

Identifying Check-Worthy Claims and going beyond Textual Claims

Text-based Claim Detection

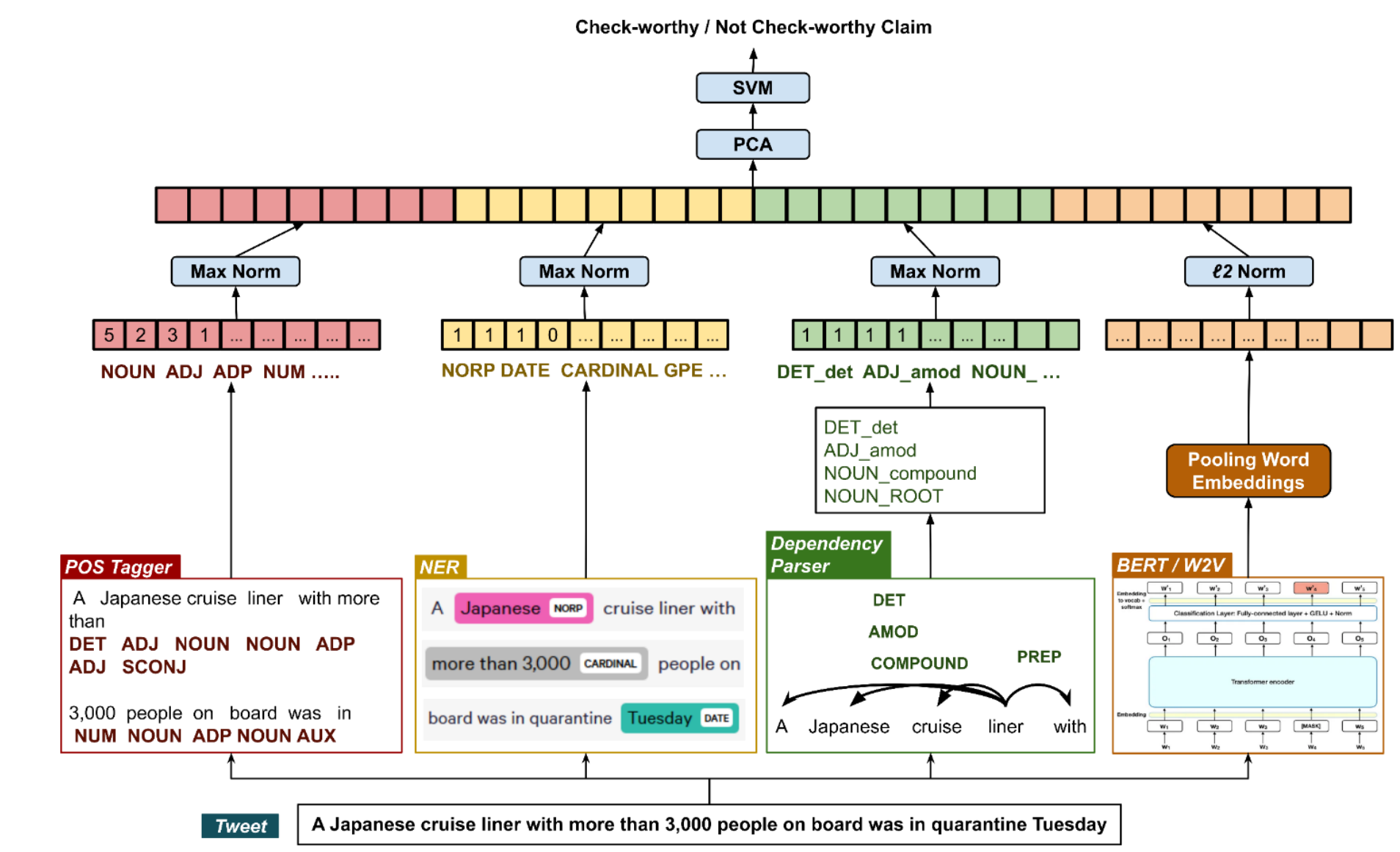
- Our approach combines [1]
 - Lexical features - named entities, PoS tags, dependency parsing
→ Provides clues for claim structure
 - Embeddings - Word2vec, BERT
→ Provides context & semantics
- **Check-worthy claims** can cause harm, spread widespread misinformation

*PoS: Parts of Speech
*BERT: Bidirectional Encoder Representations from Transformers

Not claims
Worldwide Geothermal Energy Potential: #climatechange #climateaction #environment #energy
If Bitcoin (BTC) Turns \$11,300 Into Support, Bull Run is Back On

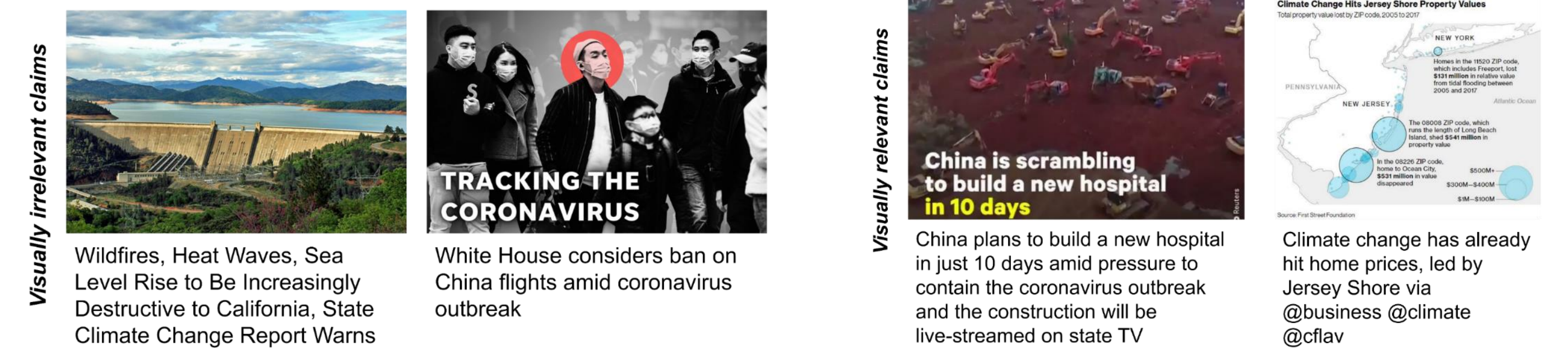
Claims but not check-worthy
The Carbon Bubble links the stock market to climate change. So what does that mean?
The Four Main Types of Cyberattack That Affect Data Center Uptime Here are the most common types of attack that bring down data centers [...]

Check-worthy claims
UK mobile operators seek 5G concession in exchange for rural broadband deal
Facebook vows to run on 100 percent renewable energy by 2020
Here's how AI identified the Coronavirus outbreak and alerted people before the UN
Merkel Pushing Back Against Higher EU Climate Change Target Forbes in 2014, the leaders of the 28 countries of the European Union [...]
China's Huawei to invest US\$800 million in new Brazil factory amid 5G push

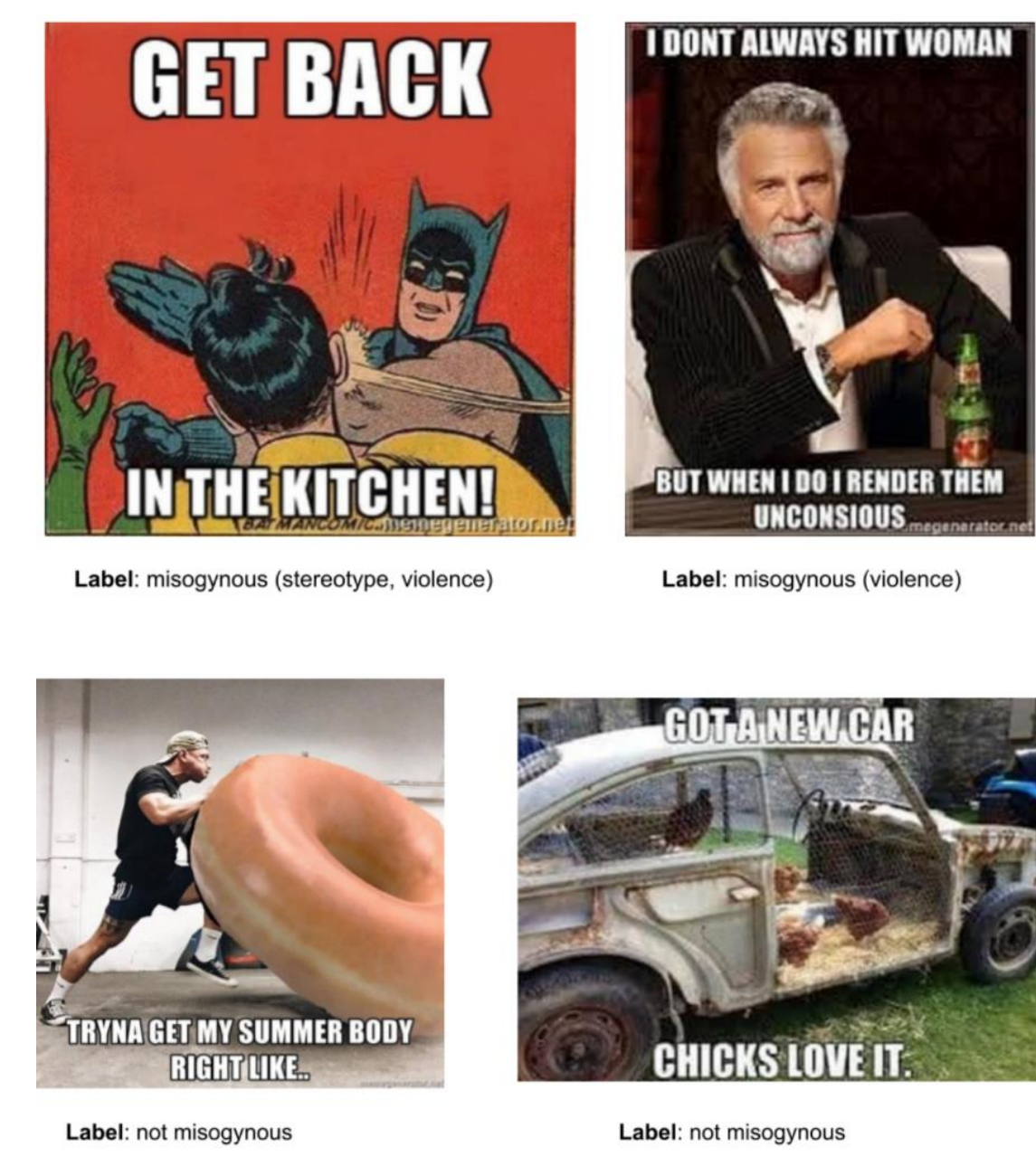
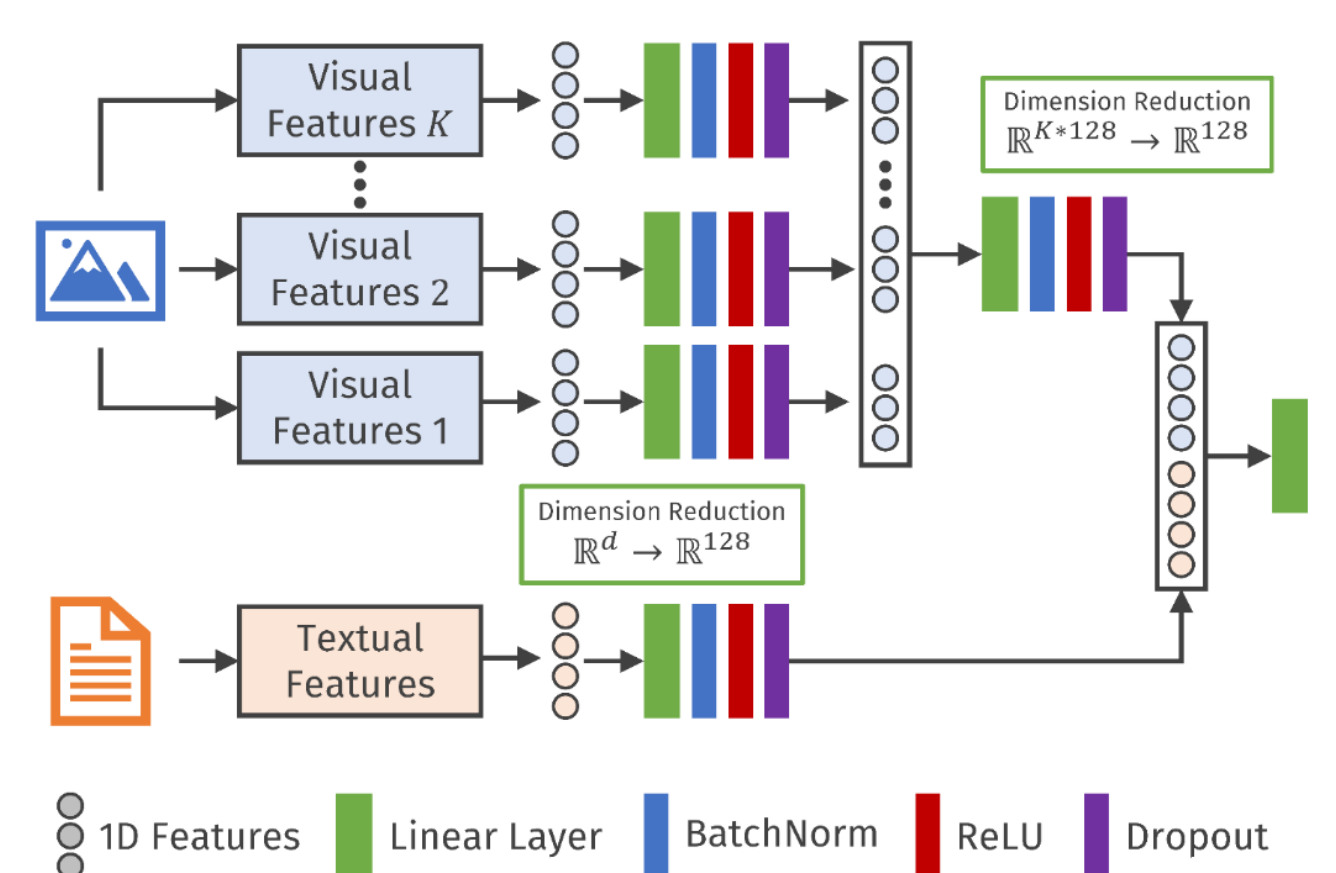
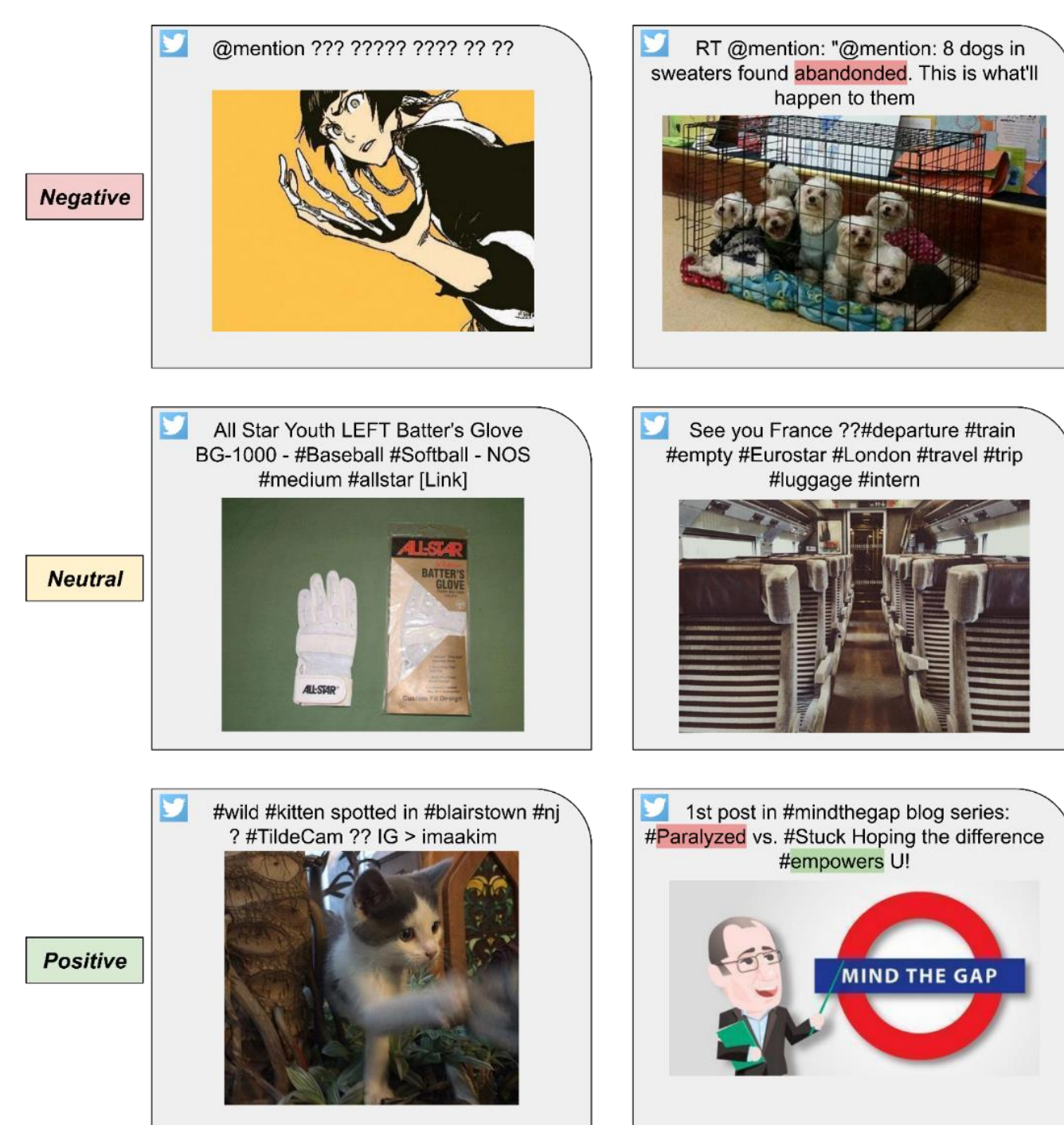


Multimodal Claim Detection

- Studied the role of images in claim detection [2]
→ Images important for importance, evidence & text-in-image claims
- Proposed a new dataset on multimodal claim detection [3]
→ Claims across multiple topics: COVID-19, Climate change, Technology
- Performed a fine-grained analysis of image-text relations
→ Understand semantics & variety of image-text pairs in claims vs not claims



Multimodal Sentiment and Hate Speech Detection



Multimodal sentiment detection

- Proposed a simple multi-layer multimodal neural network [4]
- Relies on extracting expert visual features for enriched scene information
- Object, scene, facial expressions and overall affective content specific features
- Addition of visual features improves sentiment predictions
- CLIP Image encoder features superior to any combination visual features
- Superior performance compared to complex attention-based methods

*CLIP: Contrastive Language-Image Pre-training

Misogyny detection in multimodal memes

- Proposed a multi-task model [5] that predicts if an image-text pair is:
 - Misogynous or not misogynous
 - Misogynous meme as *stereotype*, *shaming*, *objectification* and *violence*
- Relies on the powerful CLIP image encoder and,
→ an LSTM for pooling the overall textual context in an end-to-end training
- Multi-task loss function and penalty schemes used to train the best model
- Misogynous categories best detected in a multi-task setting

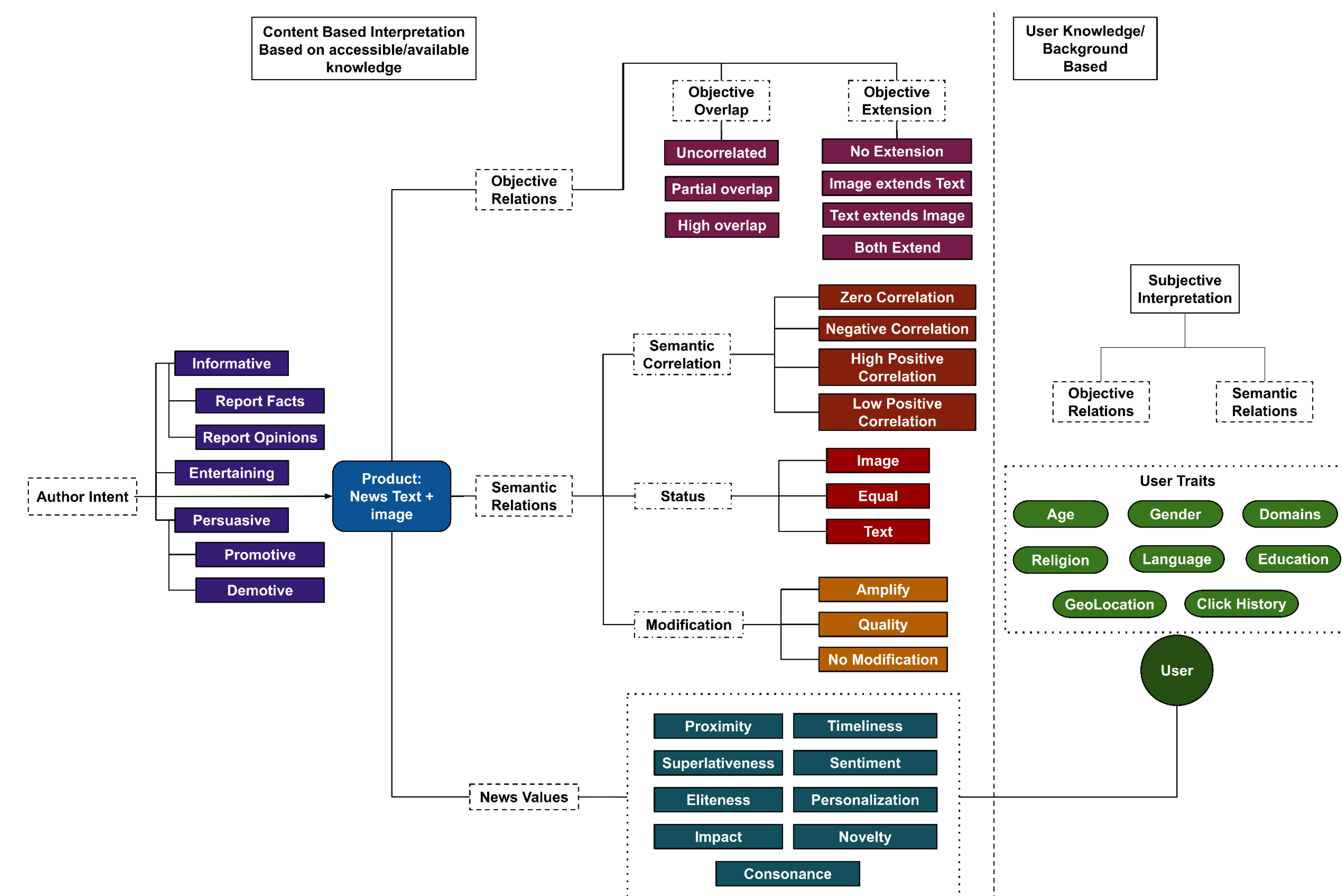
*LSTM: Long Short-Term Memory

Understanding Image-Text Relations and News Values for News Analysis

Proposed novel framework [6] for analysis of multimodal news

- Theoretical framework at the intersection of:
 - multimodal learning, multimodal analytics & computational social sciences
- A set of complex **image-text relations** motivated from:
 - semiotics and computational learning
 - to understand the use of image and text in news
- A set of news-centric attributes called **news values** adopted from:
 - journalism studies*
 - can be used to tag news articles with certain characteristics
- Covers other aspects of:
 - news production (**author intent**)
 - news consumption (**subjective interpretation and user traits**)
- Real-world news examples show applicability of the framework

*Cape, H., & Bednarek, M. (2016). Rethinking news values: What a discursive approach can tell us about the construction of news discourse and news photography. *Journalism*, 17(4), 435-455.



References:

[1] Cheema, G. S., Hakimov, S., & Ewerth, R. (2020). *Check_square at CheckThat! 2020: Claim Detection in Social Media via Fusion of Transformer and Syntactic Features*. In *Working Notes of CLEF 2020-Conference and Labs of the Evaluation Forum, September 22-25, 2020, Thessaloniki, Greece* (Vol. 2696, p. 216).

[2] Cheema, G. S., Hakimov, S., Müller-Budack, E., & Ewerth, R. (2021). *On the Role of Images for Analyzing Claims in Social Media*. In *CLEOPATRA 2021 Cross-lingual Event-centric Open Analytics 2021, April 12 2021, Ljubljana, Slovenia* (Vol. 2829).

[3] Cheema, G. S., Hakimov, S., Sittar, A., Müller-Budack, E., Otto, C., & Ewerth, R. (2022). *MM-Claims: A Dataset for Multimodal Claim Detection in Social Media*. In *Findings of the Association for Computational Linguistics: NAACL 2022* (pp. 962-979).

[4] Cheema, G. S., Hakimov, S., Müller-Budack, E., & Ewerth, R. (2021). *A fair and comprehensive comparison of multimodal tweet sentiment analysis methods*. In *Proceedings of the 2021 Workshop on Multi-Modal Pre-Training for Multimedia Understanding* (pp. 37-45).

[5] Hakimov, S., Cheema, G. S., & Ewerth, R. (2022). *TIB-VA at SemEval-2022 Task 5: A Multimodal Architecture for the Detection and Classification of Misogynous Memes*. In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)* (pp. 756-760).

[6] Cheema, G. S., Hakimov, S., Müller-Budack, E., Otto, C., Bateman, J. A., & Ewerth, R. (2023) Understanding Image-Text Relations and News Values for Multimodal News Analysis. *Frontiers in Artificial Intelligence*, 6, 29.