Detecting Check-worthy Factual Claims in Presidential Debates
Naeemul Hassan, Chengkai Li, Mark Tremayne

Motivation

People make claims all the time

- “I built a net worth of more than $10 billion”
  - Donald Trump in 2016 GOP presidential candidate debate. factcheck.org checked this claim as False.¹
- “... our Navy is smaller than it’s been since 1917.”
  - Mitt Romney in 2012 Presidential candidate debate. Politifact.org checked this claim as Pants on Fire.²

Fact-checking is growing, but human challenges remain
- 64 active fact-checking sites worldwide³, up from 44 in 2014.
- Falsehoods growing faster than fact-checkers.

Limitations of current approach

- Significant time gap between speech and reporting times.
- Audience does not get correct information.
- Requires advanced writing skills to persuade readers.
- Such skilled writers are sparse.

Lack of Structured Journalism and use of old publishing frameworks hinders Semantic Web applications.

Dataset

- U.S. Presidential Debate Transcripts [1960 - 2012]
- 11 Presidential Elections. 30 Debates
- 20788 Sentences [excluding short sentences (less than 5 words)]

Problem Formulation

Three categories of sentences
- Non-Factual Sentence (NFS): (Opinions, beliefs, declarations)
  - But I think it’s time to talk about the future.
  - You remember the last time you said that?
- Unimportant Factual Sentence (UFS):
  - Next Tuesday is Election Day.
  - Two days ago we ate lunch at a restaurant
- Check-worthy Factual Sentence (CFS):
  - He voted against the first Gulf War.
  - Over a million and a quarter Americans are HIV-positive.

Goal: Given a sentence, find if it belongs to CFS category. A supervised learning problem.