NLP Approaches to Fact Checking and Fake News Detection

Andreas Hanselowski, Iryna Gurevych
Outline:

1. Fake News Detection

2. Automated Fact Checking
Outline:

1. Fake News Detection

2. Automated Fact Checking
Fake News Detection: Fake News Challenge

General idea:

“The Fake News Challenge is a competition in the spirit of the Netflix Prize designed to foster the development of AI technology to help solve the fake news problem.”

Dean Pomerleau

- A number of competitions to explore AI approaches for different problems in fake news detection
Fake News Detection: Fake News Challenge

General idea:

“The Fake News Challenge is a competition in the spirit of the Netflix Prize designed to foster the development of AI technology to help solve the fake news problem.”

Dean Pomerleau

- A number of competitions to explore AI approaches for different problems in fake news detection

- Fake News Challenge Stage One: Stance classification
  → Given: An article headline and an article body
  → To be determined: Stance of the article with respect to the headline
  Classes: agrees, disagrees, discusses, unrelated
Fake News Detection: Fake News Challenge

General idea:
“The Fake News Challenge is a competition in the spirit of the Netflix Prize designed to foster the development of AI technology to help solve the fake news problem.”
Dean Pomerleau

- A number of competitions to explore AI approaches for different problems in fake news detection

- Fake News Challenge Stage One: Stance classification
  → Given: An article headline and an article body
  → To be determined: Stance of the article with respect to the headline
    Classes: agrees, disagrees, discusses, unrelated
## Fake News Detection: Fake News Challenge

### Results: Top 10 out of 50

<table>
<thead>
<tr>
<th>Team</th>
<th>% FNC-1 score</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOLAT in the SWEN</td>
<td>82.02</td>
</tr>
<tr>
<td>Athene UKP</td>
<td>81.97</td>
</tr>
<tr>
<td>UCL Machine Reading</td>
<td>81.72</td>
</tr>
<tr>
<td>Chips Ahoy!</td>
<td>80.21</td>
</tr>
<tr>
<td>CLUlings</td>
<td>79.73</td>
</tr>
<tr>
<td>Unconscious bias</td>
<td>79.69</td>
</tr>
<tr>
<td>OSU</td>
<td>79.65</td>
</tr>
<tr>
<td>MITBusters</td>
<td>79.58</td>
</tr>
<tr>
<td>DFKI LT</td>
<td>79.56</td>
</tr>
<tr>
<td>GTRI - ICL</td>
<td>79.33</td>
</tr>
</tbody>
</table>
Our takeaway:

- Data set relatively small only ≈1500 individual articles → Poor generalization to new corpora

- Approaches which work well:
  Topic models, bag of words features, n-grams, tf-idf

- Approaches which do not work:
  Features based on word embeddings, RNNs (LSTMs)

Problem setting is relatively simple, however for better performance larger corpora required
Outline

1. Fake News Detection

2. Automated Fact Checking
Automated Fact Checking
(Claim Validation)

Claim: Israel caused flooding in Gaza by opening river dams

Evidence documents:
The Daily Mail:
…Hundreds of Palestinians left homeless after Israel opens river dams and floods houses …

Al Jazeera:
"Israel opened water dams, without warning, last night, causing serious damage to Gazan villages near the border, " General Al-Saudi told Al Jazeera. …

The Jerusalem post:
…The Daily Mail published a story on Monday that originally accused Israel of intentionally opening dams …. in order to flood Gaza.
The only problem is, there are no dams in southern Israel. … after amending the article's headline the pretense for the charge remained in the Daily Mail article…

→ Claim is false
Automated Fact Checking (Claim Validation): NLP-Pipeline

NLP Pipeline

Document retrieval

Large corpus, Web

Extraction of evidence

Evidence 1
Evidence 2

Claim validation

Evidence 1
Evidence 2

Claim

Verdict
Israel caused flooding in Gaza by opening river dams

**Claim and Verdict**

Israel caused flooding in Gaza

In the wake of a recent severe winter storm in the region, Israeli authorities opened the floodgates to discharge the accumulated water flooding villages in Gaza. No casualties were reported as a result, but more than 80 families had to flee after their homes filled with water levels sometimes reaching more than three meters, the Gaza Ministry of Interior said in a statement. The flood caused by Israel forced the closure of the main road connecting al-Mughraqa district and Nusseirat refugee camp south of Gaza city.
Israel reportedly flooded Gaza. No casualties were reported as a result, but more than 80 families had to flee after their homes filled with water levels sometimes reaching more than three meters, the Gaza Ministry of Interior said in a statement. The flood caused by Israel forced the closure of the main road connecting al-Mughraqa district and Nusseirat refugee camp south of Gaza city.

**Evidence text snippets**

**Corpus statistics:**
Number of claims: ≈ 3800
Number of documents: ≈ 14000

**Corpus will be released**
First Results (work-in-progress)

Evidence extraction:
- Random baseline: F1 = 10%
- Deep Learning method (LSTMs): F1 = 20%
  → Further Deep Learning structures need to be explored
  → Upper bound low due to annotation by one single fact checker

Israel caused flooding in Gaza

After the last Israeli–Palestinian conflict, tensions between Israelis and the Palestinians have been rising. In the wake of a recent severe winter storm in the region, Israeli authorities opened the floodgates to discharge the accumulated water flooding villages in Gaza. No casualties were reported as a result, but more than 80 families had to flee after their homes filled with water levels sometimes reaching more than three meters, the Gaza Ministry of Interior said in a statement. The flood caused by Israel forced the closure of the main road connecting al-Mughraqa district and Nusseirat refugee camp south of Gaza city.
First Results (work-in-progress)

Evidence extraction:
- Random baseline: F1 = 10%
- Deep Learning method (LSTMs): F1 = 20%
→ Further Deep Learning structures need to be explored
→ Upper bound low due to annotation by one single fact checker

Claim validation:
- Majority baseline: F1 = 45%
- Deep Learning method (LSTMs): F1 = 63%
→ Further DL structures to be explored
→ Reasoning, world knowledge needs to be incorporated

Israel caused flooding in Gaza by opening river dams

After the last Israeli–Palestinian conflict, tensions between Israelis and the Palestinians have been rising. In the wake of a recent severe winter storm in the region, Israeli authorities opened the floodgates to discharge the accumulated water flooding villages in Gaza. No casualties were reported as a result, but more than 80 families had to flee after their homes filled with water levels sometimes reaching more than three meters, the Gaza Ministry of Interior said in a statement.

The flood caused by Israel forced the closure of the main road connecting Al-Mughraqa district and Nusseirat refugee camp south of Gaza city.
Conclusions:

- Stance classification for fake news detection works relatively well, fact checking is more challenging

Problems:
- Not enough data
  - Snopes corpus is a starting point
- For comprehensive claim validation, reasoning, world knowledge is required

Our approach:
- Break down the problem into several sub-problems and tackle them individually
  - In addition to state-of-the-art deep learning approaches, incorporation of world knowledge and reasoning techniques
Automated Fact Checking
(Claim Validation)

Claim: Israel caused flooding in Gaza by opening river dams

Evidence documents:
The Daily Mail:
…Hundreds of Palestinians left homeless after Israel opens river dams and floods houses …

Al Jazeera:
"Israel opened water dams, without warning, last night, causing serious damage to Gazan villages near the border, " General Al-Saudi told Al Jazeera. …
Automated Fact Checking (Claim Validation)

Claim: Israel caused flooding in Gaza by opening river dams

Evidence documents:

The Daily Mail:
…Hundreds of Palestinians left homeless after Israel opens river dams and floods houses …

Al Jazeera:
"Israel opened water dams, without warning, last night, causing serious damage to Gazan villages near the border, " General Al-Saudi told Al Jazeera. …

The Jerusalem post:
…The Daily Mail published a story on Monday that originally accused Israel of intentionally opening dams …. in order to flood Gaza.
The only problem is, there are no dams in southern Israel. … after amending the article's headline the pretense for the charge remained in the Daily Mail article…
Challenging Cases:

Entity matching:
- Claim: A restaurant in South Africa was shut down for serving human meat.
- Evidence: A Nigerian restaurant has been shuttered by authorities for serving human flesh.

Inference:
- Claim: Fritz Jahr introduced the concept of bioethics.
- Evidence:
  - E1: The concept of bioethics was introduced in the article “Wissenschaft von Sittenlehre”.
  - E2: Fritz Jahr has authored “Wissenschaft von Sittenlehre”.

Fake-News Detection: Fake News Challenge

Model structure

Model: Multi Layer Perceptron

Features:

- Basic features
  - Bag-of-words
  - N-grams
  - …

- Topic models
  - Latent semantic indexing (LSA)
  - Latent Dirichlet allocation (LDA)
  - …

- Lexica
  - Refuting words (fake, hoax, …)
  - …

24.07.2017 | Ubiquitous Knowledge Processing (UKP) Lab | Andreas Hanselowski, Prof. Dr. Iryna Gurevych |
PDIG2 was ist bag-of-words unigrams? wie unterscheiden sich diese von "normalen" unigrams?
Prof. Dr. Iryna Gurevych; 21.07.2017
Automated Fact Checking (Claim Validation): Extraction of Evidence using LSTMs

Israel caused flooding in Gaza

After the last Israeli–Palestinian conflict, tensions between Israelis and the Palestinians have been rising. In the wake of a recent severe winter storm in the region, Israeli authorities opened the floodgates to discharge the accumulated water flooding villages in Gaza. No casualties were reported as a result, but more than 50 families had to flee after their homes filled with water levels sometimes reaching more than three meters, the Gaza Ministry of Interior said in a statement. The flood caused by Israel forced the closure of the main road connecting al-Mughira district and Nusseirat refugee camp south of Gaza city.

Al-Saudi reports in an interview. Israeli authorities opened river damps. Israel caused flooding in Gaza

Word embeddings:

Sentence representations:

Claim representation

Not an Evidence

Evidence

LSTM-1

LSTM-2
Automated Fact Checking (Claim Validation): Claim Validation using LSTMs

Verdict: False/True

Multi Layer Perceptron

Claim

LSTM

Evidence 1
Source 1
Evidence n
Source n

Israeli authorities opened river dams to flood Gaza.
www.dailymail.co.uk

There are no dams in southern Israel.
www.jpost.com

Source 1
Evidence 1

Israeli authorities opened river dams to flood Gaza.
www.dailymail.co.uk

There are no dams in southern Israel.
www.jpost.com