

FAKE NEWS

Laurent 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.

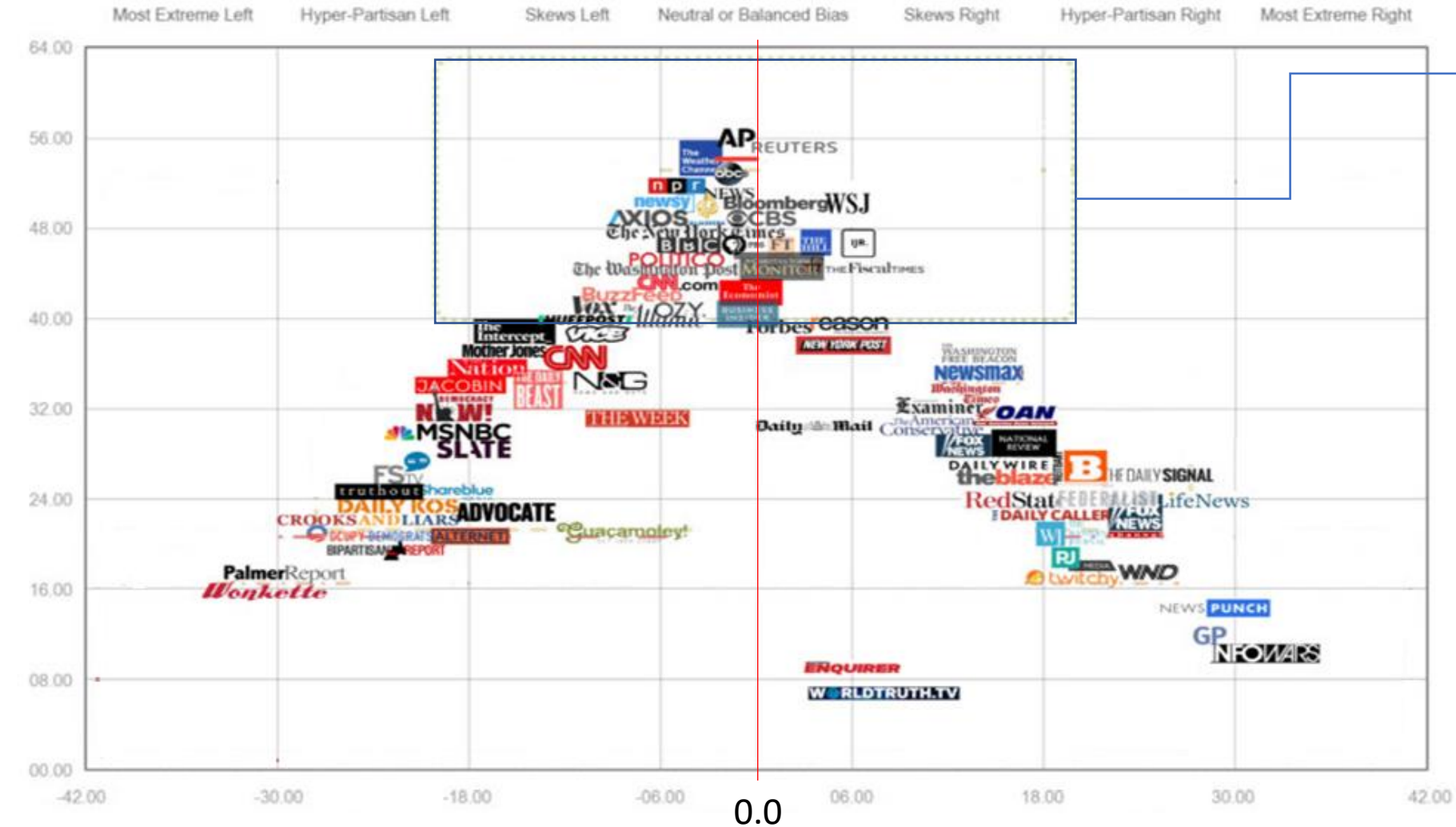
How Articles Were Chosen & Classified

- 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

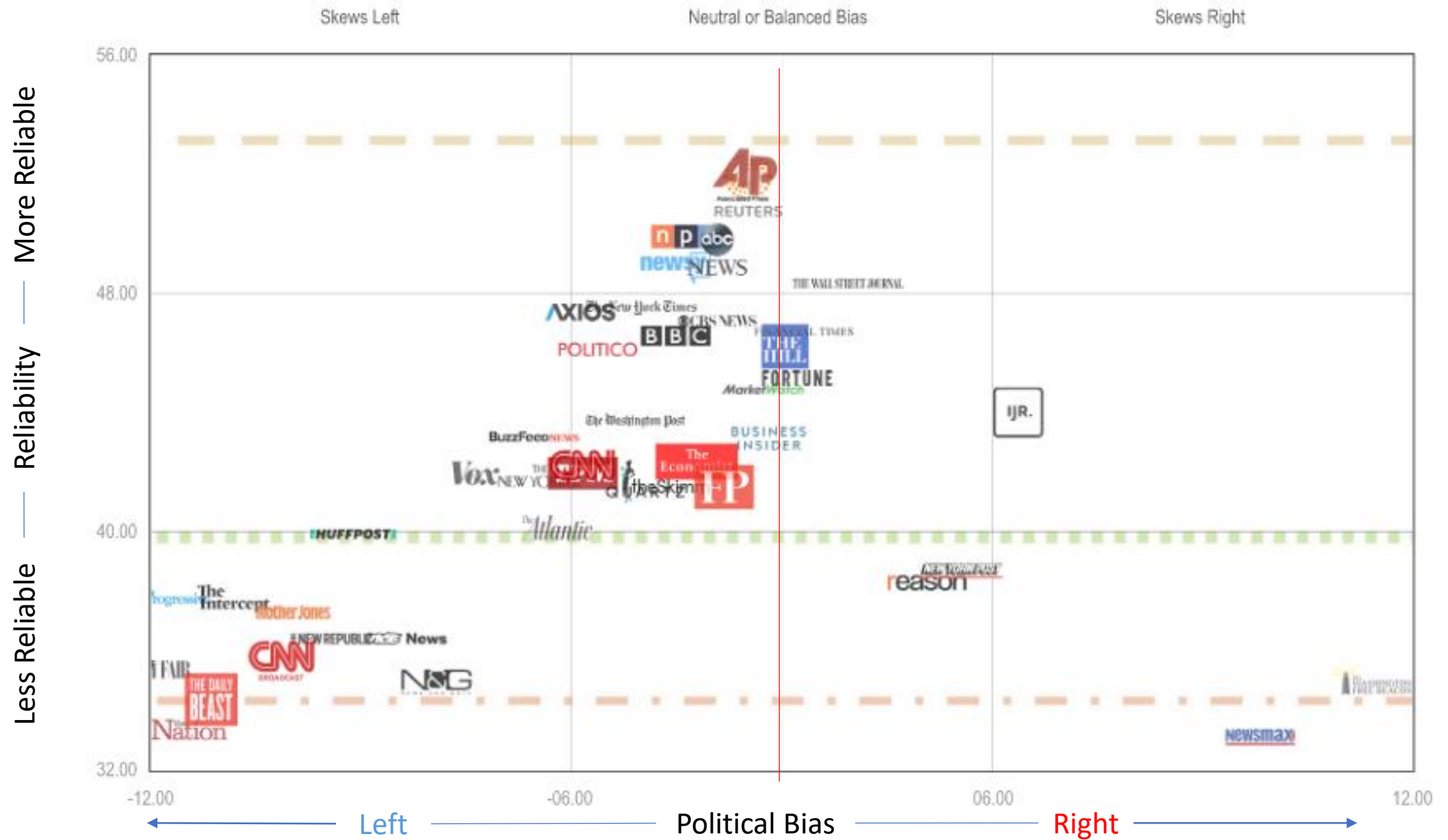
Version 5.1

Less Reliable — Reliability — More Reliable



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

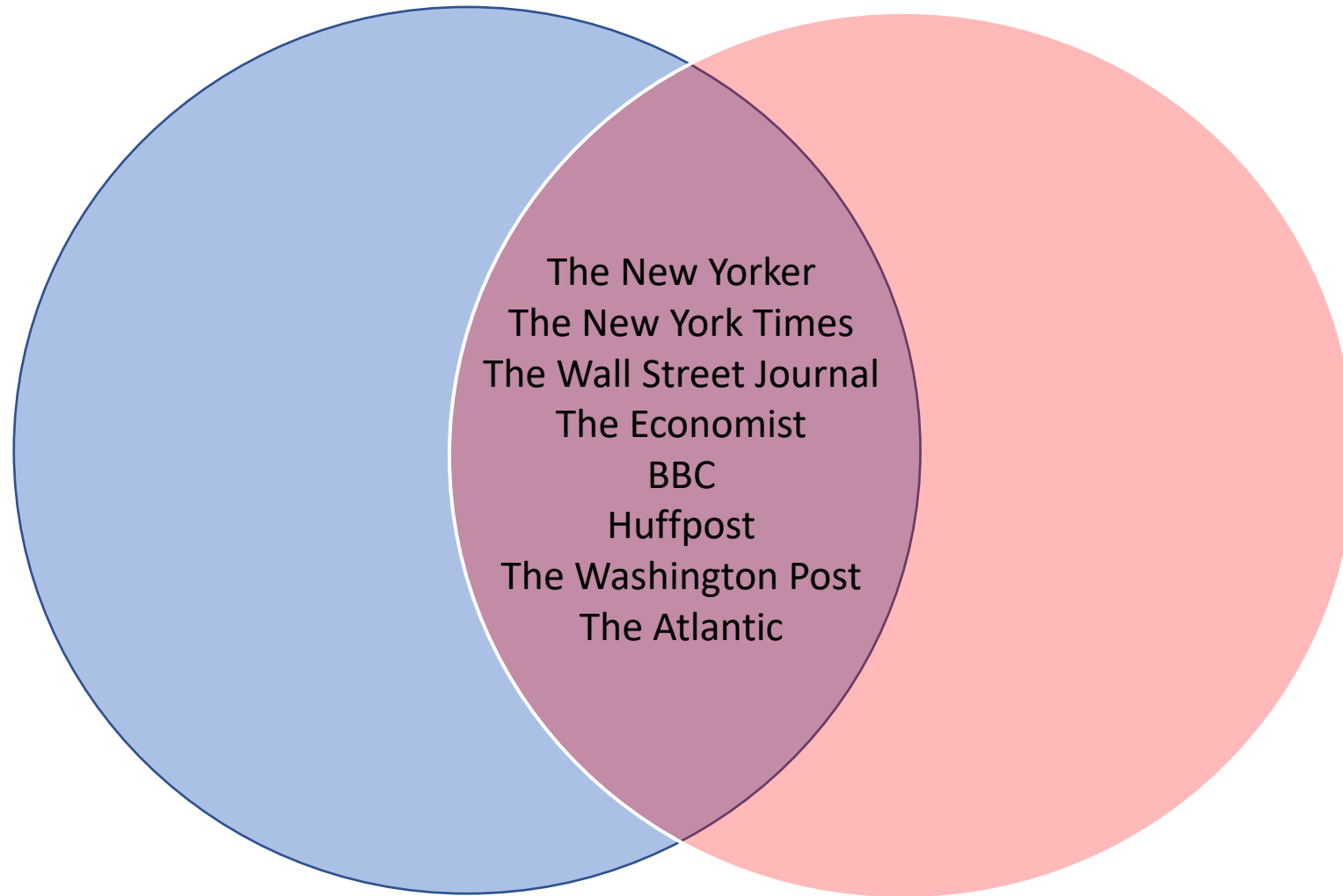




Forbes



8 Sources for Reliable News



Ad Fontes Media



- 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/>

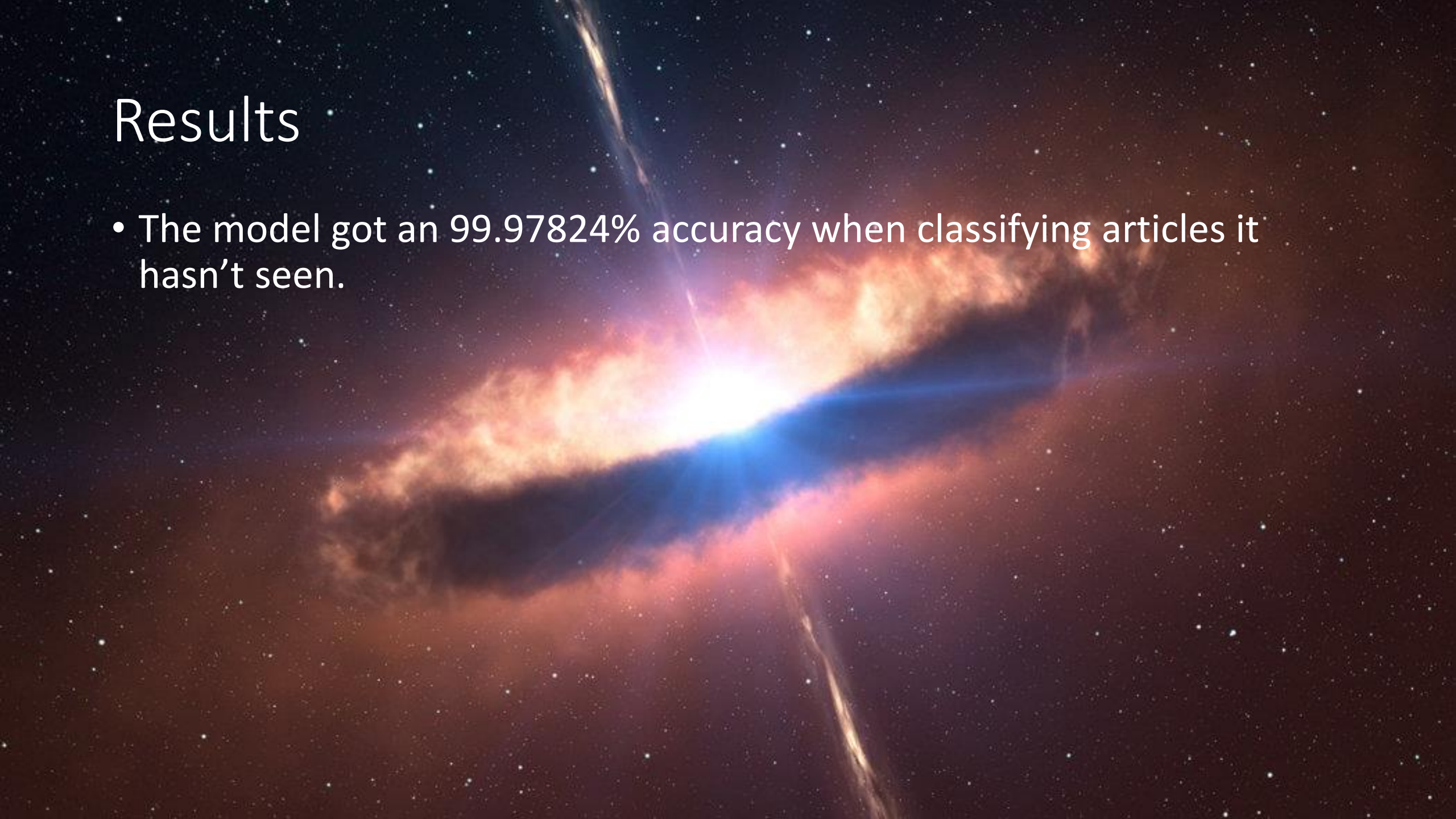
DEMO



shift

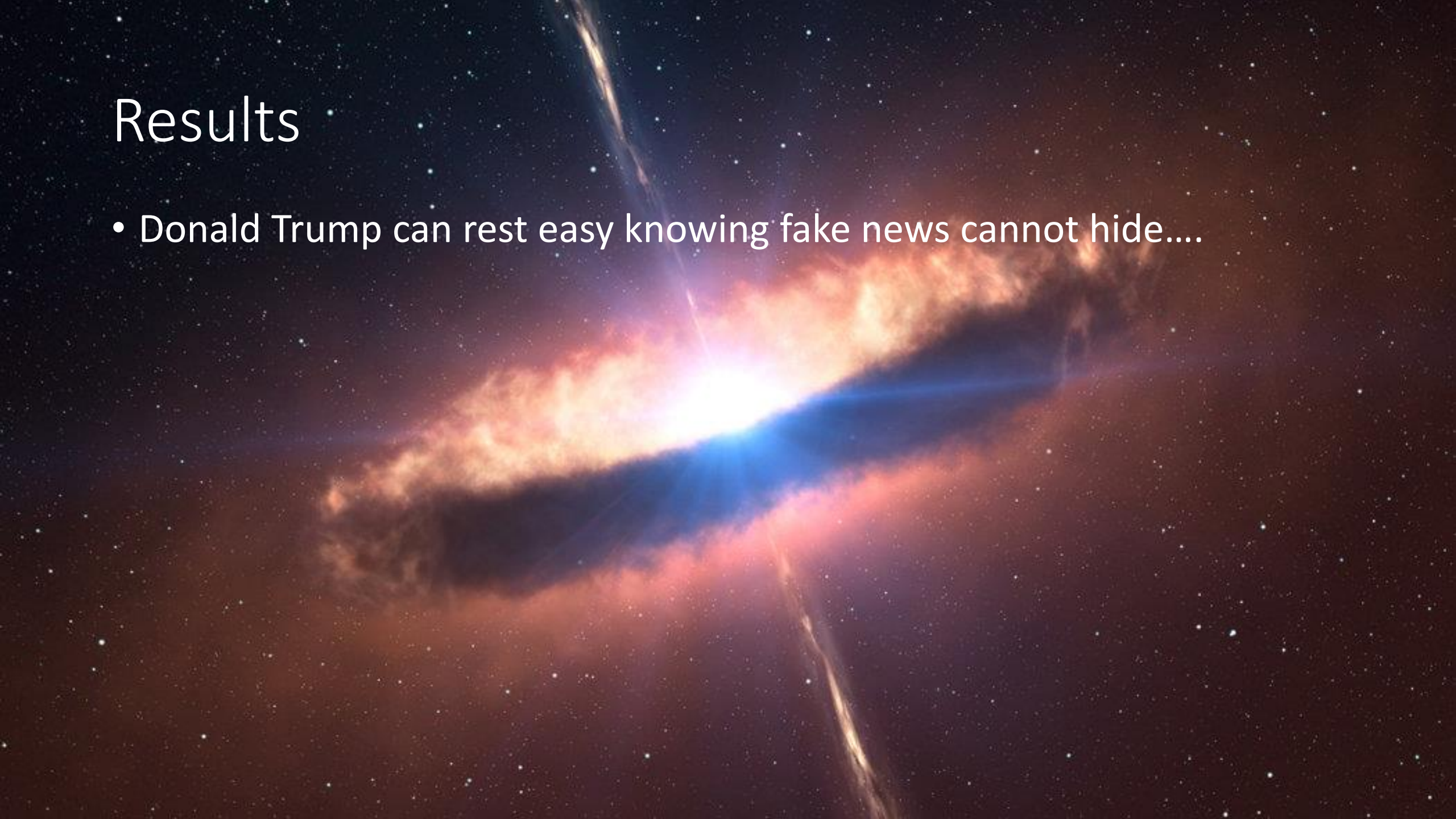
Results

- The model got an 99.97824% accuracy when classifying articles it hasn't seen.



Results

- Donald Trump can rest easy knowing fake news cannot hide....



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>