

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



Naeemul Hassan<sup>1</sup> Fatma Arslan<sup>2</sup> Chengkai Li<sup>2</sup> Mark Tremayne<sup>3</sup>

<sup>1</sup>Department of Computer and Information Science, University of Mississippi

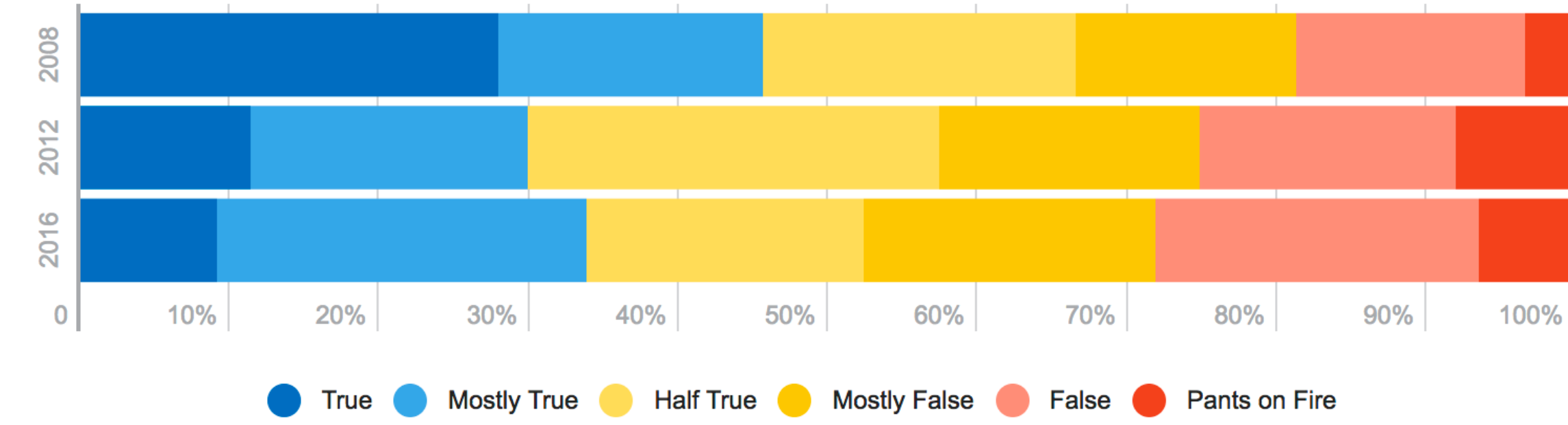
<sup>2</sup>Department of Computer Science and Engineering, University of Texas at Arlington

<sup>3</sup>Department of Communication, University of Texas at Arlington



## The Quest to Automate Fact-checking

### Politicians make false and misleading claims



### National security threats

- Russian government interfered with the 2016 election. Fake-news websites and bots used.
- Pizzagate: conspiracy theory led to shooting

### Fake-news floods social media (“filter bubbles” and “echo chambers”)

- Facebook trending topic algorithms promoted fake-news.
- A sample of 140,000 Twitter users in the battleground state of Michigan shared as many junk news items as professional news during the final ten days of the 2016 election. <http://politicalbots.org/?p=1064>

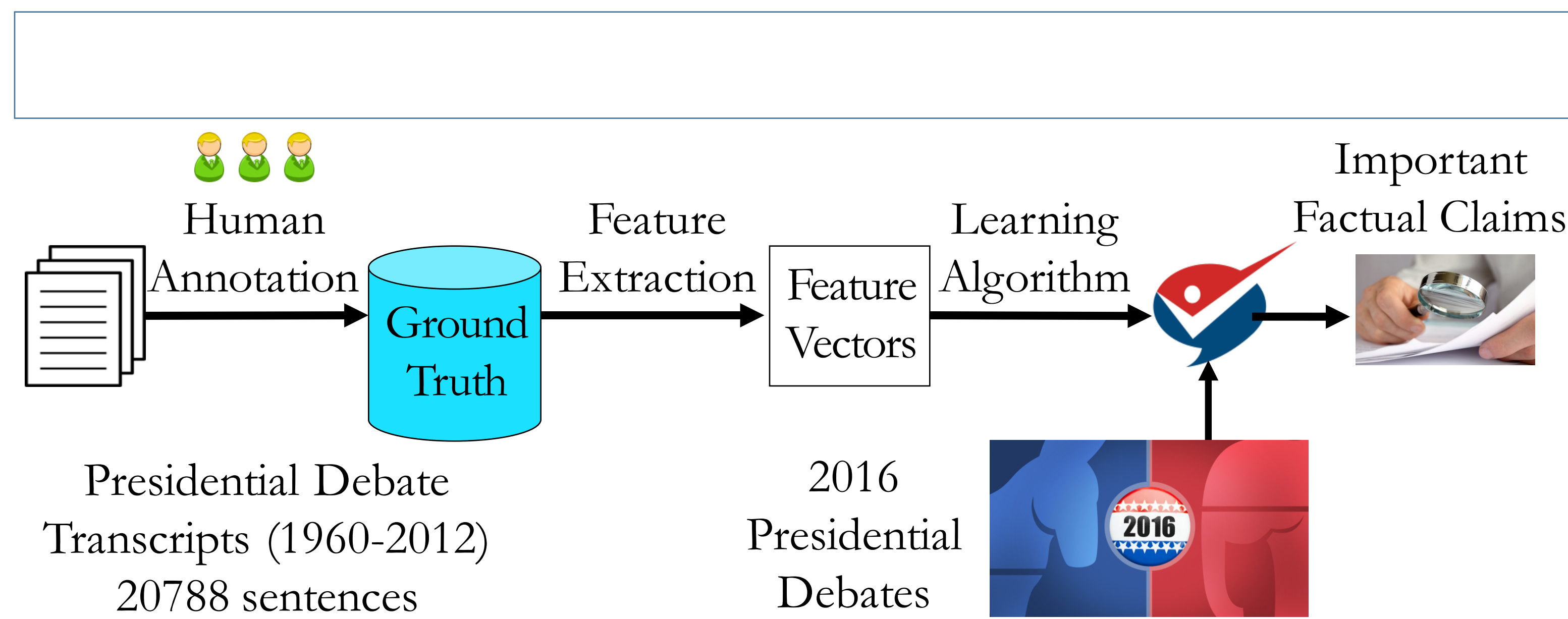
### Combating falsehoods

- 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

Fact-checks on major party presidential nominees by PolitiFact

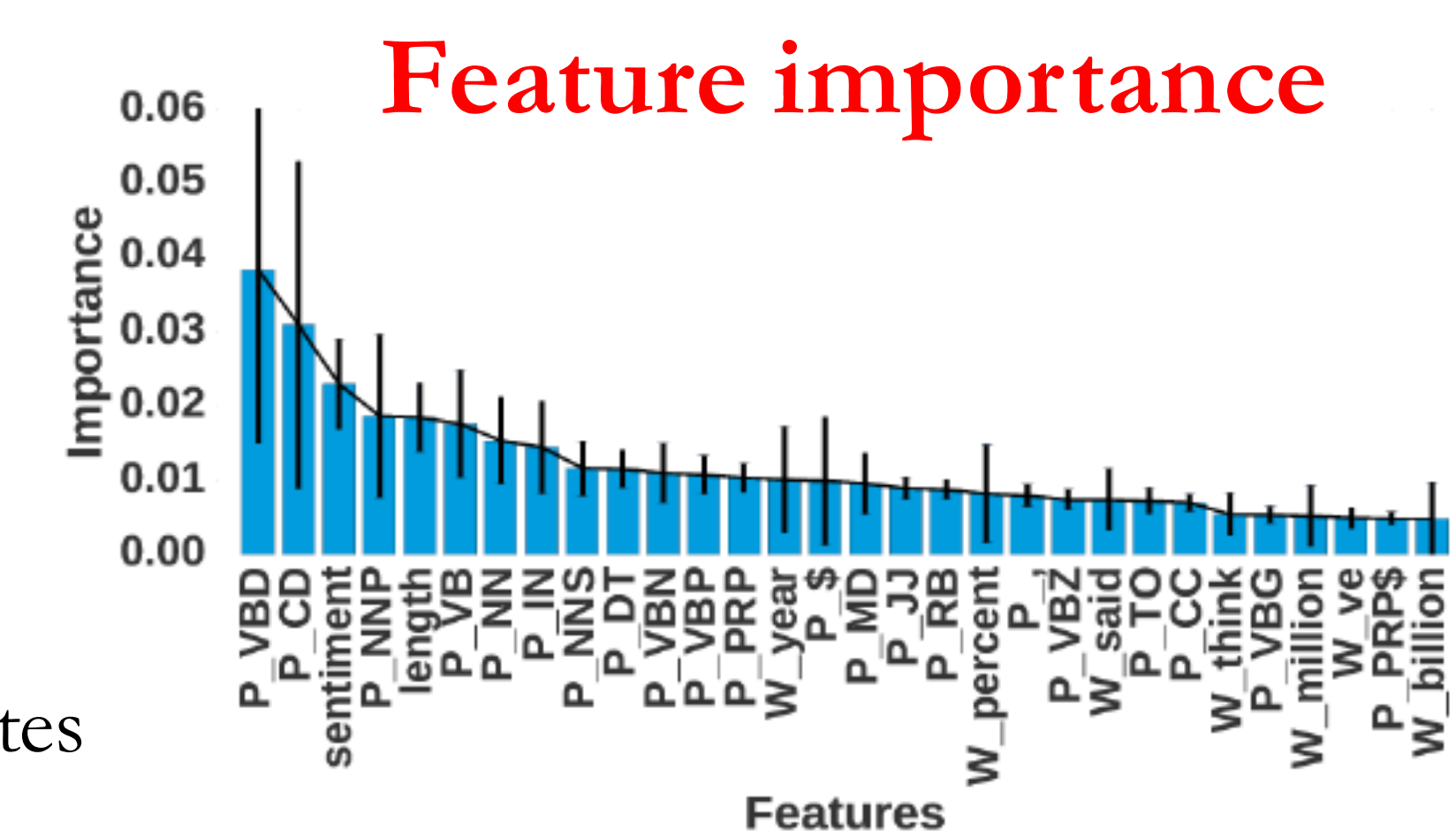
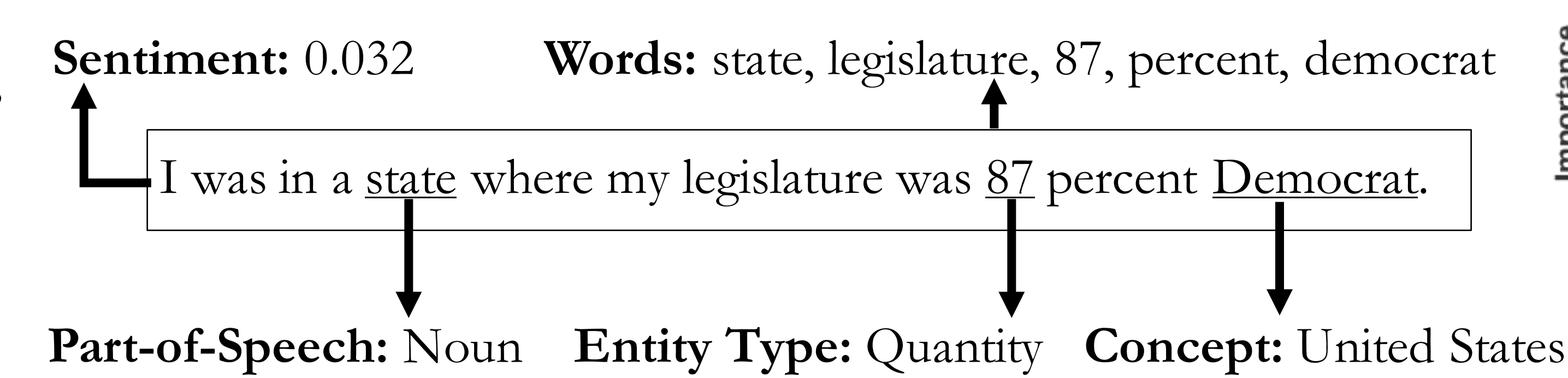


## Claim Spotting: Check-worthy Factual Claims Detection

### 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.”

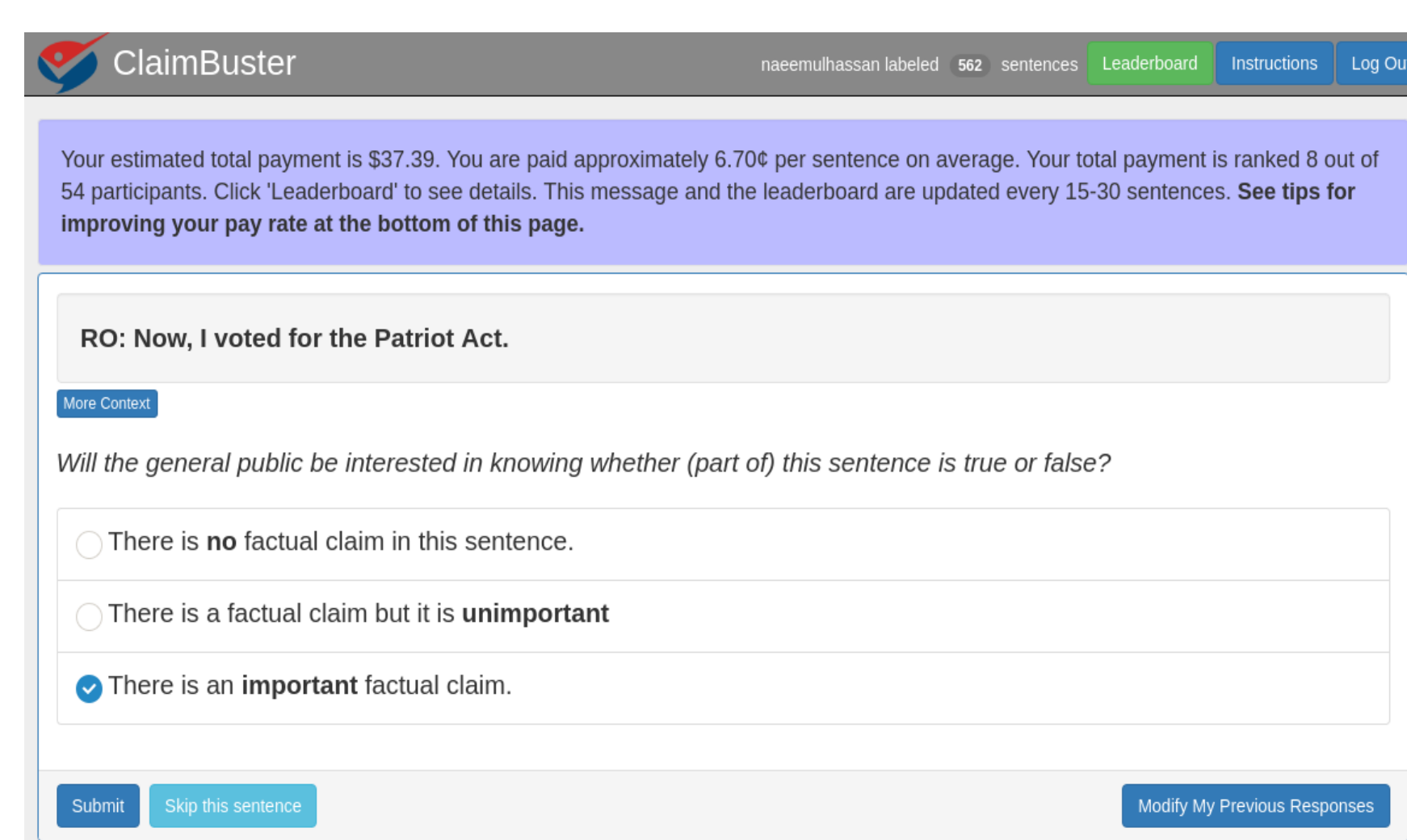
### Feature extraction and selection



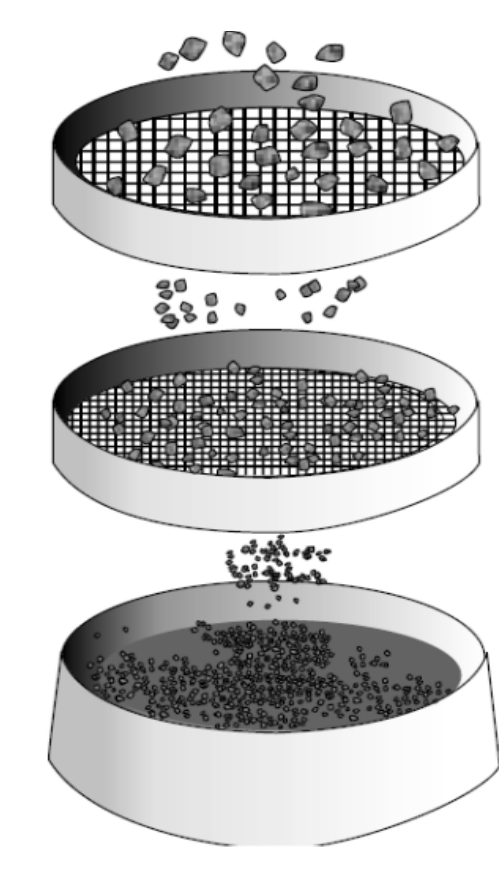
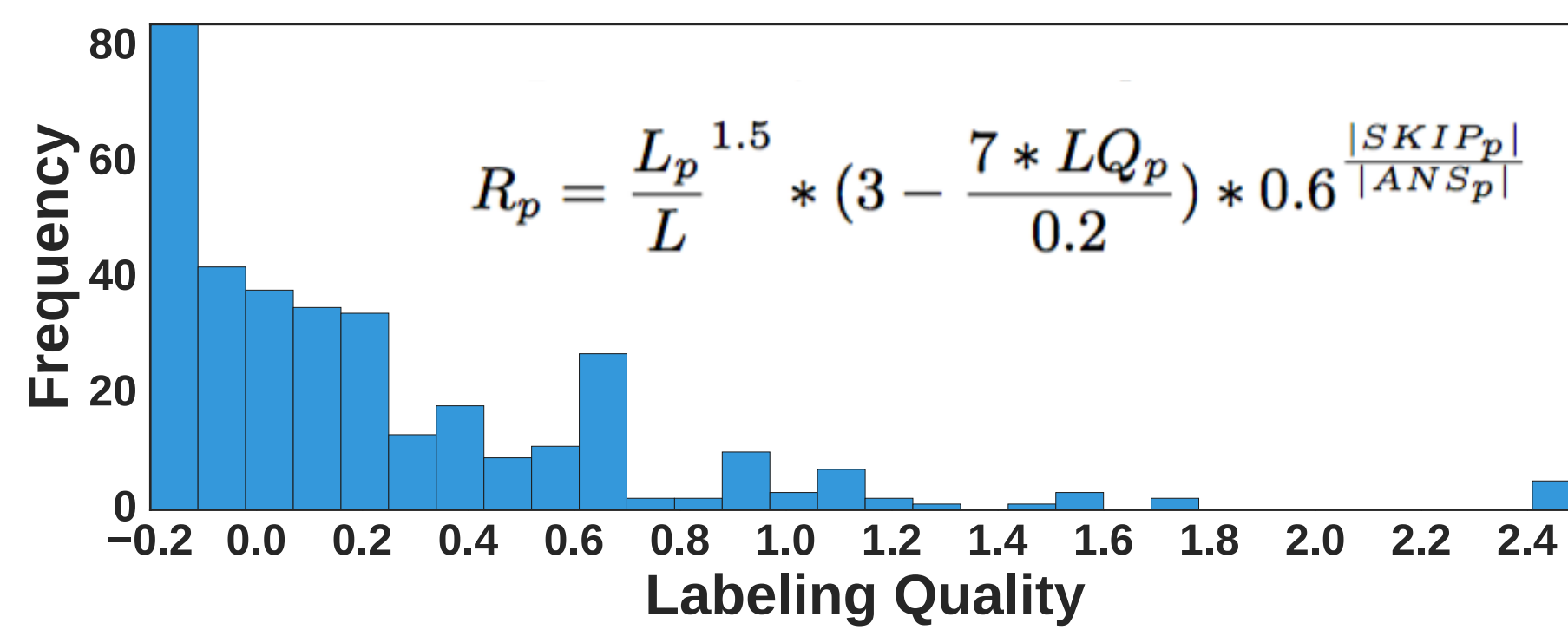
## Data Labeling and Ground-Truth Collection

### Coding website

[bit.ly/claimbusters](http://bit.ly/claimbusters)



### Coder quality



20788 sentences

374 coders

76552 labels

86 top-quality coders

52333 labels

Majority voting

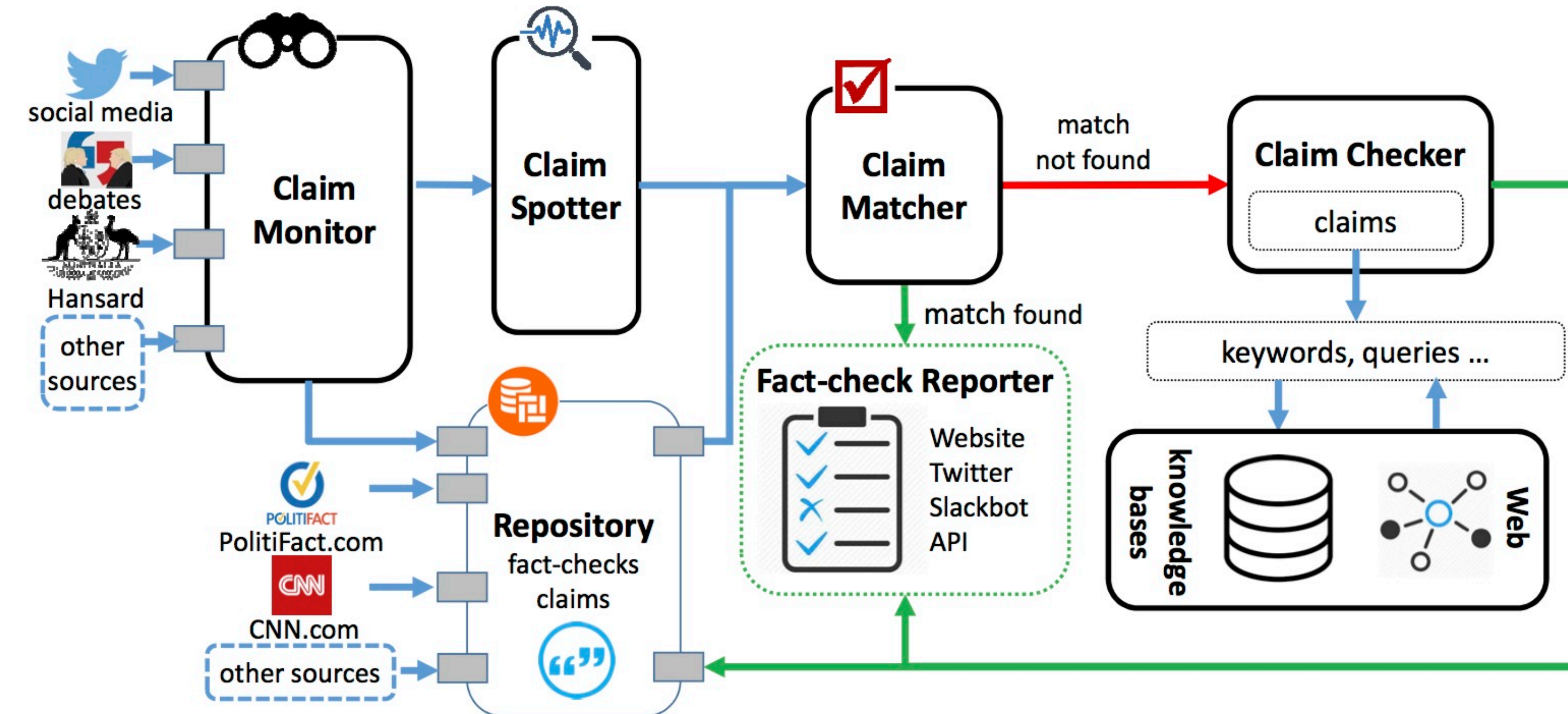
20617 admitted sentences

### Quality assurance

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

	Count	Percentage
NFS	13671	66.31
UFS	2097	10.17
CFS	4849	23.52

## End-to-End Fact-Checking



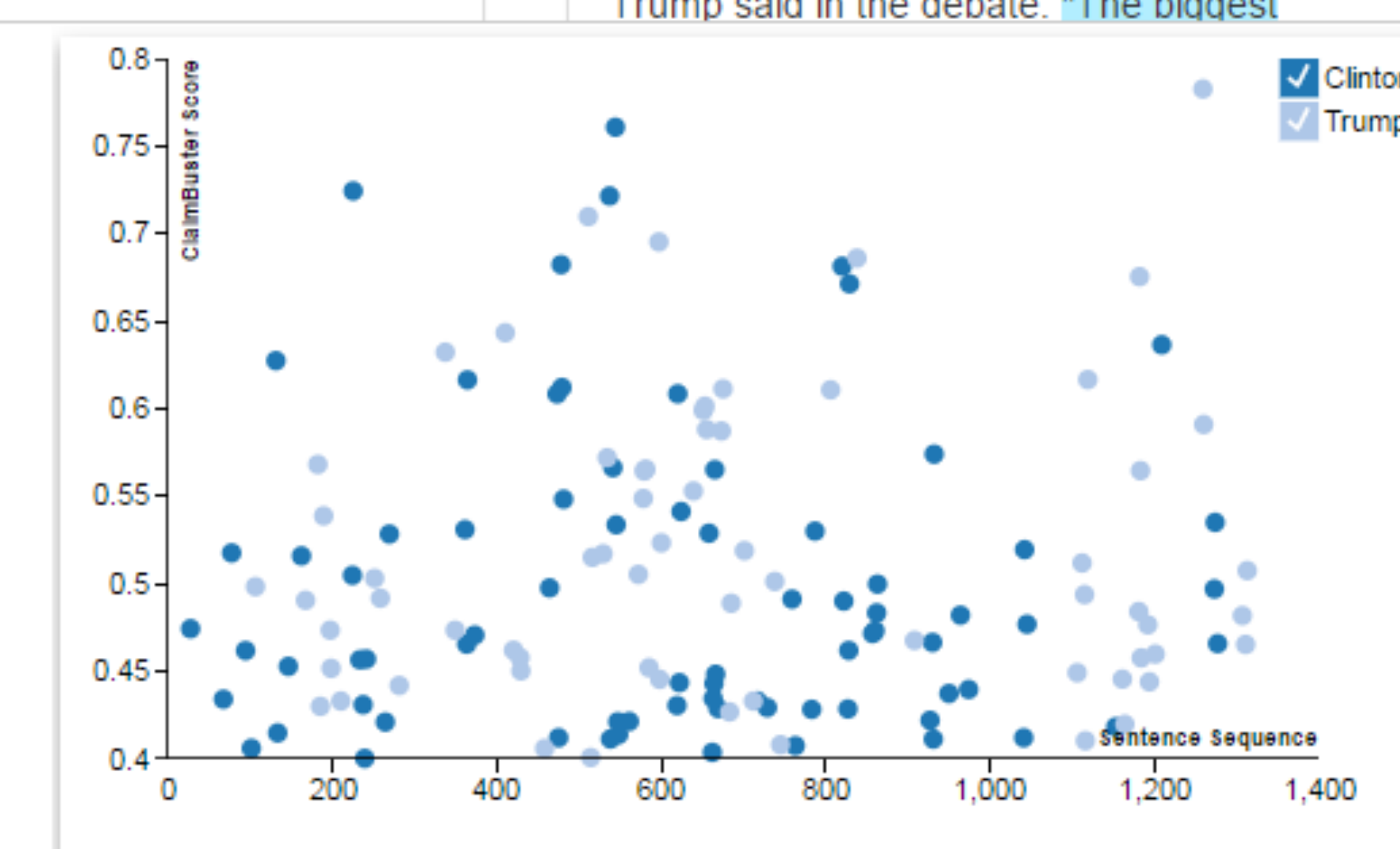
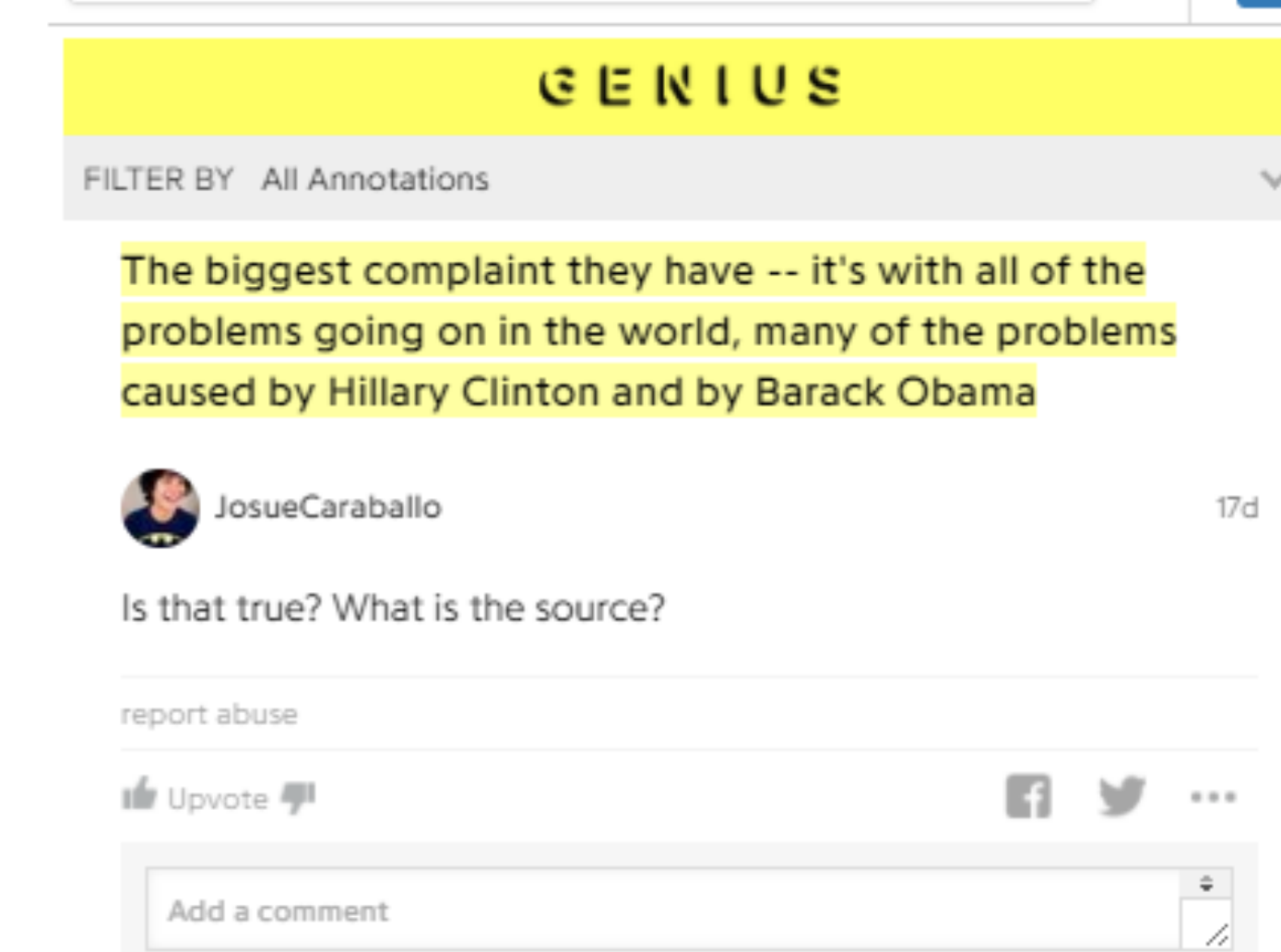
2016 Third Presidential Debate. Oct. 19, 2016, 9 p.m. EST

Chronological Order Order by Score

Least Check-worthy >=0.1>=0.2>=0.3>=0.4>=0.5>=0.6>=0.7>=0.8>=0.9>=1.0 Most Check-worthy

is tougher. But they know what's going on. They know it better than anybody. They want strong borders. They feel we have to have strong borders. I was up in New Hampshire the other day. The biggest complaint they have -- it's with all of the problems going on in the world, many of the problems caused by Hillary Clinton and by Barack Obama. All of the problems -- the single biggest problem is heroin that pours across our southern border. It's just pouring and destroying your youth. It's poisoning the blood of their youth and plenty of other people. We have to Fact-check this borders. We have to keep the drugs out of our country. We are -- right now, we're getting the drugs, they're getting the cash. We need strong borders. We need absolute -- we cannot give amnesty. Now, I want to build the wall. We need the wall.

Claim Checker - Knowledge Bases	Claim Matcher	Claim Checker - Search Engine
Consulting the knowledge bases produced the following results:  <b>Truth Rating</b> Indeterminable <b>Question Asked</b> What is all of the problems-- the single biggest problem? <b>Response Reviewed</b> The single biggest problem in communication is the illusion that it has ...	We found the following claims which have been professionally fact-checked. Check them out!  <b>Truth Rating</b> True <b>Claim</b> "Heroin .. pours across our southern borders." <b>Speaker</b> Donald Trump <b>URL</b> politifact  <b>Truth Rating</b> True	We found the following information after processing some search engine results:  All of the problems -- the single biggest problem is heroin that pours across our southern border. It's just pouring and destroying your youth. <b>Similarity Rating</b> 0.8320502943378437 <b>URL</b> source  "I was up in New Hampshire the other day." Trump said in the debate. <u>The biggest</u>

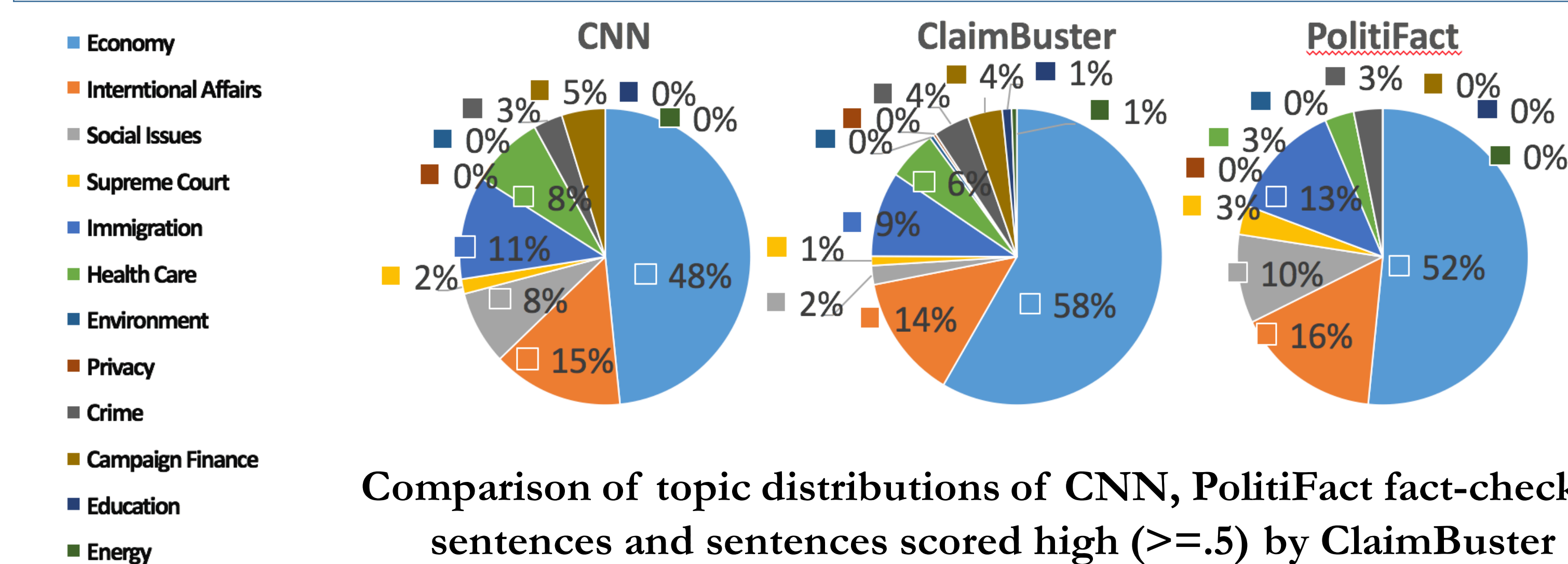


## Classification and Ranking Accuracy

algorithm	features	p_NFS	p_UFS	p_CFS	p_wavg	r_NFS	r_UFS	r_CFS	r_wavg	f_NFS	f_UFS	f_CFS	f_wavg
RFC	W	0.755	0.125	0.638	0.692	0.965	0.004	0.235	0.745	0.848	0.008	0.343	0.685
NBC	W	0.788	0	0.816	0.747	0.983	0	0.385	0.791	0.875	0	0.522	0.744
SVM	W	0.871	0.426	0.723	0.811	0.925	0.227	0.667	0.826	0.897	0.296	0.694	0.816
RFC	W_P	0.772	0.358	0.701	0.731	0.968	0.011	0.312	0.764	0.859	0.02	0.43	0.713
NBC	W_P	0.799	0	0.805	0.753	0.979	0	0.44	0.8	0.88	0	0.569	0.758
SVM	W_P	0.873	0.43	0.724	0.813	0.925	0.24	0.671	0.827	0.898	0.307	0.696	0.818
RFC	W_P_ET	0.77	0.238	0.665	0.715	0.964	0.008	0.298	0.758	0.856	0.016	0.411	0.706
NBC	W_P_ET	0.803	0	0.791	0.752	0.976	0	0.455	0.801	0.881	0	0.577	0.76
SVM	W_P_ET	0.873	0.427	0.723	0.813	0.925	0.24	0.67	0.827	0.898	0.307	0.695	0.817

k	P@k	AvgP	nDCG
10	1.000	1.000	1.000
25	1.000	1.000	1.000
50	0.980	0.995	0.987
100	0.943	0.979	0.956
200	0.916	0.955	0.931
300	0.848	0.937	0.874
400	0.764	0.915	0.801
500	0.679	0.897	0.827

## Case Study: 2016 U.S. Presidential Election Debates



System [idir.uta.edu/claimbuster](http://idir.uta.edu/claimbuster)

