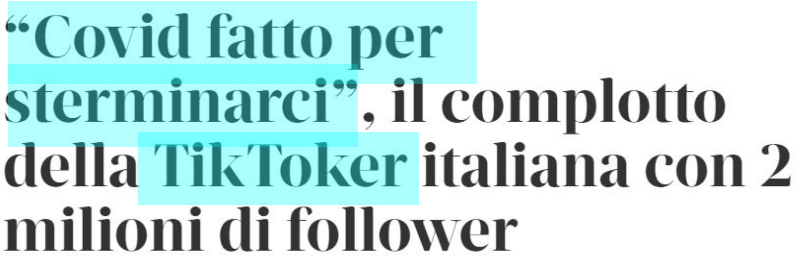
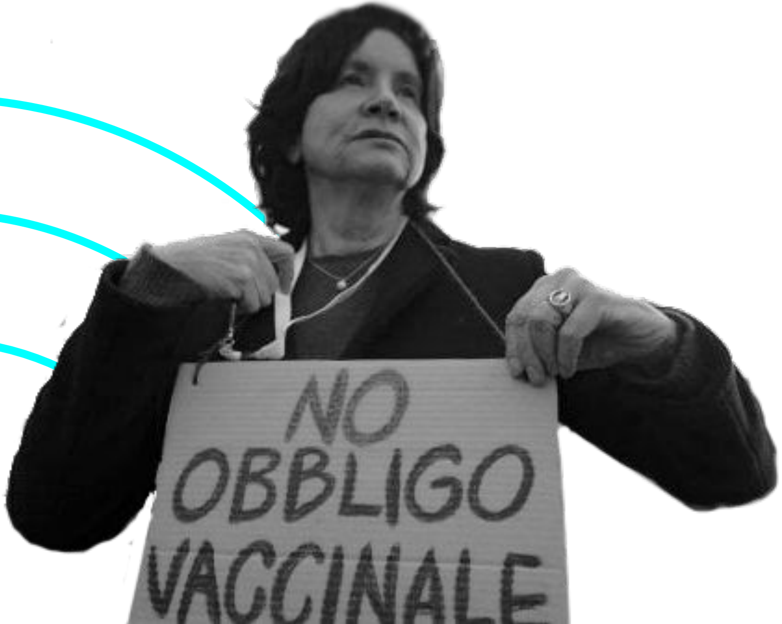
Automated Techniques for Identifying Fake News and Assisting Fact Checkers

Stefano Agresti 913079
Professor Mark J. Carman

Why is it important to fight fake news?


"Covid fatto per sterminarci", il complotto della TikToker italiana con 2 milioni di follower



Fighting Fake News

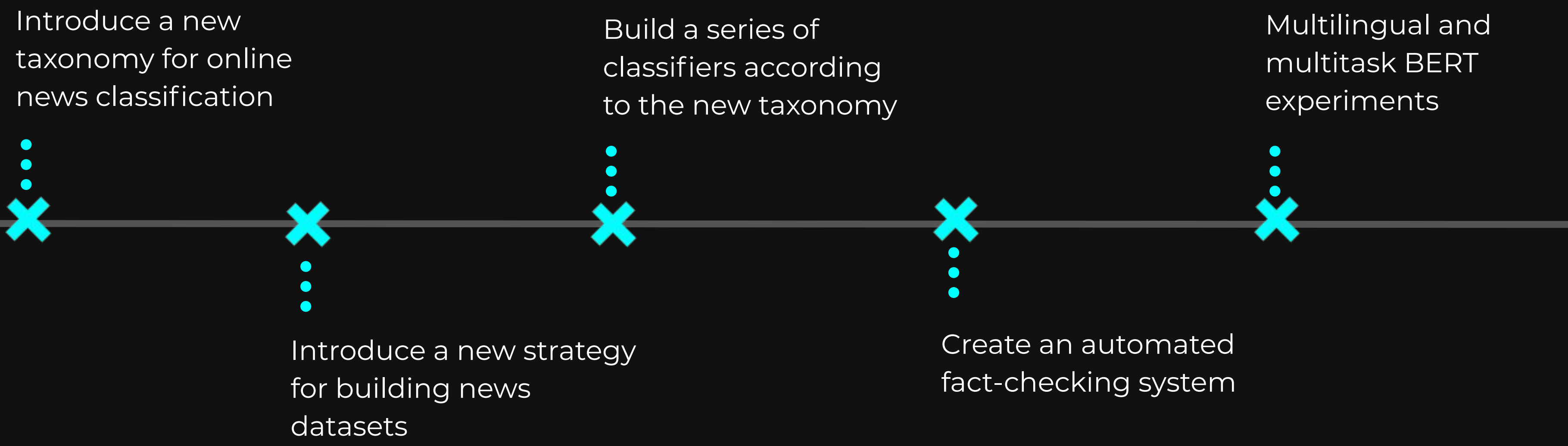
Knowledge

Propagation

Style

Source

Objectives of the thesis



Introduce a new taxonomy for online news classification

Introduce a new strategy for building news datasets

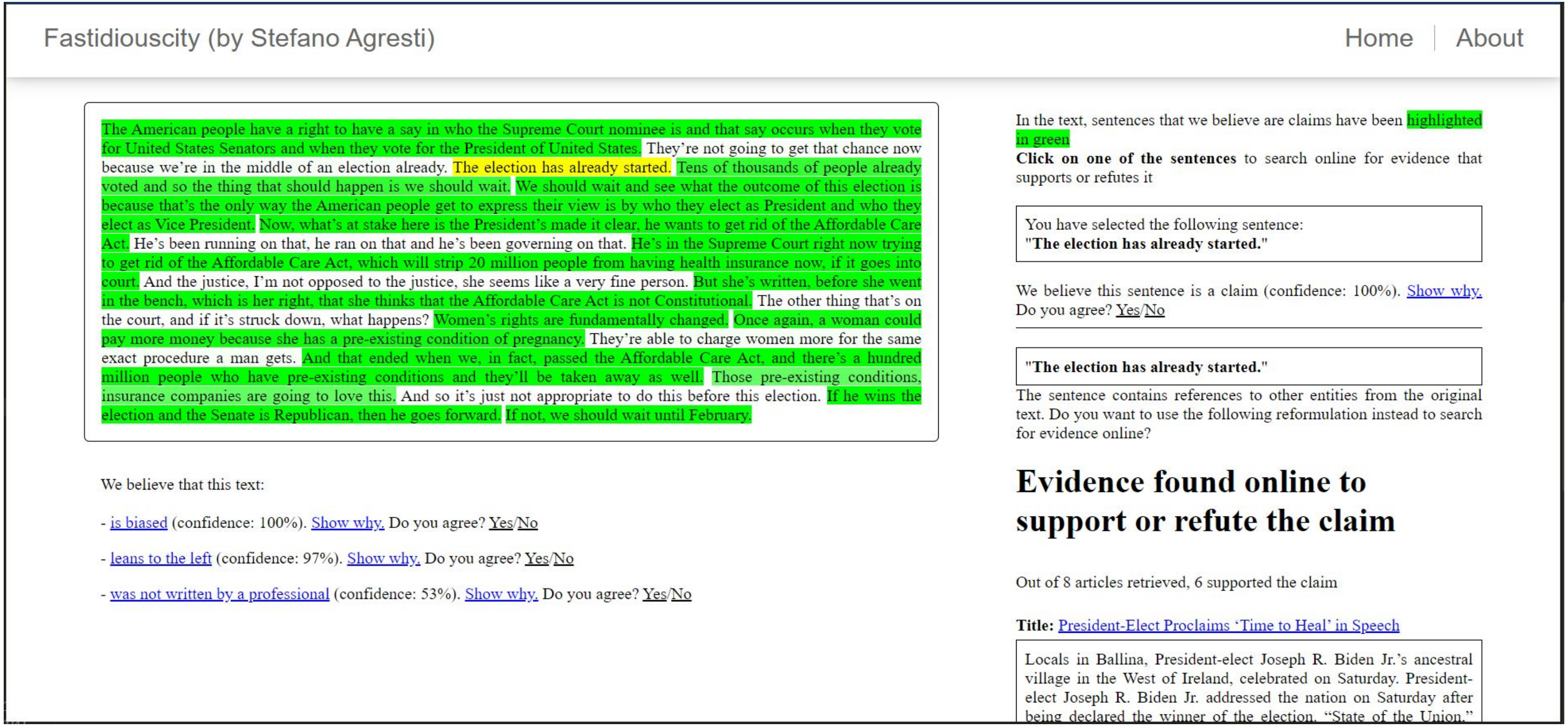
Build a series of classifiers according to the new taxonomy

Create an automated fact-checking system

Multilingual and multitask BERT experiments

Prototype available at:

www.fastidiouscity.com



Fastidiouscity (by Stefano Agresti)

Home | About

The American people have a right to have a say in who the Supreme Court nominee is and that say occurs when they vote for United States Senators and when they vote for the President of United States. They're not going to get that chance now because we're in the middle of an election already. The election has already started. Tens of thousands of people already voted and so the thing that should happen is we should wait. We should wait and see what the outcome of this election is because that's the only way the American people get to express their view is by who they elect as President and who they elect as Vice President. Now, what's at stake here is the President's made it clear, he wants to get rid of the Affordable Care Act. He's been running on that, he ran on that and he's been governing on that. He's in the Supreme Court right now trying to get rid of the Affordable Care Act, which will strip 20 million people from having health insurance now, if it goes into court. And the justice, I'm not opposed to the justice, she seems like a very fine person. But she's written, before she went in the bench, which is her right, that she thinks that the Affordable Care Act is not Constitutional. The other thing that's on the court, and if it's struck down, what happens? Women's rights are fundamentally changed. Once again, a woman could pay more money because she has a pre-existing condition of pregnancy. They're able to charge women more for the same exact procedure a man gets. And that ended when we, in fact, passed the Affordable Care Act, and there's a hundred million people who have pre-existing conditions and they'll be taken away as well. Those pre-existing conditions, insurance companies are going to love this. And so it's just not appropriate to do this before this election. If he wins the election and the Senate is Republican, then he goes forward. If not, we should wait until February.

We believe that this text:

- is biased (confidence: 100%). Show why. Do you agree? Yes/No

- leans to the left (confidence: 97%). Show why. Do you agree? Yes/No

- was not written by a professional (confidence: 53%). Show why. Do you agree? Yes/No

In the text, sentences that we believe are claims have been highlighted in green

Click on one of the sentences to search online for evidence that supports or refutes it

You have selected the following sentence:
"The election has already started."

We believe this sentence is a claim (confidence: 100%). Show why.
Do you agree? Yes/No

"The election has already started."

The sentence contains references to other entities from the original text. Do you want to use the following reformulation instead to search for evidence online?

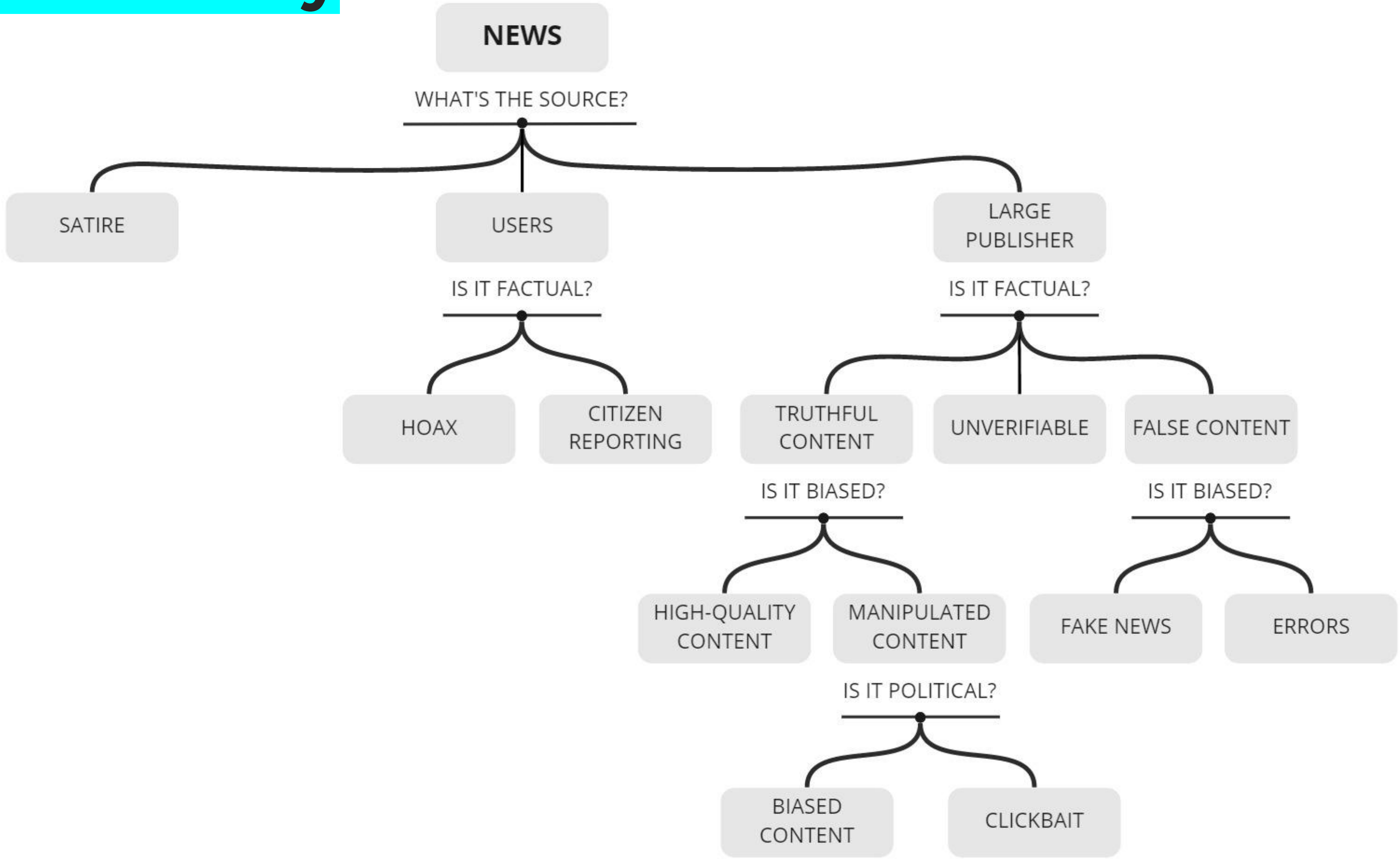
Evidence found online to support or refute the claim

Out of 8 articles retrieved, 6 supported the claim

Title: President-Elect Proclaims 'Time to Heal' in Speech

Locals in Ballina, President-elect Joseph R. Biden Jr.'s ancestral village in the West of Ireland, celebrated on Saturday. President-elect Joseph R. Biden Jr. addressed the nation on Saturday after being declared the winner of the election. "State of the Union."

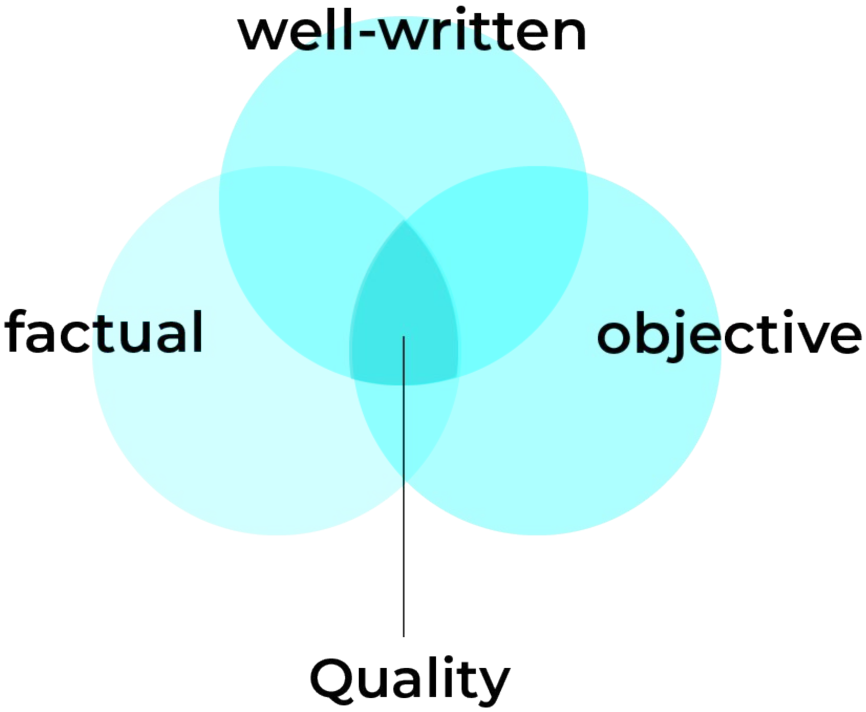
A new taxonomy



A new taxonomy

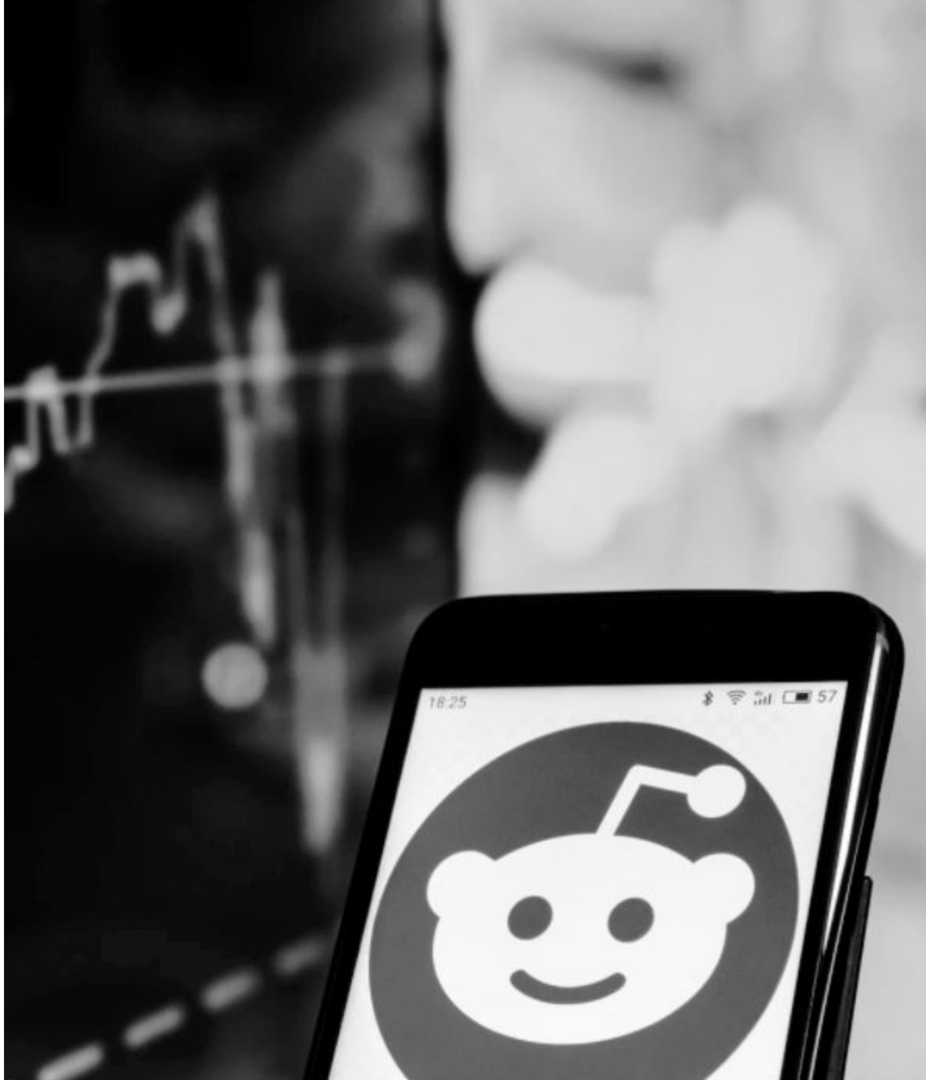
Five questions:

1. Is it newsworthy?
2. Is the source reliable?
3. Is it factual?
4. Is it objective?
5. What’s the writer’s political stance?



How to build good-quality datasets?

- Reddit's communities offer mono-thematic and variegated content
- Could it be a good source for news datasets?
- We launched a crowdsourcing experiment to find out
- In 91.2% of the cases, human and Reddit labelling were the same



Building the classifiers

“Is it newsworthy?”

• r/news

As source of news

•• r/inthenews

As source of opinions

••• Blogger corpus

For uninteresting content

Building the classifiers

“Is the source reliable?”

- **Goal**

Discriminating professional and unprofessional journalism

- **r/qualitynews**

As source of high-quality content

- **r/savedyouaclick**

As source of low-quality content

"Is it factual?"

More complex problem, required a system able to:

1

Detect claims inside a text

2

Search online for related evidence

3

Determine whether the found evidence supports the original claim

Building an automated fact-checking system

Detect claims inside a text

Model trained on Politifact claims and movie lines

Crowdsourcing experiment to evaluate its performances

Overall accuracy of 0.72

Building an automated fact-checking system

Online search and coreference resolution

System based on RoBERTa and SpaCy

“**He** said that”
becoming
“**Trump** said that”

Tested on GAP dataset, accuracy of 0.75

Building an automated fact-checking system

Agreement detection

Dataset of fact-checking articles with their claims

Multilingual BERT

The chosen model had a 0.68 overall accuracy

Building the classifiers

Is it biased?

Unbiased samples

Articles from several press agencies

Biased samples

Articles from conservative and liberal subreddits

Almost 1.00 overall accuracy

Building the classifiers

What’s the writer’s political stance?

Data source: conservative and liberal subreddits

×

Overall accuracy of 0.90

Multilingual experiment with BERT

Classification task:

- Reddit posts in 5 languages, 4 categories
- Multilingual model trained and tested on the entire dataset
- Monolingual models trained and tested on monolingual portions

Results:

- Monolingual models slightly outperformed the multilingual one
- The multilingual model obtained similar results on all languages

Useful information when building a system

Experiment on BERT's multi-task training

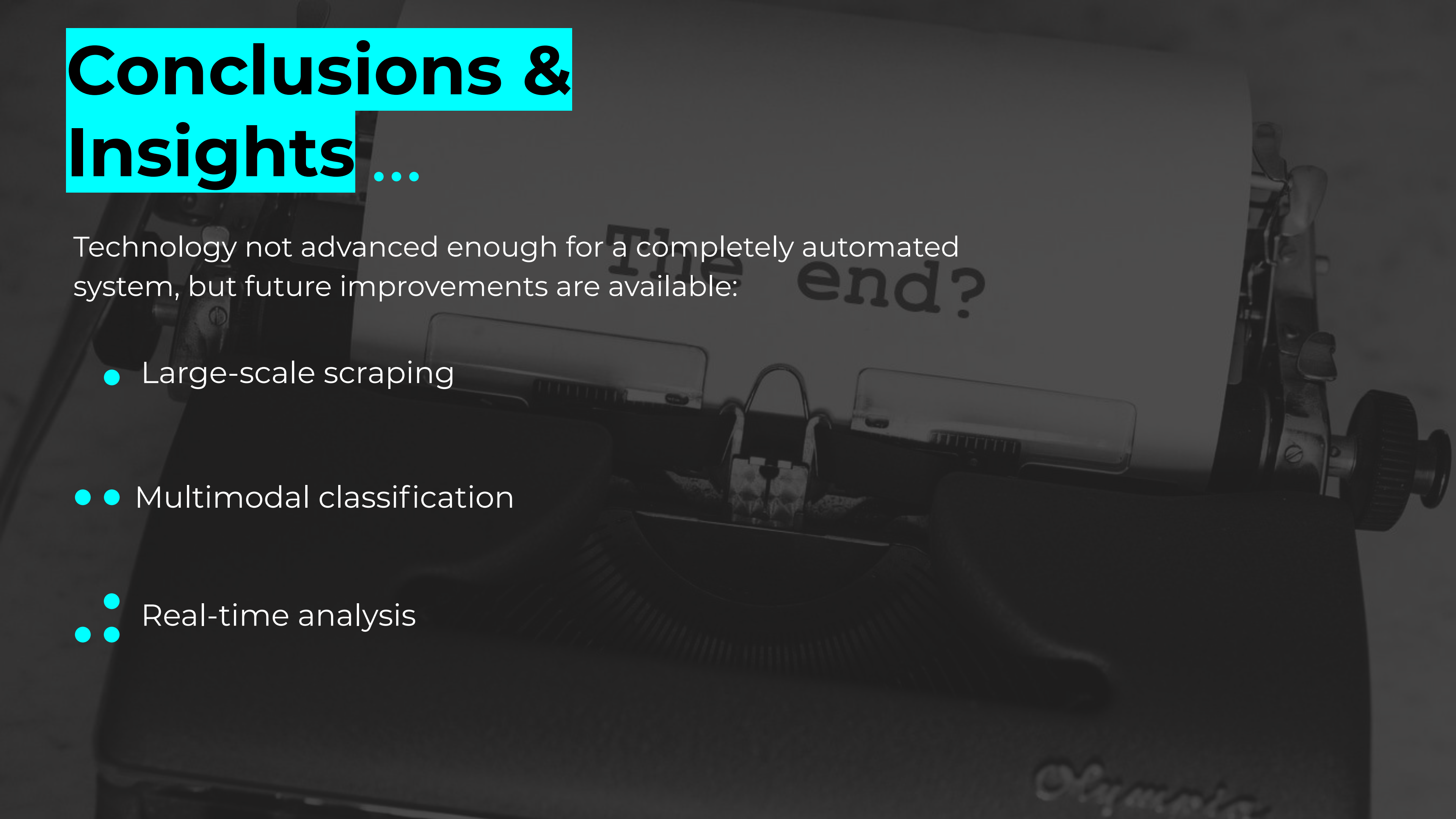
Three different tasks:

1. Related/Unrelated
2. Agreement/Disagreement
3. Entailment/Contradiction

Three different settings:

1. Baseline
2. Hard-coded correlation
3. Parallel training

Multi-task strongly improved training performances



Conclusions & Insights …

Technology not advanced enough for a completely automated system, but future improvements are available:

- Large-scale scraping
- Multimodal classification
- Real-time analysis

Main contributions

- ● New taxonomy
- ●● Creating news datasets
- ●●● Methodology for helping fact-checkers
- ●●●● Building a working prototype

References

Multi-Task Learning for Multi-Lingual Claim Checking, Luca Favano and Mark J. Carman, 2019

r/Fakeddit: A New Multimodal Benchmark Dataset for Fine-grained Fake News Detection, Nakamura et al, 2019

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, Devlin et al, 2019

A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities, Xinyi Zhou and Reza Zafarani, 2018

Datasets introduced

Newsworthiness classifier	r/news, r/inthenews
Professionality classifier	r/qualitynews, r/savedyouaclick, selected publishers
Claim detection	Politifact claims, dataset from crowdsourcing
Agreement detection	Dataset of fact-checking articles
Political bias and ideology classifier	Articles from conservative and liberal subreddits
Multilingual experiment	Reddit posts grouped by four categories over five different languages

Thank you!

Questions?