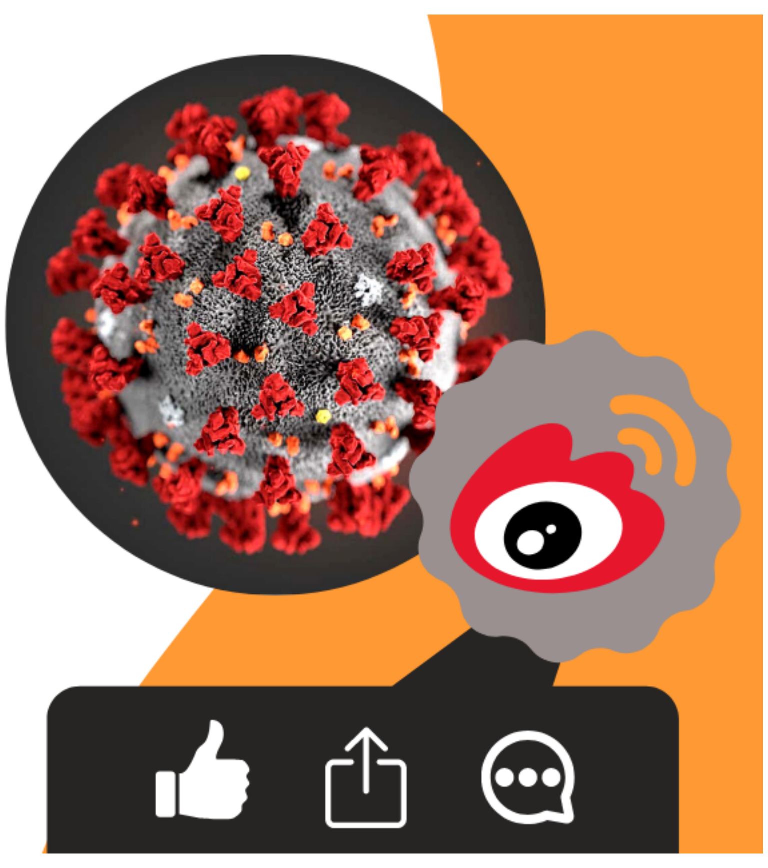
CAN WE PREDICT WHICH COVID-19 MICROBLOGS ARE "FAKE" ON WEIBO?

能不能预测新冠肺炎的假微博?

BY KATHERINE MOOSE HUNT



Overview

Social media as a primary news source increased dramatically during the COVID19 pandemic relative to prior years. However, growing evidence suggests social media platforms are also at the crux of misinformation and disinformation about the COVID19 pandemic. While research has been conducted on popular US-based social media platforms about fake COVID19 posts, little research has been done on China's most popular social media platform, Weibo. What can policymakers and technology administrators do to stop the spread of misinformation and disinformation about COVID19 on social media, including on platforms like Weibo? One method is deploying machine learning models to predict which posts are fake.

A team from Syracuse University researched fake Weibo microblogs about COVID19 on Weibo. Then, the team published their source code and datasets on GitHub for further research. This project adapated their code to answer three research questions.

This project aims to explore three research questions:

1. How are fake and real COVID19 microblogs' in Chinese different on Weibo?
2. For fake COVID19 microblogs, which Chinese words are more or less likely to have a greater reach (e.g., # of likes and shares) on Weibo?
3. Are the original researchers' predictive models reproducable?

Summary of Findings and Implications

Disaggregating fake data from the total dataset surfaced surprising differences. For example, the distribution of keywords and reach was different for fake versus real microblogs. Further, a smaller subset of fake microblogs had a higher reach. Higher reach fake microblogs could be prioritized for administrative review if we can predict 1. which microblogs are fake and 2. which fake microblogs have a greater liklihood of a higher reach to spread disinformation and misinformation.

Researchers originally used 5 neural-network-based machine learning methods: FastText, TextCNN, TextRNN, Attention-based TextRNN (Att-TextRNN), and Transformer. This project replicated the Macro F1 scores for each model, validating these models likely predict fake news with some accuracy. However, the project also recommends running predictive models with additional Chinese data and a new comparison dataset from a US-based social media platform.

Analysis

How are fake and real COVID19 microblogs' in Chinese different on Weibo?

After replicating the original exploratory data analysis, this paper disggregated fake microblogs for additional analysis. The project surfaced peaks and valleys in keywords' frequencies emerged when fake COVID19 microblogs were disaggregated from the total dataset. This finding implies additional research to better understand the significance of differences real and fake microblogs keywords and other terms.For example, we could conduct further natural language processing (NLP) analysis to learn if there are keywords more frequent in fake versus real microblogs.

Image 1: Distribution of Selected Keywords in All Microblogs (n = 2,104)

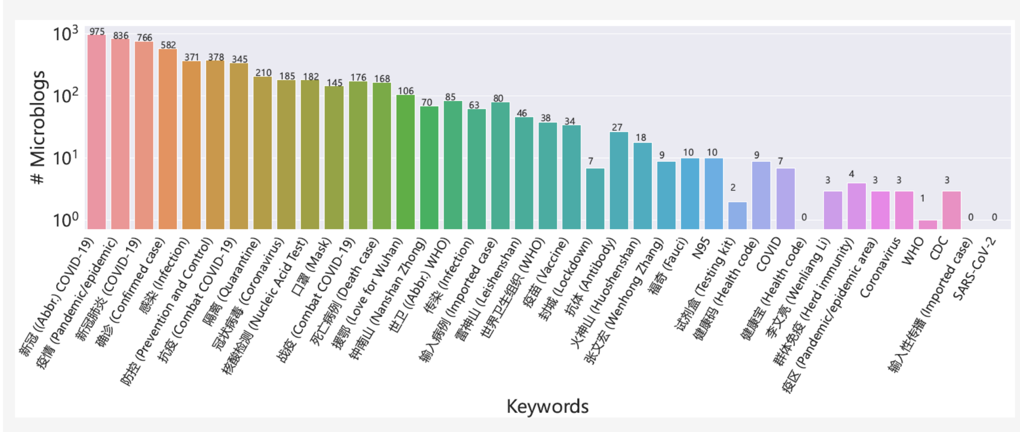
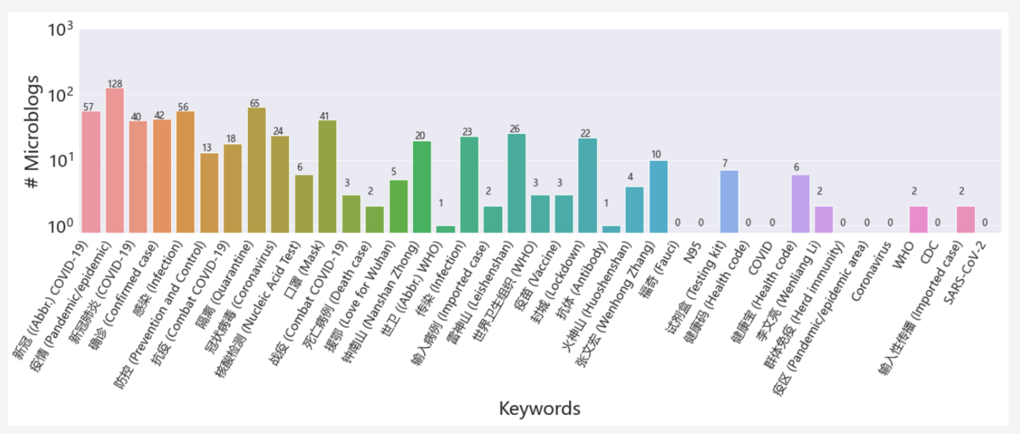


Image 2: Distribution of Selected Keywords in Fake Collected Microblogs (n = 344)



For fake COVID19 microblogs, which Chinese words are more or less likely to have a greater reach (e.g., # of likes and shares) on Weibo?

What do we mean by reach, and how are microblogs' reach distributed?

Overall, the aggregated data on microblogs have a power-law-like distribution. This means most microblogs have a smaller reach (e.g. likes, shares, comments) and there is a consistently declining number of microblogs with greater reach. However, disaggregating fake microblogs did not result in the same reach distribution - there are a small number of inconsistent outliers with much greater reach.

Image 3: Distribution of Comments in CHECKED Fake Microblog Data

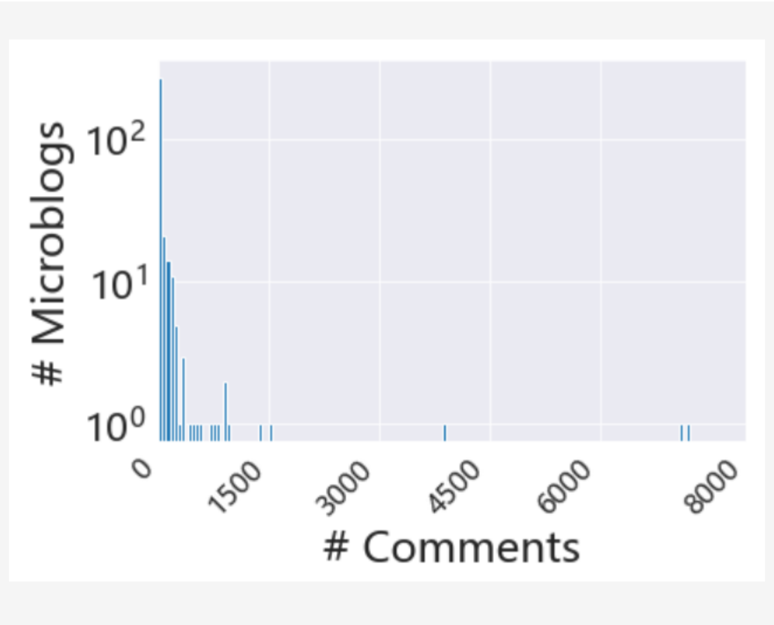


Image 4: Distribution of Likes in CHECKED Fake Microblog Data

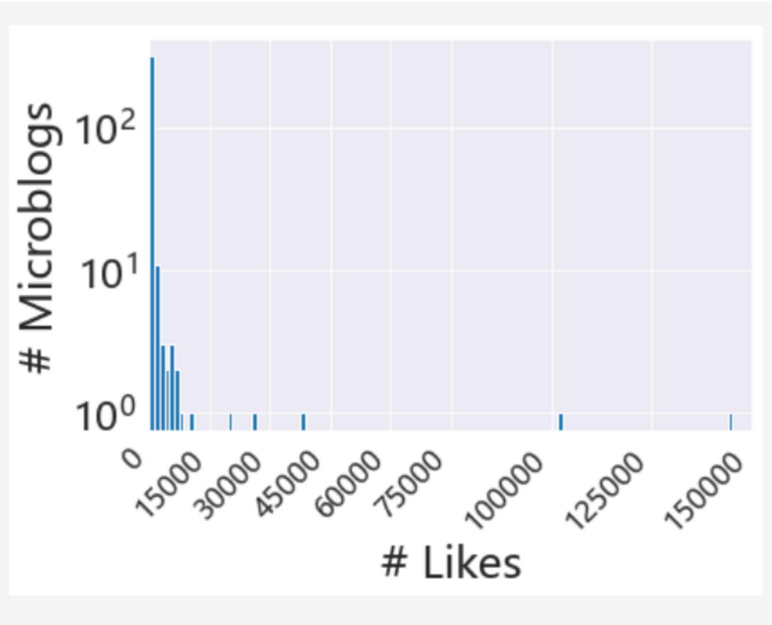
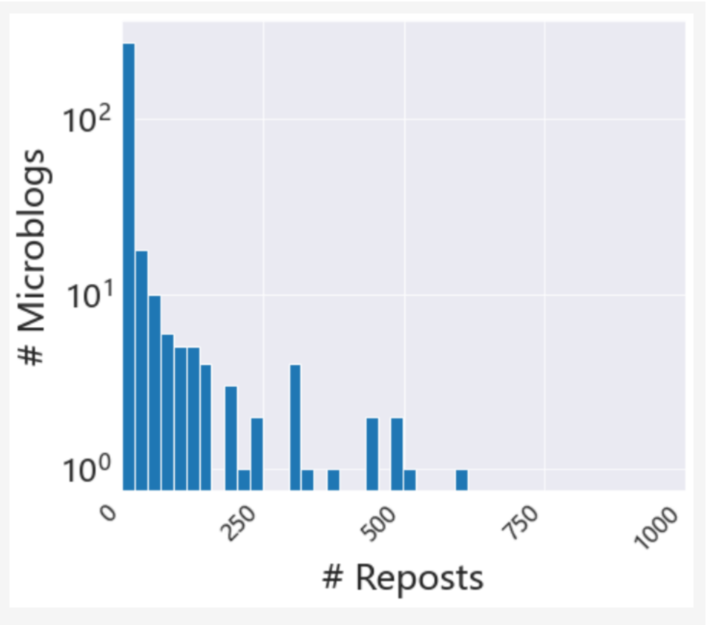


Image 5: Distribution of Reposts in CHECKED Fake Microblog Data



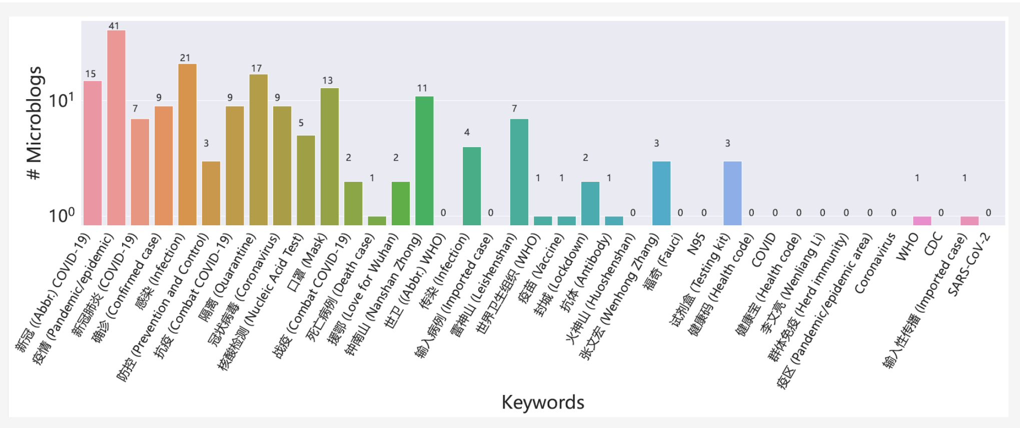
Which microblogs have a greater reach and what are their keywords?

This project divided fake microblog's reach into low and high. High reach fake microblogs represent the 75% quartile with over 36 comments, 14 reposts, and/or 128 likes. High reach fake microblogs surfaced additional peaks and valleys in keywords' frequencies. Additional research is recommended to better understand how keywords with higher reach are influenced by time of publication, additional non-keyword NLP analysis, and other key variables.

Table 1: Statistics of CHECKED "Fake" Microblog Data

Fake COVID-19 Microblogs on Weibo			
	# of Comments	# of Reposts	# of Likes
Count	344	344	344
Mean	110	162	1,294
Std	601	1,625	9,953
Min	0	0	0
25%	2	0	2
50%	10	2	17
75%	36	14	128
Max	7,224	27,199	145,108

Image 6: Distribution of Selected Keywords in Highest Reach "Fake" Microblogs (n = 112)



Are the original researchers' predictive models reproducable?

Researchers originally used 5 neural-network-based methods: FastText, TextCNN, TextRNN, Attention-based TextRNN (Att-TextRNN), and Transformer. This paper used the original research parameters to divided data by posting date for training, validation, and testing with a 70%:10%:20% split. This project reproduced benchmark results for each model below to test their predictive power with Macro F1 scores. Macro F1 scores were used to evaluate the number of times the model correctly predicted fake news. In short, Macro F1 scores close to 1 can be interpreted as correctly predicting fake news. Given the consistency in Macro F1 scores from the original and reproduced projects, this project recommends running models with additional Chinese data and a new US-based social media comparison dataset to test the models' predictive power with additional data.

Table 2: Benchmark Results using CHECKED data to Detect Fake Microblogs

Benchmark Results using CHECKED data to Detect Fake News					
	FastText	TextCNN	TextRNN	Att-TextRNN	Transformer
Reproduced Macro F_1	0.839	0.904	0.674	0.871	0.927
Original Macro F_1	0.839	0.938	0.700	0.871	0.927

Troubleshooting Code

Adapting the original code to run in the computer's environment required some troubleshooting. Two examples are outlined below .

Enabling Chinese Text Segmentation

Analysis code produced an error, "[ModuleNotFoundError: No module named 'jieba'], even after installing the jieba package.

In [1]:
```
# install the conda package for jieba per https://stackoverflow.com/questions/57887947/python-3-cannot-find-a-module
#import sys
#!conda install -c conda-forge jieba
```

Chinese Font Updates

Most operating systems come with standard Chinese fonts. However, Chinese text segmentation for the NLP library used in this project required MicrosoftYaHei font style, which should not be assumed as enabled in the local environment. The project developed a sanity test to confirm font availability. Several tutorials were used to construct the code, including https://stackoverflow.com/questions/42097053/matplotlib-cannot-find-basic-fonts.

In [2]:
```
#import matplotlib.font_manager

#sanity tests to ensure Chinese char enabeled font is rebuilt into the Matplotlib Cache, but code will need to be modified to insert the correct fi
#matplotlib.font_manager.FontManager().addfont(path='<INSERT FILE PATH TO ANNOCONDA INSTANCE\\Lib\\site-packages\\matplotlib\\mpl-data\\fonts\\ttf\\
#[f for f in matplotlib.font_manager.fontManager.ttflist if 'MicrosoftYaHei' in f.name]
#matplotlib.font_manager.findfont('Microsoft YaHei', fontext='ttf', rebuild_if_missing = True)
```

Conclusions

The TextCNN, Attention-based TextRNN, FastText, and Transformer predictive power (i.e. Macro F1 scores) implies these models can predict fake microblogs with some accuracy. However, a larger fake COVID19 data set is needed to understand if there is a statistical difference between real and fake microblogs' impact (e.g. # of comments, likes, and reposts) and addintional NLP analysis on non-keywords. Specifically, explore high reach fake microblogs and identify a US-based comparison dataset to continue refining prediction models.

Reference

{yang2020checked, title={CHECKED: Chinese COVID-19 Fake News Dataset}, author={Yang, Chen and Zhou, Xinyi and Zafarani, Reza}, journal={Social Network Analysis and Mining (SNAM)}, doi={10.1007/s13278-021-00766-8} year={2021} }

In [ ]: