Combating Misinformation on Social Media: From Detection to Mitigation

Part 2: Misinformation Detection

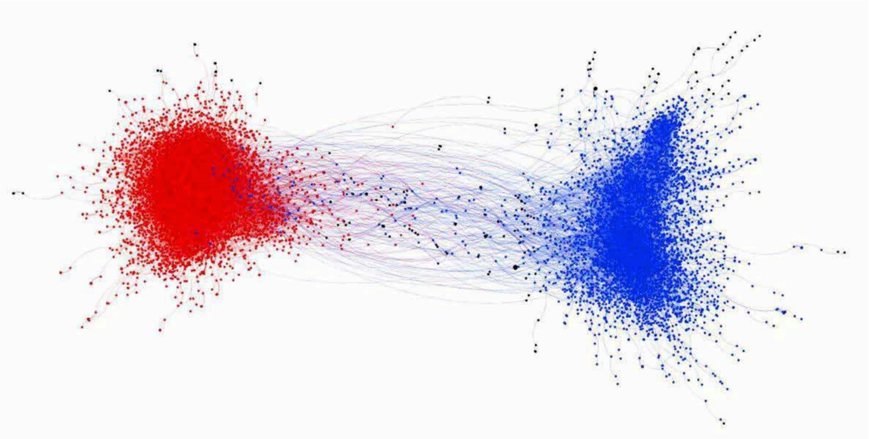
ALTA 2022 Tutorial

The University of Melbourne
Jey Han Lau

14 December 2022

Misinformation

- COVID-19 infodemic: Ilustrate how misinformation can cause panic and social division
- Social media the catalyst
  - “Free” platform for anyone to speak
  - Amplify misinformation propagation due to **filter bubbles** and **echo chambers** — driven by aggressive recommendation system



Why Detect?

- Call to develop automated methods to identify misinformation, so as to alert:
  - Users of misleading content
  - Journalist of stories that require fact-checking
  - Defense agencies of mass influence activities
- Ultimately, detection constitutes the first step to enable *other forms of corrective measures* for countering misinformation

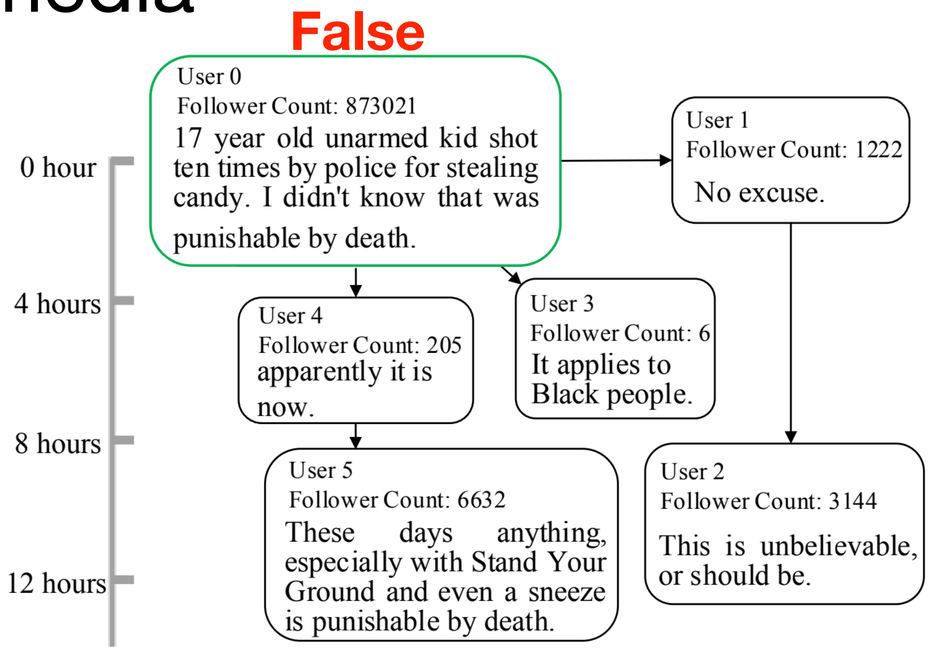
Two Tasks

Misinformation detection is typically framed as two tasks:

- **Automated fact checking**: given a claim, verify its *veracity* (truthfulness)

Sea-level rise is not accelerating → **False**

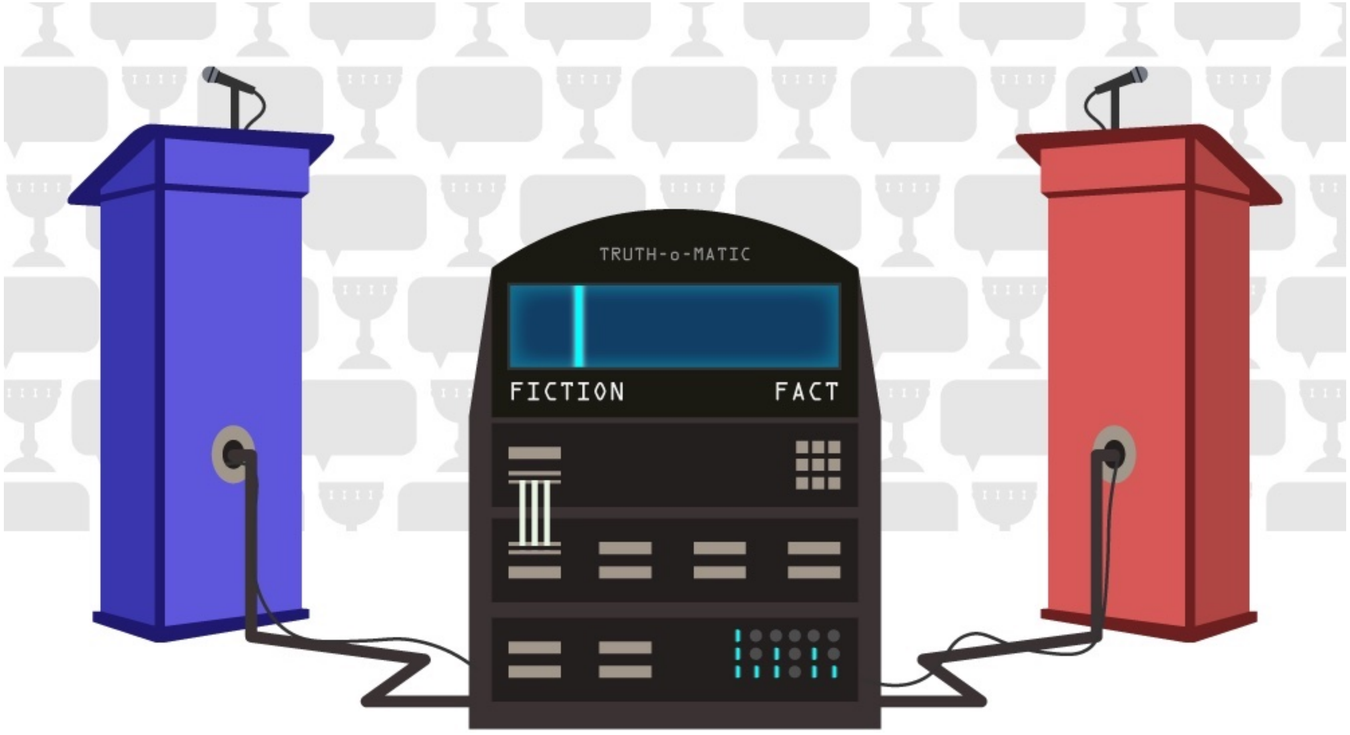
- **Rumour detection**: identify unverified stories on social media



Outline

- **Automated Fact Checking**
  - Task definition, representative methods, datasets
- **Rumour Detection**
  - Task definition, representative methods, datasets
- **Challenges**
  - Interpretability, multimodality, multilinguality, disinformation

Automated Fact Checking



Pipeline

Sea-level rise is not accelerating

↓

Evidence Retrieval

Climate-change driven accelerated sea-level rise detected in the altimeter era

Antarctica ice melt has accelerated by 280% in the last 4 decades

However scientists have found that ice is being lost, and at an accelerating rate

↓

Verdict Prediction

False / Refute

Pipeline

Sea-level rise is not accelerating

↓

Evidence Retrieval

Climate-change driven accelerated sea-level rise detected in the altimeter era

Antarctica ice melt has accelerated by 280% in the last 4 decades

However scientists have found that ice is being lost, and at an accelerating rate

↓

Verdict Prediction

False / Refute

Evidence Retrieval

- Aim: find additional information beyond the claim to determine whether it is truthful or not
- “Additional information” can be metadata, text data, structured data (e.g. table or tuples), images, etc
- Metadata such as publication source and author profile can be very informative
- But does not explain *how* or *why* a claim is (un)truthful

Textual Evidence

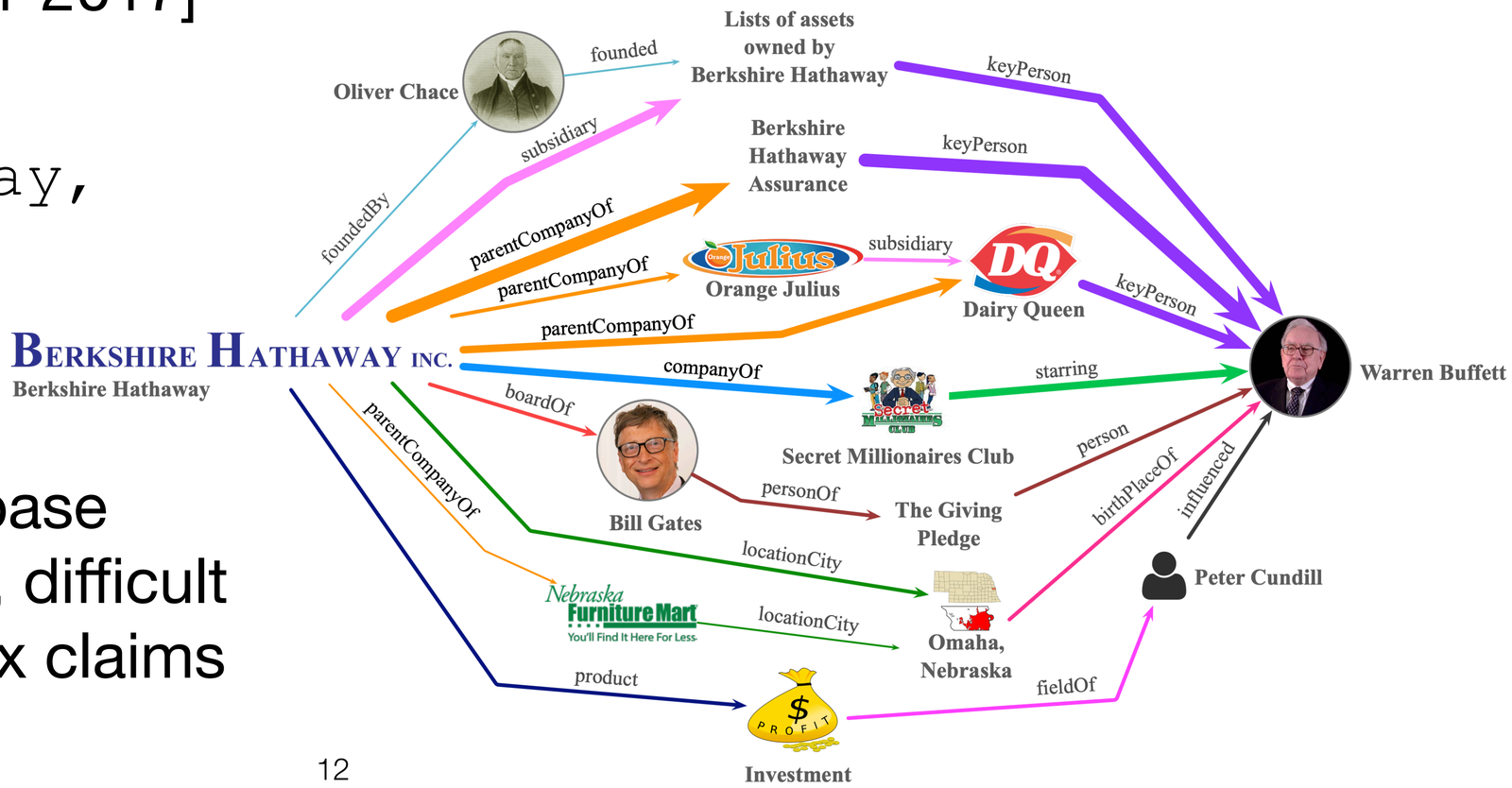
- External knowledge source = document collections
  - News headlines [Ferreira-2016]
  - News articles [Pomerleau-2017]
  - Wikipedia [Thorne-2018]
  - Web [Baly-2018]
  - Fact checking sites (e.g. Snope) [Hanselowski-2019]

Textual Evidence Retrieval

- Standard IR to retrieve documents/sentences [Thorne-2018b]
- Dense retrievers [Karpukhin-2020]
  - Encode query and documents using BERT
  - Distance between query and document = cosine similarity
  - Trained on QA datasets; objective = learn representations that produce high similarity for real QA pairs
- To further improve precision, use another model to re-rank retrieved evidence, e.g. stance detection [Thorne-2018b, Hanselowski-2019]

Structured Knowledge

- External knowledge source = knowledgebases (DBpedia)
- Given a claim expressed in a triple, find paths in the knowledgebase that support it [Shiralkar-2017]
- E.g.
  `(Berkshire Hathaway, keyPerson, Warren Buffett)`
- Drawback: knowledgebase is not always complete, difficult to handle more complex claims



Pipeline

Sea-level rise is not accelerating

↓

Evidence Retrieval

Climate-change driven accelerated sea-level rise detected in the altimeter era

Antarctica ice melt has accelerated by 280% in the last 4 decades

However scientists have found that ice is being lost, and at an accelerating rate

↓

Verdict Prediction

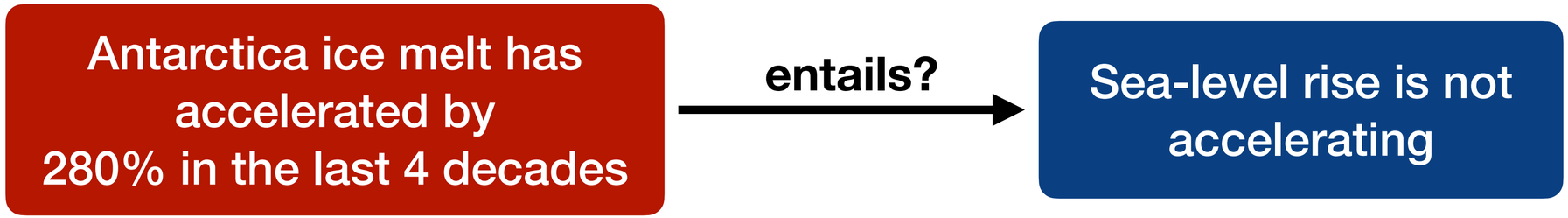
False / Refute

Verdict Prediction - Labels

- Binary: *true* or *false*
- Multiple classes: *true*, *mostly true*, *mostly false, false*
- Nowadays: *refute* or *support* or *not enough information* (based on evidence) [Thorne-2018]

Verdict Prediction - Methods

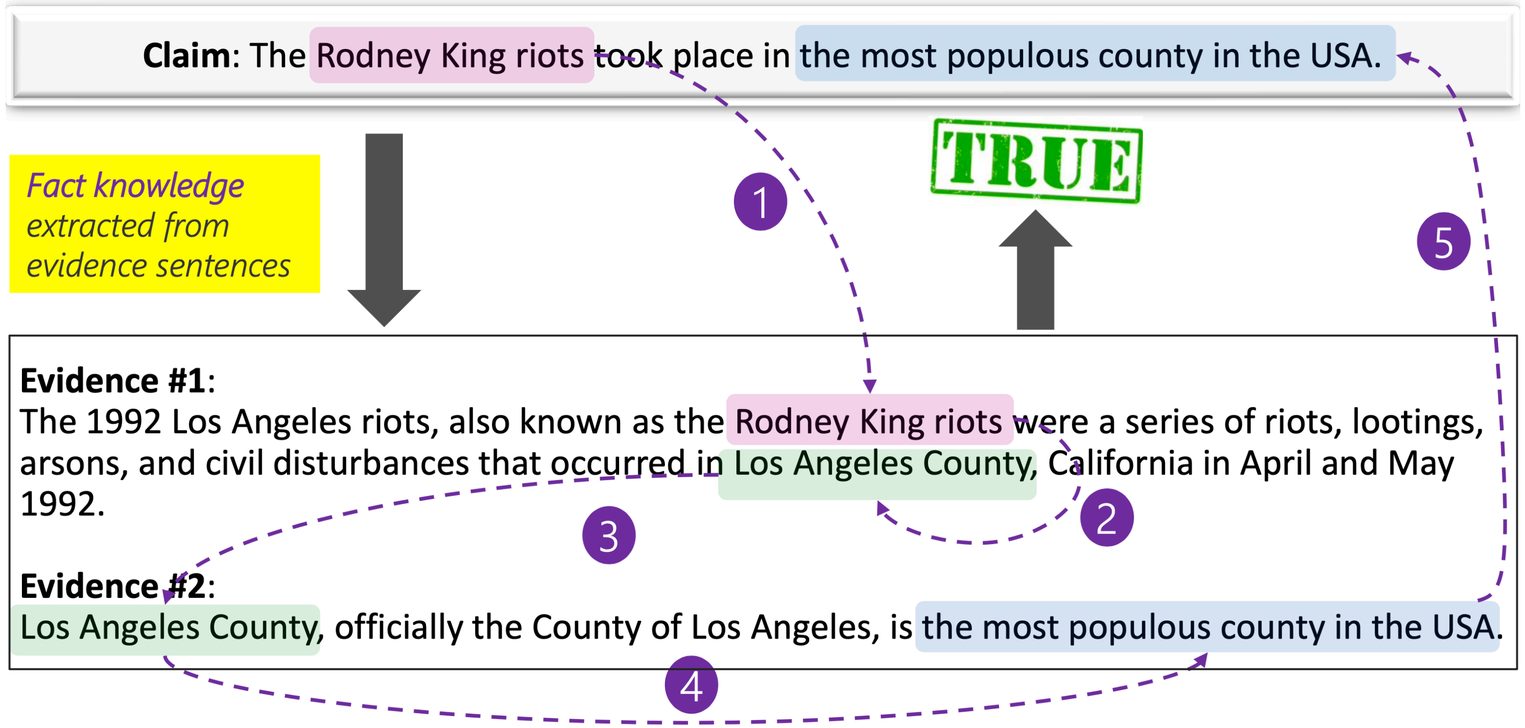
- Given evidence, verdict prediction can be seen as a form of textual entailment or natural language inference task



- But we usually have multiple pieces of evidence
- Concatenate them together into a single input string [Thorne-2018]

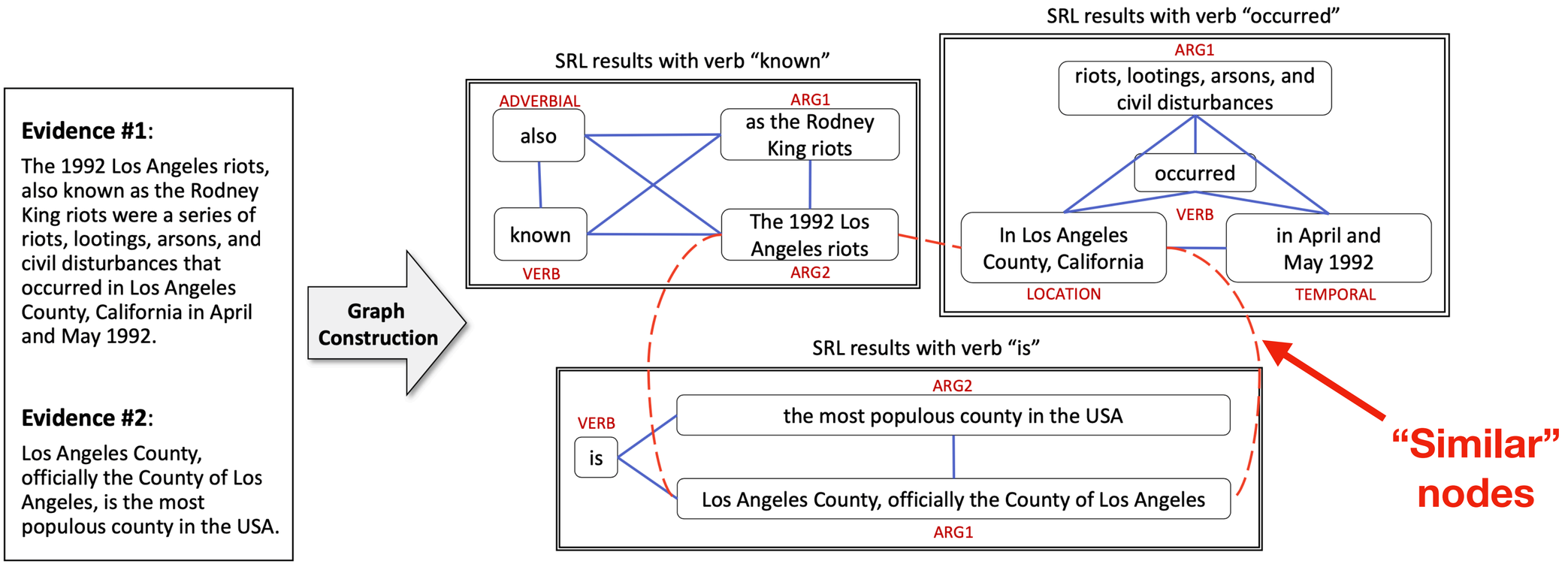
Graph-based Reasoning

- Allows verification of more complex claims where we need to combine information from multiple evidence



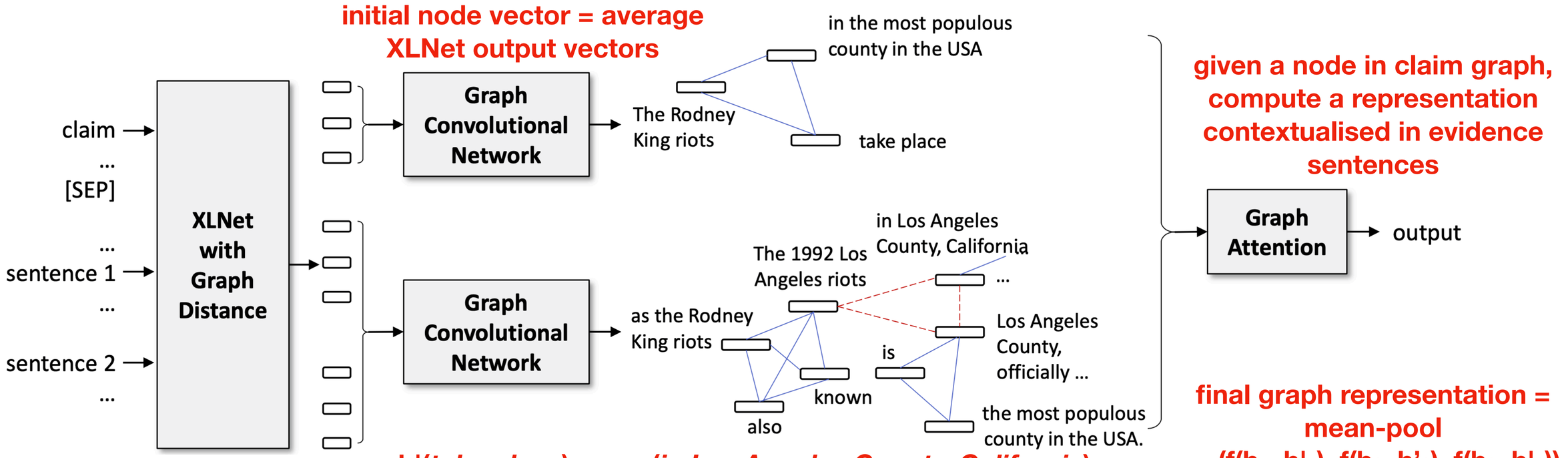
Graph-based Reasoning

- Reason over graphs based on semantic role labels [Zhong-2020]
- First step: construct the SRL graph



Graph-based Reasoning

- Second step: process SRL graphs from claim and evidence with graph networks to predict veracity



Automated Fact Checking: Datasets

Datasets - Natural Claims

Dataset	Input	#Inputs	Evidence	Verdict	Sources	Lang
CrimeVeri (Bachenko et al., 2008)	Statement	275	✗	2 Classes	Crime	En
Politifact (Vlachos and Riedel, 2014)	Statement	106	Text/Meta	5 Classes	Fact Check	En
StatsProperties (Vlachos and Riedel, 2015)	Statement	7,092	KG	Numeric	Internet	En
Emergent (Ferreira and Vlachos, 2016)	Statement	300	Text	3 Classes	Emergent	En
CreditAssess (Popat et al., 2016)	Statement	5,013	Text	2 Classes	Fact Check/Wiki	En
PunditFact (Rashkin et al., 2017)	Statement	4,361	✗	2/6 Classes	Fact Check	En
Liar (Wang, 2017)	Statement	12,836	Meta	6 Classes	Fact Check	En
Verify (Baly et al., 2018)	Statement	422	Text	2 Classes	Fact Check	Ar/En
CheckThat18-T2 (Barrón-Cedeño et al., 2018)	Statement	150	✗	3 Classes	Transcript	En
Snopes (Hanselowski et al., 2019)	Statement	6,422	Text	3 Classes	Fact Check	En
MultiFC (Augenstein et al., 2019)	Statement	36,534	Text/Meta	2–27 Classes	Fact Check	En
Climate-FEVER (Diggelmann et al., 2020)	Statement	1,535	Text	4 Classes	Climate	En
SciFact (Wadden et al., 2020)	Statement	1,409	Text	3 Classes	Science	En
PUBHEALTH (Kotonya and Toni, 2020b)	Statement	11,832	Text	4 Classes	Fact Check	En
COVID-Fact (Saakyan et al., 2021)	Statement	4,086	Text	2 Classes	Forum	En
X-Fact (Gupta and Srikumar, 2021)	Statement	31,189	Text	7 Classes	Fact Check	Many
cQA (Mihaylova et al., 2018)	Answer	422	Meta	2 Classes	Forum	En
AnswerFact (Zhang et al., 2020)	Answer	60,864	Text	5 Classes	Amazon	En
NELA (Horne et al., 2018)	Article	136,000	✗	2 Classes	News	En
BuzzfeedNews (Potthast et al., 2018)	Article	1,627	Meta	4 Classes	Facebook	En
BuzzFace (Santia and Williams, 2018)	Article	2,263	Meta	4 Classes	Facebook	En
FA-KES (Salem et al., 2019)	Article	804	✗	2 Classes	VDC	En
FakeNewsNet (Shu et al., 2020)	Article	23,196	Meta	2 Classes	Fact Check	En
FakeCovid (Shahi and Nandini, 2020)	Article	5,182	✗	2 Classes	Fact Check	Many

[Guo-2022]

Datasets - Artificial Claims

Dataset	Input	#Inputs	Evidence	Verdict	Sources	Lang
KLinker (Ciampaglia et al., 2015)	Triple	10,000	KG	2 Classes	Google/Wiki	En
PredPath (Shi and Weninger, 2016)	Triple	3,559	KG	2 Classes	Google/Wiki	En
KStream (Shiralkar et al., 2017)	Triple	18,431	KG	2 Classes	Google/Wiki/WSDM	En
UFC (Kim and Choi, 2020)	Triple	1,759	KG	2 Classes	Wiki	En
LieDetect (Mihalcea and Strapparava, 2009)	Passage	600	✗	2 Classes	News	En
FakeNewsAMT (Pérez-Rosas et al., 2018)	Passage	680	✗	2 Classes	News	En
FEVER (Thorne et al., 2018a)	Statement	185,445	Text	3 Classes	Wiki	En
HOVER (Jiang et al., 2020)	Statement	26,171	Text	3 Classes	Wiki	En
WikiFactCheck (Sathe et al., 2020)	Statement	124,821	Text	2 Classes	Wiki	En
VitaminC (Schuster et al., 2021)	Statement	488,904	Text	3 Classes	Wiki	En
TabFact (Chen et al., 2020)	Statement	92,283	Table	2 Classes	Wiki	En
InfoTabs (Gupta et al., 2020)	Statement	23,738	Table	3 Classes	Wiki	En
Sem-Tab-Fact (Wang et al., 2021)	Statement	5,715	Table	3 Classes	Wiki	En
FEVEROUS (Aly et al., 2021)	Statement	87,026	Text/Table	3 Classes	Wiki	En
ANT (Khouja, 2020)	Statement	4,547	✗	3 Classes	News	Ar
DanFEVER (Nørregaard and Derczynski, 2021)	Statement	6,407	Text	3 Classes	Wiki	Da

[Guo-2022]

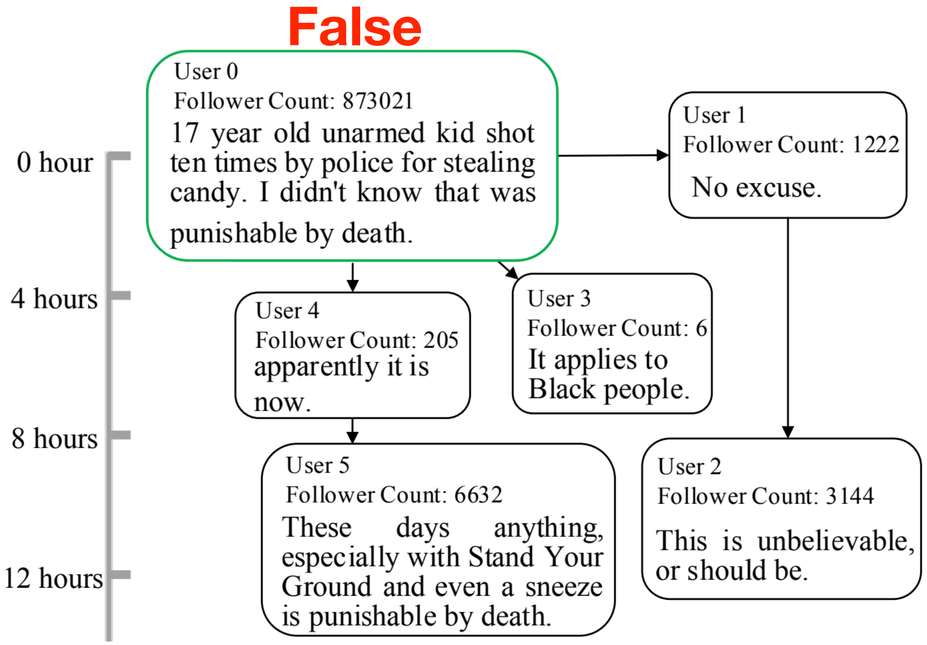
Discussion

- Enabled the development of large/complex models for automated fact checking
- Inspired a new wave of fact check datasets based on Wikipedia
- Applicability to fact checking real-world claims remains to be seen...

Rumour Detection

Rumour Detection

- Identify unverified stories that are spread on social media
- Uses **social features** for detection
  - e.g. comments/reactions and patterns of spread
- Verdict classes:
  - *rumour* vs. *non-rumour*
  - *true* vs. *false* vs. *unverified*



Early work

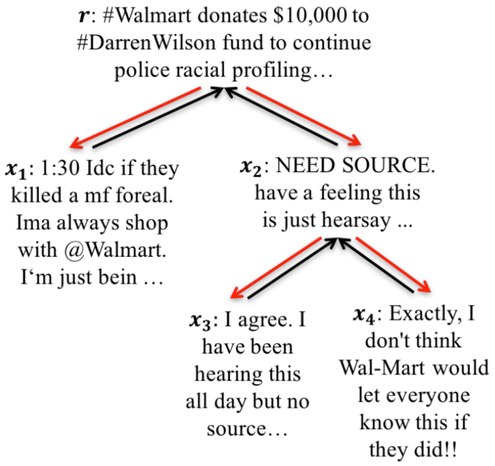
- [Ma-2015] train supervised classifiers using hand-engineered features from post content, user profile and propagation pattern
- A story = a set of tweets related to a topic (e.g. *Hillary Clinton announces 2016 campaign for president*)

Table 1: Description of features $f_{t,k}$ on microblogs from time 0 to time interval t of an event

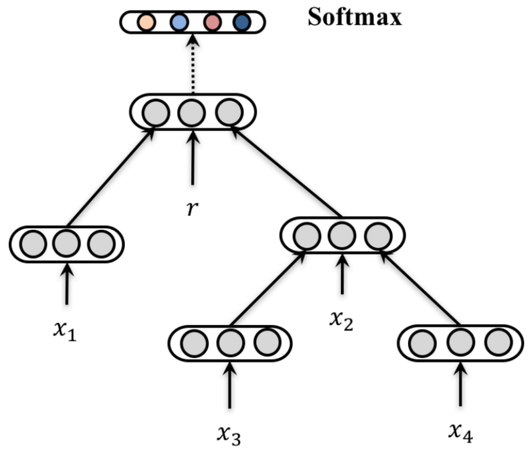
Content-based features
LDA-based topic distribution of microblogs with 18 topics [10]
Average length of microblogs [2]
of positive (negative) words in microblogs [2]
Average sentiment score of microblogs [2, 10]
% of microblogs with URL [2, 10, 11]
% of microblogs with smiling (frowning) emoticons [2]
% of positive (negative) microblogs [2]
% of microblogs with the first-person pronouns [2]
% of microblogs with hashtags [2, 11]
% of microblogs with @ mentions [2]
% of microblogs with question marks [2]
% of microblogs with exclamation marks [2]
% of microblogs with multiple question/exclamation marks [2]
User-based features
% of users that provide personal description [2, 10, 11]
% of users that provide personal picture in profile
% of verified users [2, 10, 11]
% of verified users of each type, e.g., celebrities [10, 11]
% of male (female) users [10, 11]
% of users located in large (small) cities
Average # of friends of users [2, 10, 11]
Average # of followers of users [2, 10, 11]
Average # of posts of users [2, 10, 11]
Average days users' accounts exist since registration [2, 10, 11]
Average reputation score of users (i.e., followers/followees ratio)
Diffusion-based features
Average # of retweets [2, 10, 11]
Average # of comments for Weibo posts [10, 11]
of microblogs [2]

Graph-based Method

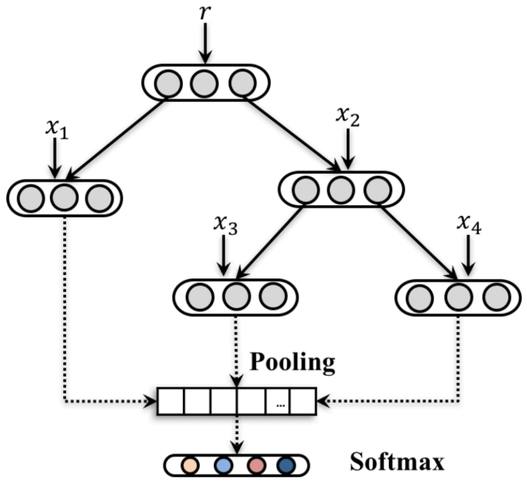
- [Ma-2018] explore recursive networks to model the conversations triggered by a tweet



(a) Bottom-up/Top-down tree



(b) Bottom-up RvNN model



(c) Top-down RvNN model

$$x_2' = \textbf{GRU}(x_2, \textbf{mean-pool}(x_3, x_4))$$

$$r' = \textbf{GRU}(r, \textbf{mean-pool}(x_1, x_2'))$$

$$x_1' = \textbf{GRU}(x_1, r)$$

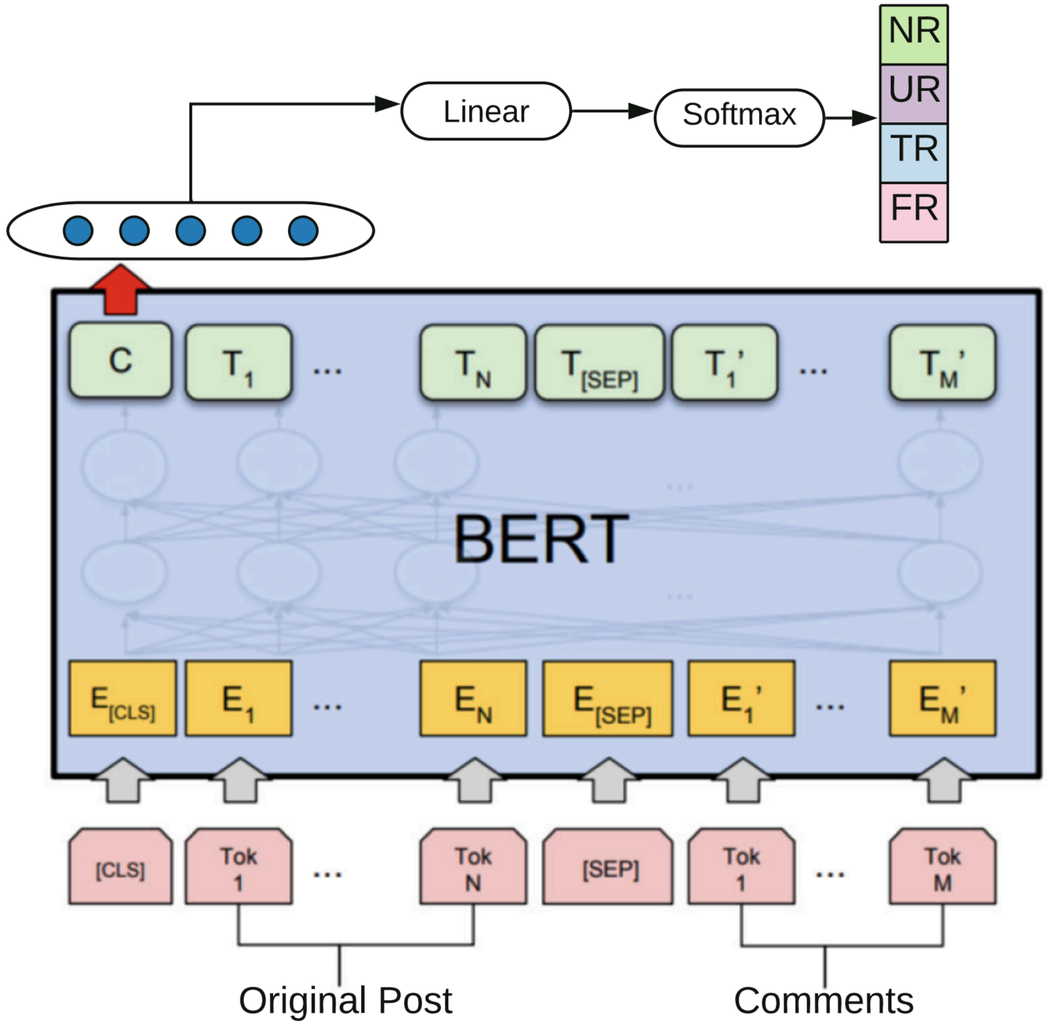
$$x_2' = \textbf{GRU}(x_2, r)$$

$$x_3' = \textbf{GRU}(x_3, x_2')$$

$$x_4' = \textbf{GRU}(x_4, x_2')$$

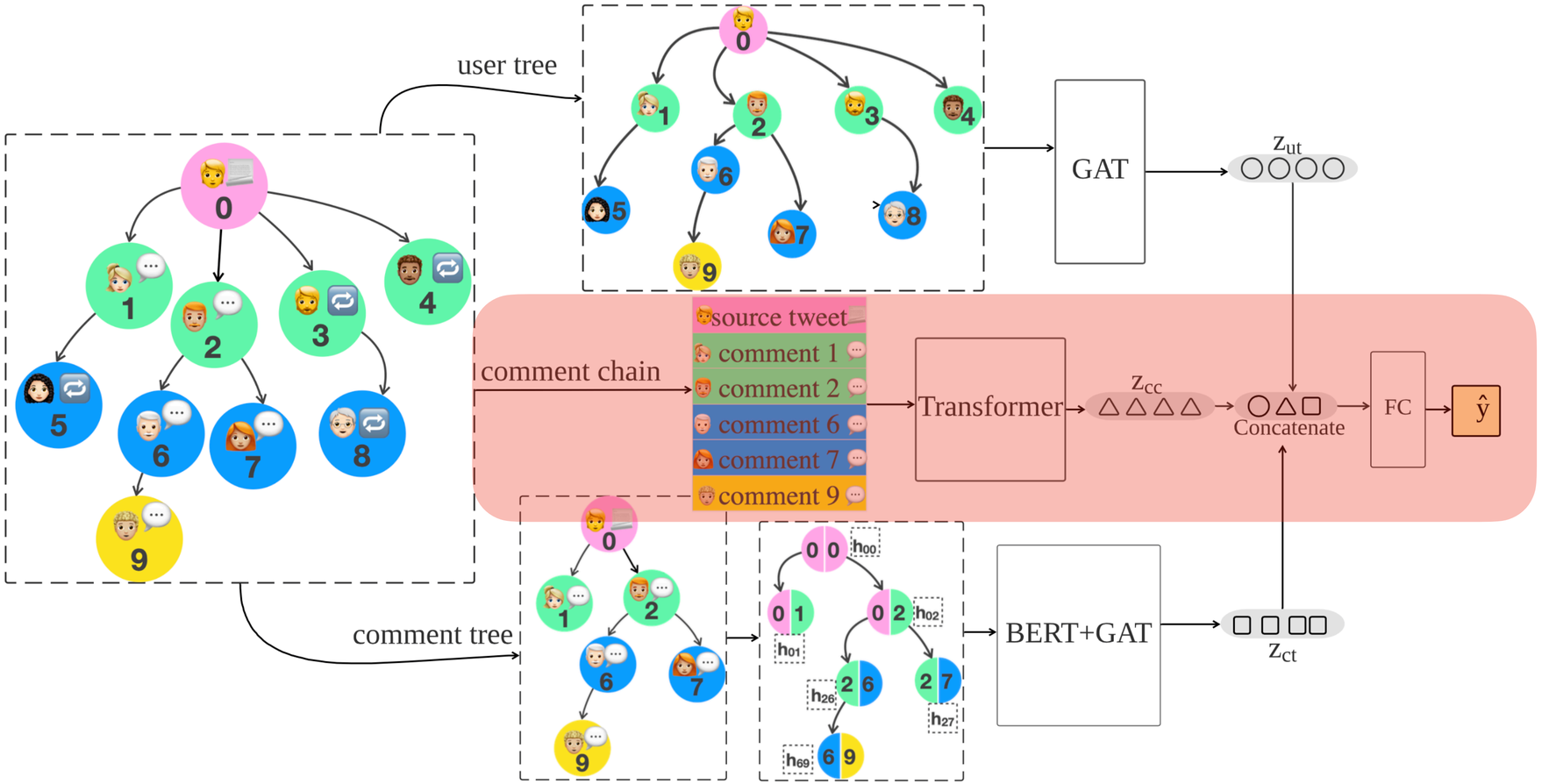
BERT-based Method

- [Tian-2020] use BERT to model the conversation as a chain
- Capture the order of appearance, but not reply-to structure
- Tried fine-tuning BERT first for stance prediction before rumour detection and saw minor benefits



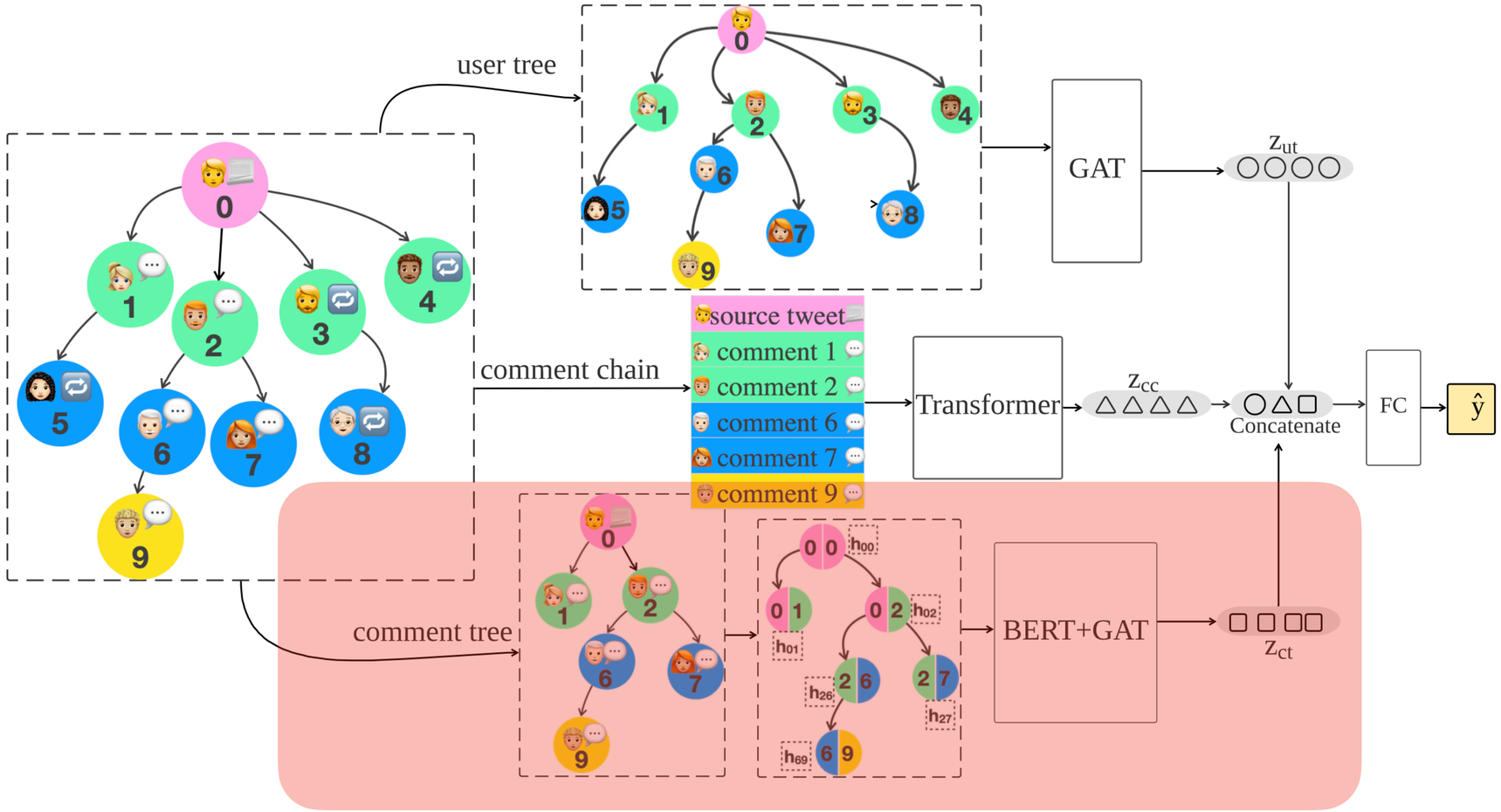
DUCK

- DUCK: Rumour detection with user and comment networks [Tian-2022]
- Captures *who* engages with a story (user network)
- Captures *how* they react to it (comment network)
  - Combination of BERT and graph networks


user tree
comment chain
comment tree
source tweet
comment 1
comment 2
comment 6
comment 7
comment 9
GAT
Transformer
BERT+GAT
Concatenate
FC
z_{ut}
z_{cc}
z_{ct}
$\hat{y}$

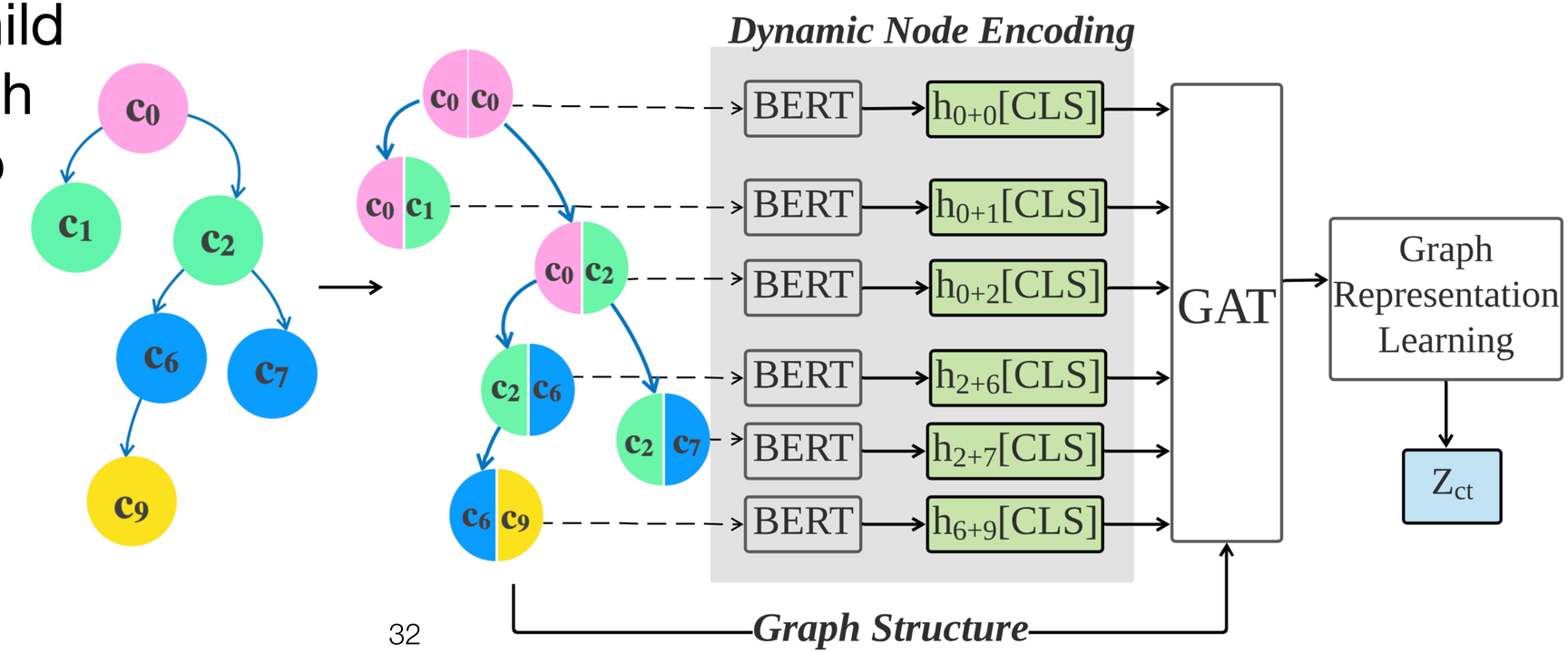
Comment Chain

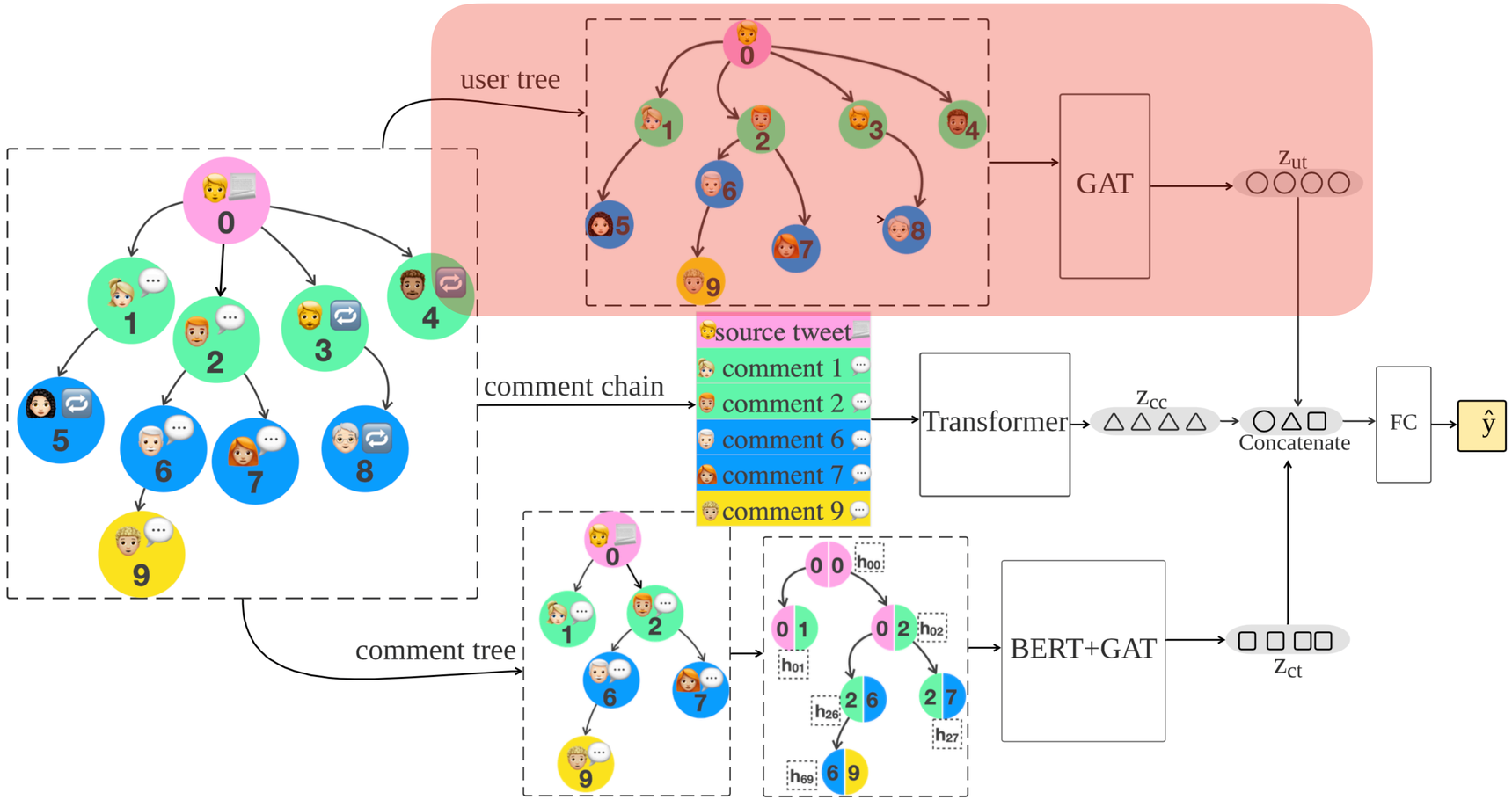
- Models the comments as a stream.
- **One-tier transformer**: concatenate source post + comments and feed to BERT
- **Longformer**: Same but use longformer instead to accommodate much longer sequence (no truncation of late comments)
- **Two-tier transformer**: first level (BERT) handles the concatenated string, second level takes in the [CLS] vector (from the first level) for each post


user tree
comment chain
comment tree
source tweet
comment 1
comment 2
comment 6
comment 7
comment 9
GAT
z_{ut}
Transformer
z_{cc}
Concatenate
FC
$\hat{y}$
BERT+GAT
z_{ct}

Comment Tree

- Models the conversation structure
- Use both pretrained language models and graph
- Key idea: use BERT to model a parent-child posts
- Once each parent-child is encoded, use graph attention networks to model the structure




user tree
GAT
z_{ut}
source tweet
comment 1
comment 2
comment 6
comment 7
comment 9
comment chain
Transformer
z_{cc}
Concatenate
FC
ŷ
comment tree
BERT+GAT
z_{ct}

User Tree

- Use graph attention networks to model the user network
- Key question: how to initialise each user node
- **GAT$_{rnd}$**: initialise randomly
- **GAT$_{prf}$**: initialise based on user profile info (e.g. username, user description, number of posts)
- **GAT$_{prf+rel}$**: initialise using representations learned by a variational graph autoencoder based on user profile **and** their social network (“followers”)

Findings

- Comment Chain: Two-tier transformer works best
- Comment Tree: Using BERT to process parent-child posts is helpful
- User Tree: Incorporating social relations gives substantial performance gain
- Full model creates a new state-of-the-art for rumour detection across multiple datasets in different languages

Rumour Detection: Datasets

Rumour Detection - Datasets

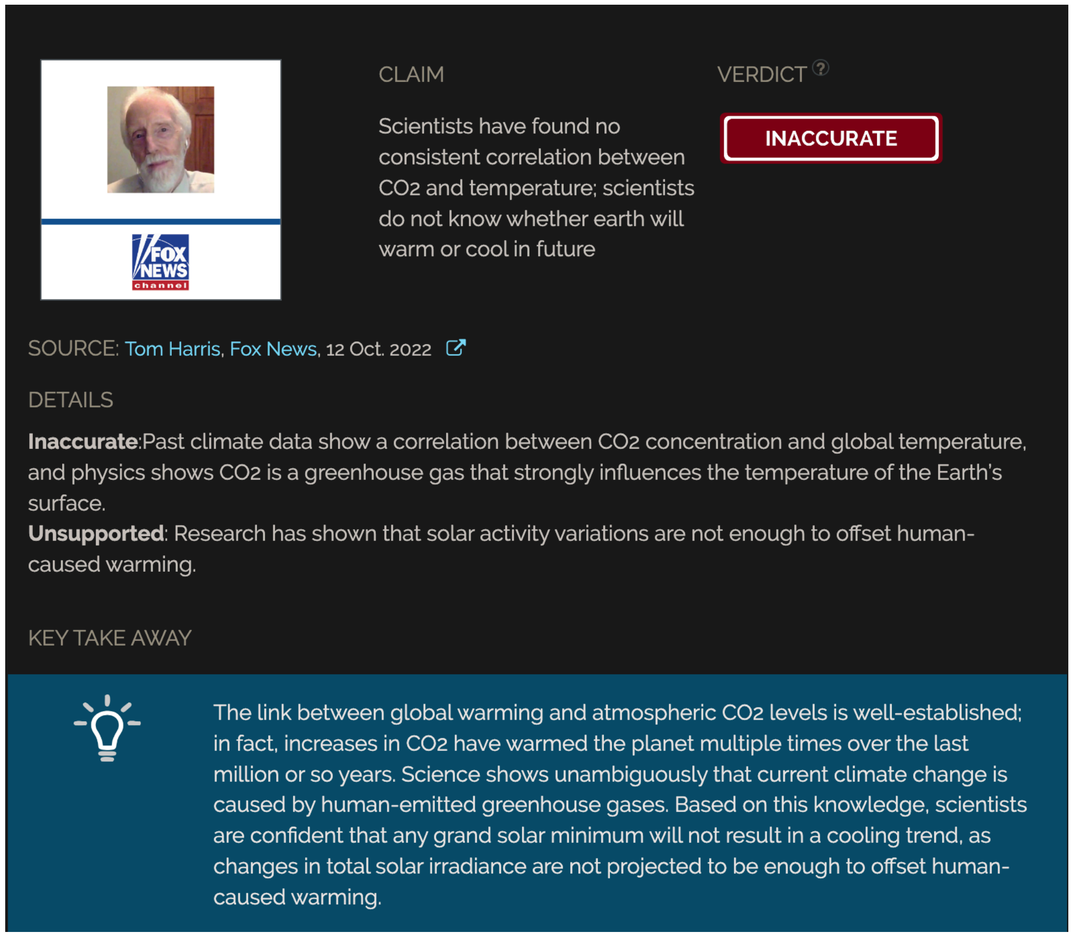
Dataset	Domain	#Inputs	#Users	Reactions	Verdict	Language
Twitter15 [Ma-2016]	Tweet	1,490	426,501	Comment IDs, Retweet IDs	True, False, Unverified, Non-rumour	English
Twitter16 [Ma-2016]	Tweet	818	251,799	Comment IDs, Retweet IDs	True, False, Unverified, Non-rumour	English
Weibo [Ma02017]	Weibo	4,664	2,746,818	Metadata, Comments, User profiles	True, False	Chinese
PHEME II [Kochkina-2018]	Tweet	6,425	50,593	Metadata, Comments, Links, User profiles	Rumour, non-rumour	English
SemEval2019 [Gorrell-2019]	Tweet, Reddit	446	5,666	Metadata, Comments, User profiles	True, False, Unverified	English
CoAID [Cui-2020]	Tweet	143,009	114,484	Comment IDs	True, False	English, Spanish, Chinese, Portuguese

Credit: Lin Tian

Challenges

Interpretability

- Classifying a story as “false” is unlikely to be persuasive in real applications
- We also need to provide some **explanation** or **justification** why the story is false (or true)



ClimateFeedback.org

Justification Production

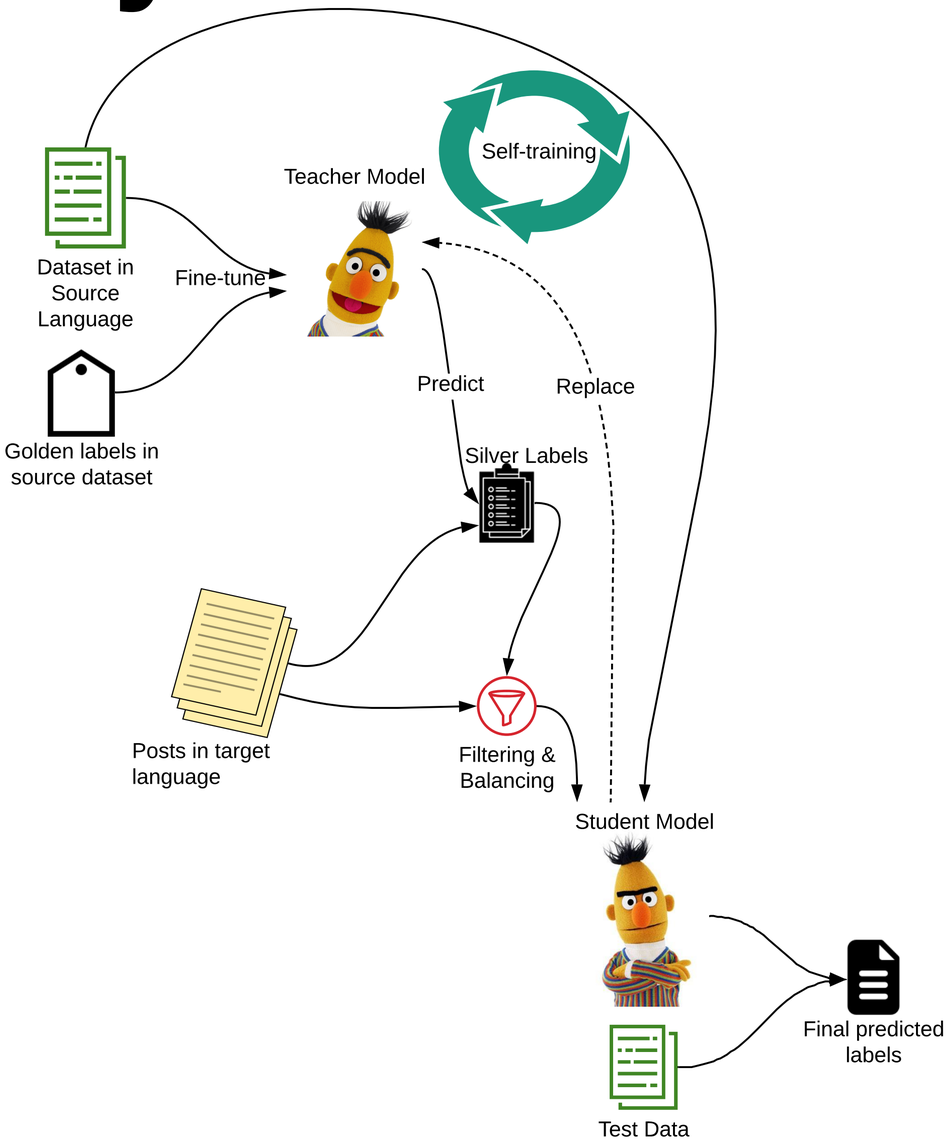
- Most fact checking or rumour detection models are black-box models; not inherently interpretable
- Examine the attention weights to highlight important words in posts and salient user features [Shu-2019, Lu-2020]
  - Problematic as studies found that removing words with high attention has little effect on the final prediction [Pruthi-2020]
- Frame justification production as a generation task [Xing-2022]
  - Generated justification may not be faithful to its prediction

Multimodality

- Information to detect misinformation comes in different modality: tables, images, videos, network propagation
- NLP studies largely focus on using just text for detection
- Most datasets focus on providing text as evidence; exceptions:
  - FEVEROUS includes tables as evidence [Aly-2021]
- That said, if tweet IDs are provided then images can be recovered (though many links may no longer be valid)

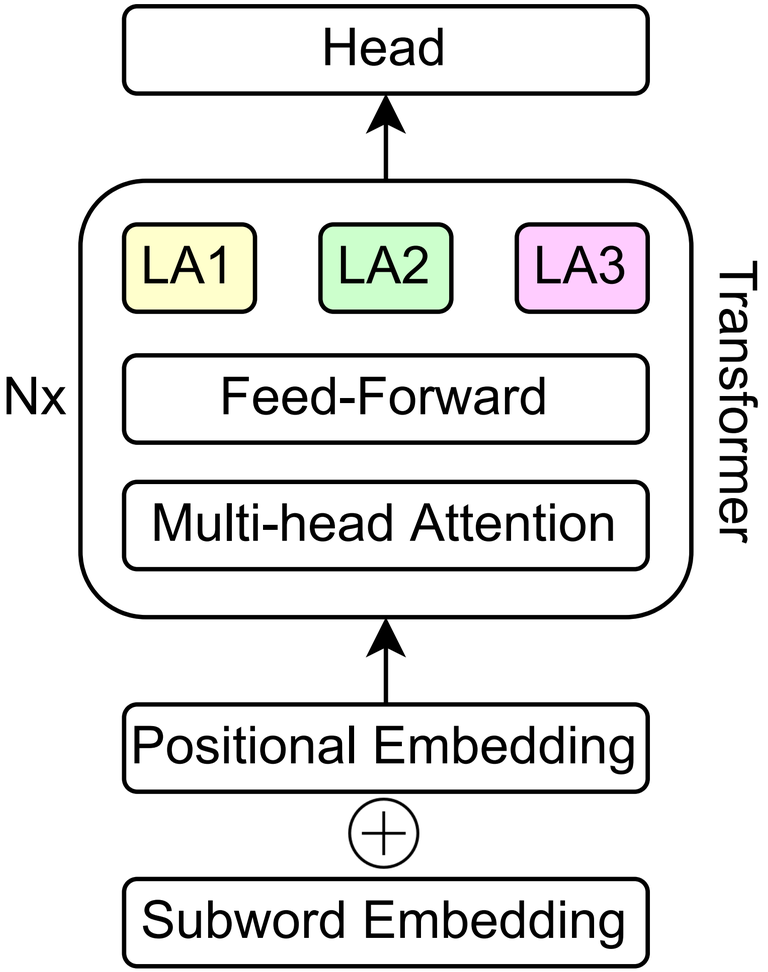
Multilinguality

- Vast majority of datasets and studies focus on English
  - Election misinformation crafted to target a particular community in the US and Australia [NYT-2022, Guardian-2022]
- [Tian-2021] explore self-training to zero-shot transfer monolingual detection system to other languages



Rumour Detection via Zero-shot Cross-lingual Transfer Learning

- Works for bilingual detection, but won't scale for true multilingual detection for many languages
- Due to "curse of multilinguality"
  - Performance degrades when adapted to many languages because of model has fixed capacity
- Multilingual pretrained models with language-specific adapters [Pfeiffer-2020]



Disinformation

- **Misinformation:** general term that refers to any false or inaccurate information
- **Disinformation:** misinformation created *deliberately* to deceive
- Current research doesn’t distinguish them, as it’s difficult to determine intent

Twitter Information Operations Archive

- Data released by twitter than collects a set of users suspected of engaging in *mass influence campaigns* [Twitter-2018]
- Rich dataset with over 10 campaigns from different countries and languages

Category	Event Time	Top-3 Hashtags
Iran-2018-Palestine	Feb 2018 – Aug 2018	#realiran #SavePalestine #InternationalQudsDay2018
Russia-2016-MAGA	Aug 2015 – Feb 2016	#MAGA #QAnon #ReleaseTheMemo
Iran-2018-Pakistan	May 2018 – Nov 2018	#pakonlinenews #SachTimes #DeleteIsrael
Venezuela-2018-Trump	Jun 2018 – Dec 2018	#TrumpTrain #MAGA #RT
Nigeria-2019-Racism	Aug 2019 – Feb 2020	#racism #BlackLivesMatter #PoliceBrutality
Iran-2020-BLM	Jul 2020 – Jan 2021	#black_lives_matter #Oscars #EEUU
GRU-2020-NATO	Jun 2020 – Dec 2020	#Syria #Idib #StopTerrorismInSyria
IRA-2020-Russia	Jun 2020 – Dec 2020	#valdaiclub #Russia #Ukraine
Uganda-2021-NRM	Jul 2020 – Jan 2021	#SecuringYourFuture #M7UGsChoice #StopHooligansim
China-2021-Xinjiang	Jul 2020 – Jan 2021	#Xinjiang #XinjiangOnline #Urumqi

Summary

- Misinformation detection as:
  - Automated fact checking
  - Rumour detection
- Great surveys on this topic: [Guo-2022], [Zubiaga-2018]

References

- **[Aly-2021]:** “FEVEROUS: Fact Extraction and VERification over unstructured and structured information." Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks. 2021.
- **[Baly-2018]:** Baly, Ramy, et al. "Integrating stance detection and fact checking in a unified corpus." Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers). 2018.
- **[Cui-2020]:** Cui, Limeng, et al. "Coaid: Covid-19 healthcare misinformation dataset." arXiv preprint arXiv:2006.00885 (2020).
- **[Ferreira-2016]:** Ferreira, William, et al. "Emergent: a novel data-set for stance classification." Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2016.
- **[Gorrell-2019]:** Gorrell, Genevieve, et al. "SemEval-2019 task 7: RumourEval, determining rumour veracity and support for rumours." Proceedings of the 13th International Workshop on Semantic Evaluation. 2019.
- **[Guardian-2022]:** https://www.theguardian.com/australia-news/2022/may/09/unattributed-attack-ads-targeting-labor-on-chinese-language-wechat-fuel-fears-of-misinformation
- **[Guo-2022]:** Guo, Zhijiang, et al. "A survey on automated fact-checking.” Transactions of the Association for Computational Linguistics 10 (2022): 178-206.
- **[Hanselowski-2019]:** Hanselowski, Andreas, et al. "A richly annotated corpus for different tasks in automated fact-checking." Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL). 2019.
- **[Karpukhin-2020]:** Karpukhin, Vladimir, et al. "Dense passage retrieval for open-domain question answering." Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2020.
- **[Kochkina-2018]:** Kochkina, Elena, et al. "All-in-one: Multi-task learning for rumour verification." Proceedings of the 27th International Conference on Computational Linguistics. 2018.
- **[Lu-2020]:** Lu, Yi-Ju, and Cheng-Te Li. "GCAN: Graph-aware co-attention networks for explainable fake news detection on Social Media." Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020.
- **[Ma-2015]:** Ma, Jing, et al. "Detect rumors using time series of social context information on microblogging websites." Proceedings of the 24th ACM international on conference on information and knowledge management. 2015.
- **[Ma-2016]:** Ma, Jing, Wei Gao, and Kam-Fai Wong. "Detect rumors in microblog posts using propagation structure via kernel learning." Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017.
- **[Ma-2017]:** Ma, Jing, et al. "Detecting rumors from microblogs with recurrent neural networks." Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. 2016.
- **[Ma-2018]:** Ma, Jing, et al. "Rumor detection on twitter with tree-structured recursive neural networks." Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018.

References

- **[NYT-2022]:** https://www.nytimes.com/2022/10/12/business/media/midterms-foreign-language-misinformation.html
- **[Parikh-2016]:** Parikh, Ankur, et al. "A decomposable attention model for natural language inference." Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016.
- **[Pfeiffer-2020]:** Pfeiffer, Jonas, et al. "MAD-X: An Adapter-Based Framework for Multi-Task Cross-Lingual Transfer." Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2020.
- **[Pomerleau-2017]:** Pomerleau, Dean, et al. "The fake news challenge: exploring how artificial intelligence technologies could be leveraged to combat fake news". Fake News Challenge. 2017.
- **[Pruthi-2020]:** Pruthi, Danish, et al. "Learning to deceive with attention-based explanations." Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020.
- **[Shiralkar-2017]:** Shiralkar, Prashant, et al. "Finding streams in knowledge graphs to support fact checking." 2017 IEEE International Conference on Data Mining (ICDM). IEEE, 2017.
- **[Shu-2019]:** Shu, Kai, et al. "defend: Explainable fake news detection." Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. 2019.
- **[Thorne-2018]:** Thorne, James, et al. "FEVER: A large-scale dataset for Fact Extraction and VERification." Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). 2018.
- **[Thorne-2018b]:** Thorne, James, et al. "The Fact Extraction and VERification (FEVER) shared task." Proceedings of the First Workshop on Fact Extraction and VERification (FEVER). 2018.
- **[Tian-2020]:** Tian, Lin, et al. "Early detection of rumours on twitter via stance transfer learning." European Conference on Information Retrieval. Springer, Cham, 2020.
- **[Tian-2021]:** Tian, Lin, et al. "Rumour detection via zero-shot cross-lingual transfer learning." Proceedings of The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases. 2021
- **[Tian-2022]:** Tian, Lin, et al. "DUCK: rumour detection on social media by modelling user and comment propagation networks." Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2022.
- **[Twitter-2018]:** https://transparency.twitter.com/en/reports/moderation-research.html
- **[Xing-2022]:** Xing, Rui, et al. "Automatic explanation generation for climate science claims." Proceedings of The 20th Annual Workshop of the Australasian Language Technology Association. 2022.
- **[Zhong-2020]:** Zhong, Wanjun, et al. "Reasoning over semantic-level graph for fact checking." Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020.
- **[Zubiaga-2018]:** Zubiaga, Arkaitz, et al. "Detection and resolution of rumours in social media: A survey." ACM Computing Surveys (CSUR) 51.2 (2018): 1-36.

Thank You!