 2017 RumourEval Task A (**winning system**), IP&M 2017

Stance Detection for Conversational Structures



SemEval 2017 RumourEval Task A (**winning system**), IP&M 2017

Stance Detection for Conversational Structures

Example Rumours (10 in total, 2 of those only in test)

- **Putin missing**: from March 2015 - Russian president Vladimir Putin did not appear in public for 10 days. Rumours emerged he had been ill or killed. *Denied* by Putin himself on 11th day.
- **Gurlitt collection**: from November 2014 - Bern Museum of Fine Arts to accept a collection of modernist masterpieces kept by the son of a Nazi-era art dealer. *Confirmed*.

SemEval 2017 RumourEval Task A

Stance Detection for Conversational Structures

Training data

	Supporting	Denying	Querying	Commenting
Development	69	11	28	173
Testing	94	71	106	778
Training	841	333	330	2734

SemEval 2017 RumourEval Task A

Stance Detection for Conversational Structures

Confusion matrix

Prediction	Gold			
	Supporting	Denying	Querying	Commenting
Supporting	**26**	2	1	6
Denying	0	**0**	0	0
Querying	1	1	**36**	12
Commenting	6	68	69	**760**

SemEval 2017 RumourEval Task A (**winning system**)

Stance Detection for Conversational Structures

Examples misclassifying *denying*

[As querying] @username Weren't you the one who abused her?

[As supporting] "Go online & put down 'Hillary Clinton illness,'" Rudy says. Yes – but look up the truth – not health smears https://t.co/EprqiZhAxM

[As supporting] @username I demand you retract the lie that people in #Ferguson were shouting "kill the police", local reporting has refuted your ugly racism

[As commenting] @FoxNews six years ago... real good evidence. Not!

SemEval 2017 RumourEval Task A (**winning system**)

Summary: Stance Detection

- Relationship between sequences can be modelled effectively with deep recurrent neural models
- Even more complicated structures (conversational threads) can be modelled effectively

- Many challenges
  - Hard to collect data, especially with balanced labels
  - Hard to train deep neural NLP models with little, imbalanced data
  - Complex sentence constructions

Beyond Sentiment Analysis: Affect

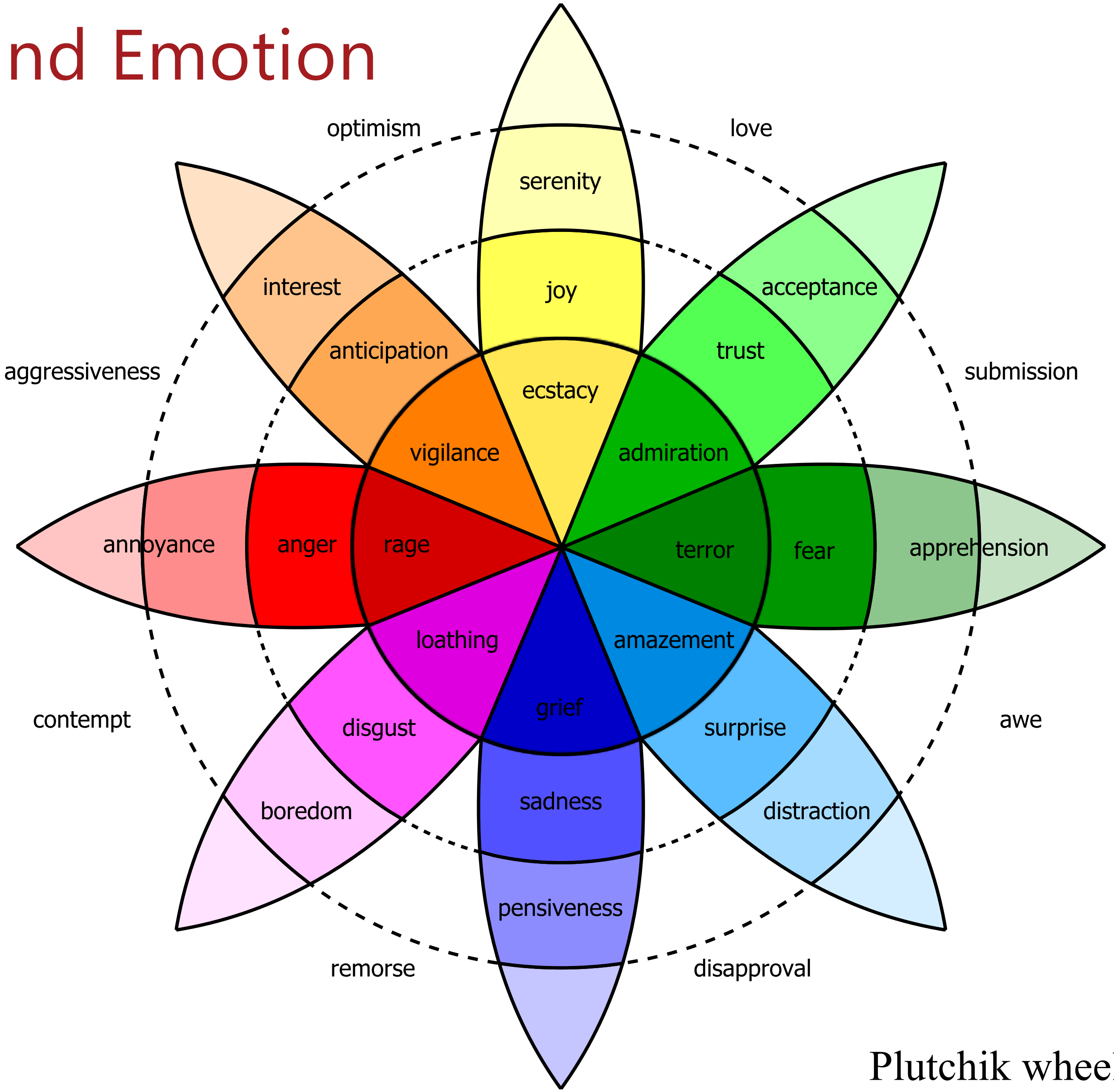
- **Emotion**: Relatively brief episode of response to the evaluation of an external or internal event as being of major significance. (*angry, sad, joyful, fearful, ashamed, proud, elated, desperate*)
- **Mood**: Diffuse affect state, most pronounced as change in subjective feeling, of low intensity but relatively long duration, often without apparent cause. (*cheerful, gloomy, irritable, listless, depressed, buoyant*)
- **Interpersonal stance**: Affective stance taken toward another person in a specific interaction, colouring the interpersonal exchange in that situation. (*distant, cold, warm, supportive, contemptuous, friendly*)

Beyond Sentiment Analysis: Affect

- **Attitude**: Relatively enduring, affectively colored beliefs, preferences, and predispositions towards objects or persons. (*liking, loving, hating, valuing, desiring*)
- **Personality traits**: Emotionally laden, stable personality dispositions and behavior tendencies, typical for a person. (*nervous, anxious, reckless, morose, hostile, jealous*)

The Scherer typology of affective states

Affect and Emotion



Plutchik wheel of emotion

Emojis and Emotion

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
USA	❤️	😍	😂	💕	🔥	😊	😎	✨	💙	😘	📷	🇺🇸	☀️	💜	😉	💯	😁	🎄	📸	😜

SemEval 2018 Emoji Prediction Task

Sentiment and Personality

- **Extroversion vs. Introversion**: sociable, assertive, playful vs. aloof, reserved, shy
- **Emotional stability vs. Neuroticism**: calm, unemotional vs. insecure, anxious
- **Agreeableness vs. Disagreeableness**: friendly, cooperative vs. antagonistic, faultfinding
- **Conscientiousness vs. Unconscientiousness**: self-disciplined, organized vs. inefficient, careless
- **Openness to experience**: intellectual, insightful vs. shallow, unimaginative

"Big Five" dimensions

Sentiment and Personality

(a)

(b)

Word clouds from Schwartz et al. (2013), showing words highly associated with **introversion (left)** or **extroversion (right)**.The size of the word represents the *association strength* (the regression coefficient), while the color (ranging from cold to hot) represents the *relative frequency* of the word/phrase (from low to high)

Take-Home Points

- Opinion mining can help us make sense of large data
- Many different more fine-grained sentiment analysis tasks: target-based, aspect-based, stance detection
- Many related tasks: rumour detection, affect, emotion, personality detection
- Different methods
  - Sequence labelling with RNNs
  - Conditional encoding
  - RNNs for conversational structures

Thank you!

isabelleaugenstein.github.io
augenstein@di.ku.dk
@IAugenstein
github.com/isabelleaugenstein