

Unsupervised Extractive Summarization of Emotion Triggers

Tiberiu Sosea, Hongli Zhan, Junyi Jessy Li, and Cornelia Caragea

ACL 2023



Covid-19 stirs people's emotions, and many take to social medias to share their feelings

Revealing “Why does the writer feel [emotion]?” is important yet remains unexplored

- *Understanding language plays a central role in **clinical psychological diagnosis** (Pennebaker et al., 2003).*
- *While emotion detection is a well-established task, **few have studied what leads to these emotions** in the scope of the text concerned in a data-driven manner.*
- ***Generic summaries tend not to be informative** if users are concerned with specific emotions expressed [Zhan et al., 2022].*

Previous work [Zhan et al., 2022] introduced Covid-ET, a dataset annotated with emotions and an **abstractive** summary of triggers that lead to expressed emotions in text.

Covid-ET

Reddit Post

- 1: My sibling is 19 and she constantly goes places with her friends and to there houses and its honestly stressing me out.
- 2: Our grandfather lives with us and he has dementia along with other health issues and my mom has diabetes and heart problems and I have autoimmune diseases & chronic health issues.
- 3: She also has asthma.
- 4: Its stressing me out because despite this she seems to not care about how badly it would affect all of us if we were to get the virus.
- 5: And sadly I feel like its not much I can do she literally doesn't respect my mom and though I'm older she doesn't respect me either.
- 6: Its so frustrating.

Emotions and Abstractive Summaries of Triggers

Emotion: **anger**

Abstractive Summary of Trigger: **My sister having absolutely no regard for any of our family's health coupled with the fact that I can't do anything about it is so aggravating to me.**

Emotion: **fear**

Abstractive Summary of Trigger: **My sibling, who, in spite of our family's myriad of issues that all make us high-risk people, continuously goes out and about, which makes her likely to get infected. I am scared for all of us right now.**

However...

- *Obtaining timely and qualitative abstractive summaries is expensive and extremely time-consuming.*
- *As seen in Covid-ET, labeled data is necessary to attain good performance on our task and general summarization models struggle.*
- *Therefore, in time-sensitive, high-stake contexts, lack of data can block necessary responses.*

We extend Covid-ET with extractive trigger summaries and turn to **fully unsupervised** extractive trigger summarization.

Covid-ET-EXT

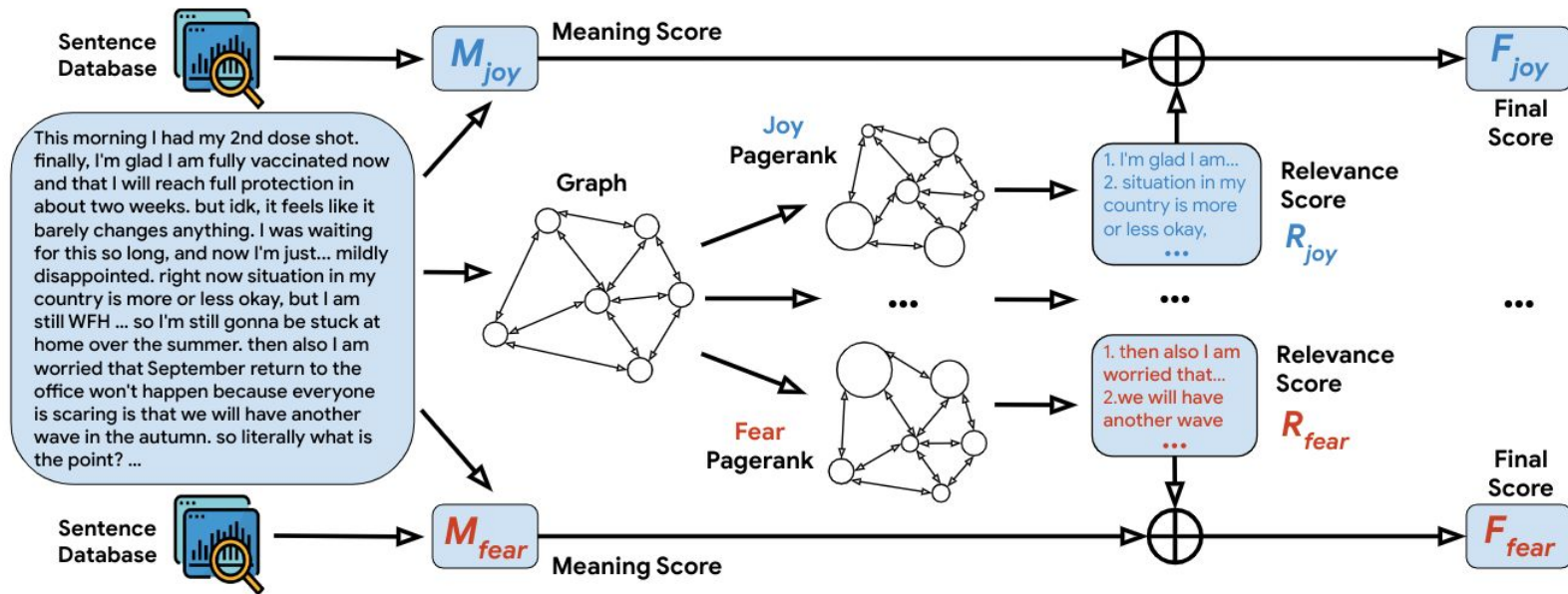
Extractive Summarization

Emotion Triggers in a Reddit Post

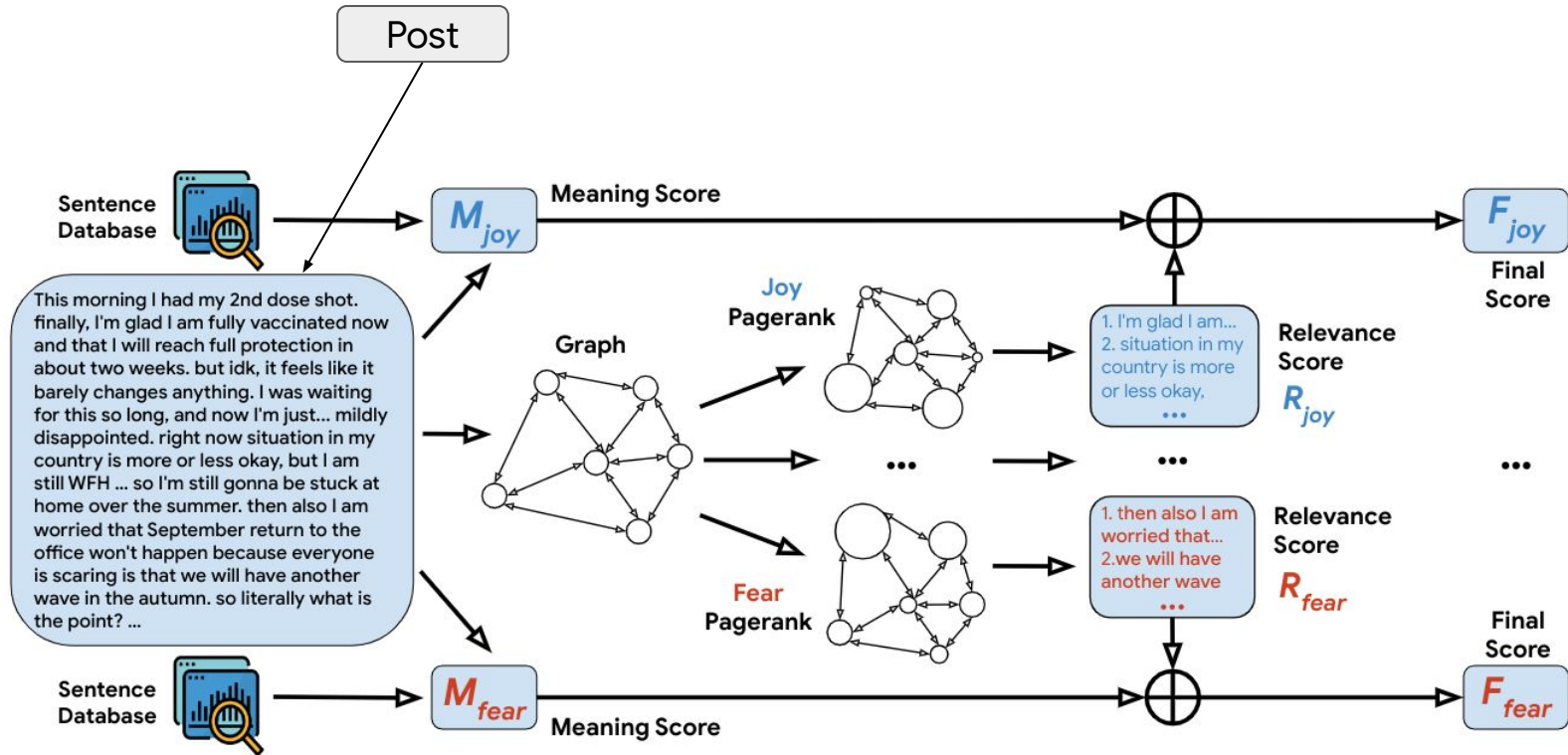
1. It finally happened.
2. Took an older relative for her first Pfizer dose.
3. Not that many people showed up so all accompanying family members were offered the shot and give papers for a second dose.
4. I wasnt due to get my shot for the next couple of months and have had some scares, many for which Ive gotten support from you lovely awesome people.
5. I now have a whole different perspective on my governments organization. **Trust**
6. They have a good oiled vaccination machine.
7. They just need more doses.
8. Best part is I was told I could come back for my second dose whenever my relative was scheduled to get Hers. **Joy**
9. I have a lil arm pain.
10. But its the slightest arm pain when moving it past certain angles.
11. Ive noticed that a lil blood drop shows on my vaccination spot (took a shower half an hour later). **Joy**
12. Is this normal? **Fear**

Emotion-aware Pagerank

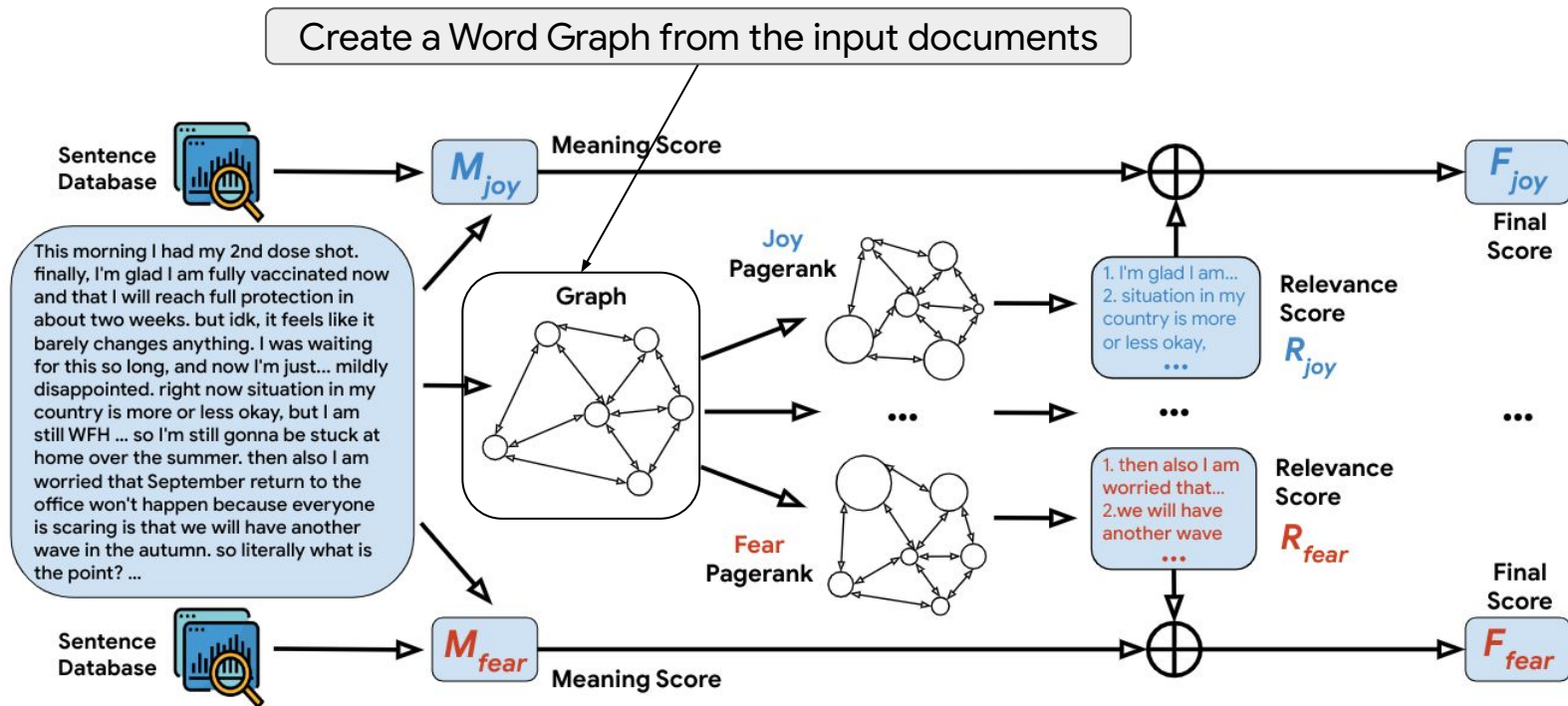
To model the new task, we introduce *Fully unsupervised Emotion Detection and Extractive Trigger Summarization*, and propose *Emotion-aware Pagerank (EAP)*



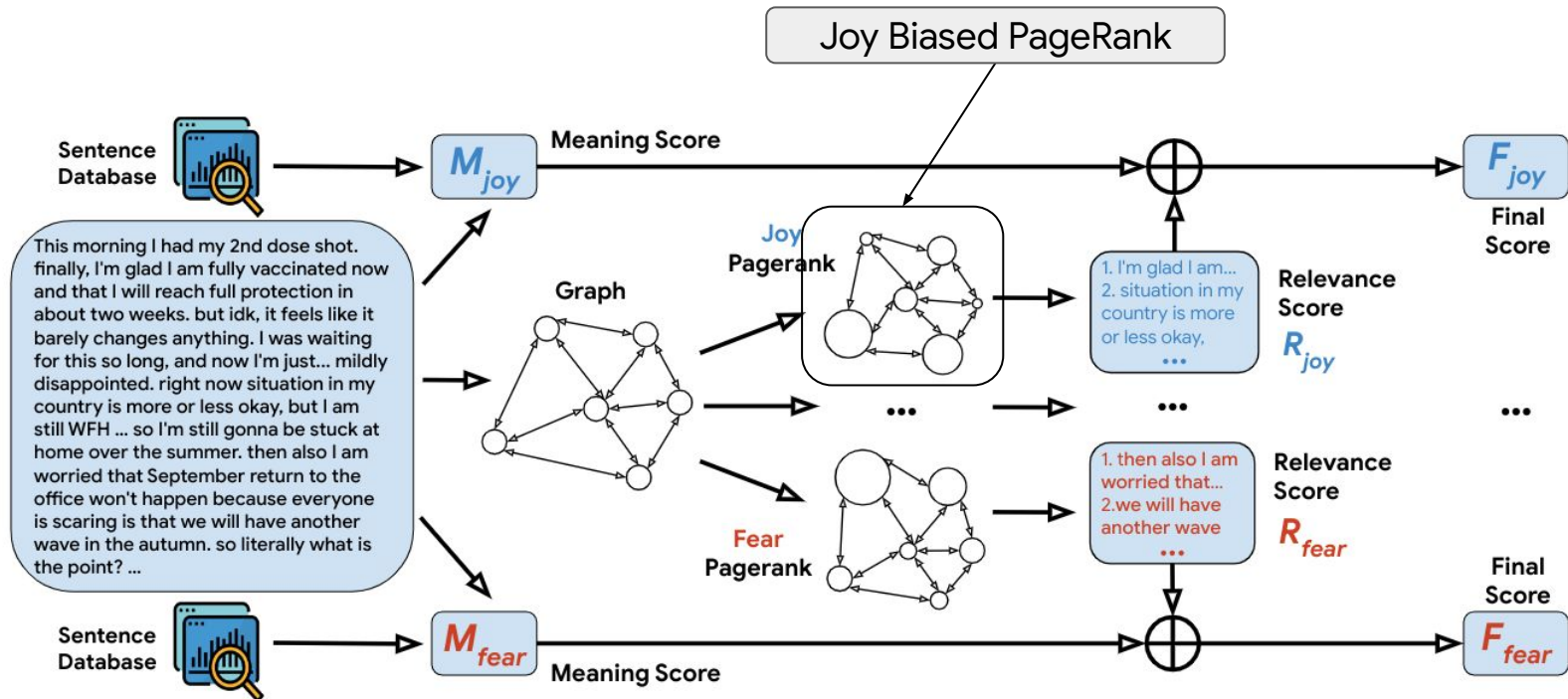
Summarizing Emotion Triggers



Summarizing Emotion Triggers



Summarizing Emotion Triggers



PageRank Relevance

$$\mathcal{R}(w_i) = \lambda \sum_{k=1}^{|V|} \beta(w_k, w_i) \mathcal{R}(w_k) + (1 - \lambda) \frac{1}{|V|}$$

Biased PageRank Relevance

$$\mathcal{R}_e(w_i) = \lambda \sum_{k=1}^{|V|} \beta(w_k, w_i) \mathcal{R}_e(w_k) + (1 - \lambda) \frac{i_e(w_i)}{N}$$

Words in the graph

PageRank Relevance

$$\mathcal{R}(w_i) = \lambda \sum_{k=1}^{|V|} \beta(w_k, w_i) \mathcal{R}(w_k) + (1 - \lambda) \frac{1}{|V|}$$

Constant random jump probability

Biased PageRank Relevance

$$\mathcal{R}_e(w_i) = \lambda \sum_{k=1}^{|V|} \beta(w_k, w_i) \mathcal{R}_e(w_k) + (1 - \lambda) \frac{i_e(w_i)}{N}$$

Biased random jump probability

PageRank Relevance

$$\mathcal{R}(w_i) = \lambda \sum_{k=1}^{|V|} \beta(w_k, w_i) \mathcal{R}(w_k) + (1 - \lambda) \frac{1}{|V|}$$

Biased PageRank Relevance

$$\mathcal{R}_e(w_i) = \lambda \sum_{k=1}^{|V|} \beta(w_k, w_i) \mathcal{R}_e(w_k) + (1 - \lambda) \frac{i_e(w_i)}{N}$$

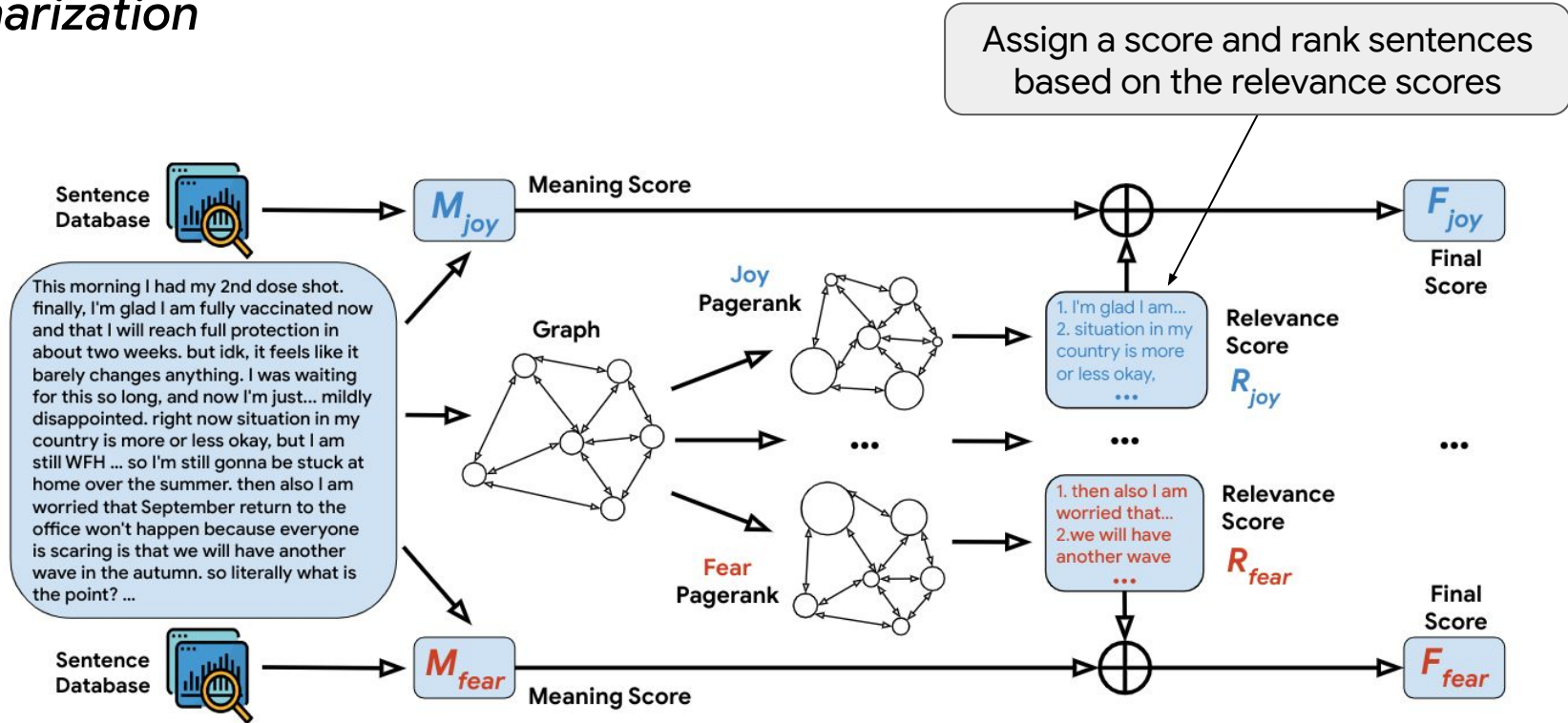
Emotion-intensity lexicon

$$i_e(w) = \begin{cases} \text{int}_e(w) & \text{if } w \in \mathcal{I}_e \\ c & \text{if } w \in V \setminus \mathcal{I}_e \end{cases}$$

We assign higher random jump probability based on the emotion intensity of words

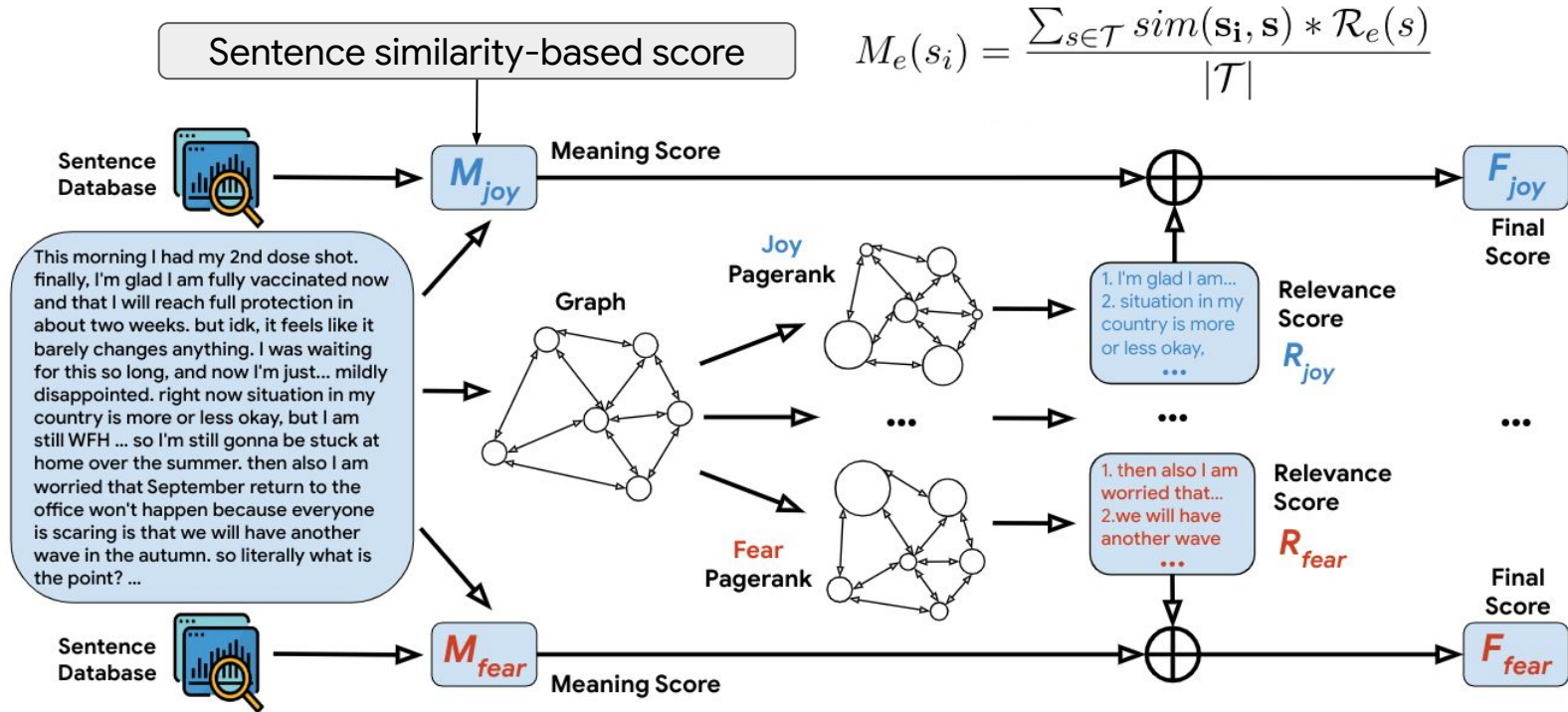
Summarizing Emotion Triggers

Solution: Fully unsupervised Emotion Detection and **Extractive** Trigger Summarization



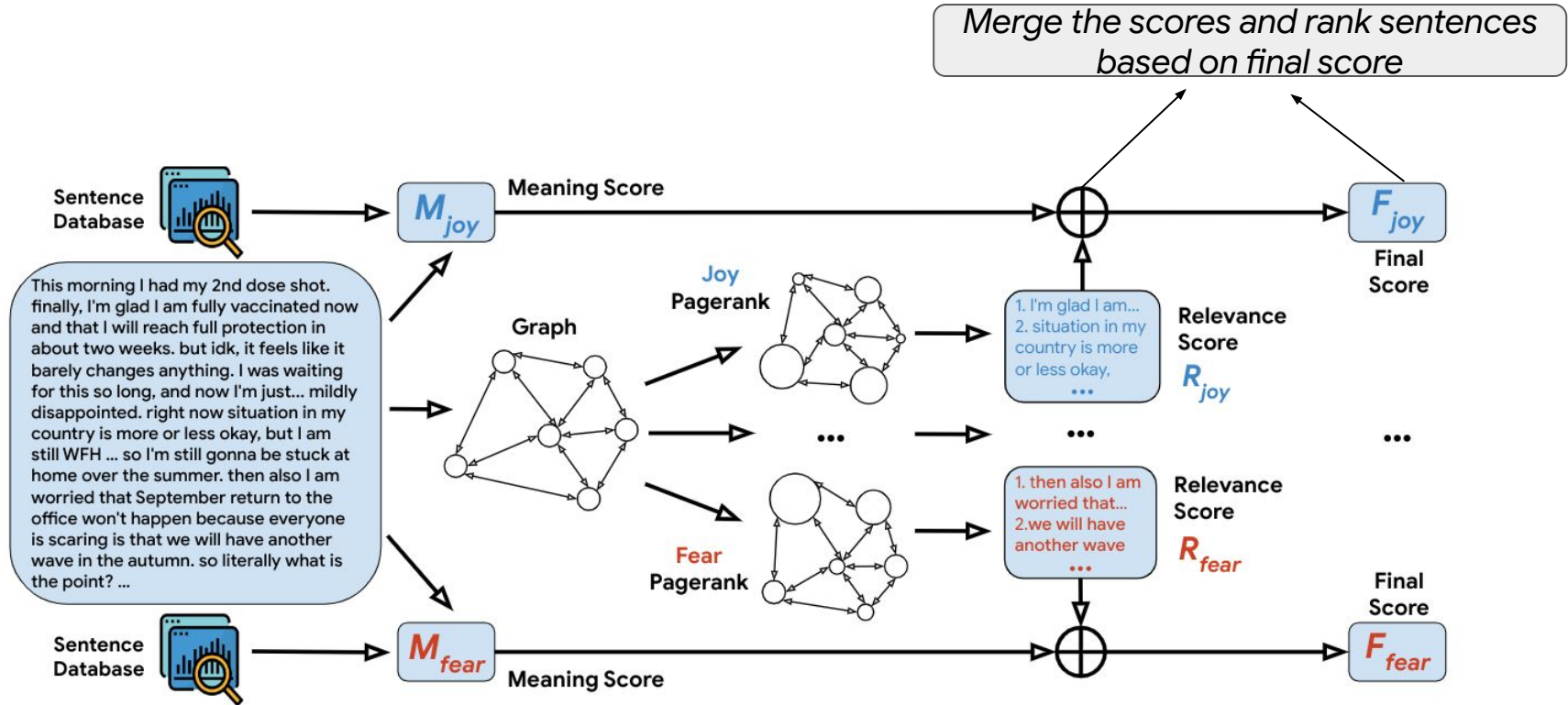
Summarizing Emotion Triggers

Create a 'meaning' score M based on similarity measures with our sentence database



Summarizing Emotion Triggers

Create final ranking



Summarizing Emotion Triggers

Results

	ANGER		DISGUST		FEAR		JOY		SADNESS		TRUST		ANTICIPATION		AVG	
	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L
1-SENT	0.174	0.240	0.095	0.170	0.202	0.256	0.119	0.179	0.110	0.177	0.189	0.236	0.160	0.220	0.149	0.211
3-SENT	0.301	0.315	0.196	0.253	0.322	0.343	0.273	0.310	0.239	0.292	0.248	0.279	0.263	0.307	0.258	0.288
PACSUM	0.308	0.314	0.210	0.218	0.327	0.331	0.276	0.282	0.287	0.304	0.225	0.234	0.283	0.295	0.273	0.282
PRESUMM	0.306	0.312	0.219	0.221	0.332	0.335	0.268	0.274	0.295	0.317	0.222	0.227	0.284	0.291	0.275	0.282
TEXTRANK	0.296	0.301	0.236	0.235	0.319	0.326	0.272	0.276	0.286	0.306	0.225	0.231	0.218	0.221	0.264	0.270
EMOLEX	0.213	0.260	0.218	0.256	0.309	0.341	0.218	0.252	0.301	0.331	0.176	0.203	0.207	0.242	0.234	0.269
EMOINTENSITY	0.307	0.322	0.269	0.281	0.342	0.355	0.222	0.235	0.329	0.341	0.227	0.242	0.295	0.310	0.284	0.298
BERT-GoEMO	0.247	0.264	0.232	0.237	0.296	0.312	0.221	0.247	0.314	0.321	0.201	0.204	0.247	0.225	0.253	0.258
EAP	0.324[†]	0.348[†]	0.285[†]	0.296[†]	0.364[†]	0.373[†]	0.285[†]	0.319[†]	0.348[†]	0.354[†]	0.258[†]	0.291[†]	0.319[†]	0.324[†]	0.309[†]	0.325[†]

Summarizing Emotion Triggers

Results

	ANGER		DISGUST		FEAR		JOY		SADNESS		TRUST		ANTICIPATION		AVG	
	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L	R-2	R-L
1-SENT	0.174	0.240	0.095	0.170	0.202	0.256	0.119	0.179	0.110	0.177	0.189	0.236	0.160	0.220	0.149	0.211
3-SENT	0.301	0.315	0.196	0.253	0.322	0.343	0.273	0.310	0.239	0.292	0.248	0.279	0.263	0.307	0.258	0.288
PACSUM	0.308	0.314	0.210	0.218	0.327	0.331	0.276	0.282	0.287	0.304	0.225	0.234	0.283	0.295	0.273	0.282
PRESUMM	0.306	0.312	0.219	0.221	0.332	0.335	0.268	0.274	0.295	0.317	0.222	0.227	0.284	0.291	0.275	0.282
TEXTRANK	0.296	0.301	0.236	0.235	0.319	0.326	0.272	0.276	0.286	0.306	0.225	0.231	0.218	0.221	0.264	0.270
EMOLEX	0.213	0.260	0.218	0.256	0.309	0.341	0.218	0.252	0.301	0.331	0.176	0.203	0.207	0.242	0.234	0.269
EMOINTENSITY	0.307	0.322	0.269	0.281	0.342	0.355	0.222	0.235	0.329	0.341	0.227	0.242	0.295	0.310	0.284	0.298
BERT-GoEMO	0.247	0.264	0.232	0.237	0.296	0.312	0.221	0.247	0.314	0.321	0.201	0.204	0.247	0.225	0.253	0.258
EAP	0.324[†]	0.348[†]	0.285[†]	0.296[†]	0.364[†]	0.373[†]	0.285[†]	0.319[†]	0.348[†]	0.354[†]	0.258[†]	0.291[†]	0.319[†]	0.324[†]	0.309[†]	0.325[†]

Conclusion and Future work

- We introduced CovidET-EXT, a new benchmark dataset composed of 1,883 Reddit posts annotated for the task emotion detection and extractive trigger summarization in the context of the COVID-19 pandemic.
- The proposed Emotion-Aware Pagerank approach yields strong results on our datasets, consistently outperforming prior work in an unsupervised learning context.
- In the future, we plan to study *abstractive* trigger summarization from a zero-shot perspective to remove the requirement for labeled data.

Thank you!



Contact: tsosea2@uic.edu
cornelia@uic.edu

honglizhan@utexas.edu
jessy@austin.utexas.edu