



FAKE NEWS

Christianlaurent Mundell

Introduction

- Since the beginning of any form of media sharing rather it be a newspaper, radio, or website there have been people who tried to trick or deceive others with lies.
- Recently popularize by Donald Trump, I set out to fix an age-old problem by detecting rather an online article is a fake or reliable news source.

Data Final Form

		total	label
0	House Dem Aide: We Didn't Even See Comey's Let...	1	
1	FLYNN: Hillary Clinton, Big Woman on Campus - ...	0	
2	Why the Truth Might Get You FiredConsortiumnew...	1	
3	15 Civilians Killed In Single US Airstrike Hav...	1	
4	Iranian woman jailed for fictional unpublished...	1	
5	Jackie Mason: Hollywood Would Love Trump if He...	0	
6	Life: Life Of Luxury: Elton John's 6 Favorite ...	1	
7	Benoît Hamon Wins French Socialist Party's Pre...	0	
8	Excerpts From a Draft Script for Donald Trump'...	0	
9	A Back-Channel Plan for Ukraine and Russia, Co...	0	
10	Obama's Organizing for Action Partners with So...	0	
11	BBC Comedy Sketch "Real Housewives of ISIS" Ca...	0	
12	Russian Researchers Discover Secret Nazi Milit...	1	
13	US Officials See No Link Between Trump and Rus...	1	

Data I can use

link
https://www.huffingtonpost.com/entry/texas-ama...
https://www.huffingtonpost.com/entry/will-smit...
https://www.huffingtonpost.com/entry/hugh-gran...
https://www.huffingtonpost.com/entry/jim-carre...
https://www.huffingtonpost.com/entry/julianna-...

Trump Turns Daily Coronavirus Briefing Into a Defense of His Record

Also at the briefing on Monday, Dr. Anthony Fauci said a comment he made that was seen as critical of the president was taken out of context, and Mr. Trump said he did not intend to fire him.



How Articles Were Chosen & Classified

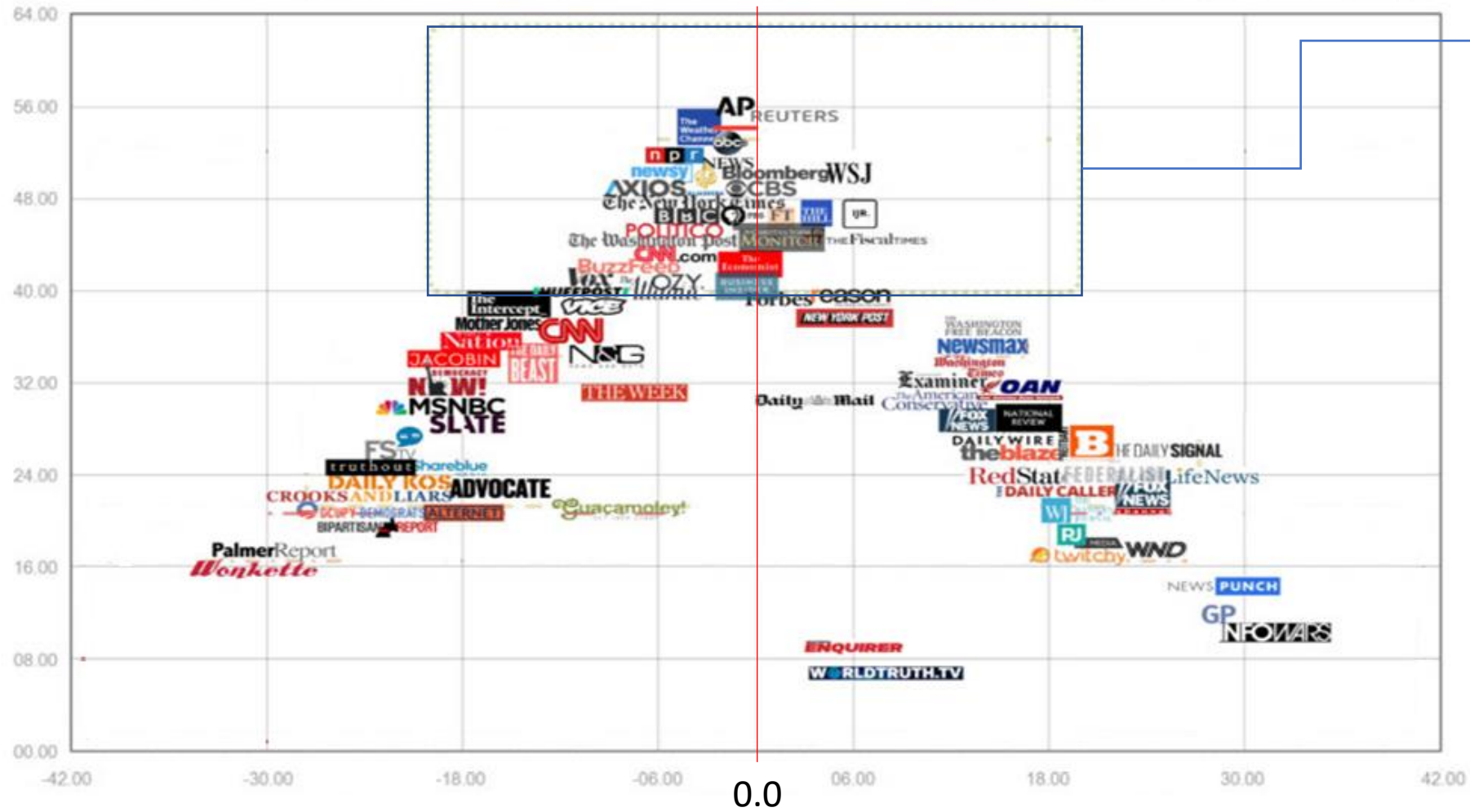
- The original dataset from Kaggle provided ~ 20000 source of equal balance of reliable and fake news articles.
- BuzzFeed provided a list of 8000 fake news articles along with impact metrics.
- Reliable news sources were picked by taking a vim diagram of two sources and only scraping from media outlets that appeared as reliable in both sources.

The Media Bias Chart

Version 5.1

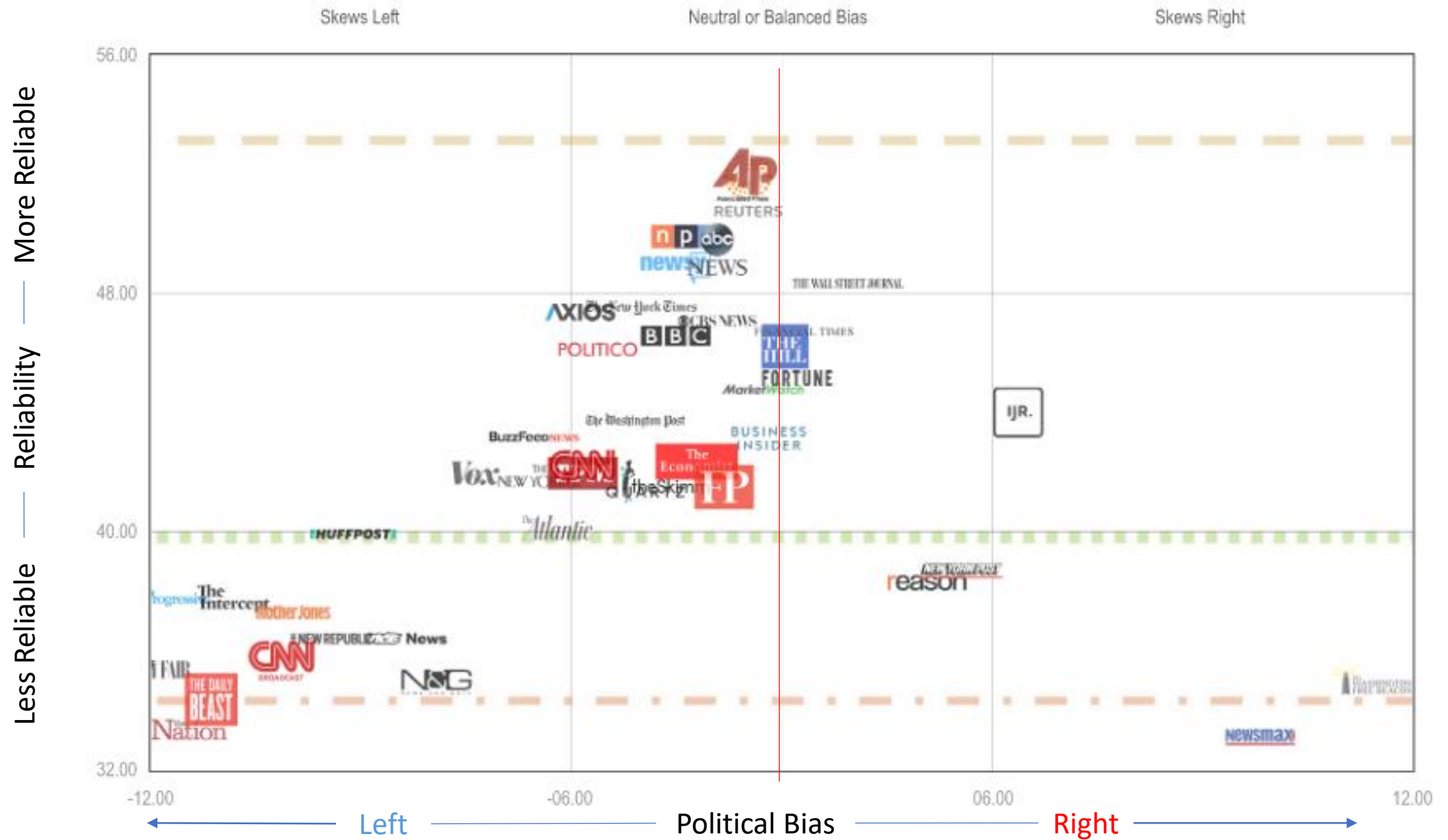
Less Reliable — Reliability — More Reliable

Most Extreme Left Hyper-Partisan Left Skews Left Neutral or Balanced Bias Skews Right Hyper-Partisan Right Most Extreme Right



Area of focus
&
most reliable
sources

Left ————— Political Bias ————— Right



10 Journalism Brands Where You Find Real Facts Rather Than Alternative Facts

1. The New York Times
2. The Wall Street Journal
3. The Washington Post
4. BBC
5. The Economist
6. The New Yorker
7. Wire Services: The Associated Press, Reuters, Bloomberg News
8. Foreign Affairs
9. The Atlantic
10. Politico

<https://www.forbes.com/sites/berlinschoolofcreativeleadership/2017/02/01/10-journalism-brands-where-you-will-find-real-facts-rather-than-alternative-facts/#146d12dae9b5>



Forbes

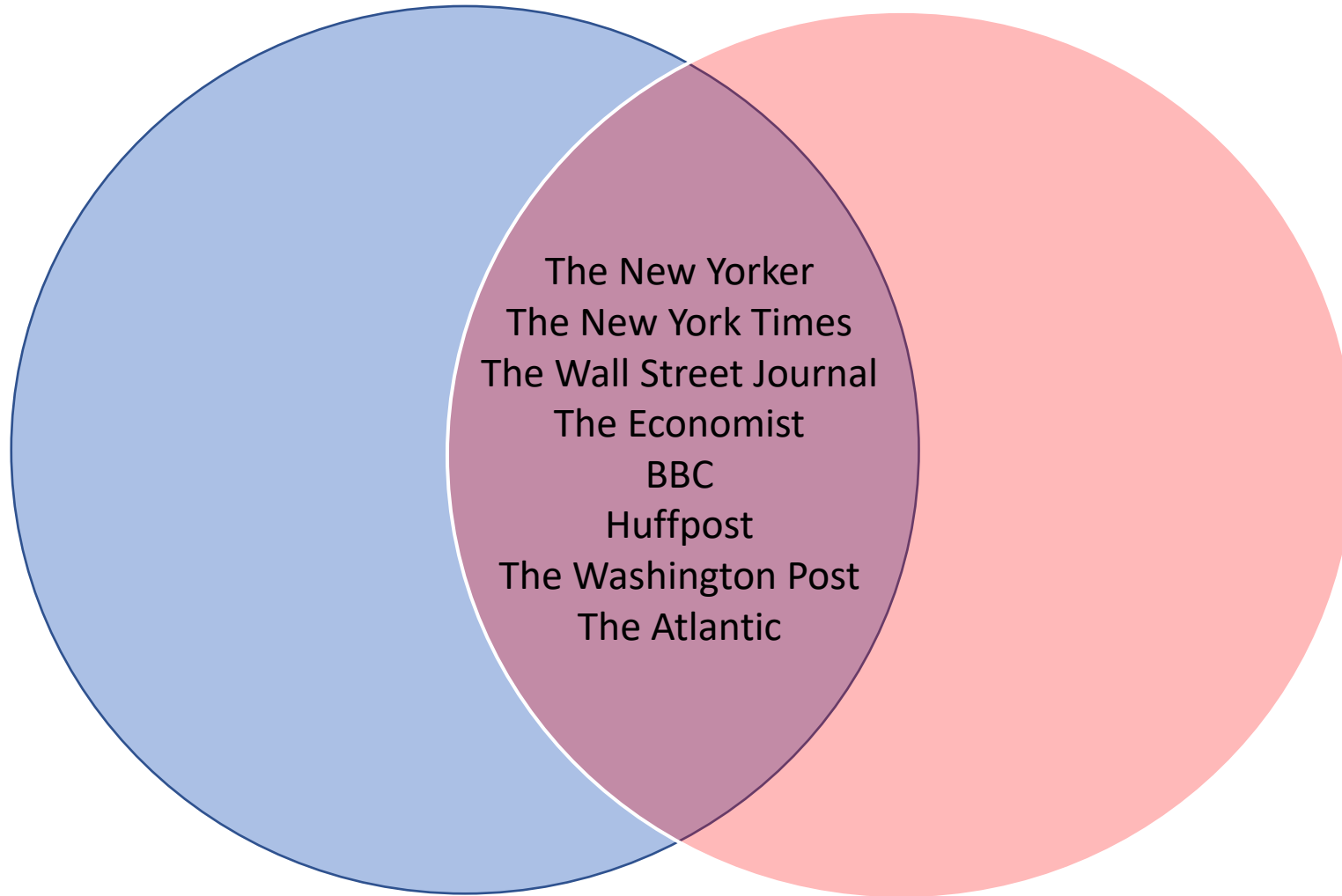




Forbes



8 Remaining Sources for Reliable News



Forbes



Ad Fontes Media

Results

99.98% accuracy!

Results

- The model made the right prediction 99.98% on test data and shows good signs of being able to be used on future data.
- It is possible to detect fake news. In the near future I'm hopeful fake news will have less of an impact on the general population.



- Step one: Provide a URL to a web article.

Example: <https://www.nytimes.com/2020/04/13/us/politics/trump-daily-coronavirus-briefing.html>

- Step 2: Press the enter / submit button

URL

Submit

The article is real.

Results

<https://i-found-you-faker.herokuapp.com/>

DEMO



Shift

QUESTIONS?

Links

- <https://www.kaggle.com/c/fake-news/data>
- <https://www.kaggle.com/rmisra/news-category-dataset>
- <https://libguides.library.kent.edu/c.php?g=278400&p=1854632>
- <https://www.forbes.com/sites/berlinschoolofcreativeleadership/2017/02/01/10-journalism-brands-where-you-will-find-real-facts-rather-than-alternative-facts/#82973b6e9b5a>
- <https://www.adfontesmedia.com/interactive-media-bias-chart/?v=402f03a963ba>

Links

- <https://github.com/BuzzFeedNews/2018-12-fake-news-top-50>
- <https://www.buzzfeednews.com/article/craigsilverman/facebook-fake-news-hits-2018>
- <https://www.kaggle.com/nzalake52/new-york-times-articles>
- <https://www.kaggle.com/plarmuseau/minimalistic-logistic-ngram-tfidf-lb-0-975>

Data Sources

- Kaggle

- <https://www.kaggle.com/c/fake-news/data>
- <https://www.kaggle.com/rmisra/news-category-dataset>
- <https://www.kaggle.com/nzalake52/new-york-times-articles>

- BuzzFeed

- <https://github.com/BuzzFeedNews/2018-12-fake-news-top-50>