A Zero-Shot Claim Detection Framework using Question Answering

Revanth Gangi Reddy[1], **Sai Chetan Chinthakindi**[1], **Yi Ren Fung**[1], **Kevin Small**[2], **Heng Ji**[1]

[1]University of Illinois at Urbana-Champaign [2]Amazon Alexa AI

Background and Motivation

- NewsClaims (Reddy et al 2021) extends the claim detection task to extract additional attributes relating to the claim, such as the claimer, claim object etc.
- The benchmark mainly contains claims about COVID-19 from 143 news articles.

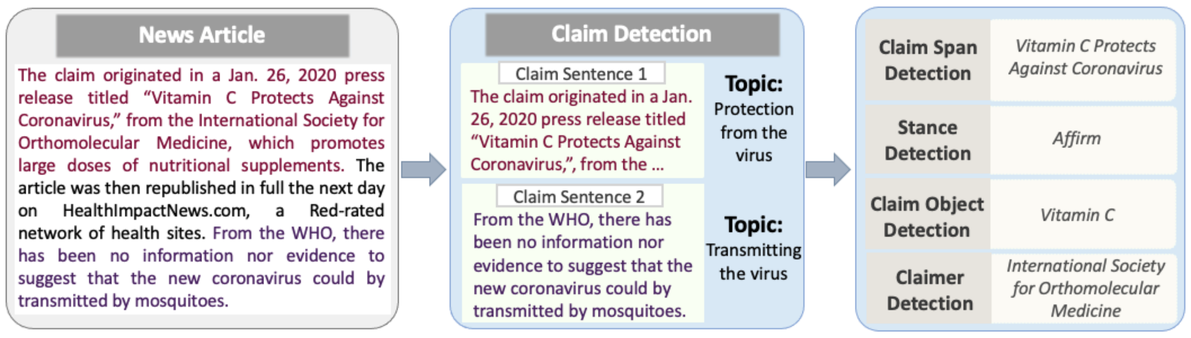

Figure 1: An example from NewsClaims.

- Claim detection approaches need to be able to ported to new scenarios, without access to much training data.
- Further, the claimer detection subtask within NewsClaims requires considerable document-level reasoning, making it harder for existing attribution models which mainly involve sentence-level reasoning.
- We hypothesize that identifying claim topics and extracting corresponding claim attributes can be formulated as a Question Answering (QA) task.

QA-based Claim Detection Framework

- The same extractive QA model solves multiple sub-tasks within claim detection, without the need of any task-specific training data.
- This involves:
  - Filtering claims relating to specific topics
  - Identifying claim objects associated with such topics
  - Attribution for identifying the claimers making these claims

CODE

- We realise this by using directed questions to help solve connected sub-tasks.

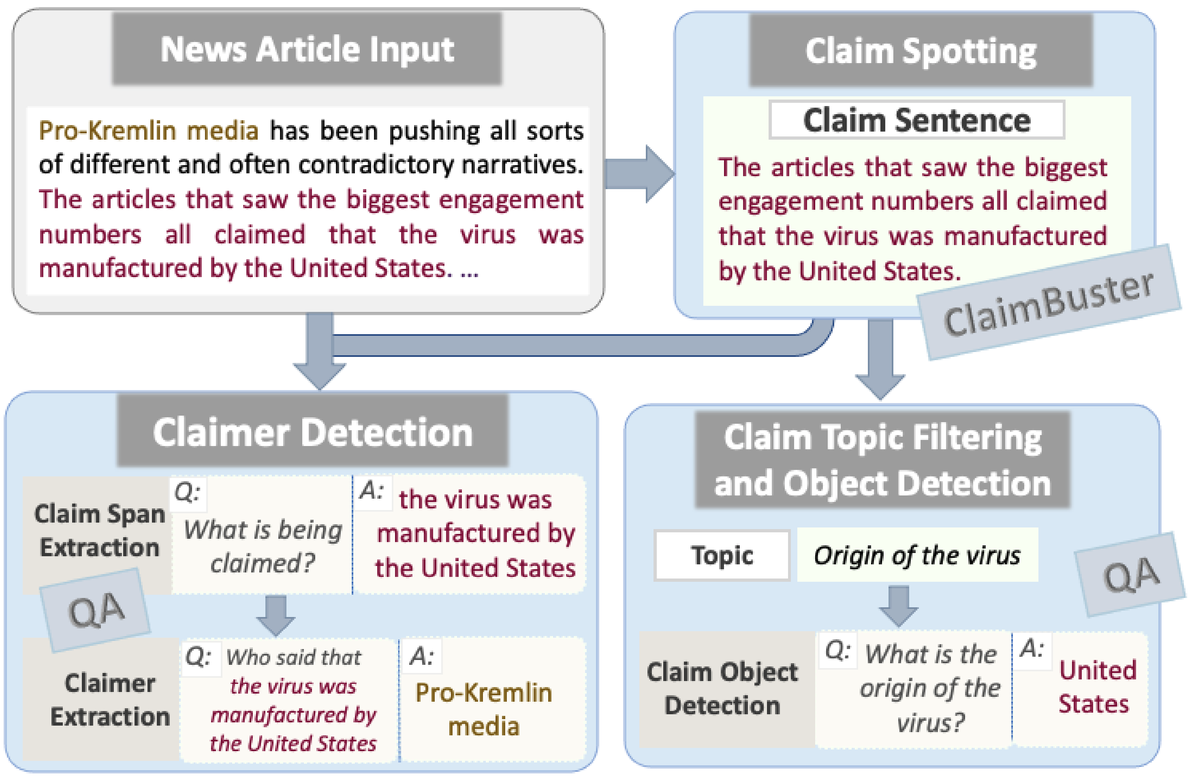

Figure 2: Proposed QA-based claim detection framework.

Claim Topic Filtering and Claim Object Detection

Claim Topic Filtering

- Topic filtering by measuring topic relevance as the answer confidence from a QA model.
- Input to the QA model comprises of the claim sentence passed as context along with the question corresponding to the individual topic.
- Claims are filtered based the highest topic score using a threshold.

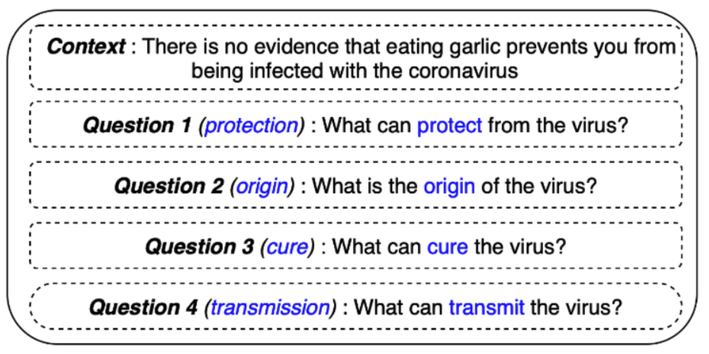

Figure 3: Questions corresponding to individual topics.

Model	P	R	F1
ClaimBuster	13.0	**86.5**	22.6
ClaimBuster + NLI	21.8	53.3	30.9
ClaimBuster + QA	**30.7**	43.4	**36.0**

Table 1: Performance for detecting claims about COVID-19.

Claim Object Detection

- Claim object detection involves identifying what is being claimed in the claim sentence with respect to the topic.
- The answer span from the QA model for the question corresponding to the claim topic is used as the claim object.

Approach	Model	Type	F1
Prompting	GPT-3	Zero-shot	15.2
Prompting	T5	Zero-shot	11.4
In-context learning	GPT-3	Few-Shot	51.9
Prompt-based fine-tuning	T5	Few-Shot	51.6
QA	BERT	Zero-shot	**57.0**

Table 2: F1 score for claim object detection.

Claimer Detection

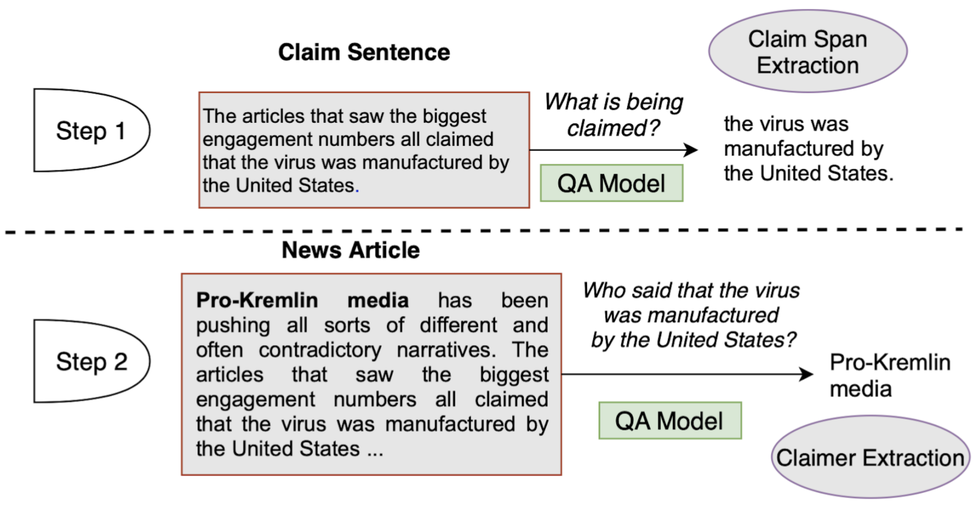

Figure 4: Two-step pipeline for identifying the claimer.

Model	Overall F1	Reported	Journalist
SRL	41.7	23.5	**67.2**
PolNeAR	42.3	25.5	65.9
QA	**50.1**	**39.8**	64.4

Table 3: F1 for identifying the claimer.

Model	In-sentence	Out-of-sentence
SRL	35.8	2.4
PolNeAR	38.9	2.7
QA	**46.2**	**29.0**

Table 4: F1 for claimer detection for when it is present within or outside the claim sentence.

Claims from journalists are direct assertions and do not have attribution. Thus, they can be identified from just sentence-level information.

Claim Sentence	Type
It is not yet known if remdesivir is safe for the treatment of COVID-19.	Journalist
Inhaling bleach fumes is dangerous and will not kill viruses that are already inside.	Journalist