Politicians make false and misleading claims

- Russian security threats
- Pizzagate: conspiracy theory led to shooting

Fact-checks on major party presidential nominees by PolitiFact

Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims by ClaimBuster

Naeemul Hassan1 Fatma Arslan2 Chengkai Li3 Mark Tremayne3
1Department of Computer and Information Science, University of Mississippi
2Department of Computer Science and Engineering, University of Texas at Arlington
3Department of Communication, University of Texas at Arlington

The Quest to Automate Fact-checking

- 100+ active fact-checking sites in 2017 (PolitiFact.com, FullFact.org, CNN, Washington Post, …)
- Google and Bing include fact-checks in search results.
- Facebook lets users report false items and flags items disputed by fact-checkers.

Lack of automated tools that assist fact-checkers
- "The Holy Grail": fully automated fact-checking

Claim Spotting: Check-worthy Factual Claims Detection

Classification and ranking by check-worthiness
- Non-Factual Sentence (NFS) (Opinions, beliefs, declarations): "But I think it's time to talk about the future."
- Unimportant Factual Sentence (UFS): "Two days ago we ate lunch at a restaurant."
- Check-worthy Factual Sentence (CFS): "He voted against the first Gulf War."

Data Labeling and Ground-Truth Collection

Coding website
bit.ly/claimbusters

Coder quality

Feature extraction

Important Factual Claims

- 20788 sentences
- 374 coders, ~$4,000 paid
- 30 training sentences
- 1032 screening sentences (731 NFS, 238 CFS) to detect spammers & low-quality coders
- 63 UFS, 238 CFS) to 1032 screening sentences
- 4,000 paid
detailed labeling

Funding by

20788 sentences
374 coders
765,522 labels
86 top-quality coders
523,333 labels
Majority voting
20,017 admitted sentences

Classification and Ranking Accuracy

| algorithm | features | p NPS | p LPS | p CFS | p wNFS | p wLPS | p CSC | p wCSC | t NFS | t LPS | t CFS | t wNFS | t wLPS | t CSC | t wCSC | l NFS | l LPS | l CFS | l wNFS | l wLPS | l CSC | l wCSC |
|-----------|----------|-------|-------|-------|--------|--------|-------|--------|-------|-------|-------|--------|--------|-------|--------|-------|-------|-------|--------|--------|-------|--------|-------|--------|--------|-------|
| RFC       | W        | 0.755 |       |       | 0.125 | 0.238 | 0.492 | 0.965 | 0.004 | 0.225 | 0.745 | 0.848 | 0.003 | 0.343 | 0.685 |
| NBC       | W        | 0.788 | 0.076 | 0.616 | 0.674 | 0.983 | 0.035 | 0.791 | 0.975 | 0.522 | 0.744 |
| SVM       | W        | 0.772 | 0.671 | 0.626 | 0.723 | 0.814 | 0.925 | 0.227 | 0.667 | 0.826 | 0.997 | 0.294 | 0.968 | 0.816 |
| RFC       | W        | 0.772 | 0.671 | 0.626 | 0.723 | 0.814 | 0.925 | 0.227 | 0.667 | 0.826 | 0.997 | 0.294 | 0.968 | 0.816 |
| NBC       | W        | 0.799 | 0.076 | 0.616 | 0.674 | 0.983 | 0.035 | 0.791 | 0.975 | 0.522 | 0.744 |
| SVM       | W        | 0.873 | 0.437 | 0.724 | 0.813 | 0.925 | 0.224 | 0.672 | 0.827 | 0.908 | 0.307 | 0.996 | 0.827 |
| RFC       | W        | 0.772 | 0.671 | 0.626 | 0.723 | 0.814 | 0.925 | 0.227 | 0.667 | 0.826 | 0.997 | 0.294 | 0.968 | 0.816 |
| NBC       | W        | 0.808 | 0.076 | 0.616 | 0.674 | 0.983 | 0.035 | 0.791 | 0.975 | 0.522 | 0.744 |
| SVM       | W        | 0.873 | 0.437 | 0.724 | 0.813 | 0.925 | 0.224 | 0.672 | 0.827 | 0.908 | 0.307 | 0.996 | 0.827 |

Classification and Ranking Accuracy

| algorithm | features | p NPS | p LPS | p CFS | p wNFS | p wLPS | p CSC | p wCSC | t NFS | t LPS | t CFS | t wNFS | t wLPS | t CSC | t wCSC | l NFS | l LPS | l CFS | l wNFS | l wLPS | l CSC | l wCSC |
|-----------|----------|-------|-------|-------|--------|--------|-------|--------|-------|-------|-------|--------|--------|-------|--------|-------|-------|-------|--------|--------|-------|--------|-------|--------|--------|-------|
| RFC       | W        | 0.755 |       |       | 0.125 | 0.238 | 0.492 | 0.965 | 0.004 | 0.225 | 0.745 | 0.848 | 0.003 | 0.343 | 0.685 |
| NBC       | W        | 0.788 | 0.076 | 0.616 | 0.674 | 0.983 | 0.035 | 0.791 | 0.975 | 0.522 | 0.744 |
| SVM       | W        | 0.772 | 0.671 | 0.626 | 0.723 | 0.814 | 0.925 | 0.227 | 0.667 | 0.826 | 0.997 | 0.294 | 0.968 | 0.816 |
| RFC       | W        | 0.772 | 0.671 | 0.626 | 0.723 | 0.814 | 0.925 | 0.227 | 0.667 | 0.826 | 0.997 | 0.294 | 0.968 | 0.816 |
| NBC       | W        | 0.799 | 0.076 | 0.616 | 0.674 | 0.983 | 0.035 | 0.791 | 0.975 | 0.522 | 0.744 |
| SVM       | W        | 0.873 | 0.437 | 0.724 | 0.813 | 0.925 | 0.224 | 0.672 | 0.827 | 0.908 | 0.307 | 0.996 | 0.827 |
| RFC       | W        | 0.772 | 0.671 | 0.626 | 0.723 | 0.814 | 0.925 | 0.227 | 0.667 | 0.826 | 0.997 | 0.294 | 0.968 | 0.816 |
| NBC       | W        | 0.808 | 0.076 | 0.616 | 0.674 | 0.983 | 0.035 | 0.791 | 0.975 | 0.522 | 0.744 |
| SVM       | W        | 0.873 | 0.437 | 0.724 | 0.813 | 0.925 | 0.224 | 0.672 | 0.827 | 0.908 | 0.307 | 0.996 | 0.827 |

End-to-End Fact-Checking

Case Study: 2016 U.S. Presidential Election Debates

Comparison of topic distributions of CNN, PolitiFact fact-checked sentences and sentences scored high (~$0.5) by ClaimBuster

System
idir.uta.edu/claimbuster

Feature importance

Feature extraction and selection

Sentiment: 0.032
Words: state, legislature, 87, percent, democrat

If I was in a state where my legislature was 87 percent Democrat...

Part-of-Speech: Noun
Entity Type: Quantity
Concept: United States