

MultiVerS

Improving scientific claim verification with weak supervision and full-document context

CS221.O12.KHCL

Instructor: PhD. Nguyen Thi Quy

Group 15

Le Gia Khang Nguyen Hoang Tan Le Duy Khang

December 15, 2023

Claim:

Advil (ibuprofen) worsens
COVID-19 symptoms

Evidence abstract:**Covid-19 and avoiding
Ibuprofen.**

...

Increased risk of COVID-19
infection was feared with
ibuprofen use

...

At this time, there are no
findings discouraging the use
of ibuprofen

Claim:

Advil (ibuprofen) worsens
COVID-19 symptoms

Label: Refuted

Evidence abstract:**Covid-19 and avoiding
Ibuprofen.**

...

Increased risk of COVID-19
infection was feared with
ibuprofen use

...

At this time, there are no
findings discouraging the use
of ibuprofen

Claim:

Advil (ibuprofen) worsens
COVID-19 symptoms

Label: Refuted

Task Outputs

- 1 Fact-checking label

Evidence abstract:

Covid-19 and avoiding Ibuprofen.

...

Increased risk of COVID-19
infection was feared with
ibuprofen use

...

At this time, there are no
findings discouraging the use
of ibuprofen

Claim:

Advil (ibuprofen) worsens COVID-19 symptoms

Label: Refuted

Task Outputs

- ① Fact-checking label
- ② Rationales justifying the label

Evidence abstract:

Covid-19 and avoiding Ibuprofen.

...

Increased risk of COVID-19 infection was feared with ibuprofen use

...

At this time, there are no findings discouraging the use of ibuprofen

Rationale

Claim:

Advil (ibuprofen) worsens COVID-19 symptoms

Label: Refuted

Task Outputs

- ① Fact-checking label
- ② Rationales justifying the label

Evidence abstract:

Covid-19 and avoiding Ibuprofen.

...

Increased risk of COVID-19 infection was feared with ibuprofen use

Context required

...

At this time, there are no findings discouraging the use of ibuprofen

Rationale

Claim:

Advil (ibuprofen) worsens
COVID-19 symptoms

Evidence abstract:

Covid-19 and avoiding Ibuprofen.

...

Increased risk of COVID-19
infection was feared with
ibuprofen use

...

At this time, there are no
findings discouraging the use
of ibuprofen

Prior work: Extract-then-label

Claim:

Advil (ibuprofen) worsens
COVID-19 symptoms

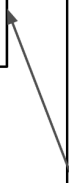
At this time, there are no
findings discouraging the use
of ibuprofen

Evidence abstract:

Covid-19 and avoiding Ibuprofen.

...
Increased risk of COVID-19
infection was feared with
ibuprofen use

...
At this time, there are no
findings discouraging the use
of ibuprofen



Prior work: Extract-then-label

Claim:

Advil (ibuprofen) worsens
COVID-19 symptoms

At this time, there are no
findings discouraging the use
of ibuprofen



Label: Refuted

Evidence abstract:

Covid-19 and avoiding Ibuprofen.

...
Increased risk of COVID-19
infection was feared with
ibuprofen use

...
At this time, there are no
findings discouraging the use
of ibuprofen

Prior work: Extract-then-label

Claim:

Advil (ibuprofen) worsens
COVID-19 symptoms

At this time, there are no
findings discouraging the use
of ibuprofen



Label: Refuted

Evidence abstract:

Covid-19 and avoiding Ibuprofen.

...

Increased risk of COVID-19
infection was feared with
ibuprofen use

...

At this time, there are no
findings discouraging the use
of ibuprofen

Drawbacks of extract-then-label:

- ① Rationales may lack context
- ② Requires rationale supervision during training

Claim

Advil (ibuprofen)
worsens COVID-19
symptoms

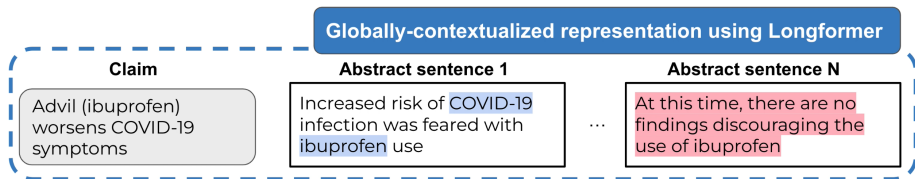
Abstract sentence 1

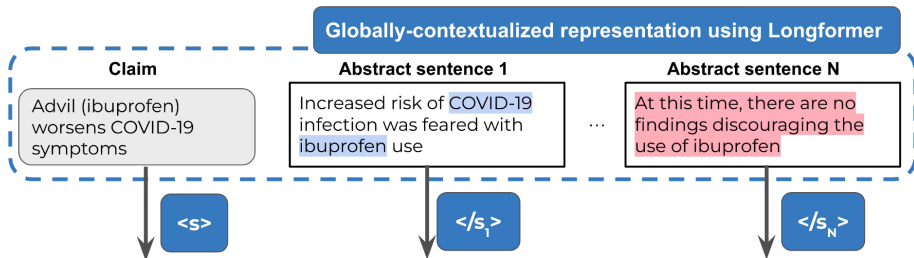
Increased risk of COVID-19
infection was feared with
ibuprofen use

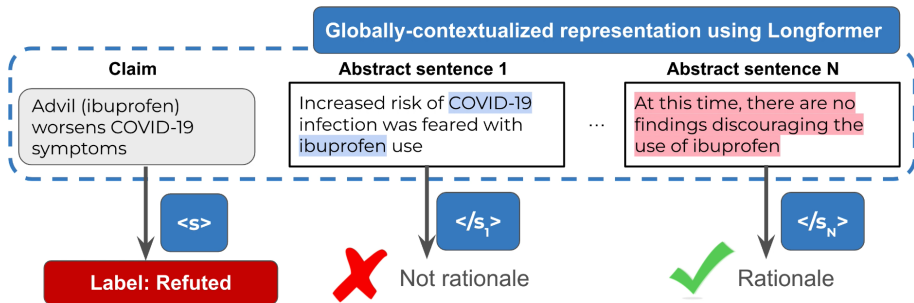
...

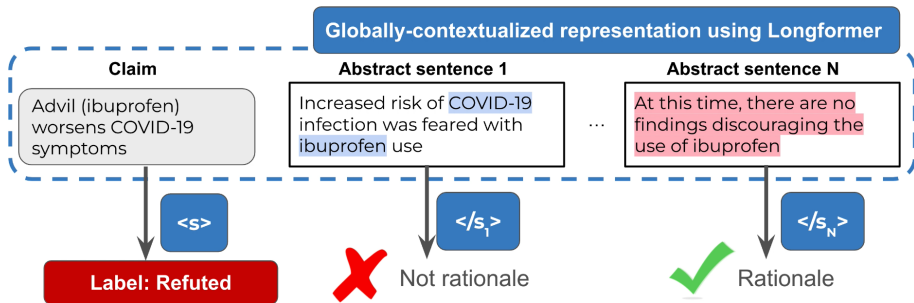
Abstract sentence N

At this time, there are no
findings discouraging the
use of ibuprofen









$$\mathcal{L} = \mathcal{L}_{label} + \lambda_{rationale} \mathcal{L}_{rationale}$$

Benefits of multitask approach:

- 1 Incorporates all relevant context
- 2 Can train on instances with no rationale annotations

Experiments

Text datasets:

- HealthVer
- COVID-Fact
- SciFact

Experiments

Text datasets:

- HealthVer
- COVID-Fact
- SciFact

Roughly 1000 claims / dataset.

Expert annotations are expensive

Experiments

Text datasets:

- HealthVer
- COVID-Fact
- SciFact

Roughly 1000 claims / dataset.

Expert annotations are expensive

Traning procedure:

- **Stage 1:** Train on a combination of *labeled out of domain data* and *weekly-labeled in-domain data*.
- **Stage 2:** Continue training on data from each target dataset.

Experiments

Text datasets:

- HealthVer
- COVID-Fact
- SciFact

Roughly 1000 claims / dataset.

Expert annotations are expensive

Traning procedure:

- **Stage 1:** Train on a combination of *labeled out of domain* data *weekly-labeled in-domain* data.
- **Stage 2:** Continue training on data from each target dataset.

Domain adaptation settings:

- **Zero-shot:** Stage 1 training only.
- **Few-shot:** 45 instances from target datasets.
- **Full-supervised:** All target data.

Supervised out-of-domain
data (FEVER)

LeBron James was born in
Ohio

Label: Supported

LeBron James is an American
basketball player. He was born
in Akron, Ohio.

Data: Stage 1

Supervised out-of-domain data (FEVER)

LeBron James was born in Ohio

Label: Supported

LeBron James is an American basketball player. He was born in Akron, Ohio.

Weekly-supervised in-domain data

Diabetes increases risk of depression

Abstract

...

...

...

Label: Supported

Data: Stage 1

Supervised out-of-domain data (FEVER)

LeBron James was born in Ohio

Label: Supported

LeBron James is an American basketball player. He was born in Akron, Ohio.

Weekly-supervised in-domain data

Diabetes increases risk of depression

} Claim: Paper title

Abstract

...
...
...

Label: Supported

Data: Stage 1

Supervised out-of-domain data (FEVER)

LeBron James was born in Ohio

Label: Supported

LeBron James is an American basketball player. He was born in Akron, Ohio.

Weekly-supervised in-domain data

Diabetes increases risk of depression

Abstract

...
...
...

Label: Supported

Claim: Paper title

Rationales likely to appear in abstract, but are not annotated

Data: Stage 1

Supervised out-of-domain data (FEVER)

LeBron James was born in Ohio

Label: Supported

LeBron James is an American basketball player. He was born in Akron, Ohio.

Weekly-supervised in-domain data

Diabetes increases risk of depression

Abstract

...
...
...

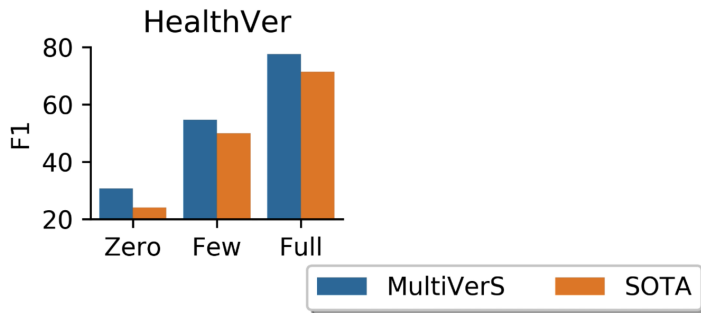
Label: Supported

Claim: Paper title

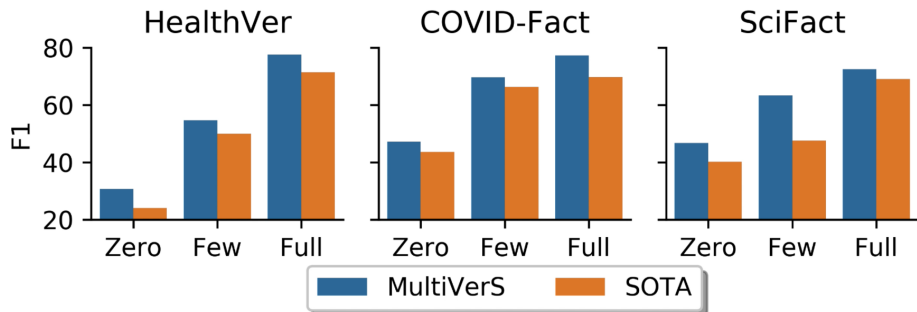
Rationales likely to appear in abstract, but are not annotated

MultiVerS can train on these examples, even though no rationale annotations are provided

Results

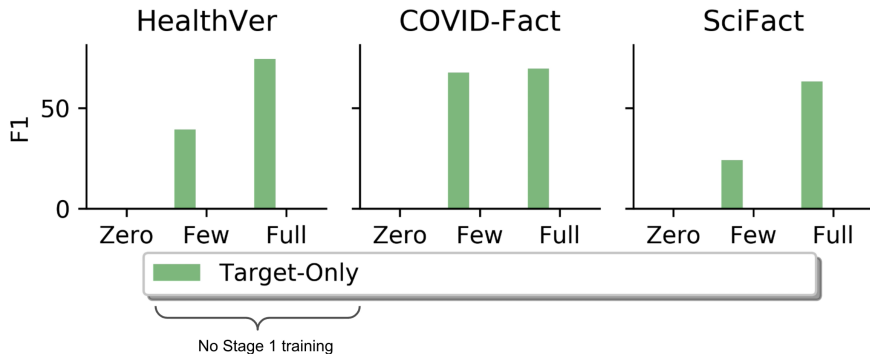


Results

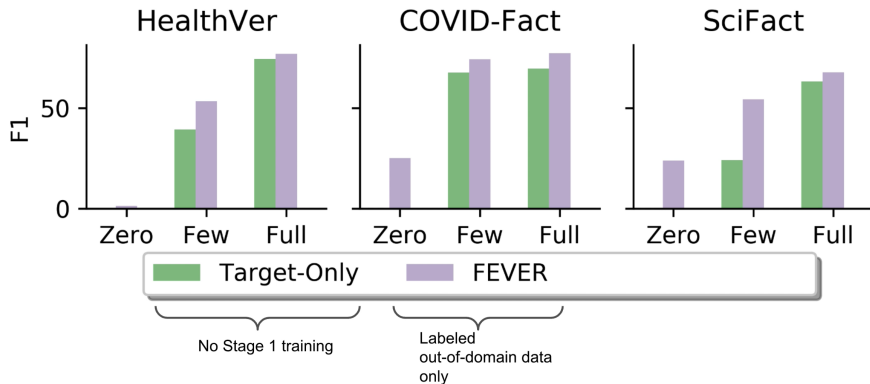


MultiVerS outperforms SOTA on all datasets

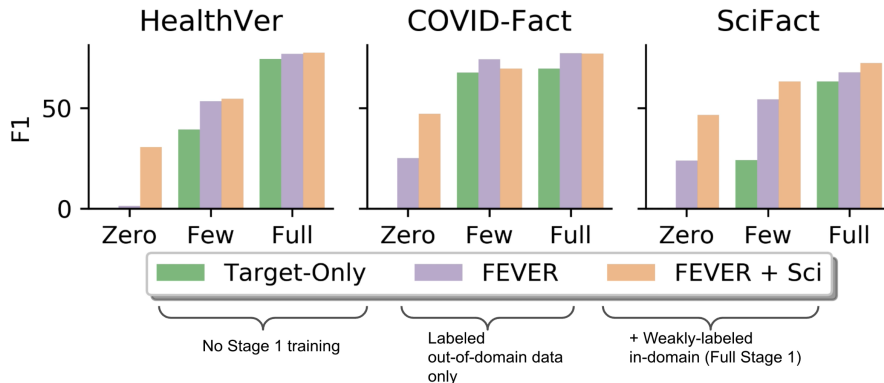
Results



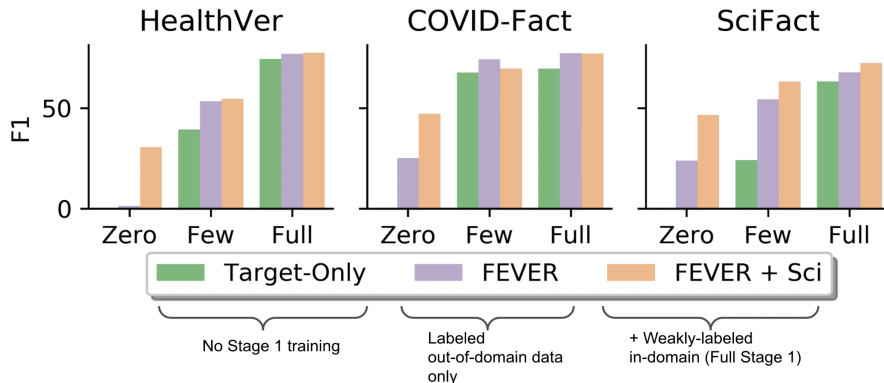
Results



Results



Results



Pretraining with weakly-supervised in-domain data improves few / zero shot performance.

Reference



MULTIVERS: Improving scientific claim verification with weak supervision and full-document context



Scientific Fact-Checking: A Survey of Resources and Approaches
Juraj Vladika and Florian Matthes



[Code and model checkpoints for the MultiVerS model](#)
dwadden/multivers

Thanks for listening!

Q&A section