Sentiment, Stance and Affective Analysis for Mis/Disinformation

Lynnette Ng

The CMU centers for:

Informed DEmocracy And Social cyber-security

Computational Analysis of Social and Organizational Systems

Carnegie Mellon University

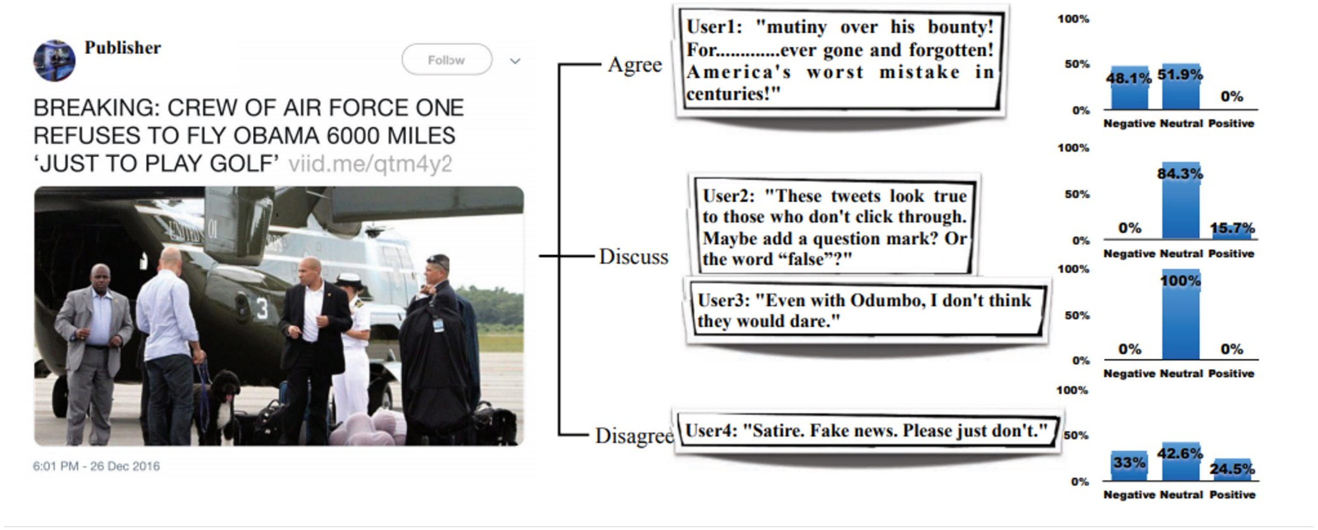
Agenda

- ❑Sentiment Analysis
  - ❑Linguistic Methods: NetMapper run through
- ❑Stance Analysis
  - ❑Linguistic Methods
  - ❑Network Methods: Stance Propagation + ORA run through
  - ❑Combining Linguistic + Network Methods
- ❑Affect Mining
  - ❑Textual Analysis
  - ❑Image Analysis

Made in IDeaS/ CASOS

Sentiment Analysis

- “a view or attitude towards a situation or event”
- Fake News Detection as a Sentiment Analysis Problem:
  - If a news piece has an standard deviation of user sentiment scores greater than a threshold, then the news is weakly labeled as fake news.



Tweet

User Opinion

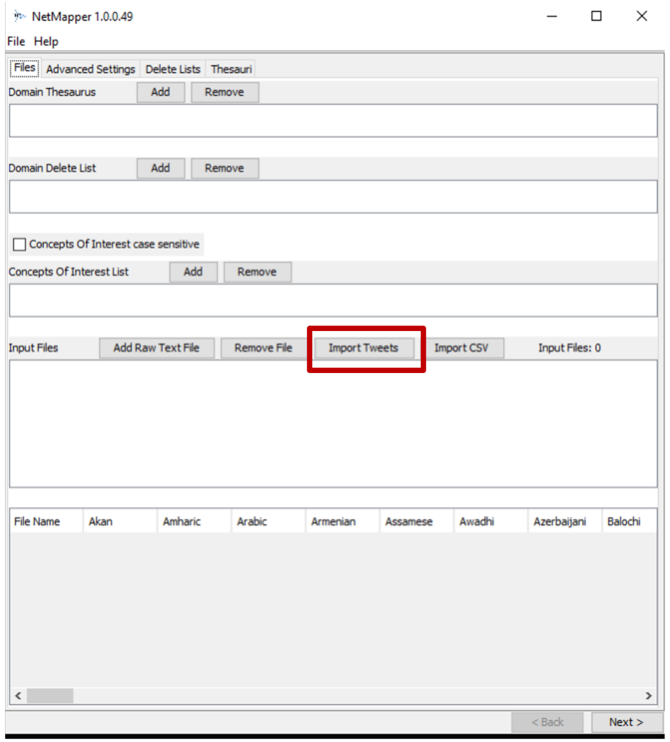
Sentiment Distributions

Limeng Cui and Suhang Wang Dongwon Lee. Same: Sentiment-aware multi-modal embedding for detecting fake news. 2019.

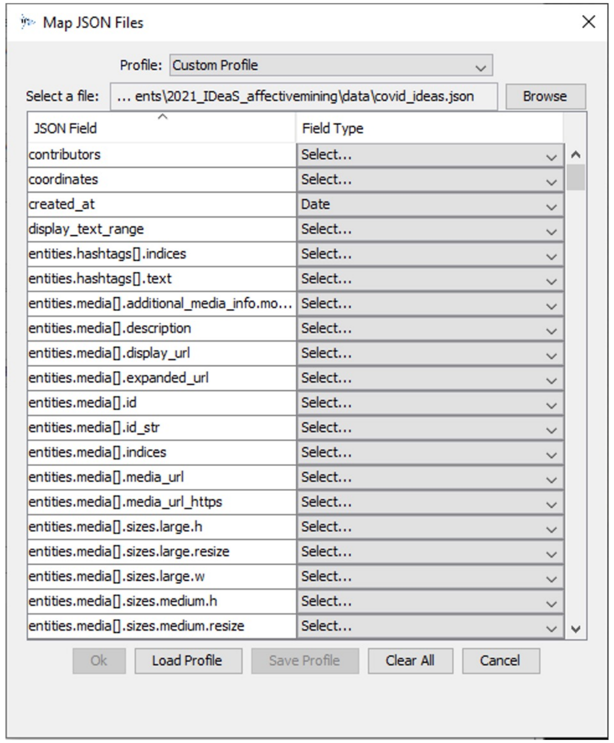
Sentiment Analysis

- Netmapper Example on COVID vaccine data

1. Import Tweets



2. Select Fields

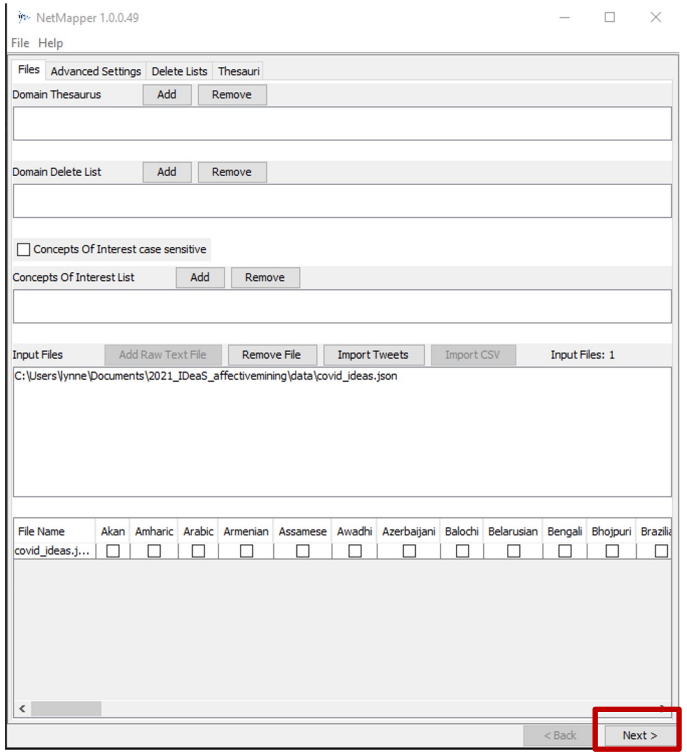


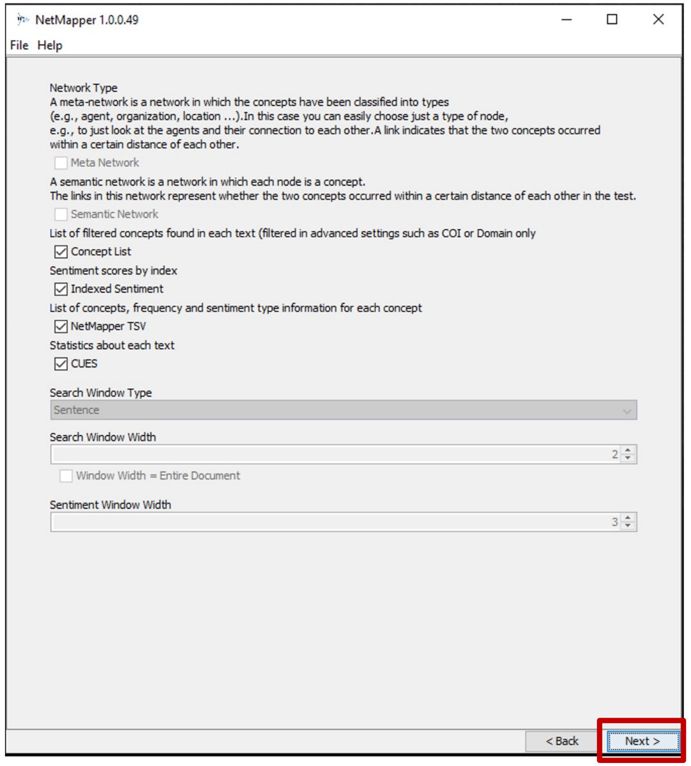
created_at: Date
id_str: Tweet ID
user.id_str: Author
extended_tweet.full_text: Text

Sentiment Analysis

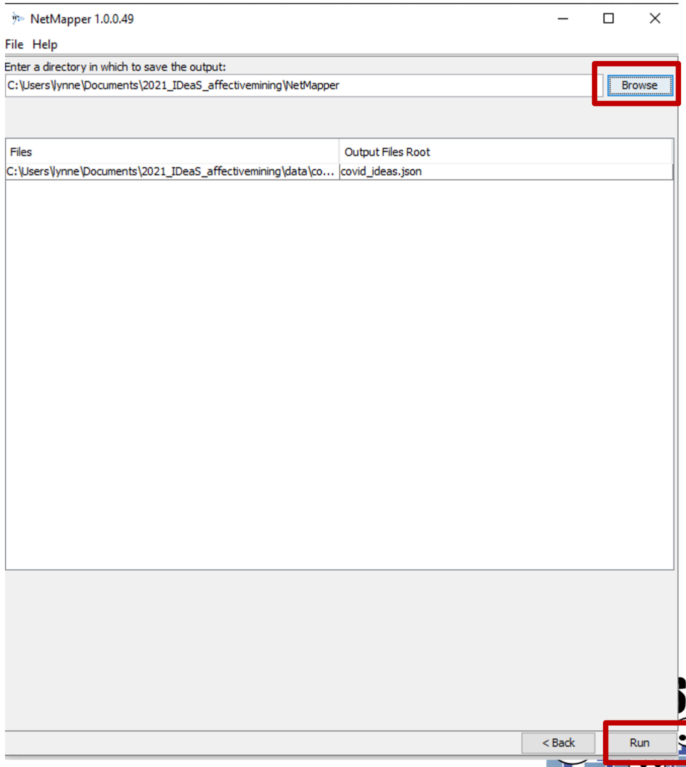
❑Netmapper Example on COVID vaccine data

3. Next





3. Select Folder and Run



Carnegie Mellon University

Sentiment Analysis

- Netmapper Example on COVID vaccine data

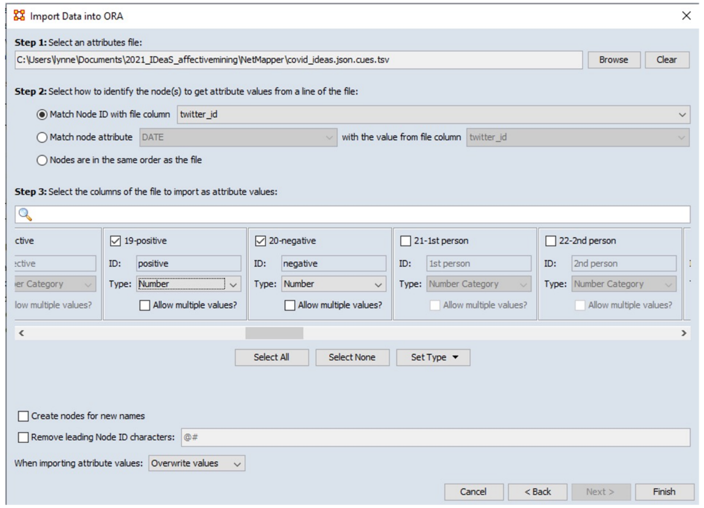
5. Sentiment is in .cues.tsv

This PC › Documents › 2021_IDeaS_affectivemining › NetMapper

Name	Date modified	Type	Size
covid_ideas.json.concepts_per_line.tsv	5/17/2021 8:36 AM	TSV File	5 KB
covid_ideas.json.cues.tsv	5/17/2021 8:36 AM	TSV File	2 KB
covid_ideas.json.emoticon.tsv	5/17/2021 8:36 AM	TSV File	1 KB
covid_ideas.json.hashtag.tsv	5/17/2021 8:36 AM	TSV File	1 KB
covid_ideas.json.indexed_sentiment.tsv	5/17/2021 8:36 AM	TSV File	3 KB
covid_ideas.json.phone_number.tsv	5/17/2021 8:36 AM	TSV File	0 KB
covid_ideas.json.rnmf.tsv	5/17/2021 8:36 AM	TSV File	5 KB
covid_ideas.json.twitter_handle.tsv	5/17/2021 8:36 AM	TSV File	1 KB
covid_ideas.json.url.tsv	5/17/2021 8:36 AM	TSV File	1 KB
covid_ideas.json.zip_code.tsv	5/17/2021 8:36 AM	TSV File	0 KB
file_map.csv	5/17/2021 8:36 AM	Microsoft Excel C...	0 KB

R	S	T	U	V	V
connectiv	positive	negative	1st persor	2nd perso	3rd p
1	5				
	1				
1	9	3	1		
	6	4			
	4	1			
2	4	5			
	5	3		1	
		1			
	5	2			

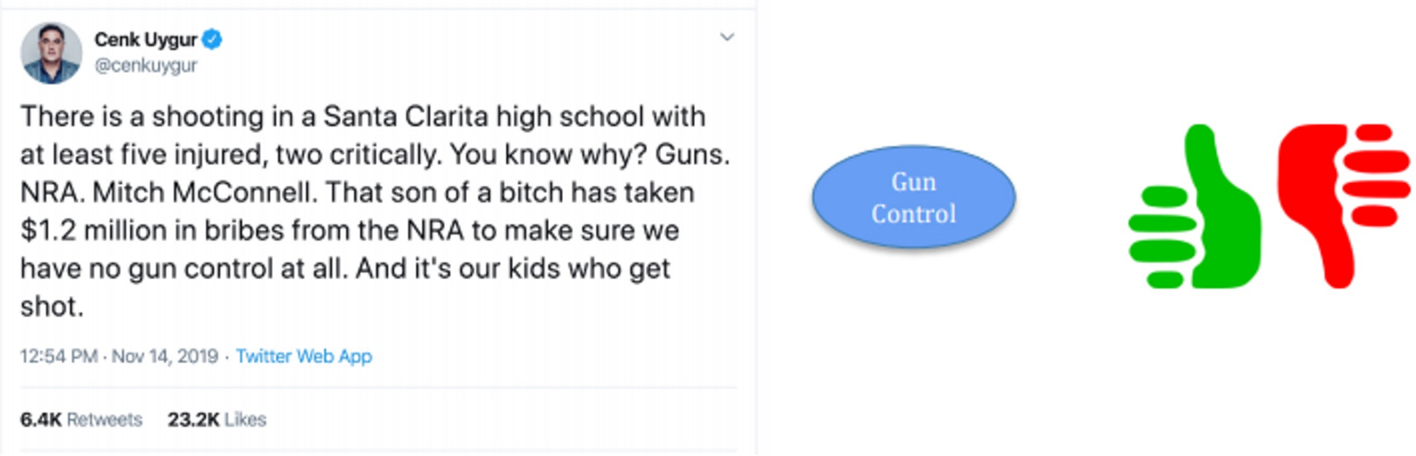
6. Can be imported as attribute in ORA



Carnegie Mellon University

Stance Analysis

- An "expression of speaker's standpoint and judgement towards a given proposition" (Biber and Finegan, 1988)
- A "mental or emotional position adopted with respect to a proposition, a person, an idea etc" (The Free Dictionary)
- Typically, a user's stance is characterized as: Pro, Con, Neutral



Tweet **Target/Topic** **Stance (Pro/Con)**

Linguistic Methods

- Fact checking as a stance analysis problem
- Fake News Challenge: detect the relatedness of a news article's body to a headline based on the stance a former takes regarding the latter, annotated by AGREE/ DISAGREE/ DISCUSSES
- FEVER: claim-evidence pairs annotated by SUPPORTED/ REFUTED/ NOT ENOUGH INFO; helps fact-checkers understand the decision models made in assessment of claim veracity

Linguistic Methods

- Key datasets in Stance Analysis for Disinformation
  - Note: Non-Comprehensive

Dataset	Source(s)	Target	Context	Evidence	#Instances	Task
English Datasets						
Rumour Has It [Qazvinian et al., 2011]	Twitter	Topic	Tweet	Multiple	10K	Rumours
PHEME [Zubiaga et al., 2016a]	Twitter	Claim	Tweet	Thread	7.5K	Rumours
Emergent [Ferreira and Vlachos, 2016]	News	Headline	Article*	Multiple	2.6K	Rumours
FNC-1 [Pomerleau and Rao, 2017]	News	Headline	Article	Single	75K	Fake news
RumourEval '17 [Derczynski et al., 2017]	Twitter	Implicit‡	Tweet	Thread	7.1K	Rumours
FEVER [Thorne et al., 2018]	W	Claim	Facts	Multiple	185K	Fact-checking
Snopes [Hanselowski et al., 2019]	Snopes	Claim	Snippets	Multiple	19.5K	Fact-checking
RumourEval '19 [Gorrell et al., 2019]	Twitter, Reddit	Implicit‡	Post	Thread	8.5K	Rumours
COVIDLies [Hossain et al., 2020]	Twitter	Claim	Tweet	Single	6.8K	Misconceptions
TabFact [Chen et al., 2020]	W	Statement	WikiTable	Multiple	118K	Fact-checking
Non-English Datasets						
Arabic [Baly et al., 2018]	News	Claim	Document	Single	3K	Fact-checking
DAST (Danish) [Lillie et al., 2019]	Reddit	Submission	Comment	Thread	3K	Rumour
Croatian [Bošnjak and Karan, 2019]	News	Title	Comment	Single	0.9K	Claim verifiability
Arabic [Khouja, 2020]	News	Claim	Title	Single	3.8K	Claim verification

Table 1: Key characteristics of the stance detection datasets for mis- and disinformation detection. *#Instances* denotes dataset size as a whole; the numbers are in thousands (K) and are rounded to the hundreds. *the article's body is summarised. ‡the stance is expressed towards a topic, which is not present in the data. *Sources*: Twitter, News, Wikipedia, Reddit. *Evidence*: Single, Multiple, Thread.

Hardalov, M., Arora, A., Nakov, P., & Augenstein, I. (2021). A survey on stance detection for mis-and disinformation identification. *arXiv preprint arXiv:2103.00242*. https://arxiv.org/pdf/2103.00242.pdf

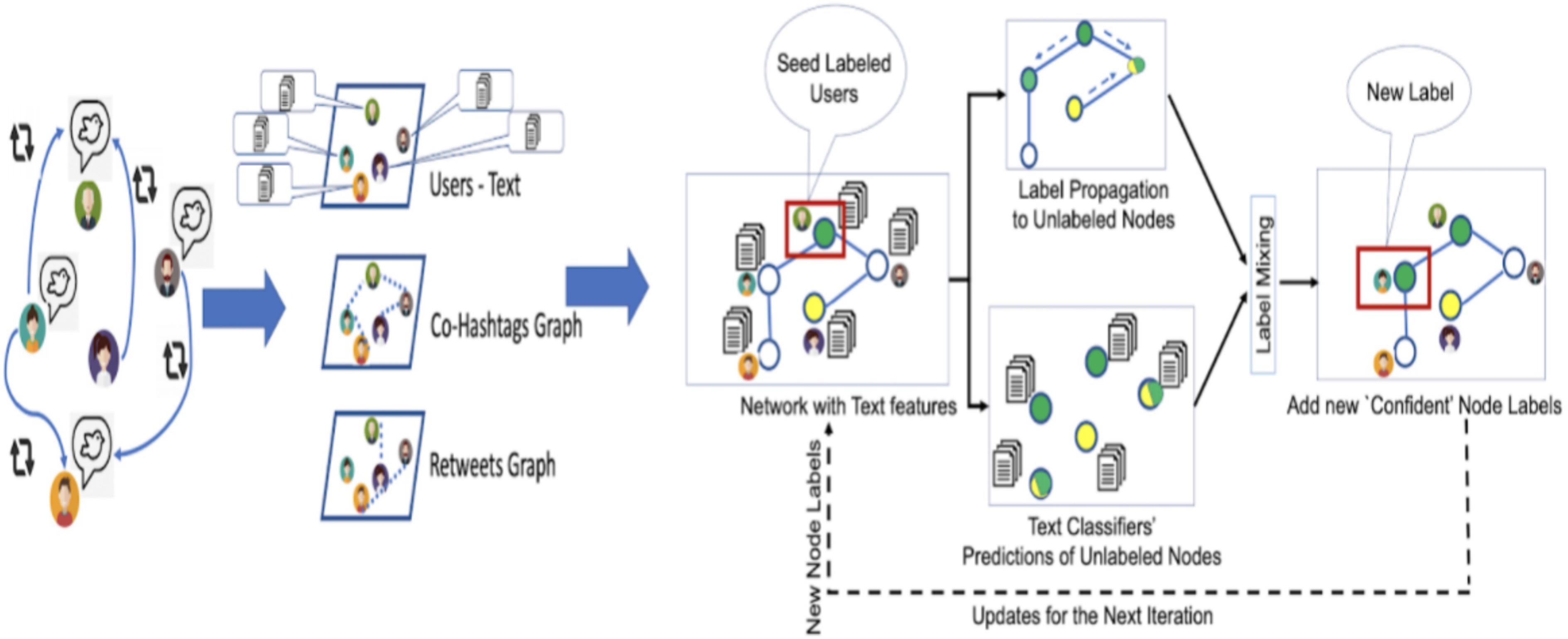
Linguistic Methods

- General methods and features used for Stance Analysis
  - Note: Non-Comprehensive

Authors	Approach	Features	Subtask
[Taulé et al. 2017]	Majority class, LDR (baselines)	Term weights	Spanish & Catalan
[Lai et al. 2017]	SVM, logistic regression, decision tree, random forest, multinominal naïve Bayes, ensemble learner combining these classifiers, majority voting	Stylistic (word and character ngrams, POS tags, lemmas), structural (hashtags/mentions, hashtag frequencies, uppercase words, punctuation marks, numbers of words and characters), contextual (language, URL) features	Spanish & Catalan
[García and Flor 2017]	SVM and ANN	TF-IDF vectors of unigram and hashtag features	Spanish & Catalan
[Vinayakumar et al. 2017]	RNN, LSTM, GRU, and logistic regression	Word embeddings	Spanish & Catalan
[González et al. 2017]	SVM, LSTM, CNN, multilayer perceptron	Character and word ngrams, word embeddings vectors, character one-hot vectors, and a sentiment lexicon feature	Spanish
[Barbieri 2017]	FastText	Word embeddings considering subword information	Spanish & Catalan
[Swami et al. 2017]	SVM	Character (1-3) and word (1-5) ngrams, and stance indicative words	Spanish & Catalan
[Wojatzki and Zesch 2017]	SVM, LSTM, and a decision tree based hybrid system	Word (1-3) ngrams, character (2-4) ngrams, and word embeddings	Spanish & Catalan
[Ambrosini and Nicolo 2017]	LSTM, bidirectional LSTM, CNN	Word embeddings	Spanish & Catalan

Dilek Küçük and Fazli Can. 2020. Stance Detection: A Survey. *ACM Comput. Surv.* 53, 1, Article 12 (February 2020), 37 pages. https://doi.org/10.1145/3369026

Network Methods – Stance Propagation

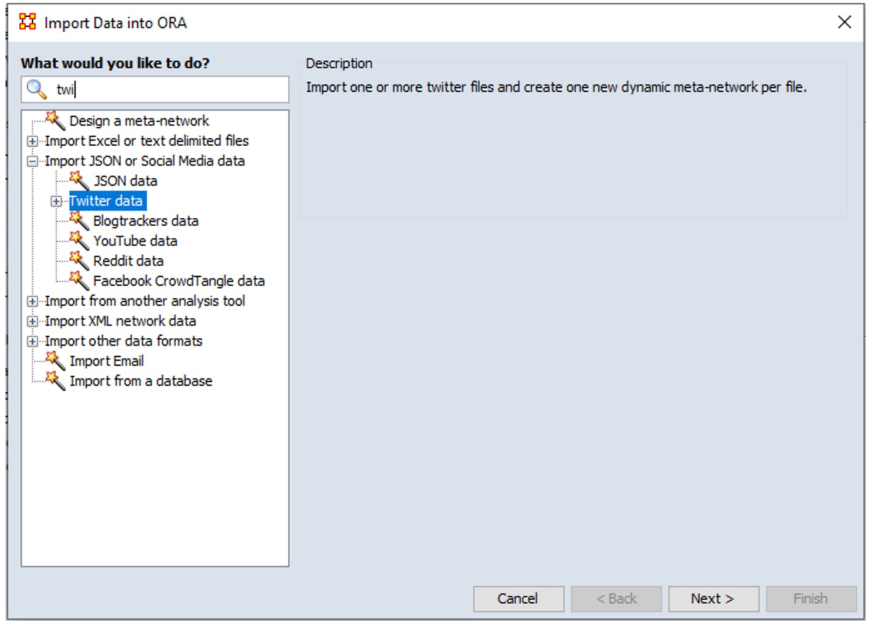


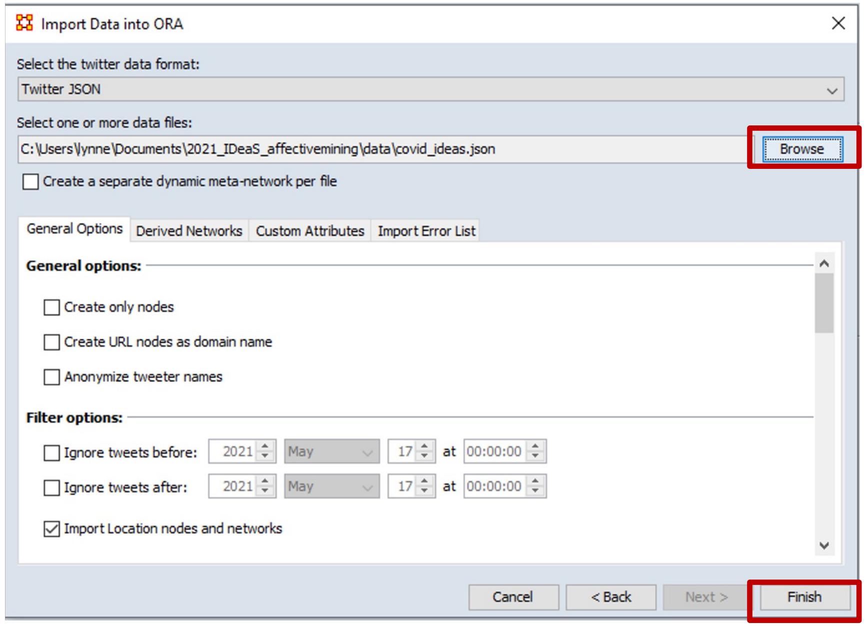
Kumar, S. (2020). *Social media analytics for stance mining a multi-modal approach with weak supervision* (Doctoral dissertation, Carnegie Mellon University).

Network Methods – Stance Propagation

- ORA Example on COVID vaccine data

1. Import Data

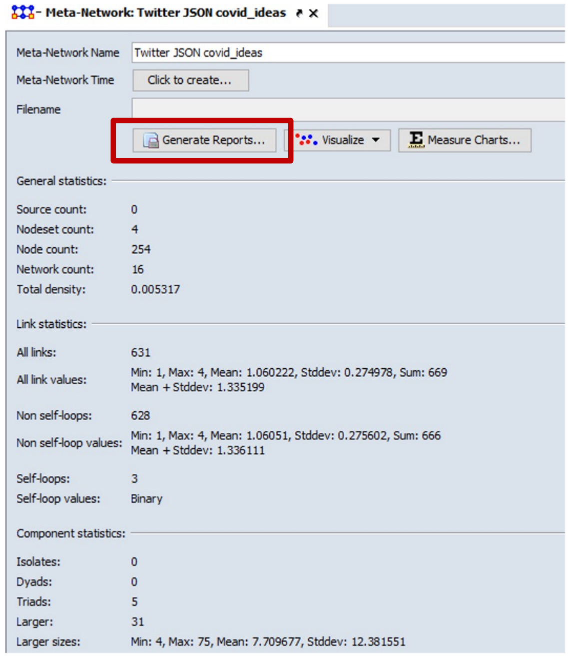




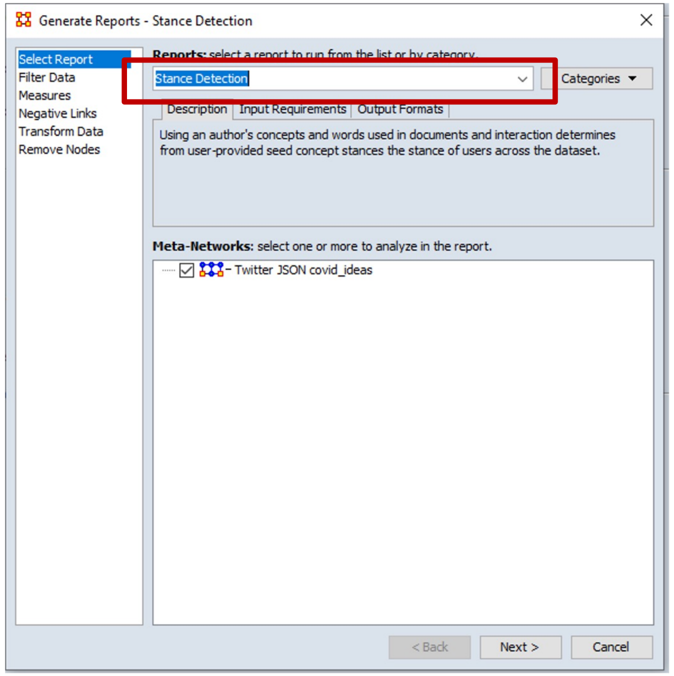
Network Methods – Stance Propagation

❑ORA Example on COVID vaccine data

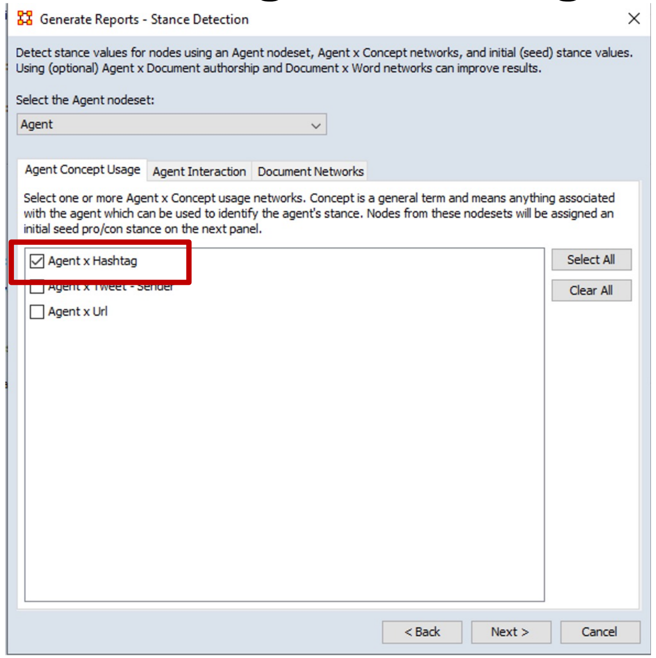
2. Generate Reports



Select Stance Detection



Select Agent x Hashtag

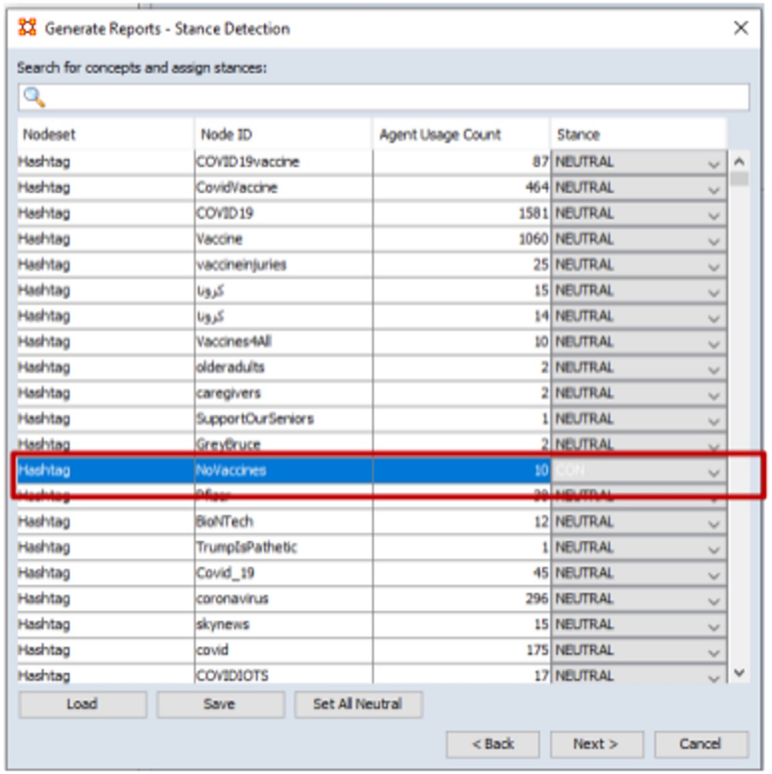


Carnegie Mellon University

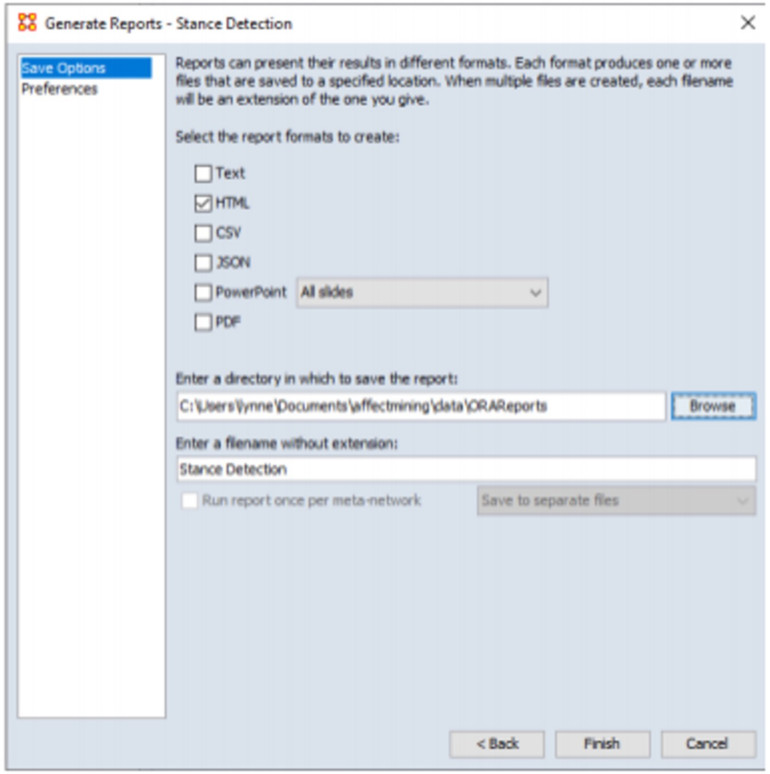
Network Methods – Stance Propagation

❑ ORA Example on COVID vaccine data

3. Assign Stance Values



Run!



Carnegie Mellon University

Network Methods – Stance Propagation

- ORA Example on COVID vaccine data

4. View Stance Detection Report

Hashtag Pro Stance

The pro-stance Hashtag nodes ranked by the confidence of the stance calculation.

If the node of interest has a higher than normal value (greater than 1 standard deviation(s) above the mean) the row is c
standard deviation of the mean. Finally, the row is colored blue if the node has a lower than normal value (less than one

Show 10 entries

Rank	Hashtag	Confidence
1	ACIP	1
2	COVID−19	1
3	CovidVaccines	1
4	FaceMasks	1
5	GTCB	1
6	GlobalTeamCoronaBusters	1
7	HealthyAtHome	1
8	IGotMyFluShot	1
9	SaveLives	1
10	TCB	1

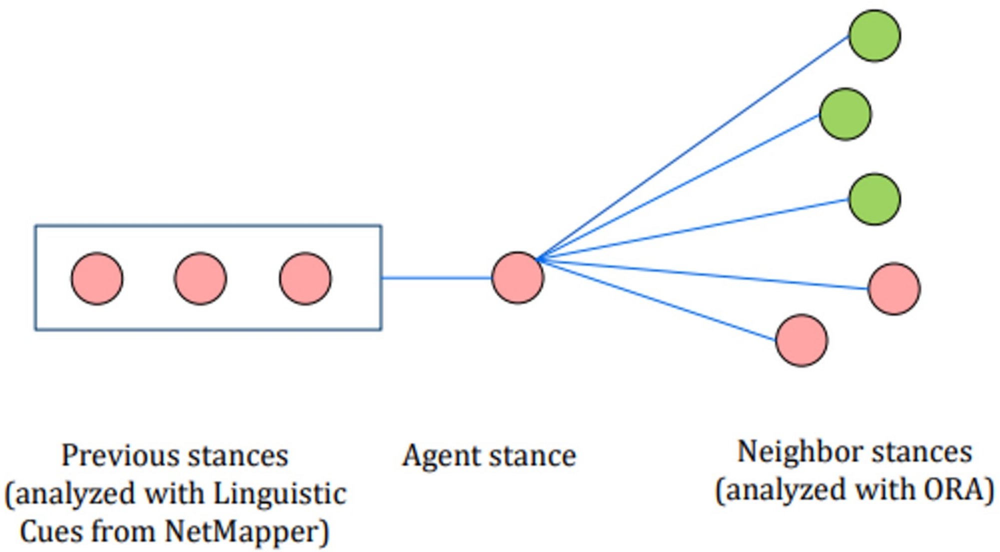
5. Agents & Tweets are annotated with stance & confidence

stance	stance-confi…
0	1
0	1
0	1
0	1
0	1
0	1
0	1
0	1
0	1
0	1
0	1

Carnegie Mellon University

Combining Linguistic & Network Methods

- Predicting stance flipping on Twitter using a Social Influence Model (ie provax > antivax)



Predicting Stance Flips with a Social Influence Model

- Predicting stance flipping using a Social Influence Model

(ie provax > antivax)

Agent stance value $$Y_{\{agent\}} = \gamma C X_* B_*$$

Stance Strength (Conviction of agent)

Connection (Proportion of neighbors that support agent's stance)

Variables (linguistic & network)

Coefficients (eg weights)

Neighbor stance value $$Y_{\{neighbor\}} = \gamma C X_* B_* + R$$

Reciprocity (Two way interaction between agents)

Predicting Stance Flips with a Social Influence Model

- Predicting stance flipping using a Social Influence Model (ie provax > antivax)

$$Y_{\{agent\}} = \gamma C X_* B_* \qquad Y_{\{neighbor\}} = \gamma C X_* B_* + R$$

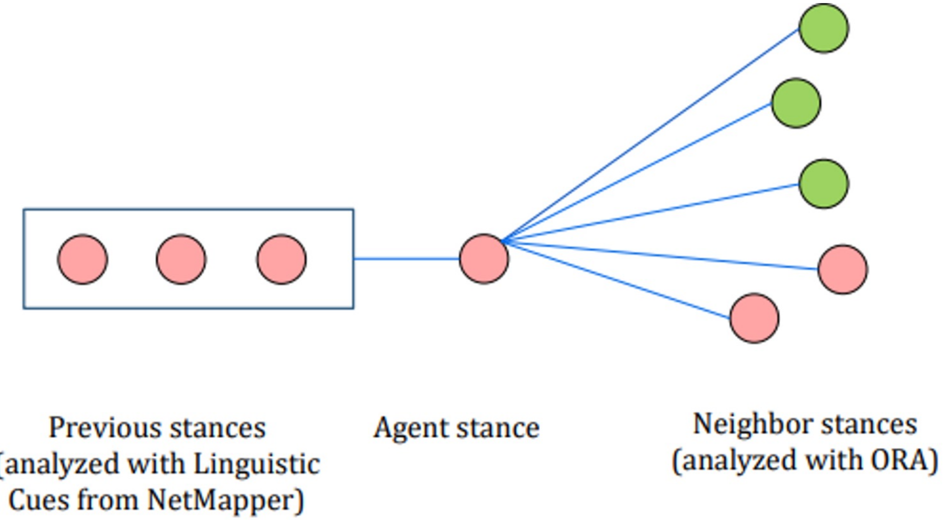
Influence on an agent

$$I = \alpha[\underbrace{\sum_{i=0}^{n} Y_{\{1st\ \deg\ neighbors\}}}_{\text{Influence from 1st degree neighbors (1 hop away)}} + \underbrace{\sum_{i=0}^{n}\sum_{j=0}^{m} \beta Y_{\{2nd\ \deg\ neighbors\}}}_{\text{Influence from 2nd degree neighbors (2 hops away)}}], \alpha = \frac{1}{n}, \beta = \frac{1}{m}$$

Influence from 1st degree neighbors (1 hop away)

Influence from 2nd degree neighbors (2 hops away)

Predicting Stance Flips with a Social Influence Model

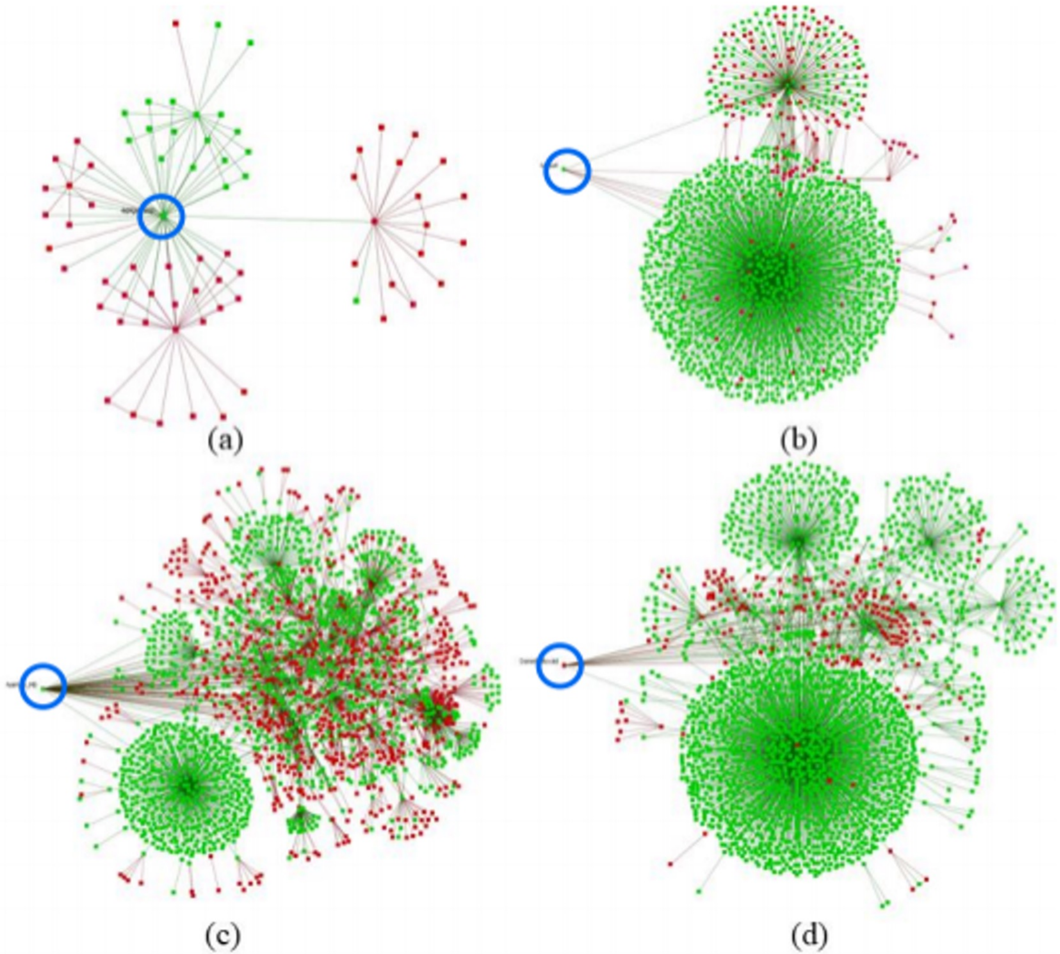


Model #	Model	Accuracy
Baseline	Decision Tree	0.53
Base - network	Base social influence model without network variables	0.47
Base - linguistic	Base social influence model without linguistic variables	0.55
Base Model 1	Base social influence model	0.50
Model 2	Model 1 + 2nd deg neighbor information	0.59
Model 3	Model 2 + stance strength	0.72
Model 4	Model 3 + connection	0.73
Model 5	Model 4 + reciprocity	0.86

Table 2: Results of Social Influence Models. The base social influence model is agent stance with 1st degree neighbor information.

Predicting Stance Flips with a Social Influence Model

- Model accurately predicts 86% of the stance flips
- Positive examples:



Green: provax
Red: antivax

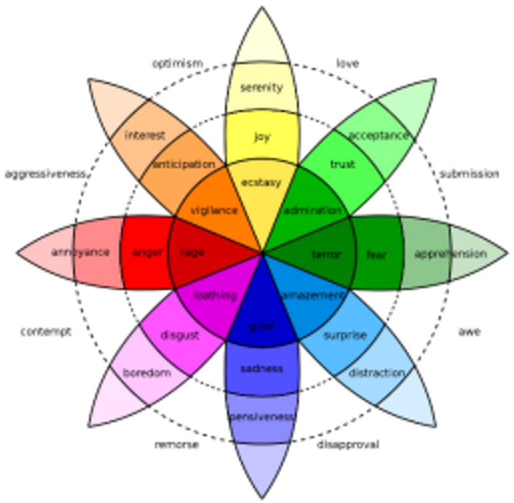
Circle: Agent in focus
Stance of agent depicted before flip

CASOS IDeaS Carnegie Mellon University

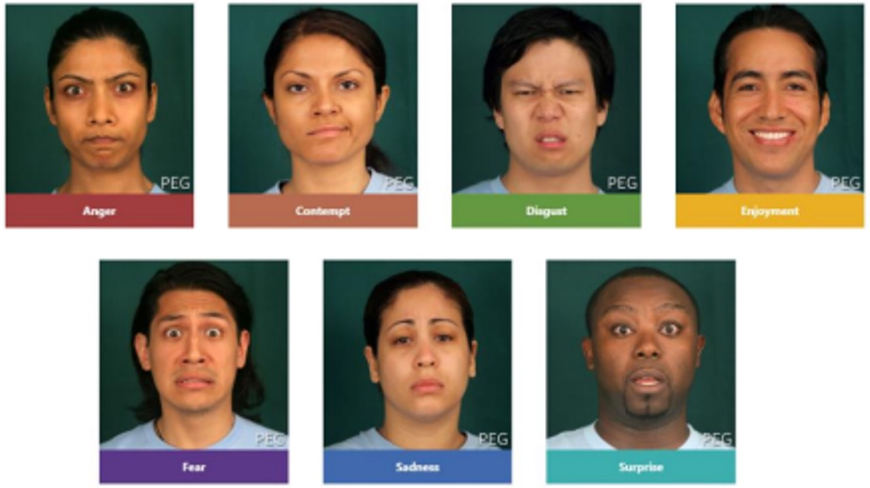
Affect

- ❑“emotion or attitude a speaker brings to an utterance” (Besiner, 1990)

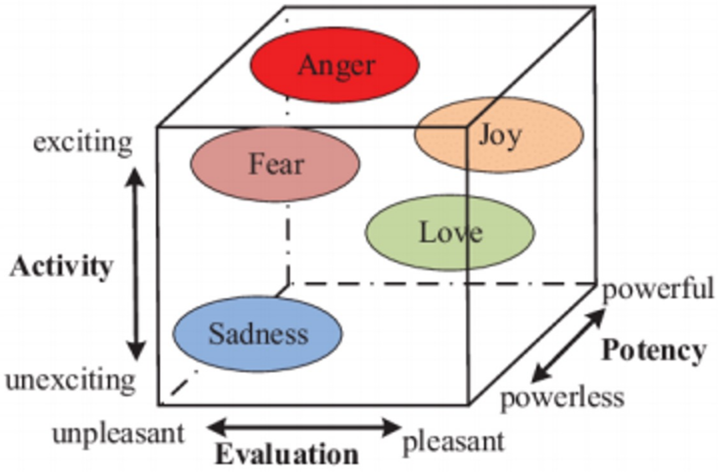
Putchnik



Ekman



Osgood



Carnegie Mellon University

Affect Control Theory

- Individuals maintain affective meanings (measured by EPA) through their actions and interpretations of events
- Deflections: distances in the EPA space, which can lead to actions
- Actions: a social behavior
- Emotions: events generate emotions for individuals
- Identity: who a person is (eg child, teacher)

Textual Affect Analysis

- Identifying dominant frames in climate change discourse
- EPA projections of frames
- Which frames are more dominant in climate change discourse ?
- What are the EPA projections of different frames
  - How does EPA projections of frames used in Climate Change news articles change with time
- Does frames with higher emotions lead to more reshare?

Carnegie Mellon University

Textual Affect Analysis

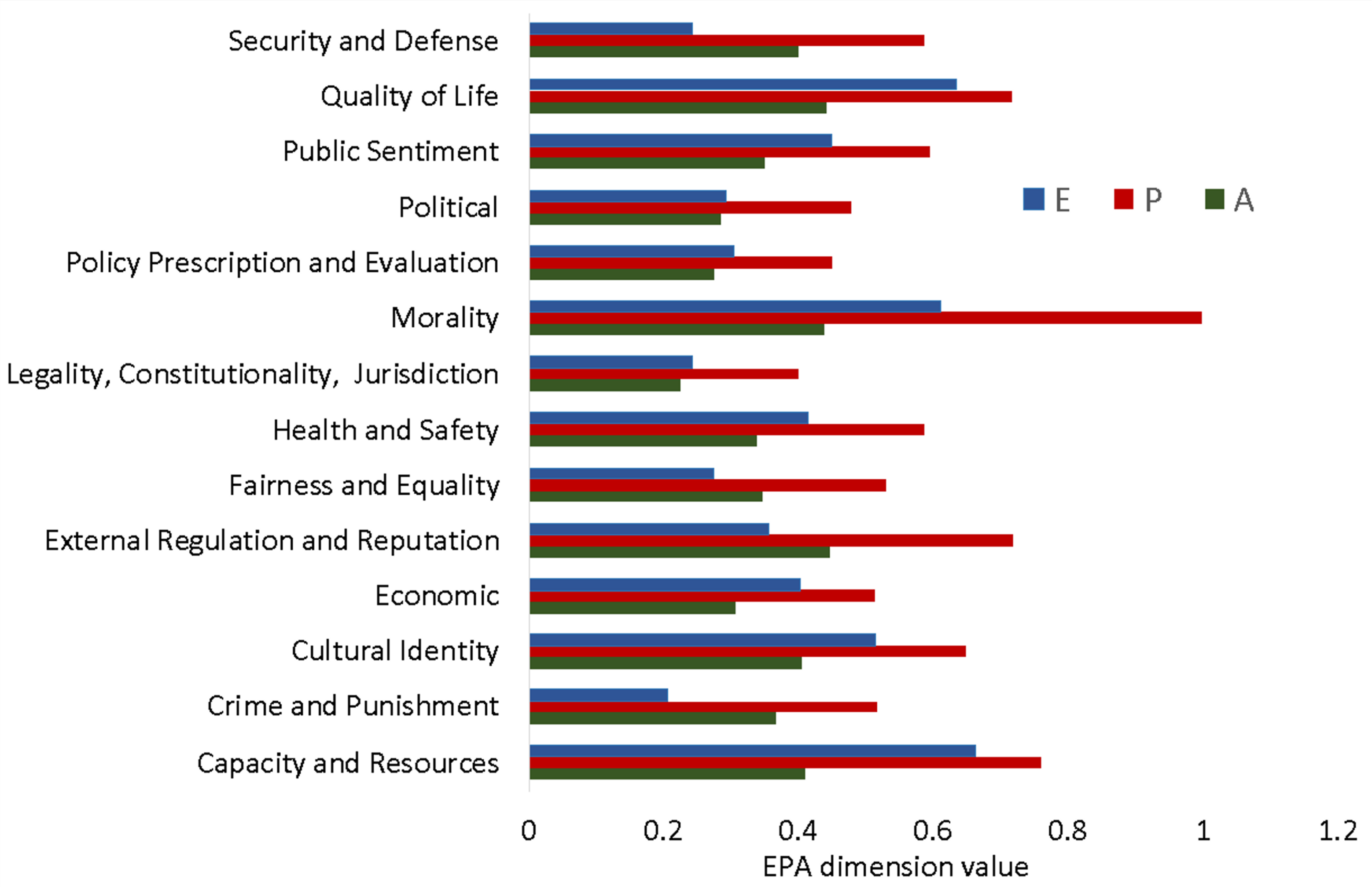
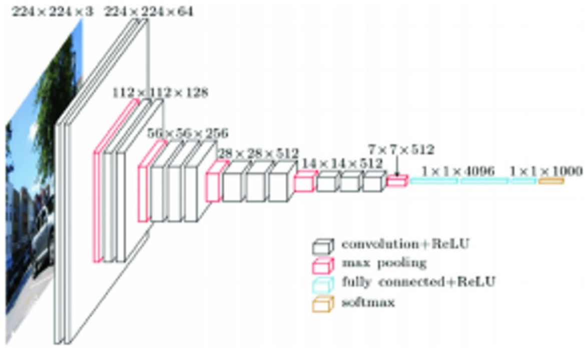


Image Affect Analysis

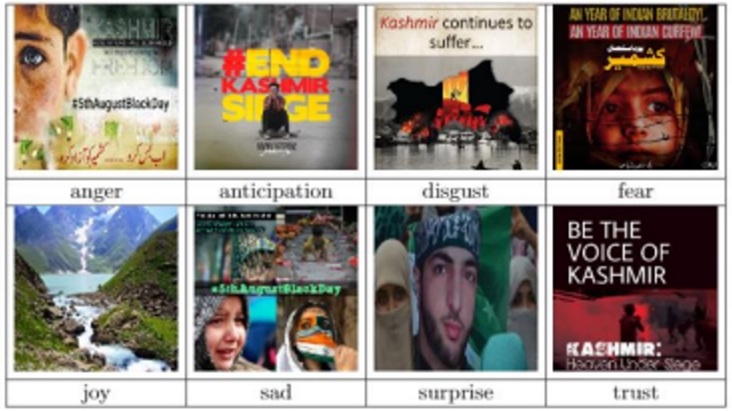
- ❑Image affect classifier based off Flickr images
- ❑Applied to recent social-cybersecurity event

1. Flickr Images (Putchnik emotional responses)

2. VGG Image Vector + CNN classifier



3. Annotate images in event



References – Sentiment Analysis

- Ajao, O., Bhowmik, D., & Zargari, S. (2019, May). Sentiment aware fake news detection on online social networks. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2507-2511). IEEE.
- Bhutani, B., Rastogi, N., Sehgal, P., & Purwar, A. (2019, August). Fake news detection using sentiment analysis. In *2019 Twelfth International Conference on Contemporary Computing (IC3)* (pp. 1-5). IEEE.
- Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. KDD exploration newsletter, 2017.
- Kumar, S. (2020). *Social media analytics for stance mining a multi-modal approach with weak supervision* (Doctoral dissertation, Arizona State University).
- Limeng Cui and Suhang Wang Dongwon Lee. Same: Sentiment-aware multi-modal embedding for detecting fake news. 2019.
- Shu, K., Wang, S., Lee, D., & Liu, H. (2020). Mining disinformation and fake news: concepts, methods, and recent advancements. In *Disinformation, Misinformation, and Fake News in Social Media* (pp. 1-19). Springer, Cham.

References – Stance Analysis

- Dilek Küçük and Fazli Can. 2020. Stance Detection: A Survey. *ACM Comput. Surv.* 53, 1, Article 12 (February 2020), 37 pages. https://doi.org/10.1145/3369026
- Hardalov, M., Arora, A., Nakov, P., & Augenstein, I. (2021). A survey on stance detection for mis-and disinformation identification. *arXiv preprint arXiv:2103.00242*. https://arxiv.org/pdf/2103.00242.pdf
- Friedkin, N. E., & Johnson, E. C. (1999). Influence networks and opinion change. *Advances in Group Processes, 16*(1), 1-29.
- Ghanem, B., Rosso, P., & Rangel, F. (2018). Stance detection in fake news a combined feature representation. In *Proceedings of the first workshop on fact extraction and VERification (FEVER)* (pp. 66-71). https://www.aclweb.org/anthology/W18- 5510.pdf
- S. Kumar (2020). "Social media analytics for stance mining a multi-modal approach with weak supervision," Ph.D. dissertation, Carnegie Mellon University.

References – Affect Mining

- Beskow, D. M., & Carley, K. M. (2019). *Social cybersecurity: an emerging national security requirement*. Carnegie Mellon University Pittsburgh United States.
- Besnier, N. (1990). Language and Affect. *Annual Review of Anthropology, 19*, 419-451. Retrieved May 10, 2021, from http://www.jstor.org/stable/2155972
- Osgood, Charles; May, W. H.; Miron, M. S. (1975), Cross-Cultural Universals of Affective Meaning, Urbana: University of Illinois Press.
- Ghanem, B., Rosso, P., & Rangel, F. (2020). An emotional analysis of false information in social media and news articles. *ACM Transactions on Internet Technology (TOIT), 20*(2), 1-18.
- Heise, D. R. (1987). Affect control theory: Concepts and model. *Journal of Mathematical Sociology, 13*(1-2), 1-33.

Carnegie Mellon University

For More Information

- Lynnette Ng & lynnetteng@cmu.edu
- Director Kathleen M. Carley - kathleen.carley@cs.cmu.edu
- IDeaS website - https://www.cmu.edu/ideas-social-cybersecurity/
- CASOS website – http://www.casos.cs.cmu.edu/
- Social Cybersecurity Working Group - http://social-cybersecurity.org
- Facebook: @IDeasCMU
- Twitter: @IDeaSCMU
- YouTube: IDeaS Center
- Email-Distro Lists

Network Affect Analysis

- Identifying emotions for deterrence of disinformation campaign with BEND maneuvers
- **Grab something from Janice**

Beskow, D. M., & Carley, K. M. (2019). *Social cybersecurity: an emerging national security requirement.* Carnegie Mellon University Pittsburgh United States.