Harnessing Abstractive Summarization for Fact-Checked Claim Detection

Long Paper (2040) - COLING 2022

Varad Bhatnagar*, Diptesh Kanojia⁺ and Kameswari Chebrolu*
*Department of Computer Science and Engineering IIT Bombay, India
⁺Surrey Institute for People-Centred AI and
Department of Computer Science, University of Surrey, United Kingdom

Index of Content

Introduction

What is Fake News?

Fake News - Any form of misinformation, spread with malicious or non malicious intent among the masses.

=> Ability to trigger events which lead to loss of life and property.

Fake News Impact:

1. US Presidential elections of 2016, 2020
2. Indian General elections of 2019
3. Tamil Nadu Lynching
4. Citizenship Amendment Act

What is Fake News Detection?

Task of detecting rumours, fake news, fake claims circulating on the web / social networks.

Manual FN Detection :

=> Journalistic process, organizations like Snopes, Politifact, Alt News, Factly. High accuracy and robust. <u>Unable to scale, timetaking.</u>
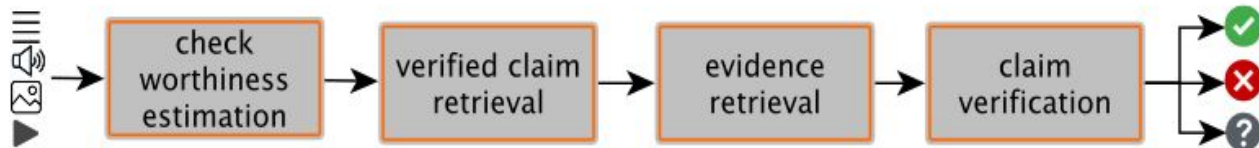
Automatic FN Detection:

NLP, DL and Web Retrieval are used to build systems. Can scale, provide timely results. <u>Suffer from domain / data overfitting, accuracy is not very high, poor generalizability.</u>
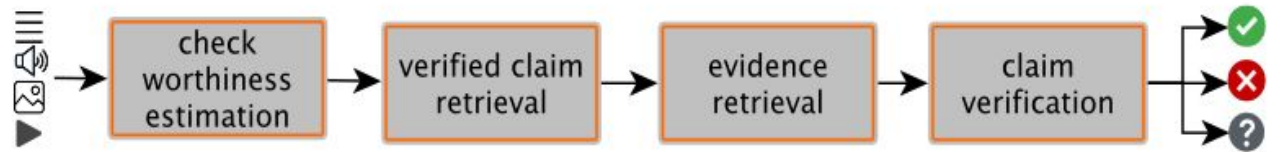
Typical Fact-checking Pipeline

Semi-Automatic FN Detection: Use human intelligence for sub-tasks which require high cognition and automate others using AI.

Focus of this work

Partial / Full Automation

Definitions and Imp. Tools

Definitions

1. **Social Media Claim (SMC):** A social media post (tweet, in this work) containing a claim.
2. **Fact Checked Article (FCA):** An article published by a fact-checking organization accepting or refuting a claim.
3. **Summary of Claim Review (SCR):** A short summary of the claim added by the publishing organization as part of the Claim Review Schema associated with every FCA.
4. **Condensed Claim Representation (CCR):** A summary of the SMC generated using trained models.

Example FCAs

Is This Photo of Windows Holding Back Hurricane Ian Floodwaters Real?

A popular tweet with more than 300,000 likes showed a picture of three windows appearing to withstand the pressure of Hurricane Ian's floodwaters.

By Jordan Liles

Published 3 October 2022, Updated 4 October 2022

Image via @bothcoasts/Twitter

12TH MAY 2021 / 1:24 AM / LAST UPDATED: 12TH MAY 2021

Fact-check: Did Reliance pass off liquid oxygen imported from Saudi Arabia as its own?

BY ARCHIT MEHTA

A video from Delhi where two men riding a motorcycle parallel to a Reliance oxygen tanker claim that the company is passing off oxygen supplied by Saudi Arabia as its own, has been widely shared on social media. The men point at Arabic text visible on the side of Reliance branding.

FOLLOW ALT NEWS

185K 414K
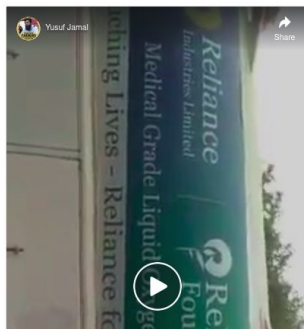
Agneepath scheme protest: Did 'Shami Aslam' pose as Hindu in chat inciting violence?

BY KALIM AHMED

Amidst the ongoing nationwide protests against the implementation of the government's Agneepath Scheme, a new military recruitment policy, a screenshot of an alleged chat on a WhatsApp group is widely circulating on social media. In the group called 'Boycott TOD (Tour of Duty)', some people can be seen conspiring to burn down the Ramadevi police post.

...shot of this alleged WhatsApp chat, Twitter user @puspendrakusav2 claimed ...ush Thakur, who can be seen in the chat inciting people to burn the police ...hami Aslam', a Muslim going by the alias Ayush. The name appears when ...Truecaller, a mobile app that helps with caller identification.

FOLLOW ALT NEWS

185K 414K
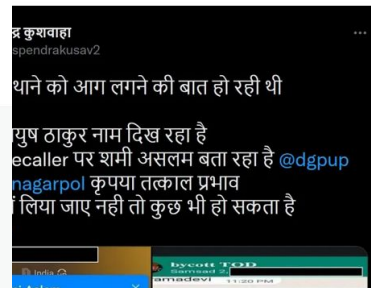
Claim Review Schema

```json
{
    "@context":"http://schema.org",
    "@type":"ClaimReview",
    "datePublished":"2021-10-22",
    "url":"https://www.altnews.in/prakash-javadekar-shares-8-month-old-video-as-govt-supplying-coal-amid-rece
    "itemReviewed":{
        "@type":"CreativeWork",
        "author":{
            "@type":"Person",
            "name":"Prakash Javadekar"
        },
        "datePublished":"2016-06-20"
    },
    "claimReviewed":"Video of 4 km long rack train with 4 engines running to supply coal to power plants",
    "author":{
        "@type":"Organization",
        "name":"Alt News",
        "url":"https://www.altnews.in"
    },
    "reviewRating":{
        "@type":"Rating",
        "ratingValue":"1",
        "bestRating":"5",
        "worstRating":"1",
        "alternateName":"False"
    }
}
```

URL

SCR

This is a web standard.
Fact Checking web pages
such as Alt News and
Boom Live add this
markup to every fact
checking article

Example SMCs and SCRs

#	SMC	SCR
1.	It is a great privilege to witness the present life of Siddhar (Yogi) who went into Jeeva Samadhi 300 years ago in Valliyur in Tamilnadu. He was found alive while digging the mud to renovate the Valliyoor temple. Siddhar is seen sitting in Yogasana. Om Namah Shivaaya 🙏	A Siddhar Yogi who went into Jeeva Samadhi (meditative state) 300 years ago found alive while digging up in Valliyur, Tamilnadu.
2.	okay i always knew they used to put cocaine in coca cola but i just found out it was about 3.5 GRAMS a bottle and oh my god ???	Coca-Cola used to (or still does) contain cocaine.
3.	Air India flight from their Chicago-Delhi flight … passengers were charged 3x Fares to "cover" social-distancing and Air India then "packed" the flight … See the actual scenario inside the plane!! @RubikaLiyaquat @s @aajtak @ABPNews @ "	Air India charged 3 times the fare and did not maintain social distancing in flight
4.	Today in Gaza, the Palestinians pretended to be at a funeral and photographed it so that they would feel sorry for them. But then an alarm sounded ... oops 🤭😕 #israel #IsraelUnderFire #IsraelUnderAttack	Video of Palestinians faking casualties amid the ongoing conflict with Israel

Google Fact Check Engine

Google Fact Check Tools hosts a search engine that can be used to search for fact checking articles

Proposed Workflow

Workflow

From Check Worthiness Estimation

If No Match Found: To Evidence Retrieval

SMC

Set of Similar Claims

Our Proposed Workflow

Source Specific Preprocessing

Abstractive Summarization

CCR

Retriever

Previously Proposed Workflows

Collection of Previously Fact-Checked Claims
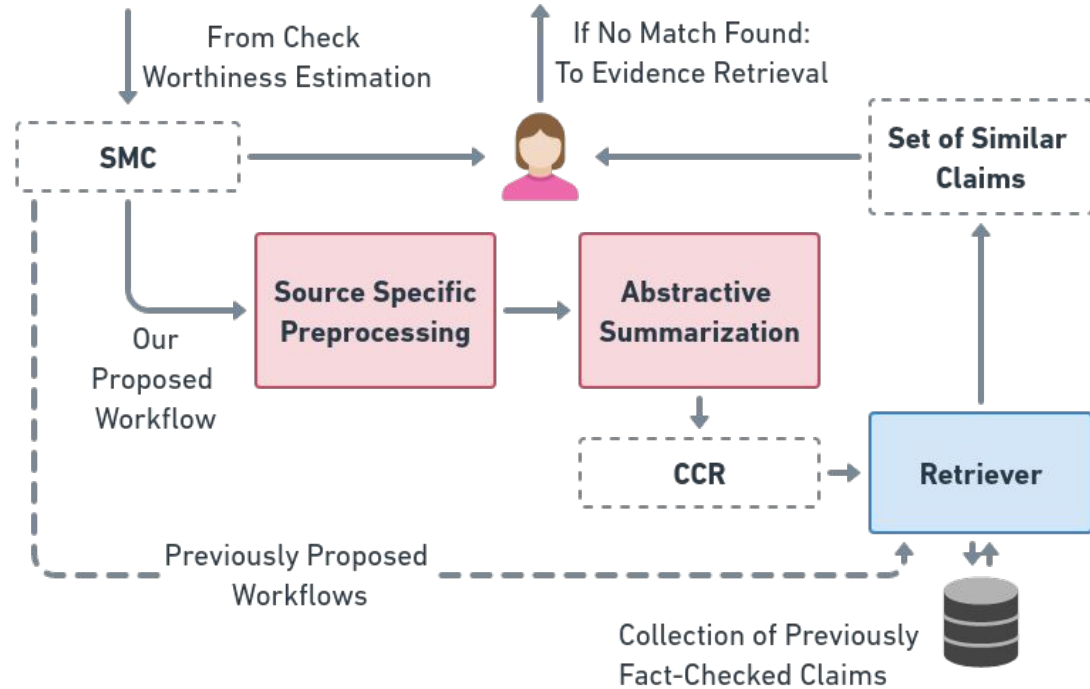
Erstwhile Work

Shaar et. al. release two datasets for this task:
1. Politifact Dataset: News article based
2. Snopes Dataset: Twitter based

Train BM25 and BERT (Devlin et al., 2019) based learning to rank models on their dataset for this task.

Mansour et. al. also solve the problem for Arabic data using BERT based rankers.

Delegate responsibility of dealing with text noise and length to the ranker

Datasets

Why Curate Own Dataset?

=> Most available summarization datasets are at document level.

=> Most datasets are clean in nature (books, news articles, blogs, etc) but Fake news thrives in noisy settings (social media, instant messaging, etc)

Curation and Annotation Strategy

1. Look for FCA URLs in comments of tweets.

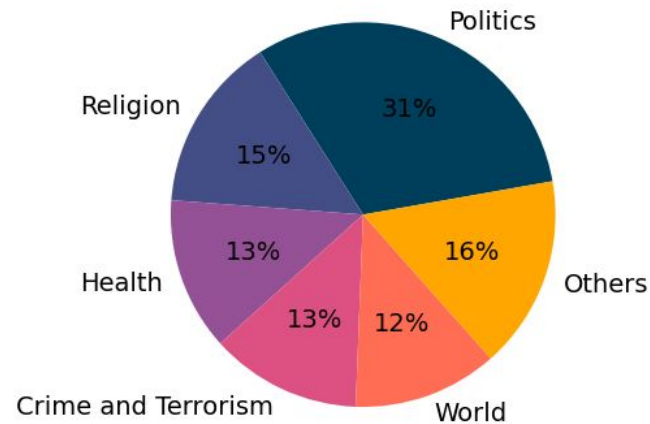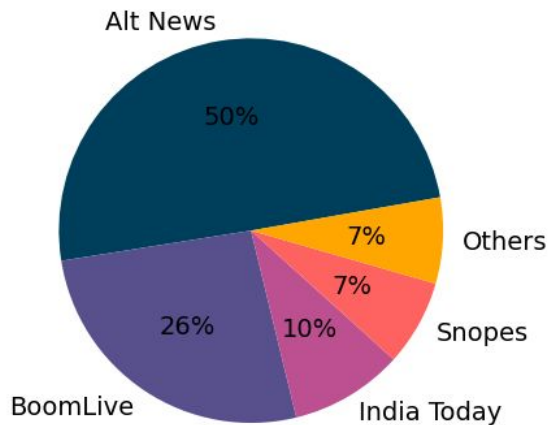   (Such tweets between 2018 and 2022 were collected for 11 FCA publishers)

2. Manually annotate if tweet summarizable to SCR (derived from the corresponding FCA).
3. Also manually annotate language and topic.

Dataset Distributions

SMC → English
SCR → English

Dataset Statistics

	Snopes	Eng-Eng Dataset
Data Entity	**Count**	
<SMC,SCR>pairs	1000	567
Unique SMC	997	531
Unique SCR	787	369
FCA Source Country	**Proportion of FCAs**	
India	0%	93%
US	100%	7%

Median Length	SMC	SCR	SMC	SCR
Chars	198	110	193	70
Words	30	17	33	11

Word Level TFIDF Cosine Similarity Threshold	**Preprocessed Set**		
	NP	**NP**	**P-H-M**
0.25	50%	59%	61%
0.5	8%	12%	13%
0.75	1%	2%	3%

Contextual Cosine Similarity Pretrained Model	**Preprocessed Set**		
	NP	**NP**	**P-H-M**
BERT	0.7	0.73	0.72
IndicBERT	0.86	0.84	0.85

Degree of non 1:1 correspondence

Target Length for CCR

Higher Word Level TF-IDF Cosine Sim.

Experiments

Experiments

1. Baselines -

   a) Verbatim Querying

   b) Truncate 11

2. Twitter Specific Preprocessing (Hashtags, Mentions, Emojis, etc)
3. Fine-tuning Pretrained Models (T5, BART, Pegasus)
4. Decoding Strategies (Greedy, Beam Search, Nucleus)
5. Parameter Scaling (T5 Base, T5 Large)
6. Recall Plateauing

Twitter Specific Preprocessing

1. Hashtag and Mention Symbol Removal
2. Hashtag and Mention Removal
3. Emoji Replacement
4. Run of Hashtag and Run of Mention Removal

   Eg: This video is going viral that if you get a call from a mobile phone number starting with 140 your mobile will be hacked & balance will disappear. Is it a genuine video? @assampolice @DGPAssamPolice @gpsinghips @mygovassam

5. Mention replaced by Twitter Handle

   Eg: @BJP4TamilNadu -> BJP Tamilnadu

Pretrained Models Used

Model	Authors	Pre-training Objective	Number of Params
T5	Google	Multi Task	220M (Base)
BART	Facebook	Denoising autoencoder	300M (Distilled)
Pegasus	Google	Gap Sentence Generation	370M (Distilled)

Transformer Encoder Decoder

Image source: https://jalammar.github.io/illustrated-transformer/
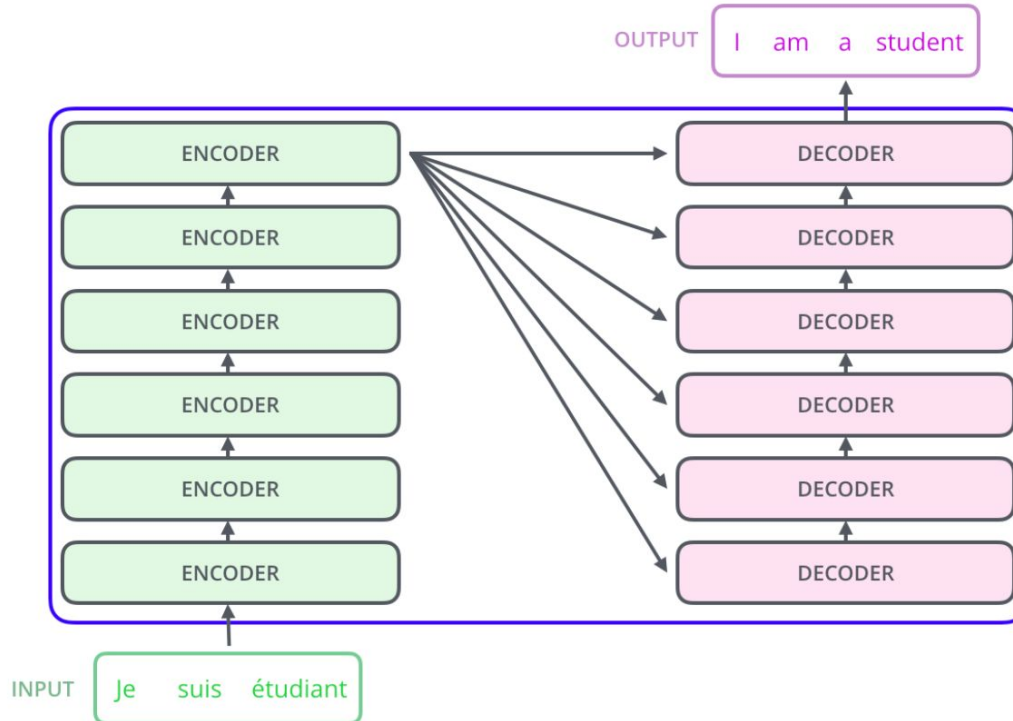
Decoding Strategies: #1 Greedy
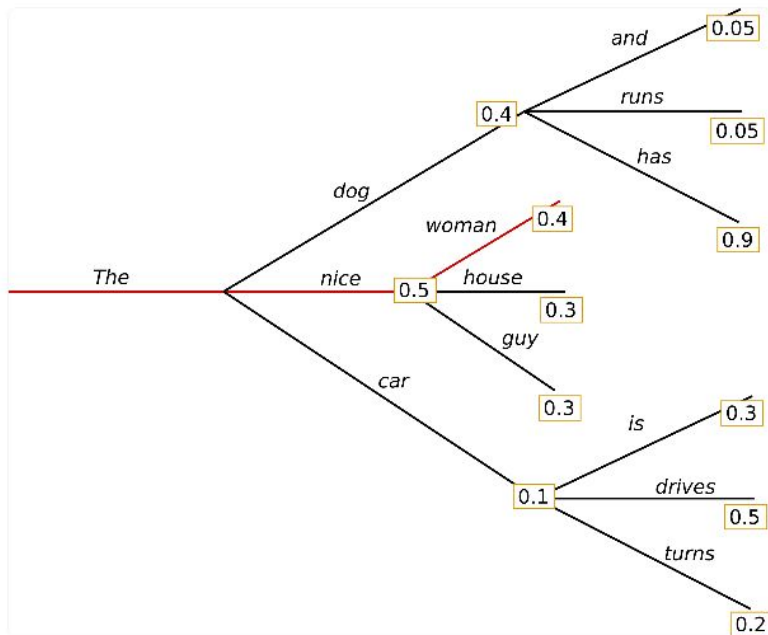
Select token with highest probability at current timestep t, conditioned on the token choices at previous time steps.

$$w_t = \text{argmax}_w P(w|w_{1:t-1})$$

All decoding images source: https://huggingface.co/blog/how-to-generate

Decoding Strategies: #2 Beam Search
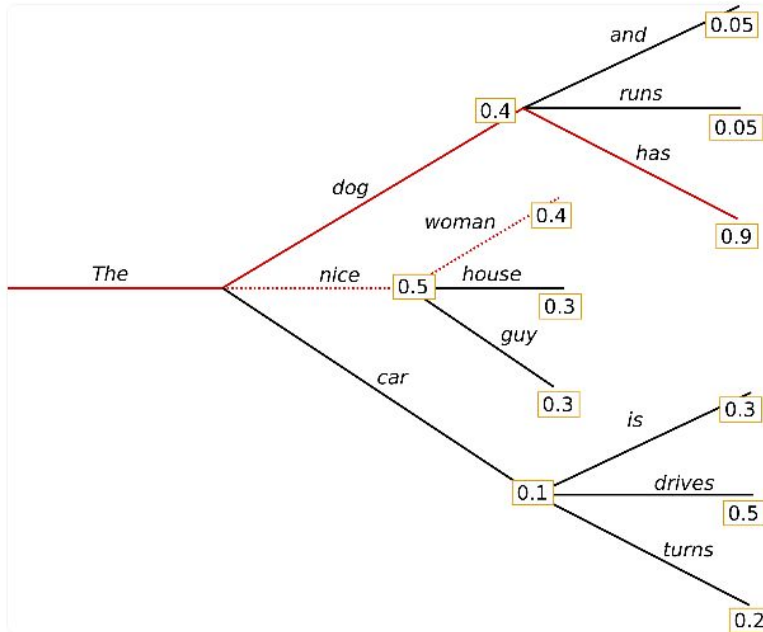
Maintain *k* most probable hypothesis at every time step and eventually choose the hypothesis with the overall highest probability.

$k \rightarrow$ Beam Width

Decoding Strategies: #3 Top K Sampling

$$\sum_{w \in V_{\text{top-K}}} P(w|\text{"The"}) = 0.68$$

$$\sum_{w \in V_{\text{top-K}}} P(w|\text{"The"}, \text{"car"}) = 0.99$$

$P(w|\text{"The"})$

nice dog car woman guy man people big house cat

$P(w|\text{"The"}, \text{"car"})$

drives is turns stops down a not the small told

Sharp Distribution
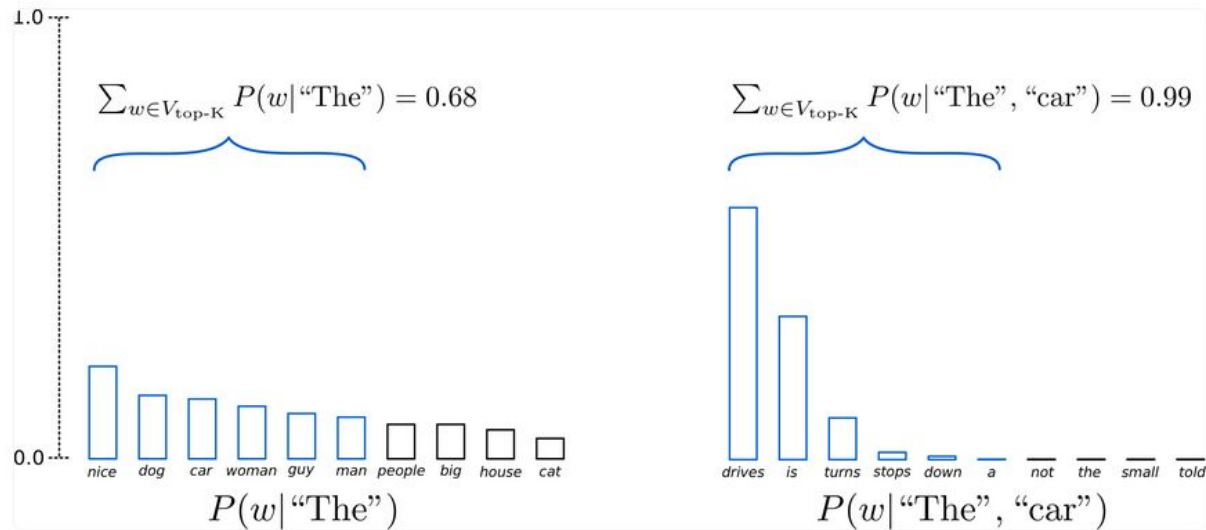
Filter the next *K* most likely words and redistribute the probability mass among them, followed by sampling.

Decoding Strategies: #4 Top P Sampling
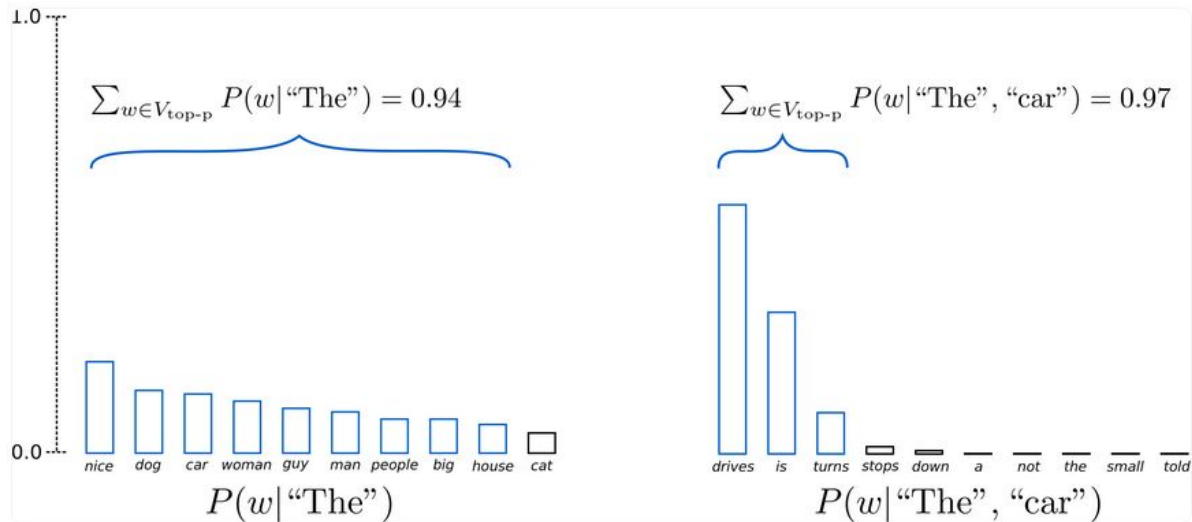
Instead of sampling from the next K words, it samples from the smallest set of words whose cumulative probability exceeds P.

Parameter Scaling

- Many variants of T5 with the same learning objective and pre-training data are available such as :

  T5-small: 60M

  T5-base: 220M

  T5-large: 770M

  T5-3B: 3B

  T5-11B: 11B

- Study retrieval performance gain as a function of parameter scaling

Evaluation Metrics

Summary Quality Evaluation

Word Level Similarity:

a. TF-IDF Cosine Similarity
b. BLEU4

Retrieval Evaluation

1. Recall @ k

$$Recall@k = \frac{|\{Relevant\ Documents\}\ \cap\ \{Top\ k\ Retrieved\ Documents\}|}{|\{Relevant\ Documents\}|}$$

2. Mean Reciprocal Rank

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i}$$

Where Q is set of queries

Results

Decoding and Param Scaling Results

Param Scaling

Decoding

T5 Large almost always better than T5 Base but slight gain offset by training time

Beam Search with beam size 6 and ngram 2 performs best

Recall Plateauing

'Elbow' is seen in all curves at Recall@5
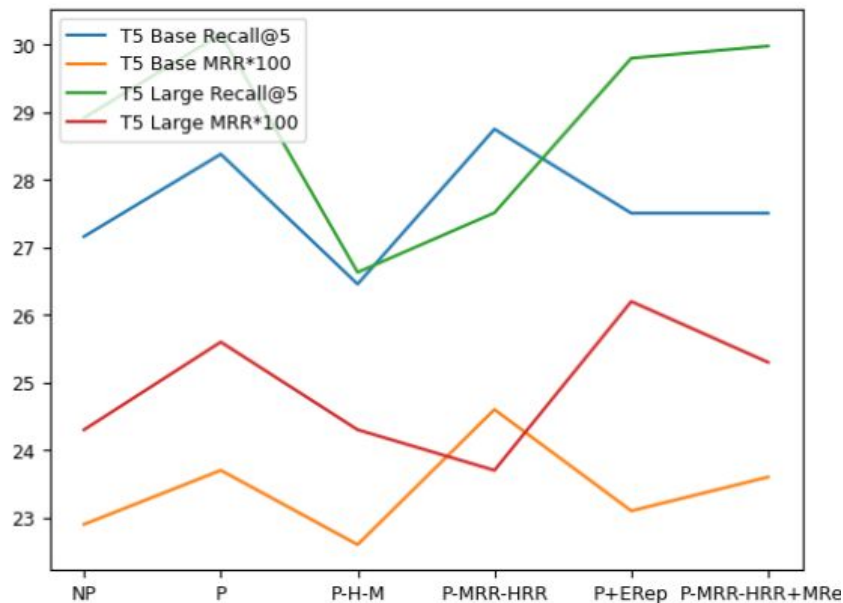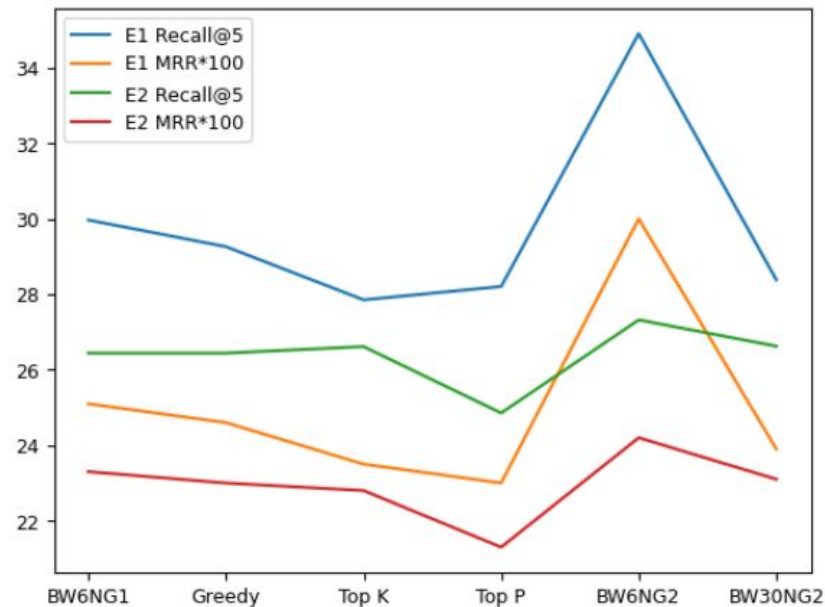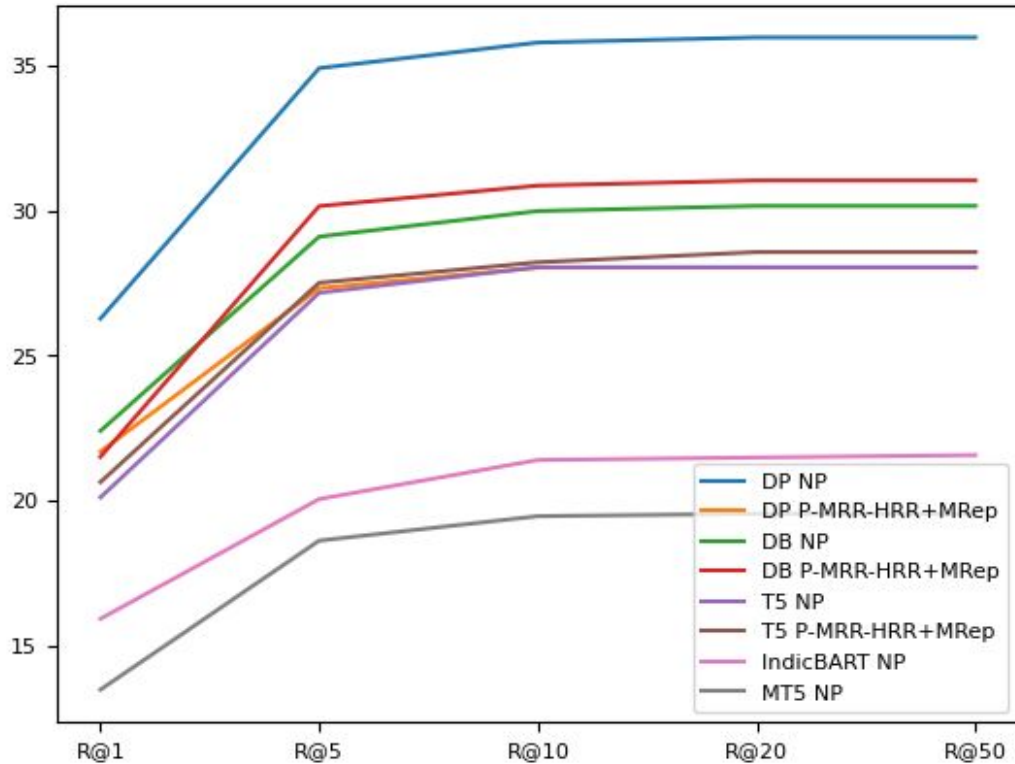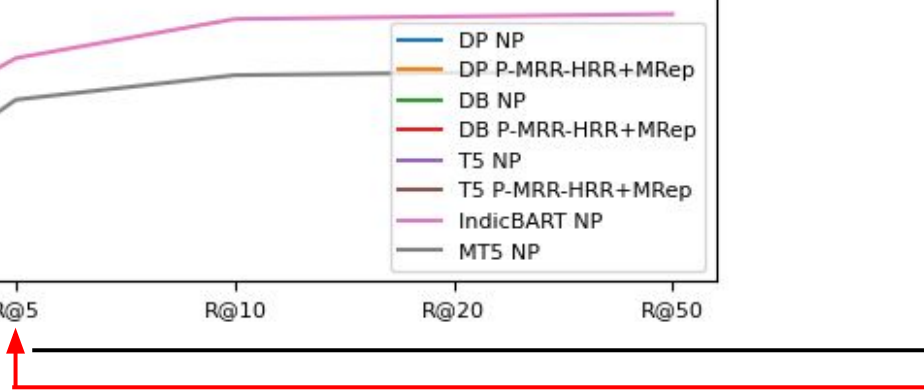
Retrieval Results

Preprocessing Strategies	No Summarization				Summarization using Out of the Box Models						Summarization using Fine Tuned Models					
	None		Truncate11		T5		D BART		D PEGASUS		T5		D BART		D PEGASUS	
	R@5	MRR	R@5	MRR	R@5	MRR	R@5	MRR	R@5	MRR	R@5	MRR	R@5	MRR	R@5	MRR
NP	9.52	.09	17.28	.14	15.52	.13	20.46	.17	**22.40**	**.19**	27.16 ±2.55	.23 ±.02	29.10 ±3.15	.26 ±.02	**34.91 ±5.91**	**.30 ±.05**
P	11.99	.11	18.34	.15	17.64	.15	17.99	.14	21.52	.17	28.38 ±6.55	.24 ±.05	28.21 ±6.88	.24 ±.06	27.69 ±2.63	.24 ±.02
-H	12.70	.12	17.99	.15	17.28	.15	17.64	.14	20.46	.16	24.87 ±5.13	.21 ±.04	30.15 ±6.08	.26 ±.06	29.61 ±6.59	.25 ±.05
-M	13.05	.12	18.69	.15	18.34	.15	17.64	.14	21.16	.17	26.80 ±5.33	.23 ±.03	*30.15 ±4.98*	.26 ±.04	29.10 ±2.45	.25 ±.02
-H-M	13.93	.13	**18.87**	**.15**	17.81	.15	17.28	.14	20.28	.16	26.46 ±4.69	.23 ±.03	27.51 ±4.09	.23 ±.04	26.28 ±2.36	.23 ±.02
+ERep	10.58	.10	17.46	.14	16.58	.14	17.99	.15	22.05	.18	27.51 ±6.13	.23 ±.05	28.20 ±5.89	.25 ±.05	30.15 ±5.72	*.27 ±.05*
-MRR-HRR	12.35	.11	17.81	.15	17.46	.15	17.81	.14	21.69	.18	28.75 ±6.38	.25 ±.05	27.33 ±5.94	.23 ±.05	25.92 ±1.98	.22 ±.01
-MRR-HRR +MRep	12.70	.12	**18.87**	**.15**	18.34	.15	17.81	.14	21.87	.18	27.51 ±5.43	.24 ±.04	30.15 ±5.98	.25 ±.05	27.32 ±3.87	.24 ±.03
Skyline	**63.85**	**.55**														

Preprocessing helping baseline models

Truncate11 performs as well as T5 OOB

DP OOB Performs well compared to Baselines

FT Gain

Best Model | Preprocessing doesn't help FT models

CCR Quality Results

Experiment	n-Gram	Cosine Similarity Threshold			BLEU4
		0.25	0.5	0.75	
E1 SMC vs CCR	1	80%	26%	2%	-
	2	83%	30%	2%	-
E1 SCR vs CCR	1	76%	38%	14%	39.7
	2	73%	38%	13%	39.3
E2 SMC vs CCR	1	83%	26%	4%	-
	2	85%	29%	3%	-
E2 SCR vs CCR	1	68%	31%	11%	38.9
	2	69%	33%	13%	39.2

Higher TF-IDF Cosine Similarity compared to the corresponding figures in Slide

High BLEU Scores for CCRs

E1 and E2 stand for Distilled PEGASUS with
NP and P-MRR-HRR+MRep experiments respectively

Qualitative Results

#	SMC	SCR	CCR
1	Congratulations to Uttarakhand CM for becoming the first CM ever to charge stranded citizens for rescue operations! Helicopter rides will now be chargeable during rescue operations in Uttarakhand. And if you can't pay, you may safely die. #AchheDin #BJP	Passengers in Uttarkhand to be charged for rescue operations	Uttarakhand CM has charged stranded citizens for helicopter rides during rescue operations
2	@AltNews We are getting various WhatsApp forward regarding as Corona has been emerges only due to 5G testing in world. Please put some light, seems ,it is only a brain shit.	5G radiation is the cause behind the second wave of coronavirus pandemic in India	Coronavirus outbreak due to 5G testing
3	This woman in Afghanistan was killed by Taliban for not wearing the proper cloth. #Afghanistan #Taliban @cnn @FoxNews @BBCWorld	Video shows a woman being shot in the head by Taliban in Afghanistan for not dressing appropriately	Woman killed by Taliban for not wearing proper cloth
4	"India is ranked 102nd in the global hunger index, out of 117 countries. We are ranked in between Niger & Sierra Leone. We are the lowest ranked South Asian country. Bangladesh is ranked 88th and Pakistan 94th. They have only recently overtaken us. Our rank was 55,only 5 years ago"	India's ranking in Global Hunger Index (GHI) has fallen from 55 in 2014 to 102 in 2019	India ranked 102nd in the global hunger index ←—— Model Unable to Reas
5	"Oxygen donated from Saudi and relabelled in india by Reliance, Share this with your contacts in Saudi and make this viral .. Let the world know the cheapness of this PM "	Oxygen sent from Saudi Arabia is being distributed in the name of Reliance	Reliance taking credit for oxygen supplied by Saudi Arabia

Goto Drawbacks

Graph View

Drawbacks

1. Reasoning: [Example](#)

2. Summary-Quality Length Tradeoff: [Example](#)

3. Lack of Cross Lingual Data: [Example](#)

4. Poor Hit Rate During Annotation: Only 1 tweet in 10 is summarizable.

Summary

- Proposed a novel workflow for detecting previously fact-checked claims using Abstractive Summarization.
- Curated and annotated 2 tweet level summarization datasets: English-English and Multilingual.
- Performed Experiments using a variety of preprocessing techniques and open–source pretrained models.
- Performed qualitative analysis of outputs at various stages in the workflow.
- Submitted English-English work to COLING 2022

Extensibility

1. Data Extensibility - Collect and annotate SMCs for other languages using our curation mechanism. X-Fact by Gupta et.al. lists FCA publishers for 25 languages.

2. Workflow Extensibility - Replace text summarization with text generation module to support image/text/video claims.

Future Work

1. Explore graph based linkage.

2. Extend work to other subcontinental languages such as Hindi.

3. Controlled generation using Names Entities ([Zhang et. al.](#))

4. White box retriever.

Derived Edges: Same Publisher

फ़ैक्ट-चेक: चित्तौड़ के इस मंदिर को मुगलों ने मस्जिद में बदल दिया था?

BY KALIM AHMED

Share   Tweet

FOLLOW ALT NEWS

f	
62K	**69K**

सोशल मीडिया पर मंदिर की तरह दिखने वाली एक स्ट्रक्चर की तस्वीर अलग-अलग कैप्शन के साथ वायरल है. मंदिर जैसी दिखने वाली इस स्ट्रक्चर के ऊपर एक गुंबद है. कैप्शन के मुताबिक, ये एक मंदिर था जिसे बाद में मस्जिद बना दिया गया था.

भाजपा सदस्य सुरेंद्र पूनिया ने ये तस्वीर ट्वीट करते हुए लिखा, "मुगलों और बाकी इनवेडर्स का आर्किटेक्चर इतना बेजोड़ और विचित्र था कि उन्होंने जो भी बनाया, उसके बेसमेंट में हमेशा मंदिर रखा." सुरेंद्र पूनिया ने ये तस्वीर इस कांस्पिरेसी के संदर्भ में ट्वीट की कि ताजमहल के तहखाने में कथित 22 कमरों में हिंदू मूर्तियां बंद हैं.

Major Surendra Poonia ✔
@MajorPoonia

मुगलों और बाकी Invaders का Architecture इतना बेजोड़ और विचित्र था कि उन्होंने जो भी बनाया उसके Basement में हमेशा मंदिर ही रखा
#KnowYourHistory

यही खूबसूरती तो है मेरे देश की ...

No, this temple in Chittor was not converted into a mosque

BY KALIM AHMED

Share   Tweet

FOLLOW ALT NEWS

f	
185K	**403K**

A photograph of a structure that appears to be a temple with a dome on top is viral on social media with various captions that suggest that the structure was a temple that was converted into a mosque.

BJP member Surendra Poonia shared the photo with a caption in Hindi which reads, "The architecture of the Mughals and the rest of the invaders was so unique and strange that whatever they built, they always kept a temple in its basement." Poonia's tweet is a reference to the conspiracy that 22 locked rooms in the basement of the Taj Mahal house Hindu idols.

Major Surendra Poonia ✔
@MajorPoonia

मुगलों और बाकी Invaders का Architecture इतना बेजोड़ और विचित्र था कि उन्होंने जो भी बनाया उसके Basement में हमेशा मंदिर ही रखा
#KnowYourHistory

चित्तौड़ में एक मंदिर को मस्जिद बना दिया गया.

Old Hindu temple in Chittor converted into a mosque by Mughals

Derived Edges: Different Publisher

Fact Check: This Picture Does Not Show Temple Turned into Mosque By Mughals

By **News Mobile Fact Check Bureau** · 2022-06-04

Subhra Gupta is 😠 feeling frustrated.
17 May at 15:56 · ⚙

Stunningly beautiful walls | with an ugly looking dome | Chittor temple converted into a mosque by Mughals...

FACT CHECK

| f Share | 🐦 Tweet | 8 Google | 🟢 Whatsapp | 🖨 Print | + |

Many social media users are sharing an image of a religious structure with a dome that looks like a temple. The image is shared with a claim that it is of a mosque, which was built on the base of a Hindu temple.

Newsmobile

2ND JUNE 2022 / 2:14 PM

No, this temple in Chittor was not converted into a mosque

BY KALIM AHMED

f Share   🐦 Tweet   ✉   ...

A photograph of a structure that appears to be a temple with a dome on top is viral on social media with various captions that suggest that the structure was a temple that was converted into a mosque.

BJP member Surendra Poonia shared the photo with a caption in Hindi which reads, "The architecture of the Mughals and the rest of the invaders was so unique and strange that whatever they built, they always kept a temple in its basement." Poonia's tweet is a reference to the conspiracy that 22 locked rooms in the basement of the Taj Mahal house Hindu idols.

FOLLOW ALT NEWS

| f 185K | 🐦 403K |

Major Surendra Poonia ✔
@MajorPoonia

मुगलों और बाकी Invaders का Architecture इतना बेजोड़ और विचित्र था कि उन्होंने जो भी बनाया उसके Basement में हमेशा मंदिर ही रखा #KnowYourHistory

Alt News

Thank You

Backup Slides

Annotation Tool

Annotation Tool

GUI based Django Webapp

Motivation: Many URLs (Twitter, Alt News, etc) need to be opened to label a sample. Reduce error due to excel sheet.
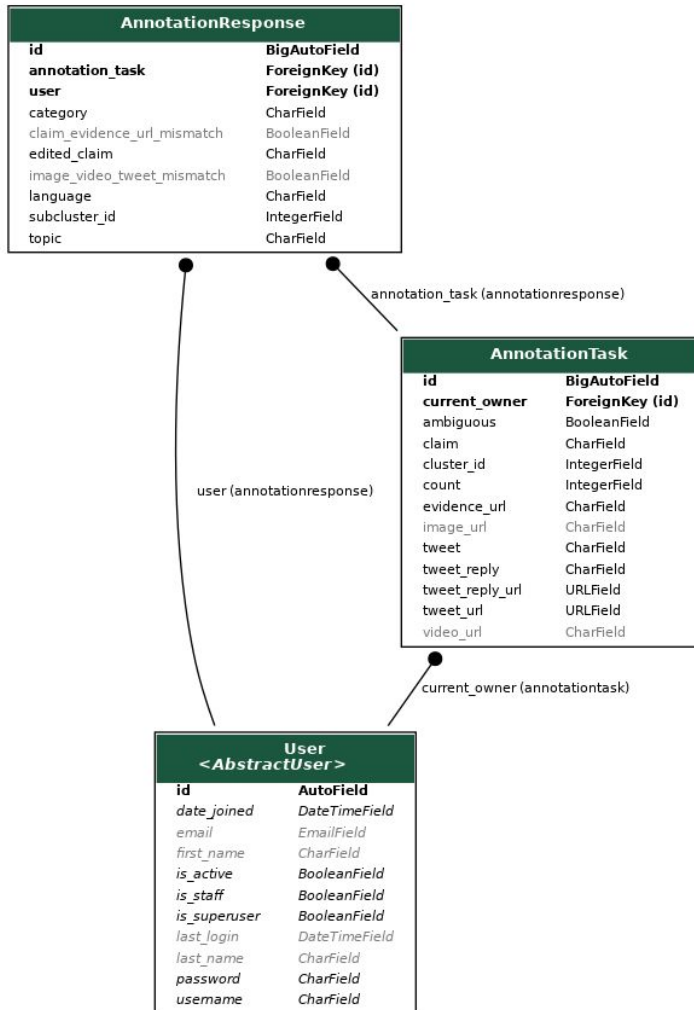
Can be used for all types of data annotation for Fake News Tasks in future as well.

Used for annotating summarization data as well.

Annotation Tool Class Diagram

AnnotationResponse

id	BigAutoField
annotation_task	ForeignKey (id)
user	ForeignKey (id)
category	CharField
claim_evidence_url_mismatch	BooleanField
edited_claim	CharField
image_video_tweet_mismatch	BooleanField
language	CharField
subcluster_id	IntegerField
topic	CharField

annotation_task (annotationresponse)

user (annotationresponse)

AnnotationTask

id	BigAutoField
current_owner	ForeignKey (id)
ambiguous	BooleanField
claim	CharField
cluster_id	IntegerField
count	IntegerField
evidence_url	CharField
image_url	CharField
tweet	CharField
tweet_reply	CharField
tweet_reply_url	URLField
tweet_url	URLField
video_url	CharField

current_owner (annotationtask)

User
<AbstractUser>

id	AutoField
date_joined	DateTimeField
email	EmailField
first_name	CharField
is_active	BooleanField
is_staff	BooleanField
is_superuser	BooleanField
last_login	DateTimeField
last_name	CharField
password	CharField
username	CharField

Annotation Tool Glimpse

Annotation Task

Annotation Task ID :	17
Tweet URL :	https://twitter.com/KapilMishra_IND/status/1280071399638364164
Tweet :	This is BIG : More than 100 Chinese Soldiers Killed in Galwan : Chinese Military Official Accepted https://t.co/D3g7sWmrcr
Image URLs:	
Video URLs:	https://kreately.in/more-than-100-chinese-soldiers-killed-in-galwan-valley-chinese-military-official-accepted/ |
Tweet Reply URL :	https://twitter.com/AbuBakar0802/status/1280500005124612100
Tweet Reply :	@KapilMishra_IND https://t.co/WAXZZJ3yZw
Evidence URLs:	https://www.boomlive.in/fake-news/no-jianli-yang-did-not-say-more-than-100-chinese-soldiers-died-in-galwan-8773 |
Cluster ID:	11

Mark as Ambiguous

Tasks in Same Cluster

Tweet	Subcluster
Yang Jianli former PLA officer claims 100 Chinese soldiers were killed on 15-16 June night. its a crime to Bury the sacrifice of your soldiers and deny them even a miltary funeral and state honours. 57 million PLA veterans are enraged with chinese govt https://t.co/uv92mALoYo	1

Annotation Tool Glimpse