





Politics Fake News Detector in LATAM (Latin America)

Participants



Andrés Pitta



Brandon Copeland



Gouri Kulkarni



Isaac Rodríguez



Gary Astorga



Ibrahim Koicha



Victor Cuspinera-Contreras

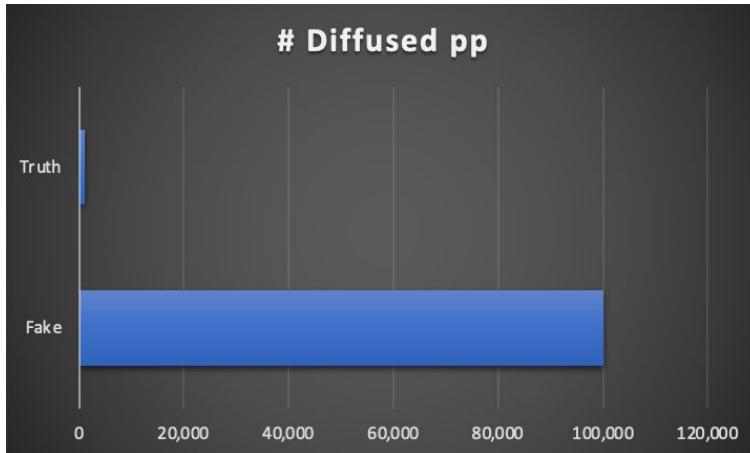


Armielyn Obinguar



Metrics

Vosoughi, et al. mentioned that in Twitter from 2006 to 2017 about **126,000 rumors** were spread by **~3 million people**





Politics Fake News Detector in LATAM (Latin America)

LET'S DO A MVP!!!

What are the benefits of detecting false/fake news?

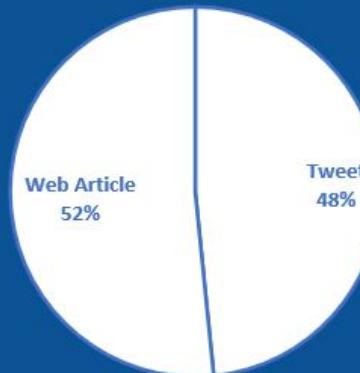
- Detect frauds
- Detect fake accounts
- Limit its spread
- Help Misinformed people



The first challenge of the project was extracting data

News articles and tweets are generally publicly available, so they were extracted by using scrapers and Twitter's APIs.

We extracted 1.500 tweets and 1.599 Web Articles, from 2013 to 2022.



Source: [Wikimedia](#)

When we receive an article, How can we tell it is **Fake**?

Is the information exaggerated?

Is the source of the data missing?

Is the article's source a reliable media outlet?

Is it showing up in fact checking web pages as **Fake**?

Is the article omitting data?

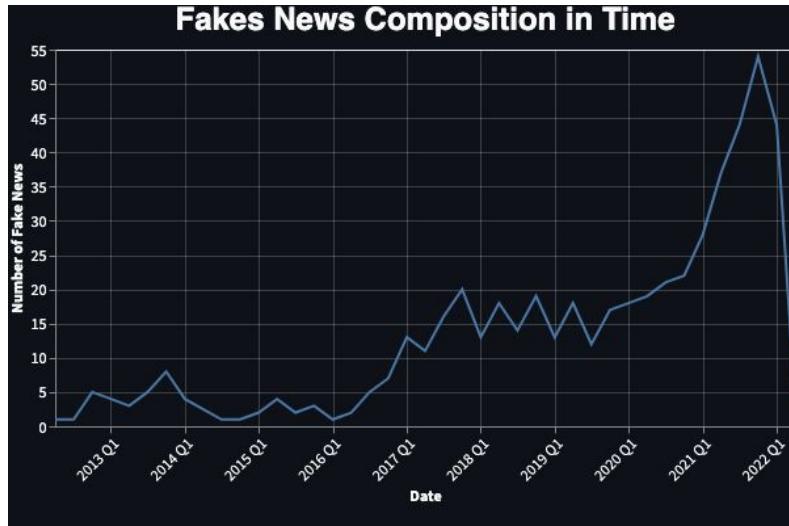


Source: [Wikimedia](#)

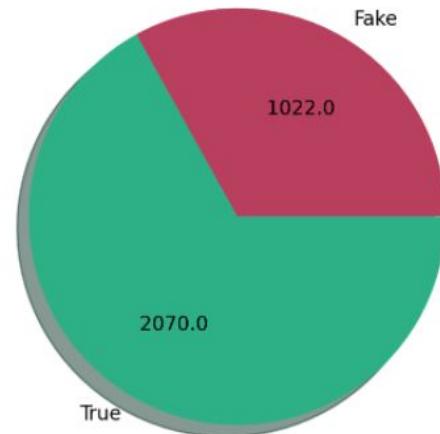
If the majority of the questions are answered by a **Yes**, then we consider the article as **Fake**



- 1,000 Web Articles
- 1,000 Tweets
- 1,000 Non-corrected Tweets & Articles



Distribution by Type of News



Model Stats Cont.

Below are some sample statistics for the highest performing model using Gradient Boost:

	precision	recall	f1-score	support
0	0.88	0.82	0.85	56
1	0.76	0.84	0.79	37
accuracy			0.83	93
macro avg	0.82	0.83	0.82	93
weighted avg	0.83	0.83	0.83	93

With support from Omdena, we were able to secure a Google Colab Pro account that aided in model training time and with a 20% jump in accuracy.

LATAM Fake News +

localhost:8501

About the Project

Dashboard

Fake News

Try it out!

Corrections

The Project Team

Politics Fake News Detector in LATAM (Latin America)

The Background

Since the Cambridge Analytica scandal a pandora box has been opened around the world, bringing to light campaigns even involving our current Latinamerica leaders manipulating public opinion through social media to win an election. There is a common and simple pattern that includes platforms such as facebook and fake news, where the candidates are able to build a nefarious narrative for their own benefit. This fact is a growing concern for our democracies, as many of these practices have been widely spread across the region and more people are gaining access to the internet. Thus, it is a necessity to be able to advise the population, and for that we have to be able to quickly spot these plots on the net before the damage is irreversible.

The Problem

Once the capacity to somewhat detect irregularities in the news activity on the internet is developed, we might be able to counter the disinformation with the help of additional research. As we reduce the time spent in looking for those occurrences, more time can be used in validating the results and

Link: [Politics Fake News Detector in LATAM](#)

Competition

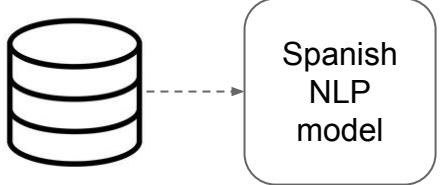


Outcomes of survey – 10 people asked & 100% validate our ML outcome on fake news detection

- “Useful when doubts are present in news, articles or websites”.
- “People could use it also for validating rumors in viral news”
- Some respondents suggest to also use audio transcription and adapt it as a plugin in Google Chrome.

Next step proposal

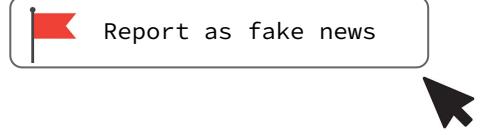
Dynamic AI Model



Checklist

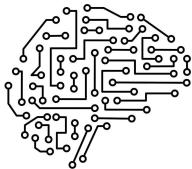
- References to other sources
- Fact Checker validation
- What have other sites written about the same event/news

User suggestion



Target solution

Heuristics + Model = Prediction



Validation's news score: **Low**



How did we get the score?

- The author included references to other sources to back up his/her claims.
- Our AI model scored a **low** probability of the news being fake.

...

[Click to find out how to spot a fake news](#)

Further actions and needed resources

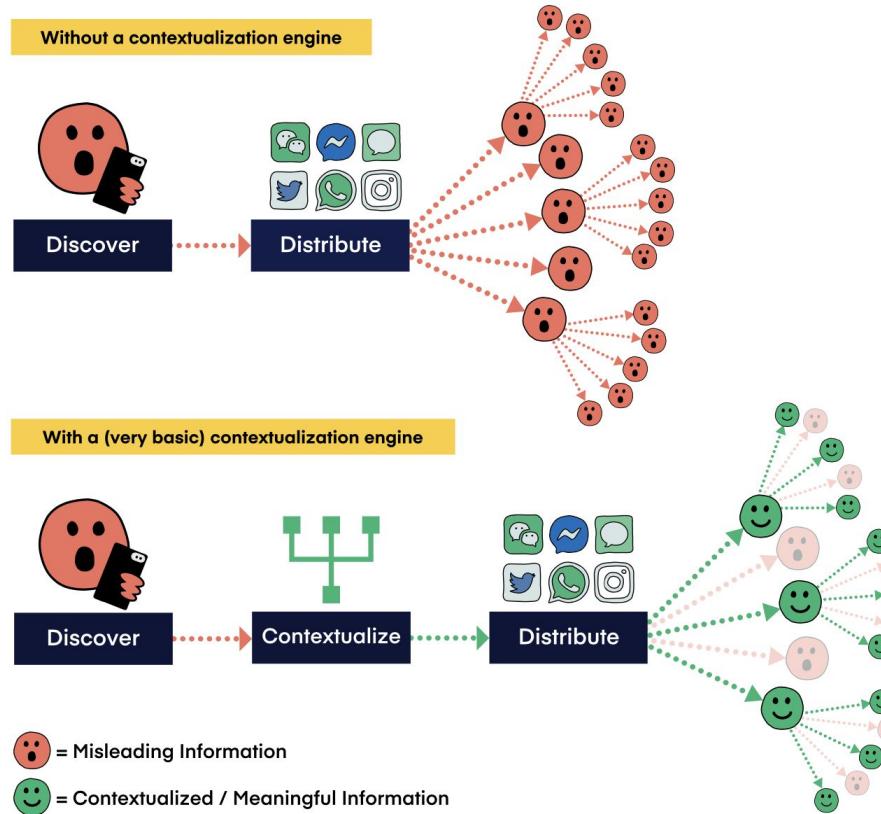
- Big Data Gathering.
- Research and experimentation on an AI model building.
- Deployment with Cloud Resources.
- Sustainable Business model for maintenance and constant improvement.

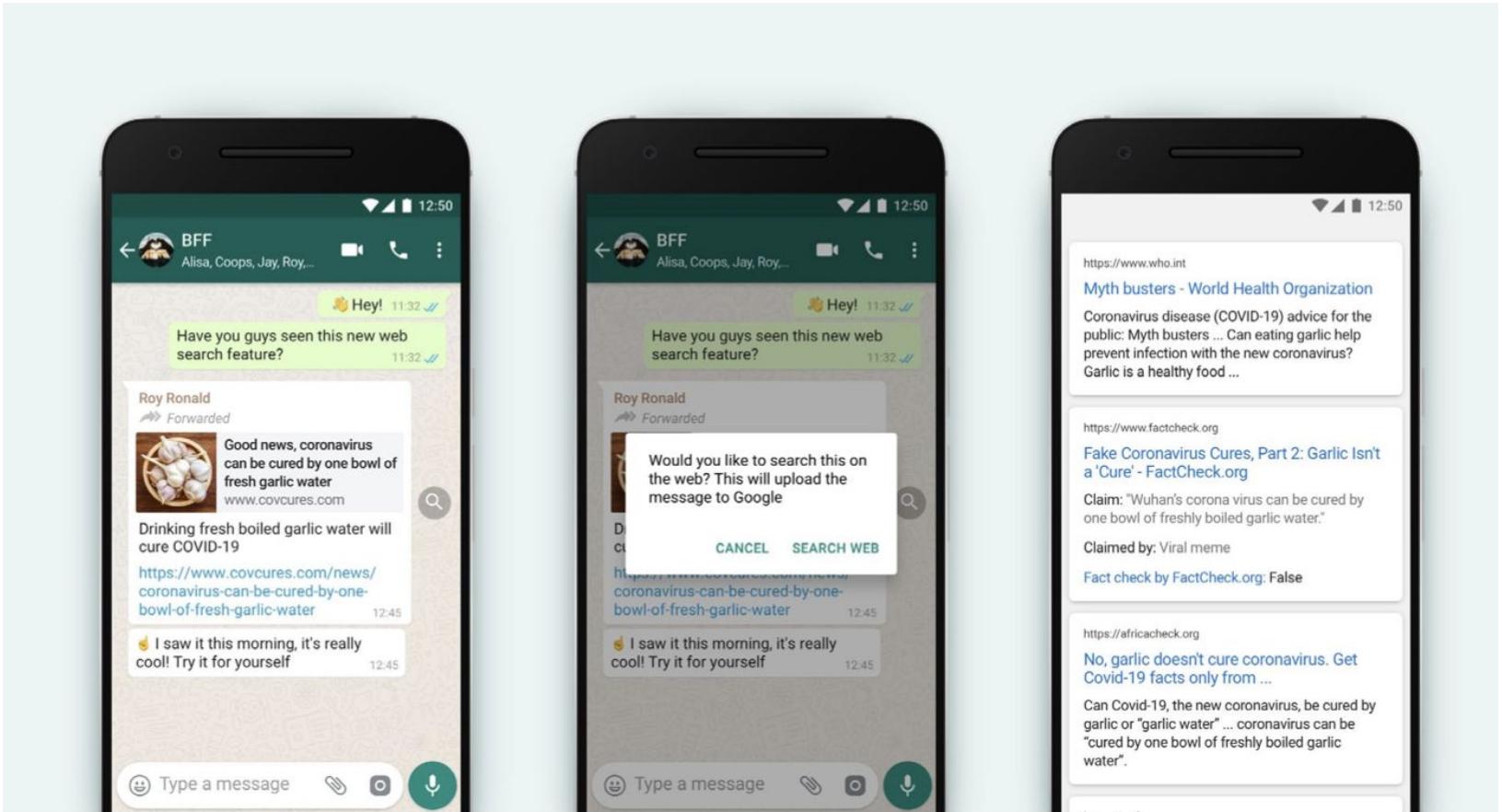


NEWS VERACITY VALIDATOR

- PHASE 1 -

WHAT'S A CONTEXT ENGINE?





SOME CHARACTERISTICS

- Contextualization engines can help people make sense of the information that they come across online by making ‘media literacy work’ less time-consuming.
- A “contextualize button” to push media to a contextualization engine could be built into everything — from social media and messaging apps to operating systems; a normal and expected part of software interfaces.
- New AI technology allows computers to *understand* media in ways that were not possible before, allowing contextualization engines to more effectively sort through information.

GOALS

- Research state-of-the-art models/algorithms of a context engine searcher.
- Implement the top context search models or algorithms.
- Automate a context search model/algorithm

TIMELINE & WORKSHOPS

EXPECTATIONS

- Playground to try any idea, solution and other things than the current timeline.
- Pretty pretty proactive (Propose any meeting, task/sub task or activity)
- Read articles, bring updates, bring new ideas!.
- Work with passion!! If it's not fun, let's talk how we can make it !!

NEXT STEPS

Slack

Github (Expected everyone to be already registered)

Task leads & benefits

Weekly Meeting - >Google Calendar



Data
Engineering
Latam



Labelbox



Google Colaboratory

THANKS TO ALL!

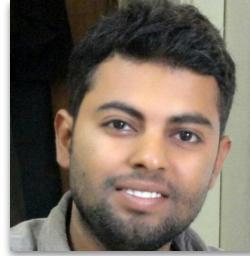
Improving labeling guidelines

From twitter Rules and Policies:

-  Is the content significantly and deceptively altered, manipulated, or fabricated?
-  Is the content shared in a deceptive manner or with false context?

Mentors/Speakers



Data gathering	Topic Modeling and Modeling
Anellee Nurkayeva	<p>Abhishek Singh Rathore</p>  A head-and-shoulders portrait of a young man with dark hair and a beard, smiling at the camera. He is wearing a light-colored shirt.

Data Labellers



Juan Carlos
Rivas



Eugenis
Villamizar



Gilberto Nava
Marcos

Ramón Iván Puón Meraz

Chapter Leads



**Giancarlo
Pablo**



**Miguel
Correa**



CALL TO ACTIONS

Mentorship

Partnership

Sponsorship

Misleading media (Twitter)

- Altered
- Manipulated
- Fabricated
- Shared with false context
- Cause confusion on public issues
- Impact public safety
- Cause serious harm

MVP

Is this news real or fake?

There is a common and simple pattern that includes platforms such as facebook and fake news, where the political actors are able to build a nefarious narrative for their own benefit.



Once the capacity to somewhat detect irregularities in the news activity on the internet is developed, we might be able to counter the disinformation with the help of additional research.

However, we did not have any indication of whether the articles were **fake** or not

News articles and tweet are generally publicly available, so they were extracted by using scrappers and Twitter's APIs.

We extracted 1.500 tweets and 1.599



Model Stats

Below are some sample statistics for the lowest performing model using Logistic Regression:

	precision	recall	f1-score	support
0	0.73	0.62	0.67	56
1	0.53	0.65	0.59	37
accuracy			0.63	93
macro avg	0.63	0.64	0.63	93
weighted avg	0.65	0.63	0.64	93



HACKERS CAN TURN YOUR HOME COMPUTER INTO A BOMB

By RANDY JEFFRIES / Weekly World News

WASHINGTON — Right now, computer hackers have the ability to turn your home computer into a bomb and blow you to Kingdom Come — and they can do it anonymously from thousands of miles away!

Experts say the recent "break-ins" that paralyzed the Amazon.com, Buy.com and eBay websites are tame compared to what will happen in the near future.

Computer expert Arnold Yabenson, president of the Washington-based consumer group National CyberCrime Prevention Foundation (NCPF), says that as far as computer crime is concerned, we've only seen the tip of the iceberg.

"The criminals who knocked out those three major online businesses are the least of our worries," Yabenson told *Weekly World News*.

"There are brilliant but unscrupulous hackers out there who have developed technologies that the average person can't even dream of. Even people who are familiar with how computers work have trouble getting their minds around the terrible."

"It is already possible for an assassin to send someone an e-mail with an innocent-looking attachment connected to it. When the receiver downloads the attachment, the electrical current and molecular structure of the central processing unit is altered, causing it to blast apart like a large hand grenade.

Sicks can wreak death and destruction from thousands of miles away.

Arnold Yabenson.

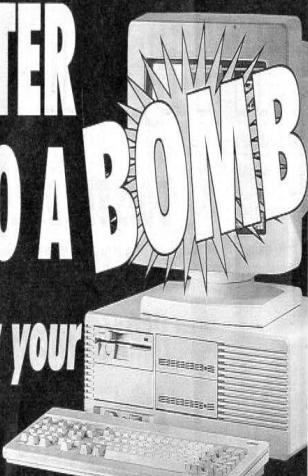
KABOOM! It might not look like it, but an innocent home computer like this one can be turned into a deadly weapon.

"As shocking as this is, it shouldn't surprise anyone. It's just the next step in an ever-escalating progression of horrors conceived and instituted by hackers."

Yabenson points out that these dangerous sociopaths have already:

- Vandalized FBI and U. S. Army websites.
- Broken into Chinese military networks.
- Come within two digits of cracking an 87-digit Russian security code that would have sent deadly missiles hurtling toward five of America's major cities.

"As dangerous as this technology is right now, it's going to get much



When we receive an article, How can we tell it is **Fake?**

1

We trained a Low Performance Model (Logistic Regression - Recall of around **64%**). Generic dataset.

Meaning: Results were unreliable but could pick up writing patterns. Baseline to correct

2

Use the Low Performance Model, fact checking web pages and a 7 Questions system to correct the data.

Source: [Flickr](#)

WHAT IS NOT A FAKE NEWS?

- Memes or satire,
- Animations, illustrations, and cartoons,
- Commentary, reviews, opinions, and/or reactions.
- Counterspeech.
- Doctored or fake Tweets, social media posts, or chat messages.

IMPACT OF A FAKE NEWS DETECTOR

- Disrupting economic incentives.
- Helping people make more informed decisions when they encounter false news.
- To find synthetic, manipulated, or out-of-context media.

NEW SOURCES

Retrieval Augmentation: <https://ruder.io/ml-highlights-2021/index.html#11retrievalaugmentation>

What's next for NPL: <https://www.amazon.science/blog/acl-what-comes-next-for-natural-language-processing>

Twitter policy: <https://help.twitter.com/en/rules-and-policies/manipulated-media>

Working to stop misinformation:

<https://www.facebook.com/formedia/blog/working-to-stop-misinformation-and-false-news>

The spread of true and false news by MIT: <https://www.science.org/doi/10.1126/science.aap9559>

FaceBook check: <https://about.fb.com/wp-content/uploads/2018/10/fake-news-trends.pdf>



THE PROBLEM

WHAT HAPPENS IN SOCIAL MEDIA IF

- News / Tweet deletion
- Labeling
- Account locks

NEWSROOM GOALS

We're committed to fight the spread of misinformation online, and promote plurality, through explainable AI and community engagement.

Our goals are to:

- Fight the spread of disinformation and hate speech;
- Highlight misinformation and bias; and
- Open up echo-chambers by providing a balanced view on the news;

- More data on another Latinamerican countries.
- More state-of-the art models implemented with different hyperparameters.
- Data labelling opened to all spanish speakers with further statistical sampling.
- Correction framework for Political Analyst to use.
- Web Scraping automation and integration for further topic modeling analysis and clustering.
- Plug-in creation for fake news detector in web browsers like Google Chrome.



DATA COLLECTION

ANNOTATIONS



**How will we define that a web article
or a tweet is true or fake?**



MODELLING



LIVE DEMO



FUTURE WORK