

SocialSec 2024: 10th International Symposium on Security and
Privacy in Social Networks and Big Data
Zayed University, Abu Dhabi, UAE | November 20-22, 2024

Investigating Influential COVID-19 Perspectives: A Multifaceted Analysis of Twitter Discourse



COLORADO STATE
UNIVERSITY

SDSU

San Diego State
University

Shadaab Kawnain Bashir (shadaab@colostate.edu)

Hossein Shirazi (hshirazi@sdsu.edu)

Noushin Salek Faramarzi (nsalekfarama@cs.stonybrook.edu)

Thomas Harris (tvharris123@gmail.com)

Ashmita Shishodia (ashishodia2546@sdsu.edu)

Hajar Homayouni (hhomayouni@sdsu.edu)

Indrakshi Ray (indrakshi.ray@colostate.edu)

Motivation



CDC Travel Health Notice: Novel Coronavirus in #China:

There is an outbreak caused by a novel coronavirus in China. People with underlying health conditions are more likely to get sick and die from COVID-19.



Vaccines Truth
@VaccinesTruth



VaccinateNow
@VaccinateNow



Marjorie Taylor Greene
@mtgreenee
Satisfactory 29%

Thousands of people are reporting very serious life changing vaccine side effects from taking covid vaccines.

5,946 deaths are reported on the CDC website.

Social media is censoring their stories & the media is silent.

Biden is going to homes to push shots.

Just say NO!

This Tweet is misleading. Find out why health officials consider COVID-19 vaccines safe for most people.

2:39 PM · Jul 11, 2021 · Twitter for iPhone



Dr. Mo
@DrMolbrahim

The vaccine was developed by scientists and not politicians. Get the vaccine if you are eligible. If you have any concerns ask your doctor.

#CovidVaccine #coronavirus
#vaccine #POLITICIAN
#Science #scientist

6:15 PM · 30 Mar 21 · Twitter for iPhone



Donald J. Trump
@realDonaldTrump · 3h

This Tweet violated the Twitter Rules about spreading misleading and potentially harmful information related to COVID-19. However, Twitter has determined that it may be in the public's interest for the Tweet to remain accessible.

[Learn more](#)

[View](#)

Problem Statement

- Influential figures' misinformation causes public confusion
- Distorts pandemic understanding
- Unclear how individual vs. organizational communication differs
- Celebrities' impact differs from organizations like WHO

Individual

Organization

The image shows two tweets from a user named 'batman' on November 16, 2021, at 3:55 AM. The top tweet is from an individual user and contains misinformation: 'Hey everyone, let's stay safe! Remember to wear your masks 🧐, wash your hands 🧼, and practice social distancing. Together, we can beat this! 🙌 #StaySafe #COVID19 #FlattenTheCurve'. The bottom tweet is from 'NewsOutlet' and contains factual information: 'FACT: Vaccines protect you and others from severe #COVID19. 📝 Let's bust myths and share facts: 🚫 COVID-19 vaccines do NOT contain microchips. 🚫 COVID-19 vaccines do NOT alter your DNA. Let's stand together for science and safety! 🌍 #HealthForAll #VaccinesWork"'. Both tweets show zero retweets, quotes, likes, and bookmarks.

Research Gap

- Current studies: general sentiment analysis of COVID-19 posts
- Lack distinction between individual & organizational influencers
- No comprehensive comparison of emotional tone/sentiment
- Scientific content analysis lacking for both influencer types

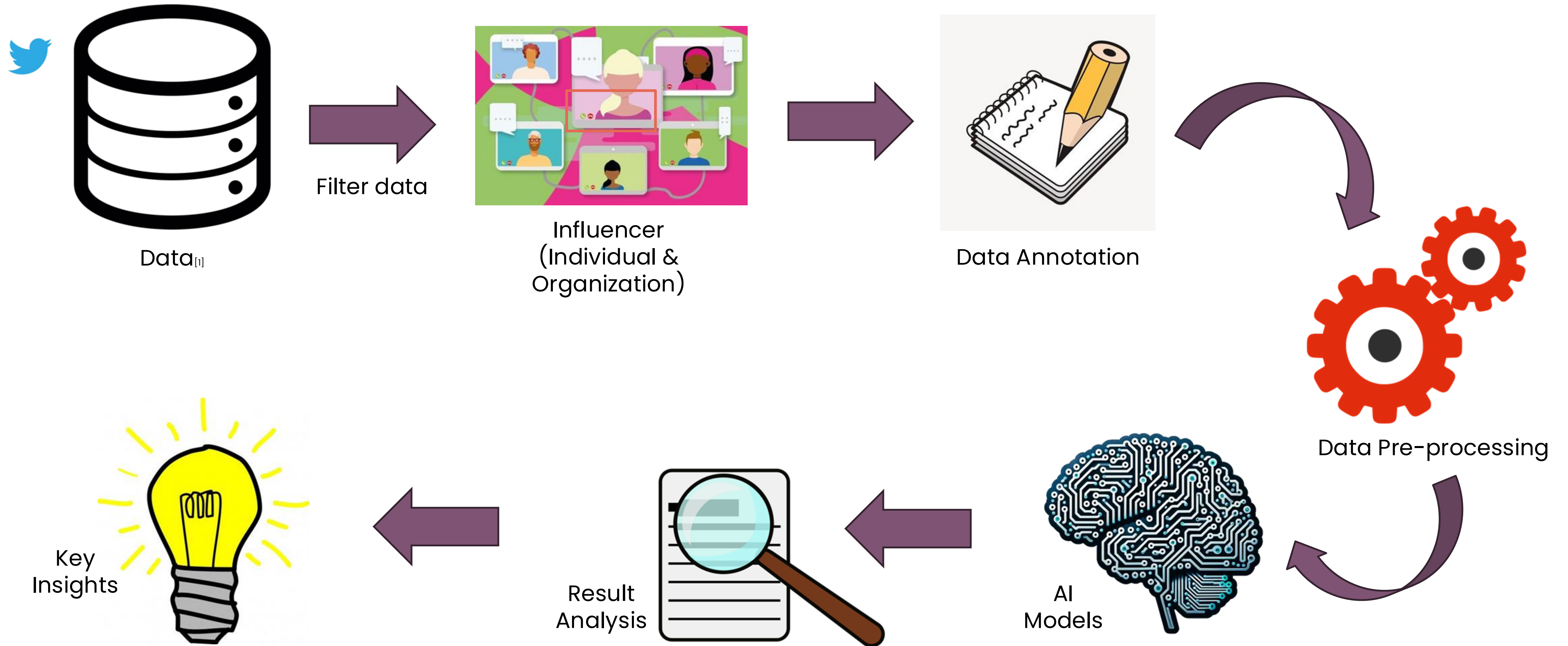
Proposed Solution

- Differentiate Influencer Types
 - Analyze individual vs. organizational style of communication
- Combat Dissemination of Misinformation
 - Use ML to detect
- Enhance Public Trust
 - Promote verified content

Contribution

- Created Dataset
 - Labeled COVID-19 Tweets: profiles, emotions, sentiments, scientific content
- Comprehensive Analysis
 - Combined profiling, emotions, sentiment, scientific content for deeper discourse insight

Workflow



[1] Zuo, C., Banerjee, R., Chaleshtori, F.H., Shirazi, H., Ray, I.: Seeing Should Probably Not be Believing: the Role of Deceptive support in COVID-19 Misinformation on Twitter. ACM Journal of Data and Information Quality 15(1), 1-26 (2022)

Workflow



Data^[1]

Original Tweet

    Jay MBE! We salute you - King of the lockdown pub quiz. #QueensBirthdayHonours #MBE

Pre-processed Tweet

: `::clapping_hands::` `::clapping_hands::`
`::clapping_hands::` `::clapping_hands::`
`::clapping_hands::` jay mbe! we salute you king of the lockdown pub quiz. queensbirthdayhonours mbe

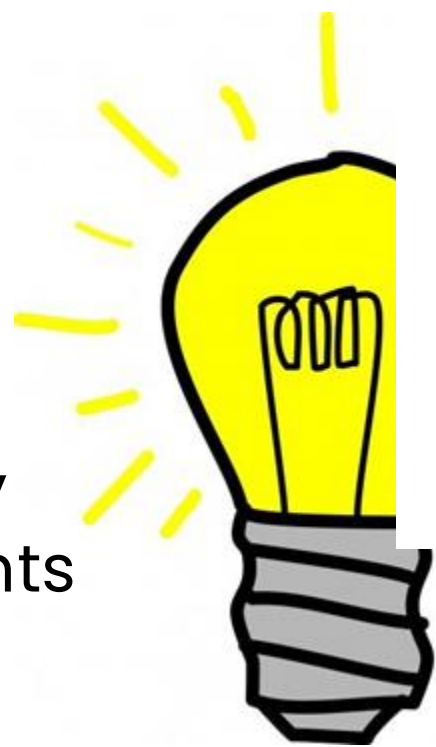
Hi @esswhywon, We have passed your comment onto our COVID 19 Social Distancing Team. Many Thanks Steve

hi USER_HANDLE, we have passed your comment onto our covid 19 social distancing team. many thanks steve

Hello, masks and social distancing are required in order to use public transit. We have been using more high capacity & extra buses on routes, and cleaning our buses with hospital-grade disinfectant, to mention a few. Please visit <https://rtcsnv.com/coronavirus/>.

hello, masks and social distancing are required in order to use public transit. we have been using more high capacity amp; extra buses on routes, and cleaning our buses with hospitalgrade disinfectant, to mention a few. please visit URL_TOKEN.

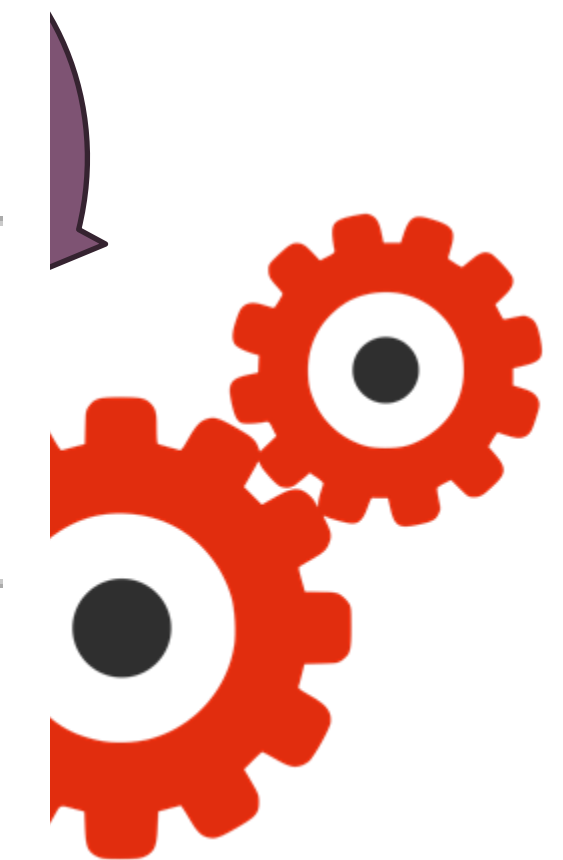
Key Insights



Result Analysis



AI Models

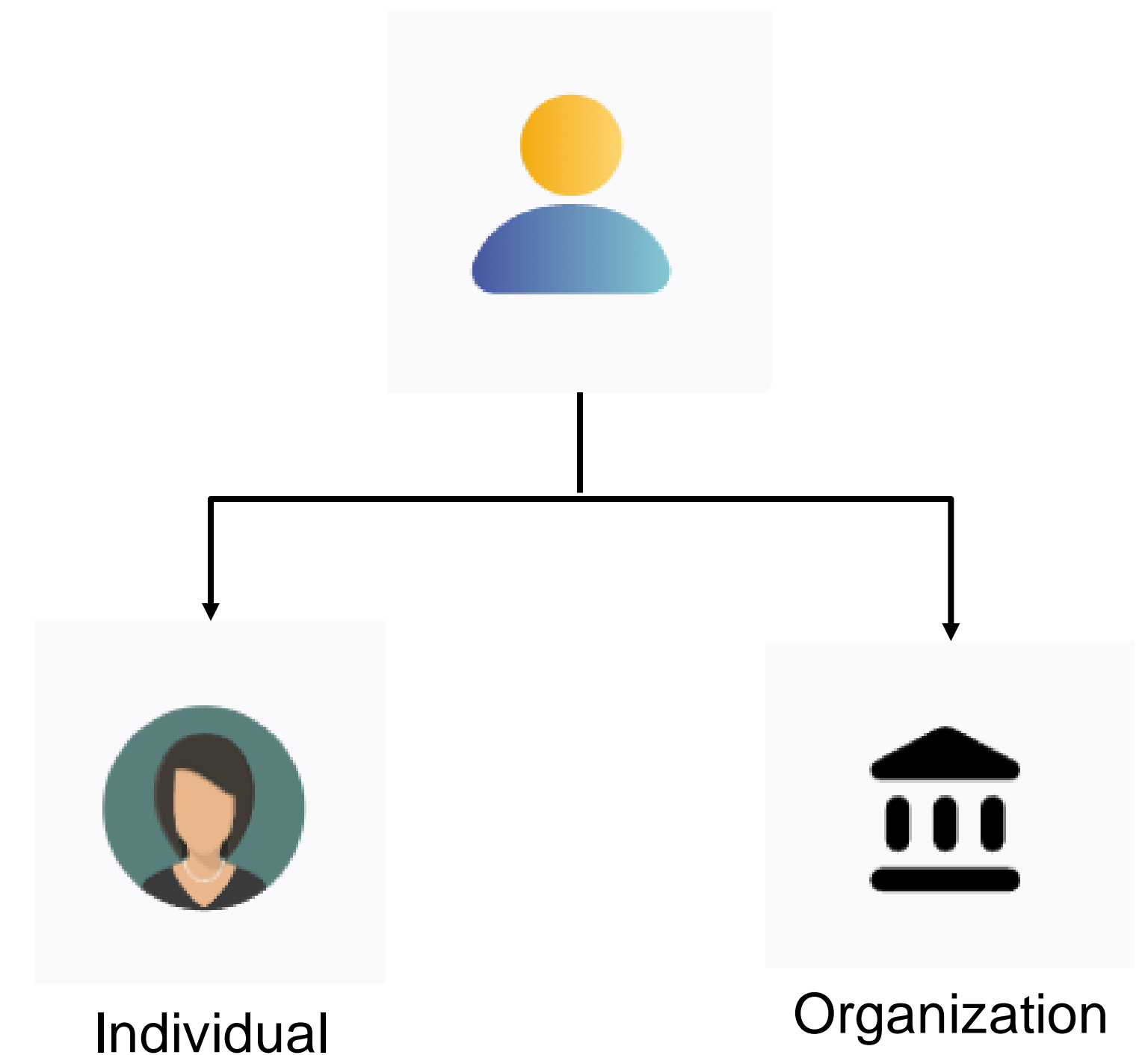


a Pre-processing

[1] Zuo, C., Banerjee, R., Chaleshtori, F.H., Shirazi, H., Ray, I.: Seeing Should Probably Not be Believing: the Role of Deceptive support in COVID-19 Misinformation on Twitter. ACM Journal of Data and Information Quality 15(1), 1-26 (2022)

User Profiling

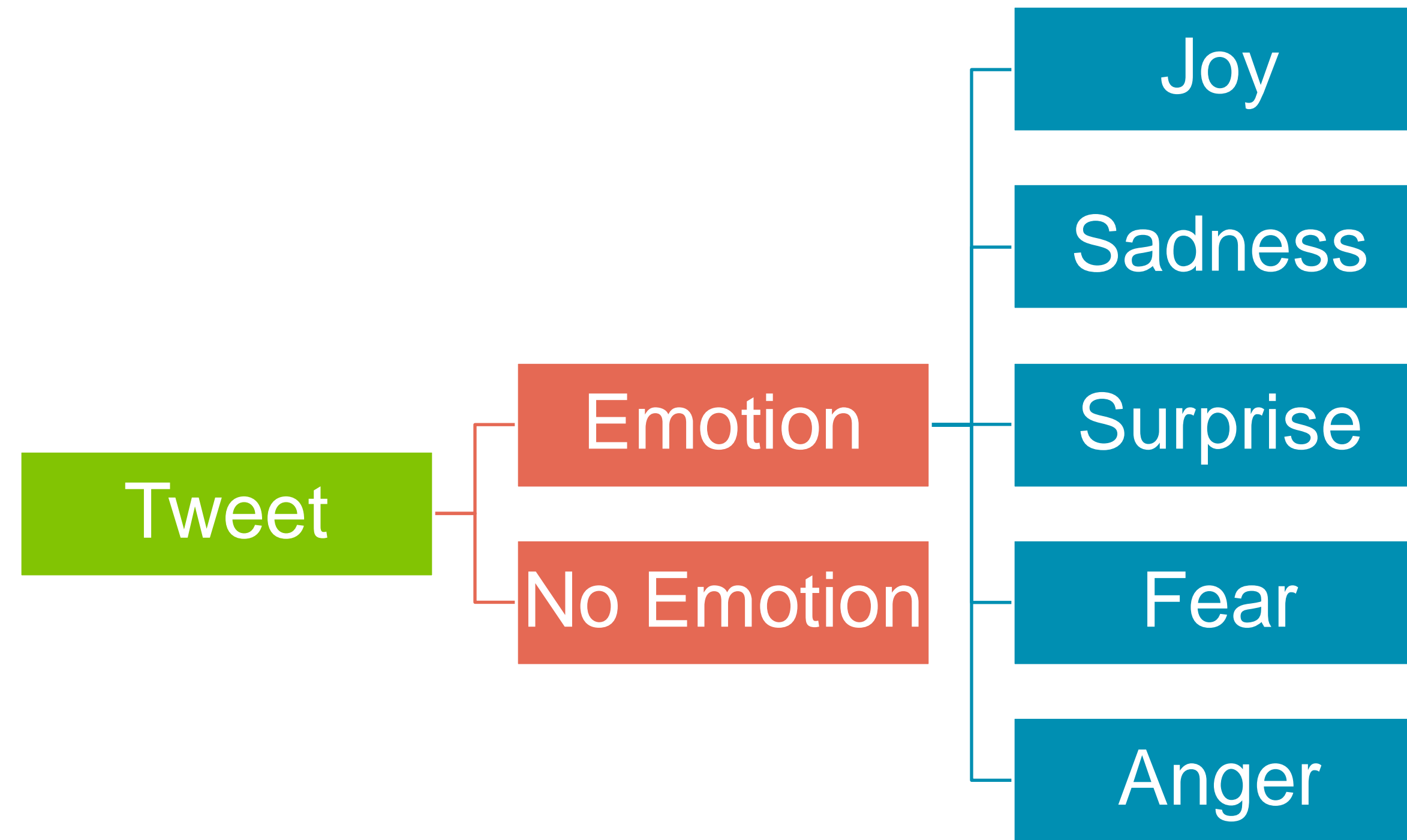
- Classification Approach
 - Used Twitter metadata (name, handle, biography)
- BERT Model Usage
 - To transform metadata into embeddings
 - Outperforms Demographer^[1]



Algorithm	Precision	Recall	Accuracy	F1
Demographer	0.88	0.60	0.79	0.71
BERT	0.95	0.94	0.96	0.95

[1] Knowles, R., Carroll, J., Dredze, M.: Demographer: Extremely Simple Name Demographics. In: Proceedings of the First Workshop on NLP and Computational Social Science. pp. 108–113 (2016)

Emotion Detection



Emotion Detection Pipeline

Emotion Detection

 **Joe** ✓
Joe

I just got vaccinated! Feeling so hopeful and grateful for science. 🇺🇸 🇺🇸 #Vaccinated #Hope

3:55 AM · Nov 16, 2021 · 0 Views

0 Retweets 0 Quotes 0 Likes 0 Bookmarks

Joy

 **Joe** ✓
Joe

People throwing parties during a pandemic? Seriously? This selfishness is putting lives at risk! 🚫 #StayHome"

3:55 AM · Nov 16, 2021 · 0 Views

0 Retweets 0 Quotes 0 Likes 0 Bookmarks

 **Joe** ✓
Joe

I lost someone I love to COVID-19. This virus is devastating. Please take it seriously. 😞 #Heartbroken

3:55 AM · Nov 16, 2021 · 0 Views

0 Retweets 0 Quotes 0 Likes 0 Bookmarks

Sadness

 **Sam** ✓
sammy

The new COVID variant sounds scary. I'm worried about what's coming next. 😬 #Anxious

3:55 AM · Nov 16, 2021 · 0 Views

0 Retweets 0 Quotes 0 Likes 0 Bookmarks

Fear

 **Sam** ✓
sammy

Wow, I just found out my coworker had COVID last week but didn't tell anyone. 😱 That's shocking and irresponsible. #Unbelievable

3:55 AM · Nov 16, 2021 · 0 Views

0 Retweets 0 Quotes 0 Likes 0 Bookmarks

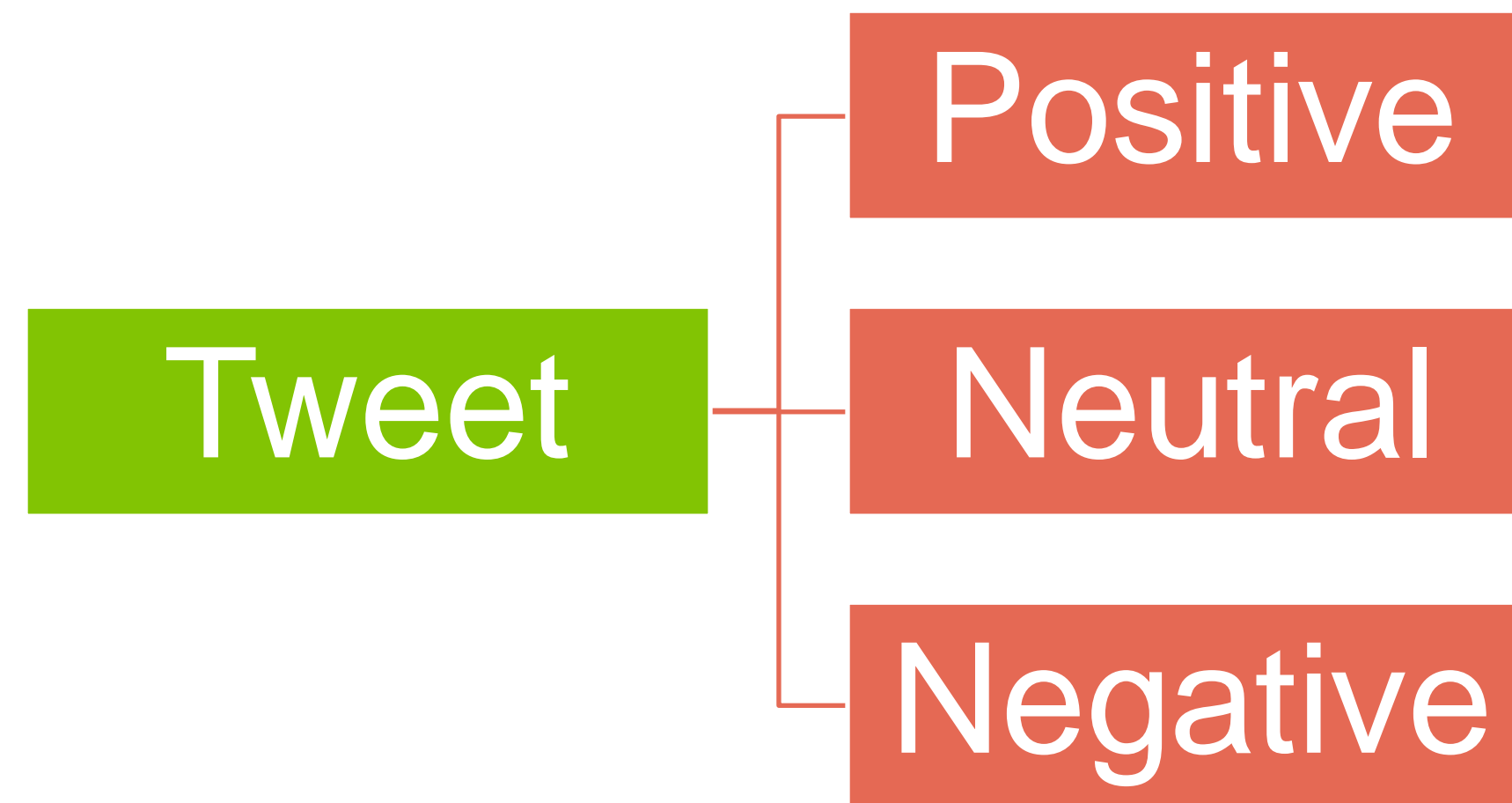
Surprise

Emotion Detection

Task	Label	BERT				SetFit				Task Specific Model				GPT-3.5				LLaMA-2-70B			
		P	R	A	F1	P	R	A	F1	P	R	A	F1	P	R	A	F1	P	R	A	F1
Emotion Detection	<i>Emotion Existence</i>	0.85	0.77	0.77	0.81	0.83	0.82	0.83	0.83	0.85	0.79	0.84	0.83	0.75	0.70	0.70	0.70	0.76	0.59	0.59	0.56
	Joy	0.79	0.81	0.93	0.80	0.93	0.91	0.94	0.92	0.85	0.79	0.90	0.82	0.91	0.91	0.91	0.91	0.91	0.81	0.81	0.84
	Sadness	0.67	0.61	0.89	0.63	0.76	0.81	0.86	0.78	0.74	0.63	0.89	0.67	0.91	0.93	0.93	0.92	0.93	0.66	0.66	0.75
	Surprised	0.45	0.26	0.88	0.33	0.79	0.79	0.87	0.79	0.77	0.48	0.86	0.59	0.91	0.87	0.87	0.89	0.91	0.66	0.66	0.75
	Fear	0.64	0.46	0.88	0.53	0.87	0.83	0.89	0.85	0.80	0.64	0.87	0.71	0.88	0.80	0.80	0.83	0.90	0.56	0.56	0.64
	Anger	0.84	0.86	0.93	0.84	0.90	0.91	0.91	0.91	0.84	0.92	0.89	0.88	0.89	0.89	0.89	0.89	0.89	0.78	0.78	0.80

- Results
 - **SetFit** achieved the top F1 scores for **emotion detection**.
 - **GPT-3.5** performed best for detecting **sadness** and **surprise**.
- Observation
 - Joy was the easiest emotion to detect.
 - Sadness and surprise posed more challenges to detect.
 - Chat-based models like GPT-3.5 excel in detecting these complex emotions.

Sentiment Analysis



Sentiment Analysis Pipeline

Sentiment Analysis



A screenshot of a tweet from a user named Joe. The tweet text is: "Amazing news! COVID-19 cases are dropping thanks to everyone doing their part. Let's keep wearing masks and getting vaccinated. We're stronger together! 🙌❤️ #StaySafe". The tweet includes a timestamp of "3:55 AM · Nov 16, 2021" and a "Views" indicator. Below the text are icons for Retweets, Quotes, Likes, and Bookmarks, and a row of interaction icons (reply, retweet, like, bookmark, share).

Positive Sentiment



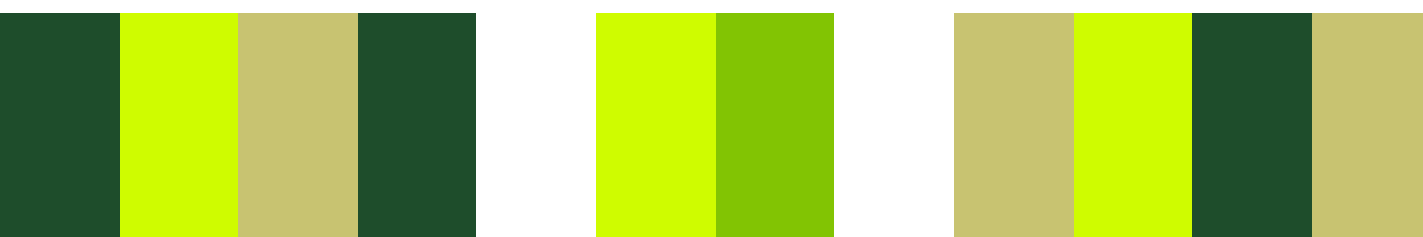
A screenshot of a tweet from WHO. The tweet text is: "Reminder: Masks are recommended in crowded spaces to reduce the spread of COVID-19. Follow local guidelines for updates. #COVID19Safety". The tweet includes a timestamp of "3:55 AM · Nov 16, 2021" and a "Views" indicator. Below the text are icons for Retweets, Quotes, Likes, and Bookmarks, and a row of interaction icons (reply, retweet, like, bookmark, share).

Neutral Sentiment



A screenshot of a tweet from a user named Sam. The tweet text is: "I'm tired of people ignoring health guidelines. This is why the virus keeps spreading. Some people are so irresponsible. 🙄 #WearAMask". The tweet includes a timestamp of "3:55 AM · Nov 16, 2021" and a "Views" indicator. Below the text are icons for Retweets, Quotes, Likes, and Bookmarks, and a row of interaction icons (reply, retweet, like, bookmark, share).

Negative Sentiment



Sentiment Analysis

Task	Label	BERT				SetFit				Task Specific Model				GPT-3.5				LLaMA-2-70B			
		P	R	A	F1	P	R	A	F1	P	R	A	F1	P	R	A	F1	P	R	A	F1
Sentiment Detection	Negative	0.68	0.53	0.53	0.60	0.64	0.66	0.66	0.65	0.77	0.68	0.67	0.72	0.47	0.73	0.73	0.57	0.35	0.69	0.46	0.69
	Neutral	0.80	0.93	0.93	0.86	0.86	0.86	0.87	0.86	0.86	0.91	0.91	0.89	0.88	0.54	0.54	0.67	0.75	0.39	0.52	0.39
	Positive	0.76	0.46	0.46	0.58	0.75	0.72	0.72	0.74	0.81	0.75	0.75	0.78	0.48	0.88	0.88	0.62	0.40	0.74	0.54	0.74

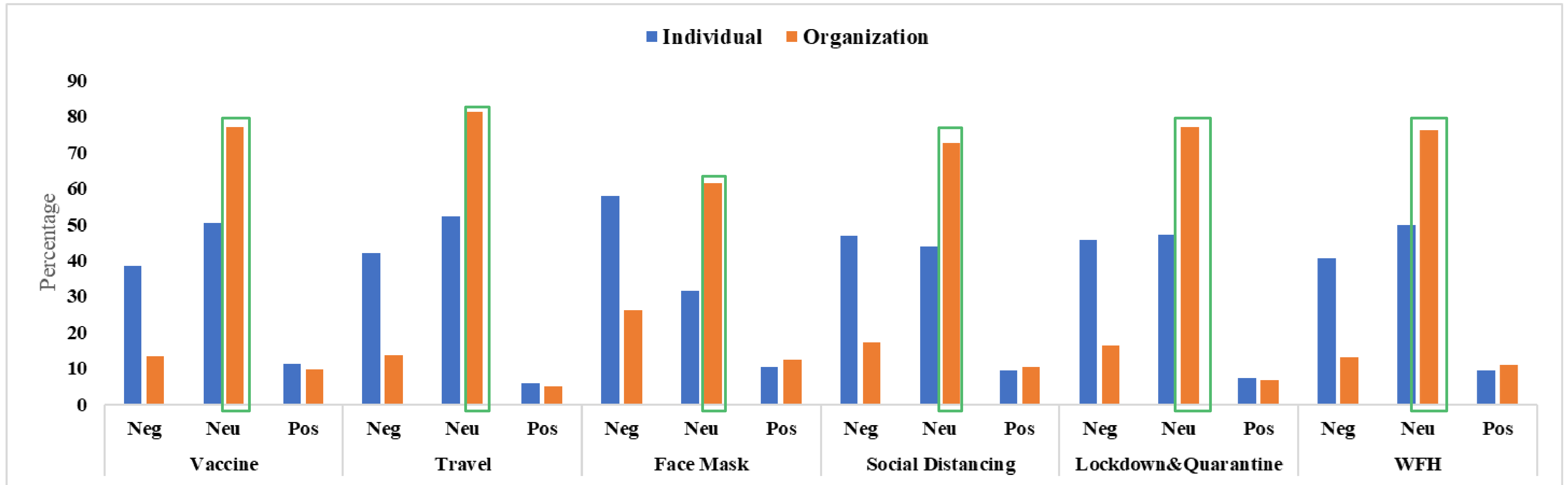
- Results

- **RoBERTa_{Twitter}** achieved the highest F1 scores for **negative, positive** and **neutral** sentiments. **RoBERTa_{Twitter}** pre-trained on Twitter dataset. Fine-tuned on our dataset

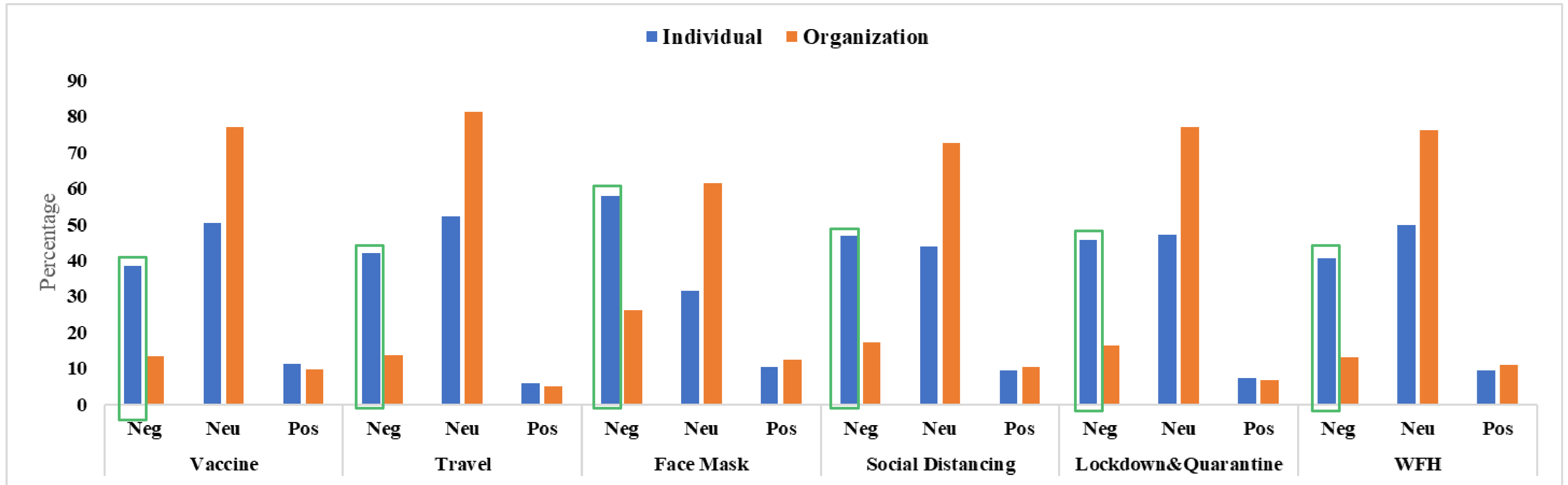
- Observation

- Neutral Tweets were classified most accurately
 - Negative sentiment was the hardest to detect

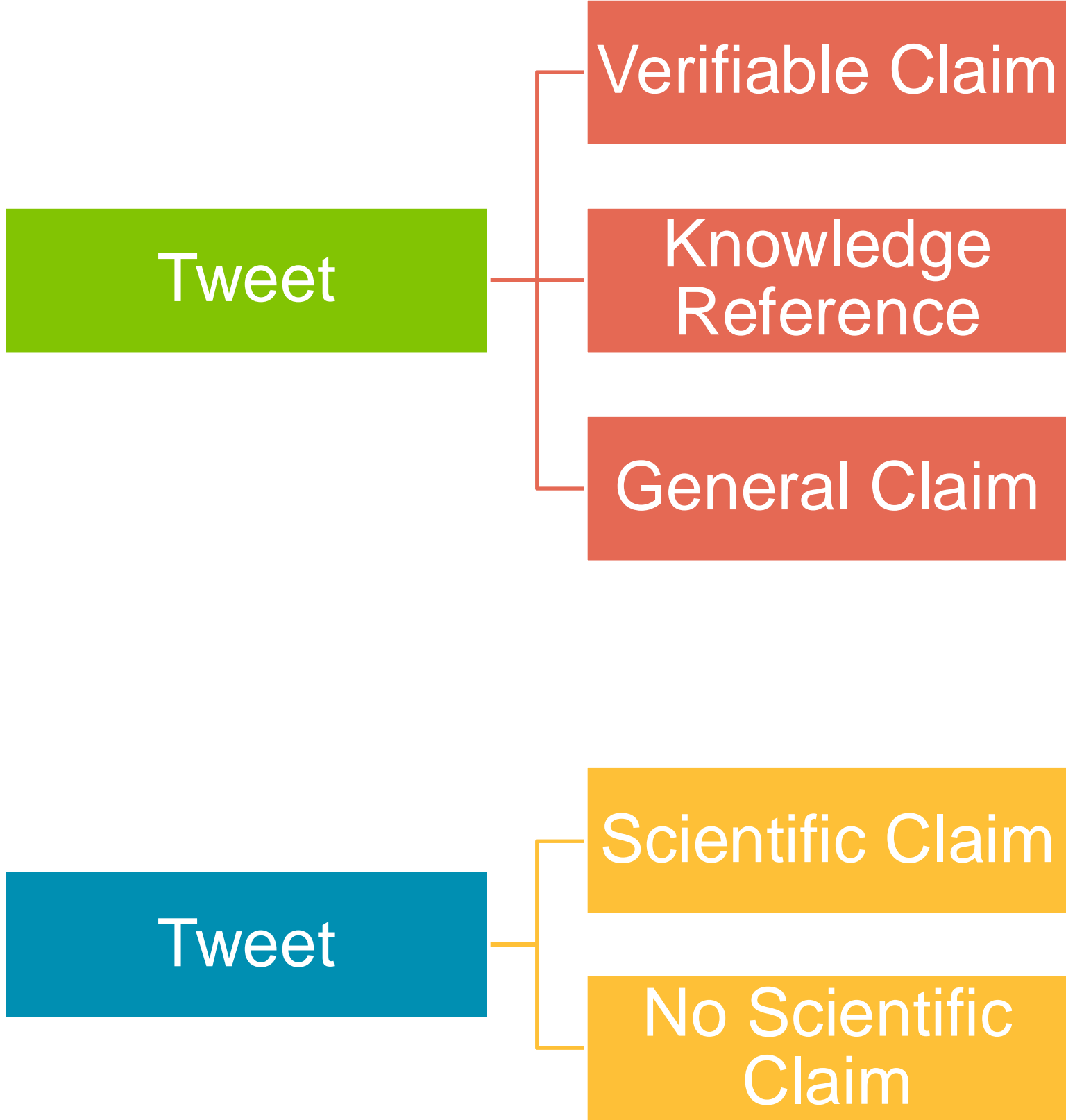
Topic-Targeted Sentiment Analysis



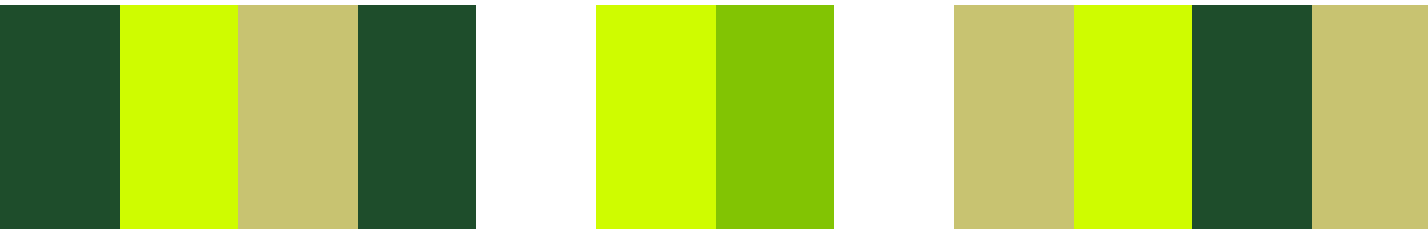
Topic-Targeted Sentiment Analysis



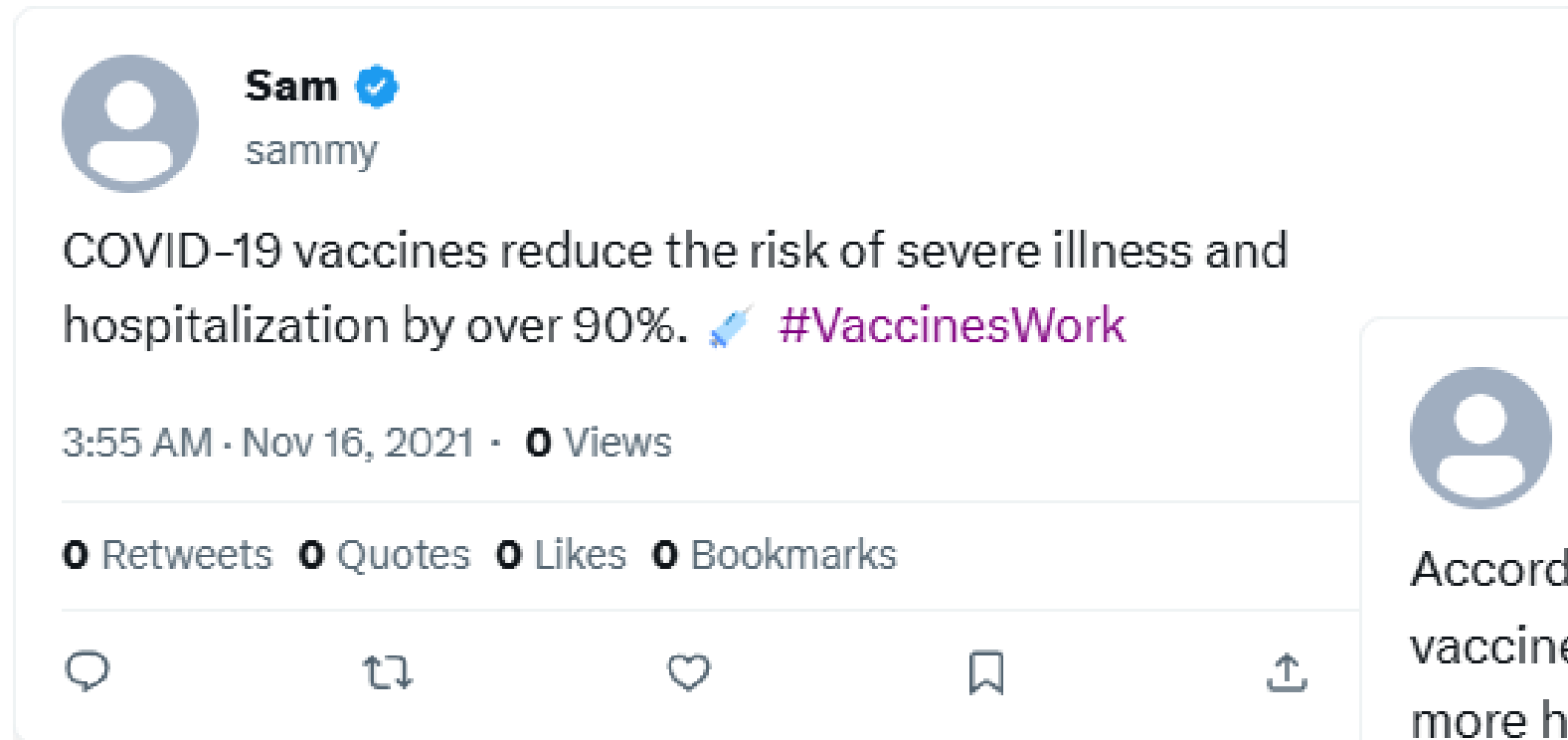
Scientific Content




Scientific Content Pipeline



Scientific Content








Sam 
sammy

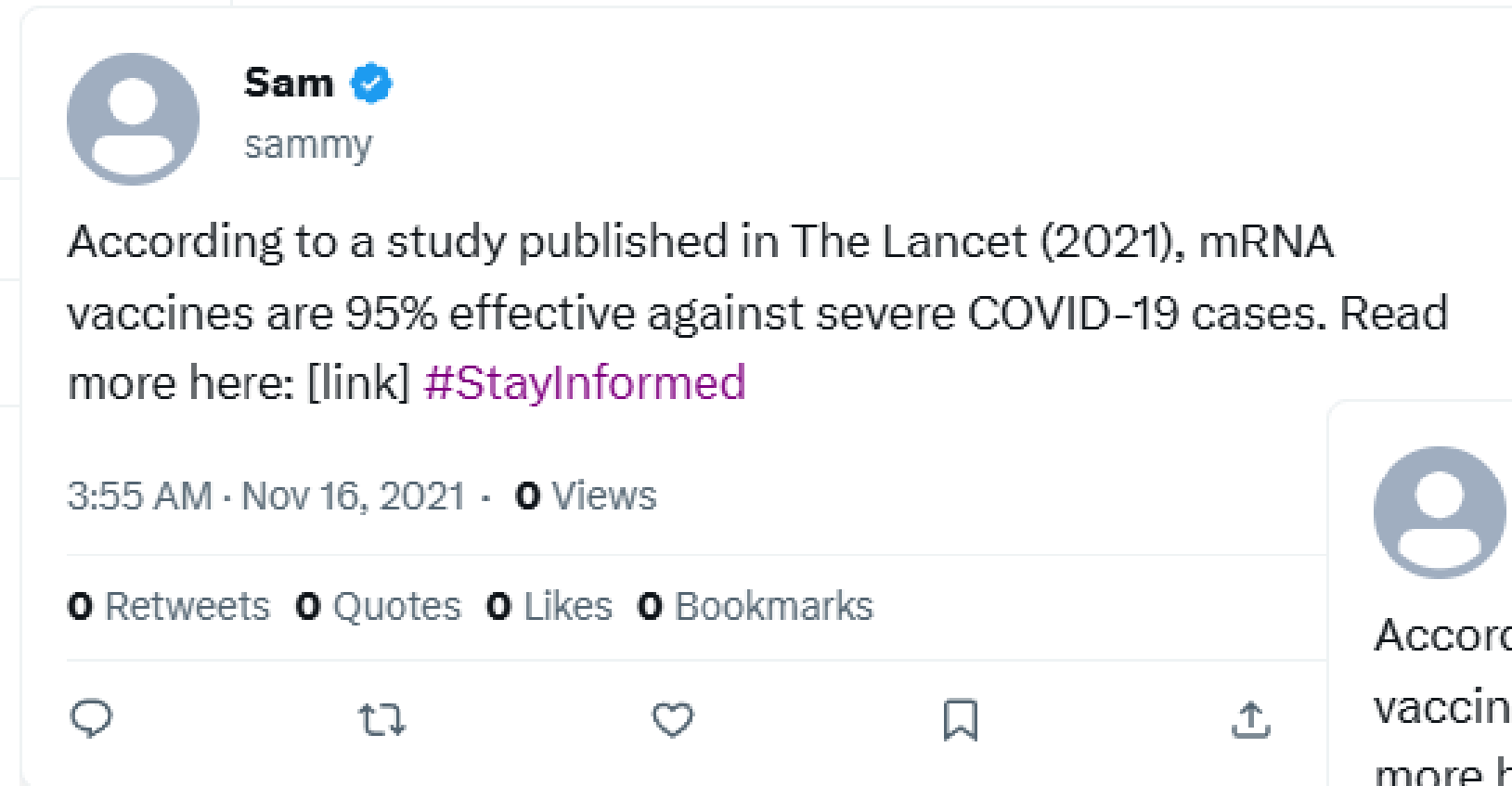
COVID-19 vaccines reduce the risk of severe illness and hospitalization by over 90%. [#VaccinesWork](#)


3:55 AM · Nov 16, 2021 · 0 Views

0 Retweets 0 Quotes 0 Likes 0 Bookmarks

Scientific Claim








Sam 
sammy

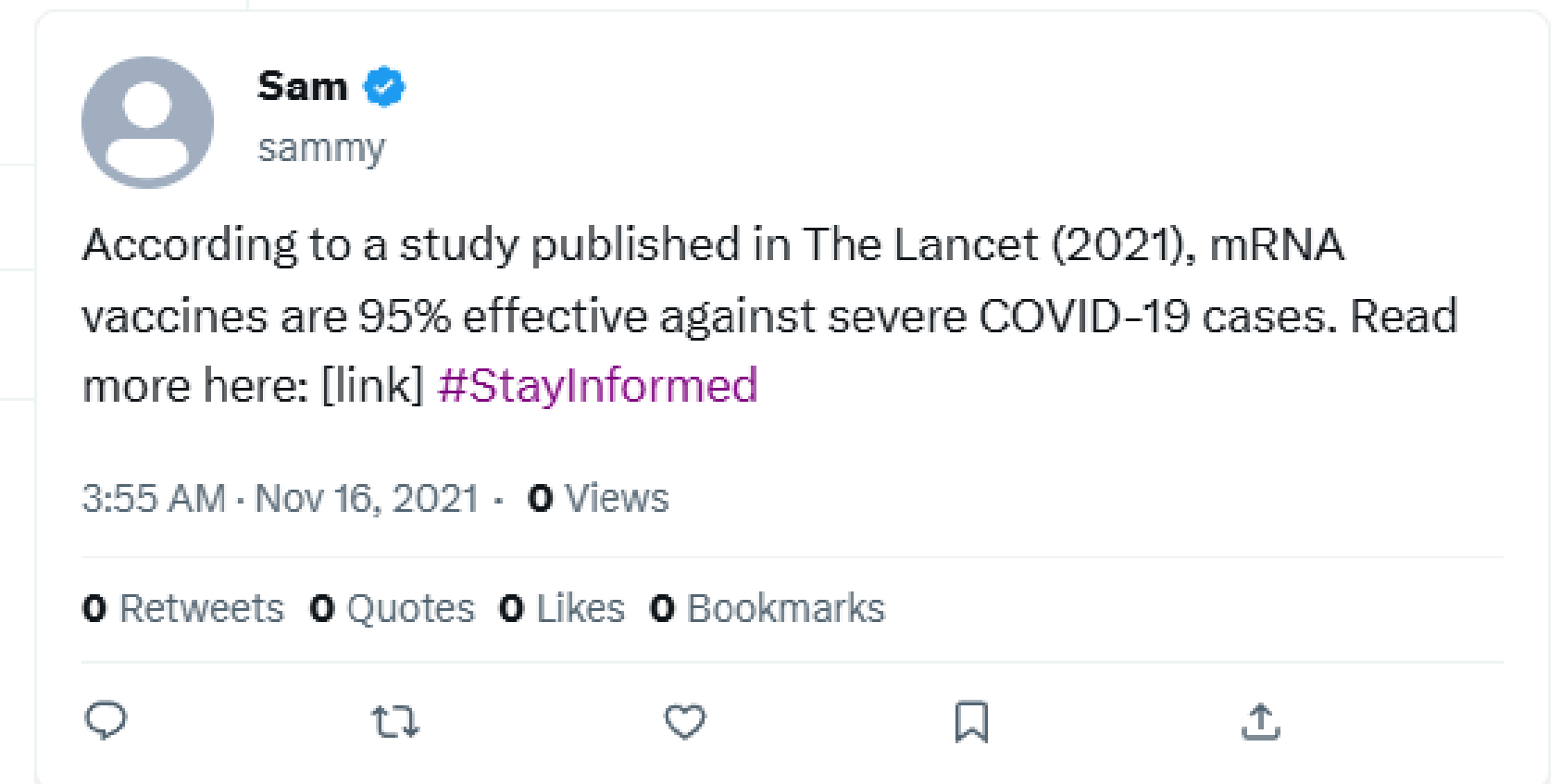
According to a study published in The Lancet (2021), mRNA vaccines are 95% effective against severe COVID-19 cases. Read more here: [\[link\]](#) [#StayInformed](#)


3:55 AM · Nov 16, 2021 · 0 Views

0 Retweets 0 Quotes 0 Likes 0 Bookmarks

Reference








Sam 
sammy

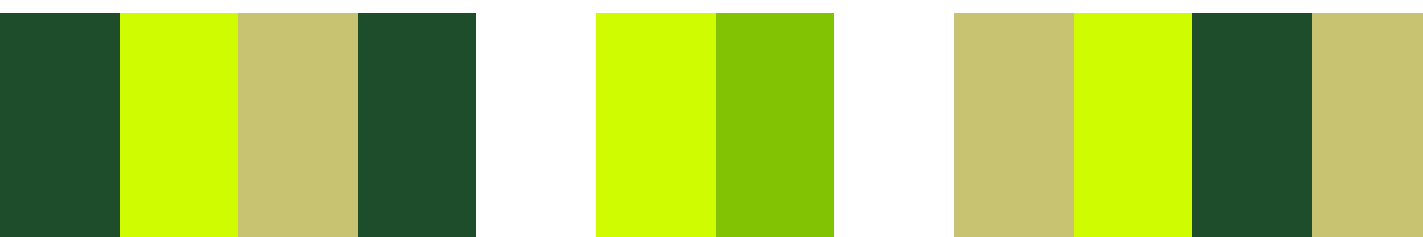
According to a study published in The Lancet (2021), mRNA vaccines are 95% effective against severe COVID-19 cases. Read more here: [\[link\]](#) [#StayInformed](#)

3:55 AM · Nov 16, 2021 · 0 Views

0 Retweets 0 Quotes 0 Likes 0 Bookmarks

General Claim



Scientific Content

Task	Label	BERT				SetFit				Task Specific Model				GPT-3.5				LLaMA-2-70B			
		P	R	A	F1	P	R	A	F1	P	R	A	F1	P	R	A	F1	P	R	A	F1
Scientific Claim	Claim Existence	0.85	0.69	0.77	0.76	0.88	0.80	0.85	0.84	0.85	0.76	0.81	0.80	0.60	1.00	0.60	0.67	0.68	0.88	0.67	0.67
	Verifiable	0.85	0.83	0.85	0.84	0.95	0.99	0.97	0.97	0.86	0.83	0.82	0.80	0.70	0.72	0.71	0.70	0.72	0.71	0.72	0.71
	Reference	0.77	0.51	0.87	0.49	0.64	0.78	0.89	0.70	0.88	0.84	0.88	0.87	0.67	0.71	0.72	0.67	0.77	0.74	0.71	0.73
	General	0.80	0.66	0.93	0.70	0.53	0.70	0.92	0.61	0.94	0.92	0.93	0.92	0.71	0.71	0.72	0.67	0.61	0.70	0.71	0.64

- Results

- **SetFit** excelled in detecting **verifiable claims**.
- **SciBERT** performed best for categorizing **scientific** and **general claims**.
SciBERT pre-trained on Scientific claim dataset. Fine-tuned on our dataset

- Observation

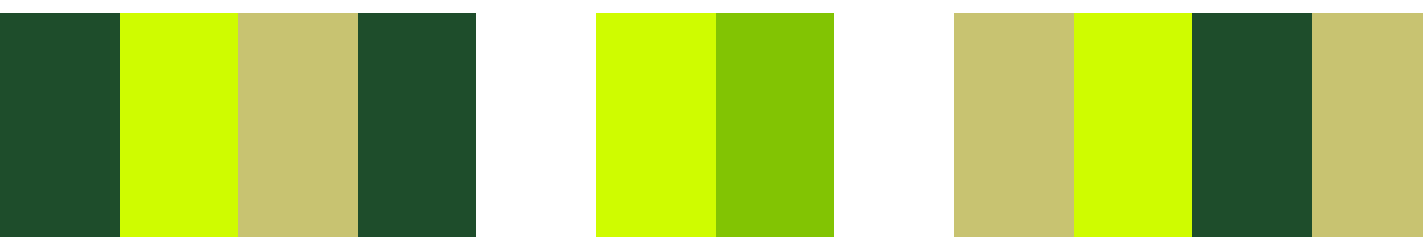
- Chat-based models had limited success in identifying scientific content

Scientific Content

Topic	Ver		Ref		Gen	
	Ind(%)	Org(%)	Ind(%)	Org(%)	Ind(%)	Org(%)
Vaccine	41.00	48.91	46.73	53.24	28.17	34.28
Travel	45.96	39.30	55.86	44.12	30.22	25.65
Face Mask	30.47	58.00	34.78	65.20	19.21	34.91
Social Distancing	37.24	44.81	41.99	47.99	25.03	29.71
Lockdown & Quarantine	39.90	47.97	47.01	52.97	26.76	32.32
Work From Home	42.75	47.75	48.54	51.44	28.88	33.52

Findings

- Individuals: more negative emotions and are subjective
- Organizations: neutral tone and are more optimistic
- Both groups: many unverified claims, highlighting misinformation issue



Limitation

- Dataset limitations:
 - English Tweets only
 - 10,000+ followers or verified accounts
 - Affects generalizability
- Challenges:
 - Subjectivity in emotion/sentiment annotation

Future Work

- Develop influencer impact metric
- Study effects of questionable posts on audience
- Analyze influencer-audience dynamics with advanced computation

Conclusion

- Deeper insight into influencer impact during crises
- Tools to track emotional/scientific social media content
- Models to inform healthier online discussion strategies



**COLORADO STATE
UNIVERSITY**

Questions?