
FAKE NEWS ANALYSIS & DETECTION

Masha Ivenskaya

THE TASK

- The term “**fake news**” has been gaining momentum in recent months
 - Many sources post news articles containing misinformation, and because in social media articles are often **decontextualized from the source**, fact can mix freely with fiction
 - What can we learn from automatic linguistic analysis of dubious sources of news? Are there **linguistic cues** that differentiate these from objective news sources?
 - Can methods that have been used for **automatic deception detection** be of use for fake news detection?
-

PAST RESEARCH

- Deception detection research:
 - Classification models with features such as n-grams, POS tags, syntactic parses, punctuation and LIWC categories
 - Cues of deceptive language vary by domain, but often include **negative emotion words, distinctive use of pronouns, fewer exclusive words, and generally less “complicated” language** because deception is a cognitively challenging task
-

PAST RESEARCH

- Fake News Research:
 - 3 categories of fake news: tabloid/yellow press, large scale hoaxes, humorous fakes
 - Characteristics of tabloid news include **headline/body dissonance** and using **questions in the headline**
 - Suggested approaches: **linguistic cues** and **network cues** (i.e. fact checking)
-

GOALS

- Linguistic approaches can likely help to identify fake news that are published in **sensationalist/tabloid style**
 - Implement **a model to classify news articles** as fake/legitimate based on features that have been used in deception detection research
 - By looking at most predictive features, learn the linguistic cues of fake/sensationalist/tabloid news
 - Automatic detection of this style could be used as a cue in a more robust fake news detection system
-

APPROACH

- News data is different from other text data in that the **lexical content of it changes** over time
 - I want to use **non-lexical features**, such as POS tags, punctuation, and LIWC categories to detect the stylistic differences between fake and legitimate news
-

DATA

- Fake news data:
 - news articles scraped from websites tagged by “**BS Detector**” Chrome extension (by Daniel Sieradski)
 - only articles from sources tagged as “fake” or “bs” by the BS Detector (not sources tagged, for example, as “junk science”)
 - only English
 - **6.5k articles** from 198 sources
-

DATA

- Objective news data:
 - news articles from **The Guardian**
 - only articles from “US News”, “World News”, and “Politics” sections
 - **6.5k articles**
-

CAVEATS

- Cannot be sure that all news sources tagged by BS Detector are fake
 - Degree of homogeneity among fake news sources - unknown
 - Ideally, would want to have multiple legitimate news sources, however this data is difficult to obtain
 - The model will try to get at stylistic differences; however, there could be fake news written in legitimate news style.
-

METHODOLOGY

- Classifiers: Logistic Regression, Linear SVC (slightly better results with Logistic Regression)
 - **Baseline Features:**
 - sentence length
 - presence of all-caps
 - presence of emoticons
 - running punctuation
-

GRAMMATICAL FEATURES

- **POS tags** on headline and text body
 - **POS bigrams** on headline and text body
 - Syntactic **parse** of headline
-

LIWC FEATURES

- LIWC - **Linguist Inquiry and Word Count**
 - Text analysis program that counts words psychologically meaningful categories
 - 2200 words and word stems, 79 categories relevant to psychological processes
 - Has been widely used in linguistics and psycholinguistics
 - Has been used in **deception detection research**
 - I used a **subset of the categories (60)**
-

RESULTS

	PRECISION	RECALL	F1 SCORE
BASELINE FEATURES	86%	56%	66%
POS TAGS	86%	83%	83%
POS BIGRAMS	89%	89%	89%
SYNTACTIC CONSTITUENTS	84%	83%	85%
LIWC FEATURES	87%	64%	69%
ALL FEATURES	90%	90%	90%

FINDINGS: GRAMMATICAL CHARACTERISTICS

- More **proper nouns** (probably because these articles are so focused on the presidential candidates)
 - More **adjectives** and **adverbs** (indicative of **sentiment**)
 - More **comparative** and **superlative** adjectives (“better”, “biggest”, “worst”)
 - More **question** words (“where”, “how”, etc.)
 - More **exclamation** marks
 - Less: semicolons, quotes, articles, apostrophes, commas (indicative of **simpler sentence structure**)
-

FINDINGS: LIWC CHARACTERISTICS

- **informal** language
 - netspeak
 - words expressing **affect**, especially **negative emotions** and **anger**
 - words expressing **certainty**
-

NEXT STEPS

- **Improve data:**
 - evaluate the “fakeness” of sources from BS Detector
 - diversify objective news data - add other sources
 - clean data to get rid of any possible User Generated Content (comments etc.)
 - **Improve features:**
 - syntactic parse of full articles
 - experiment with different subsets of LIWC categories
-

NEXT STEPS

- Fake News Challenge (<http://www.fakenewschallenge.org/>)
 - The task is broken up into stages
 - Stage I : **Stance Detection**
 - Training/testing data is provided
 - 4 categories:
 - Agrees - Disagrees - Discusses - Unrelated
-

Q & A
