Does Mainstream Media Recognize the Surge in Hate Crimes Against Asian Americans?

— Text Analysis using Twitter Data

APSTA-GE 2047 Messy Data and Machine Learning

Instructor: Ravi Shroff

Reported by
Junhui Yang
Ziyang Fu
Le Win

Table of Content

I. INTRODUCTION

Between March 2020 and February 2021, STOP AAPI Hate, a reporting center launched in response to the rising xenophobia against Asian Americans, has received nearly 3,800 reported[1] incidents of racism and discrimination against Asian-Americans. From racial slurs to violent assaults, these firsthand reports only represent a fraction of the number of AAPI hate incidents that occured. And biased media coverage isn't helping the AAPI community.

On March 16 of this year, eight people lost their lives after a man went on a shooting rampage at a spa in Atlanta[2]. Six of those people were Asian American women. News coverage of the incident attempted to whitewash it, with law enforcement agencies hesitating to label it as a racially motivated crime. Some news outlets went so far to point out the shooter's faith, highlighting his involvement in a Bible club and his quiet nature. This type of coverage perpetuates the narrative that it's shocking for a white man to commit such crimes.

News medias' flawed coverage of the Atlanta shootings highlights the industry's failure to cover the surge in anti-Asian hate crimes over the past year, further putting the pain and violence felt by the Asian American communities on edge. Downplaying the Atlanta shooting incident further highlights their disconnect to the racial communities and their lack of understanding of the long-standing struggles Asian Americans have faced since coming to the United States two centuries ago.

The goal of our study is to analyze the top 6 U.S. news media coverage on anti-Asian hate crimes and understand their efforts in combating against the rising tide of coronavirus-related discriminations and violence against Asian Americans. We first examined two research works related to Twitter sentiment analysis. Using some of the techniques mentioned in these studies, we performed some exploratory analysis on the Twitter data collected between March 2020 and March 2021 and trained two supervised machine learning models to classify whether the tweet is related to AAPI hate crime or not. We then discussed the results from these models and concluded the paper with some recommendations for future analysis related to this topic.

II. LITERATURE REVIEW

Due to the frequent racial discrimination incidents in recent years, more and more research have been conducted on the social media texts related to such incidents. Kummar and Pranesh (2020) used the Black Lives Matter[3] movement (BLM) related data collected from Twitter and

[1] https://stopaapihate.org/reports/
[2] https://www.nytimes.com/live/2021/03/17/us/shooting-atlanta-acworth
[3] More information about Black Lives Matter movement, please visit: https://en.wikipedia.org/wiki/Black_Lives_Matter#2020

performed various classification models including random forest as an attempt to develop a hate speech detection system. From the dataset, they extracted the top 10 most frequent unigrams, bi-grams, and tri-grams and the top 5 most used hashtags as features. The experiment results show that the application of their approach can be potentially used to identify and filter hateful speech contents on social media platforms. Badaoui (2020) also used Black Lives Matter movement-related data and analyzed tweets that used #BlackLivesMatter and #BLM hashtags. They observed the daily use of the hashtags and analyzed the tweet contents to check for other hashtags and names, such as "Trump", mentioned with these tweets. They found that using this way they were able to uncover the catalyst behind the surge in the usage of these hashtags and highlight the growing frustration from the Black community related to police brutality, violence, and fear.

Similar to these two studies, we used Twitter data to study news media reports on anti-Asian hate crimes, a pressing issue that has dramatically increased since the beginning of the pandemic. We applied a similar feature selection technique as Kummar and Pranesh (2020) by extracting the top 10 most frequent uni- and bi-grams from the dataset, and similar to Badaoui (2020), we analyzed the surge in the coverage of this issue by U.S. news medias in certain month, i.e. March 2021.

III. DATA

A. Obtaining and Processing the Data

In order to get enough tweets that report anti-Asian hate crimes and share resources to help the AAPI community, we decided to look at tweets of six local Asian news media for the training set. Because these six accounts had 588 tweets within the one year timeline we are examining, we can manage to manually label the tweets as related to AAPI hate crime issues or not. As such, we decided to use these tweets as the training set. The six accounts are NextShark, AAPI Data, Stop AAPI Hate, CAA San Francisco, Asian American Legal, and WashTheHate, and they all mainly focus on reporting news related to Asian Americans. As for the testing set, we selected six major U.S. news media that are most-followed on Twitter as of March 2021. These six accounts are CNN, the New York Times, Fox News, the Wall Street Journal, TIME, and the Washington Post. Using the 'academictwitteR'[4] package, we pulled the tweets of these twelve accounts that were made between March 17, 2020 and March 30, 2021.

B. Exploring the Data

Our pre-modeling hypothesis is that there are not many tweets that reported AAPI hate crime from those six news accounts until the Atlanta shooting in March 2021. Assuming that tweets

[4] 'academictwitteR' package: https://gist.github.com/schochastics/1ff42c0211916d73fc98ba8ad0dcb261

containing the word "Asian" or "Asians" are likely to be related to anti-asian hate crime reports, we examined how many times this word appears in the tweets of the six news accounts (Figure 1). The usage of "Asian(s)" in the tweets exponentially increased in March 2021. This somewhat confirms our hypothesis on the media's little attention in regards to several countless incidents that have happened to the Asian community in the US before the Atlanta shooting.

In the plot, we pointed out hate crimes including the incidents of a Chinese restaurant got vandalized back in June 2020 in New Jersey,[5] an Asian woman physically assaulted by a group of teenagers in April 2020 in New Jersey[5], a Filipino American man was slashed in the face on the Subway in February 2021 in New York[5], and not to even mention an 84-year-old immigrant from Thailand violently was shoved to the ground in January 2021 who died later in San Francisco[5]. These media accounts started putting the spotlight on the AAPI community after the Atlanta shooting in March 2021 which resulted in 8 people dead. For the later prediction, we can foresee the result of an extremely small portion of tweets that reported the AAPI related hate incidents.

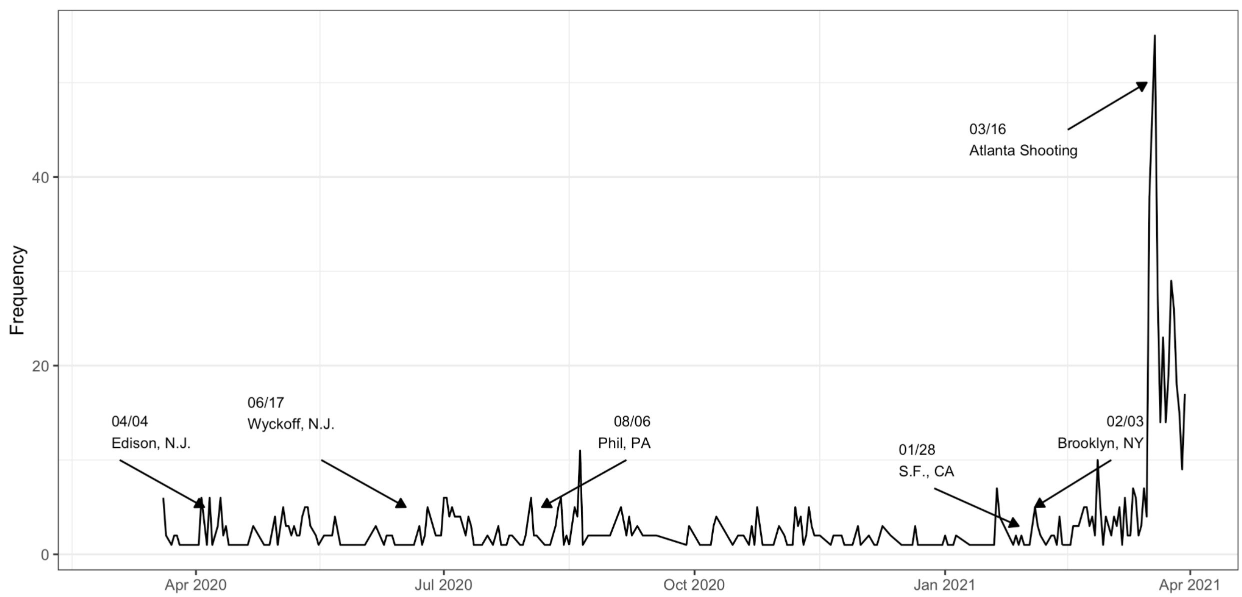

Figure 1: The Number of Times the Tweets Containing "Asian(s)" Appeared over Time

IV. METHODS AND RESULTS

In this section, we built machine learning models to predict whether a tweet is related to anti-asian hate crime or not. Using the 588 tweets collected from the six Asian News media as the training set, we manually labeled these tweets, in which the indicator serves as the response

[5] More details about Anti-Asian Violence, please visit: https://www.nytimes.com/interactive/2021/04/03/us/anti-asian-attacks.html

variable in the modelings. Of the tweets that are related to AAPI hate crimes, we extracted the top 10 most frequently used uni- and bi-grams from these tweets as binary features.

Since the response variable is a binary variable with 1 indicating a tweet related to Asian hate crimes and 0 indicating not related, logistic regression is the appropriate method for the analysis. However, as our predictors for the model are unigrams and bigrams, it may lead to our model suffering from overfitting — a model fits the training data well, but does not generalize well to a new set of data. Therefore, implementing a penalized logistic regression to perform feature selection is necessary in order to obtain more accurate predictions with many covariates. Lasso regression achieves this by automatically deleting unnecessary covariates and hence facilitates easier interpretation for our problem.

The second model we used is Random Forest, which uses trees and is less prone to overfitting. This model can tell us how much each feature contributes to class prediction using feature importance for better interpretation. These properties are desired for our study with many predictors. As such, we built two machine learning models, Logistic Lasso Regression and Random Forest, using the training dataset. Then, using the 270,000 tweets from the top six national news accounts, we determined the best performed model as our ground true model.

Within the models, the response variable, "AAPI_related_indicator", is a binary variable indicating whether the tweet is related to Asian Hate crimes or not. Predictor variables are 10 unigrams: "racist", "aapi", "report", "stopaapihate", "community", "incident(s)", "racism", "asian(s)", "china", "hate" and 10 bigrams: "asian american(s)", "aapi hate", "anti asian(s)", "hate crime(s)", "anti aapi", "china virus", "asian hate", "asian racism", "hate incident(s)", "aapi community".

A. Logistic Lasso Regression

We used the glmnet package to fit the lasso regression model. The penalty is controlled by λ, and we found the amount of penalty λ by 10-fold cross-validation, which uses misclassification error as the criterion. Based on the rule of minimum mean cross-validated error, we selected the $\lambda = 0.00978$ to estimate the coefficients of the predictors. Because the news media remained neutral in reporting the issue, the model deemed some of the bigrams including "hate_incident(s)", "anti_aapi", and "anti_asian(s)" as unnecessary features (Table 1 in Appendix). Most weights are put on "racist", "china virus", and "incident(s)". The "china virus" was unsurprisingly referenced by some of the news media due to Former President Trump's insensitive tweets.

B. Random Forest

We built the random forest model using the ranger package. We tuned a few of the hyperparameters of the model using grid search. The hyperparameters included 'mtry', the number of variables to randomly sample as candidates at each split, and 'min.node.size', the minimum number of samples within the terminal nodes. We created a grid and looped through each hyperparameter combination to evaluate the model by out-of-bag (OOB) root mean square error (RMSE).

We provided the mtry 10 values evenly spaced across the range from 2 to p, where p = 20 in our study. The minimum number of samples within the terminal nodes controls the complexity of the trees. Smaller node size allows for deeper, more complex trees and the larger node size results in shallower trees. Since the default min.node.size = 10 for probability forest, we provided a range of 3 to 10 for the parameters to tune on. Finally, we searched across 40 different models with varying mtry and minimum node size. In the models, the number of trees is 1000 and we sampled with replacement and set a random number generator seed which allows us to consistently sample the same observations for each sample size and make it more clear the impact that each change makes.

Per Table 2 in the appendix, the OOB RMSE ranges between 0.234 and 0.237. Our top 10 performing models have RMSE values around 0.235 to 0.237 and the results show that models with deeper trees (3-9 observations in a terminal node) and 6 to 8 mtry values perform the best. The best random forest model we have found uses a mtry of 6 and terminal node size of 7 observations. We set importance as 'impurity' in the above modeling, which allows us to assess variable importance. Variable importance is measured by recording the decrease in MSE each time a variable is used as a node split in a tree (Figure 2).

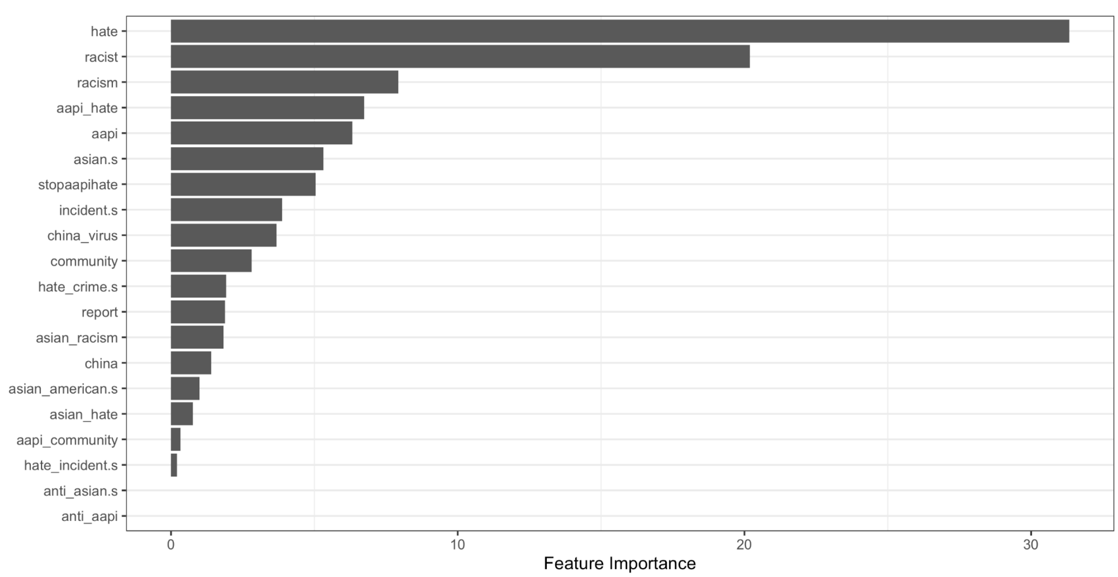

Figure 2: Random Forest model's Feature Importance of the Predictors

C. Model Selection

After we fitted the two classification models, we applied them on the testing data to predict the probability of the tweet being related to Asain Hate crimes. Using a threshold 0.5, we obtained the predicted class 0 or 1 for each tweet. We then arranged the predicted probability in descending order and subset the top 100. We manually labeled the top 100 tweets from each model to indicate whether they are actually related to AAPI hate crimes or not. The result showed that the random forest model correctly predicted 70 out of the 100 tweets while the logistic lasso regression correctly predicted 52 out of the 100 tweets. Therefore, we selected the random forest model as the better performing classifier.

V. DISCUSSION

A. Result Implications

The model result shows that out of the 270,000 tweets that were made by the top 6 U.S. news accounts, only 1717 (0.6%) of them are potentially related to Asian hate crimes. In examining the proportion of each news media's tweets that are predicted as tweets related to the AAPI issue, we see that the average monthly coverage ranges between 0% and 2% except for June 2020 and March 2021 (Figure 3).

In June 2020, the percentage of coverage dramatically increased for these news media to about 3% to 6%. However, most of these tweets were related to the Black Lives Matter movement that was ignited by the many tragic events that resulted from police violence. This proves to our main research question that the mainstream news media have extremely low coverage of anti-Asian hate crimes. Additionally, because some of the model features include "racism", "hate", and "community", tweets that are related to these types of content were included within the pool of potential AAPI related tweets. As such, we see that tweets related to BLM movements and ones discussed the overarching issue of racism in America and police violence in certain communities.

The model also confirms our pre-modeling hypothesis on the potential increase in news medias' coverage on anti-Asian hate crime only after the Atlanta shooting event. Per Figure 3, the increase in March 2021 was small compared to that of June 2020. The Atlanta shooting event may have brought the hate crime issue to the forefront of many media outlets, but this again proves the alarming state of media poor coverage on this prevalent issue.

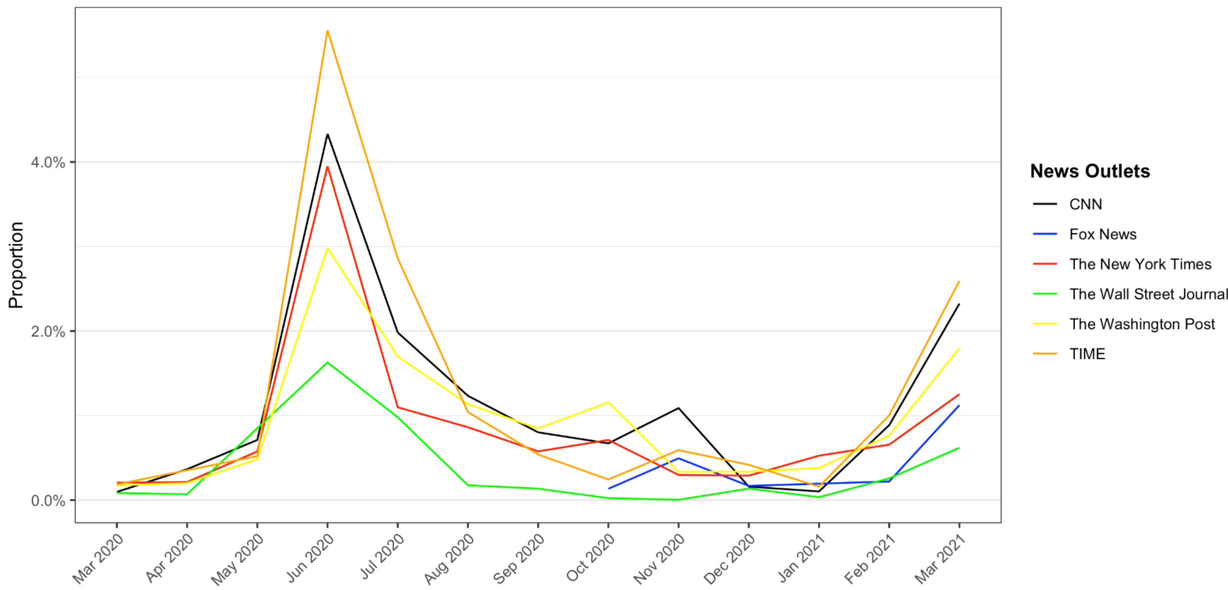

Figure 3: Proportion of Predicted Relevant Tweets by News Outlet over Time

Furthermore, we demonstrated the tweets that are predicted as AAPI hate crime related in two visualizations, frequency bar chart and word cloud. Figure 4 shows the top 5 frequently used words by each news outlet on the tweets that are predicted as AAPI hate crime related. The two of the most used words in these tweets are “racism” and “racist” for all the six news oulets. Overall, based on the counts of the words, CNN has the most tweets that are potentially related to AAPI hate crimes. On the other hand, because TIME was the only news outlet that used the word “anti” and “asian” as its top frequently used word, we conclude that TIME gives more attention to the Aisan community related to hate crimes compared to the other five news outlets. For the remaining news outlets except Fox News, we noted many of them include “black” and “police" as one of the most used words. Fox News on the flip side gives the least attention to the Aisan community related to hate crimes.

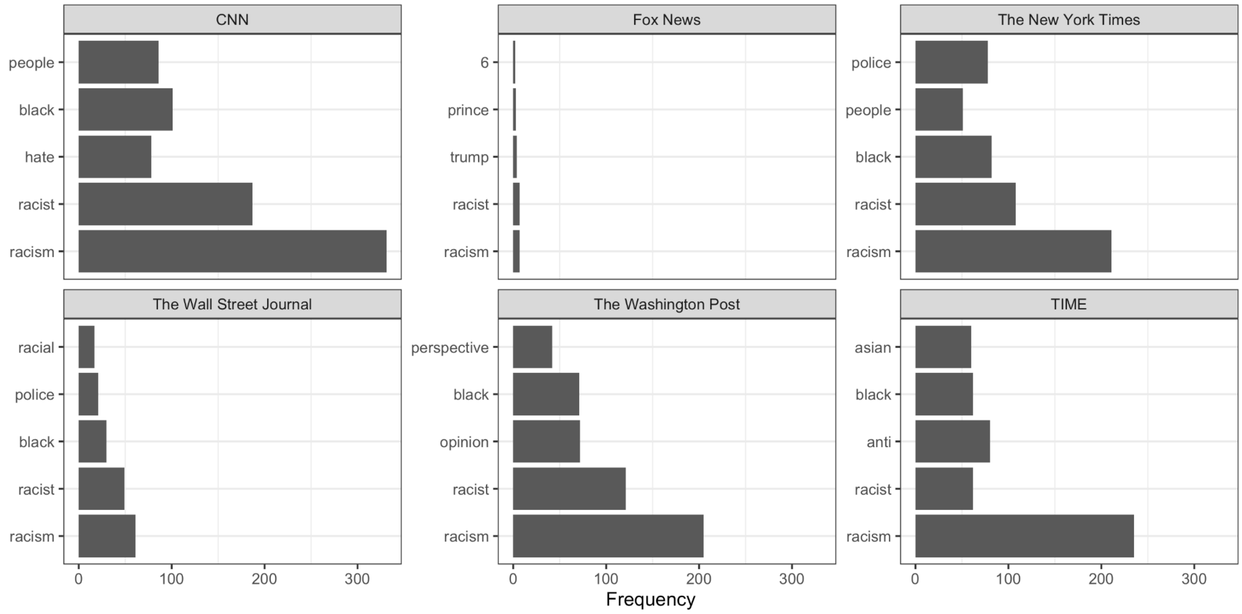

Figure 4: Five Most Used Words From the Tweets predicted as Anti-Asian Hate Crime Related by the News Media

Figure 5 is composed of the top 200 words most used in the tweets predicted as related to AAPI hate crimes, in which the size of each word indicates its frequency or importance. The most used word is “racism”. As noted before, many of the tweets are related to the BLM movements containing words like “black” and “police”. Other words like “white”, “president”, “anti”, and “hate” could be used in tweets that talk about the BLM movement and anti-asian hate crimes. As discussed previously, such discrepancy in the prediction result and the reality is inevitable when these two events that are targeted both minorities about the racism topic in a text analysis.

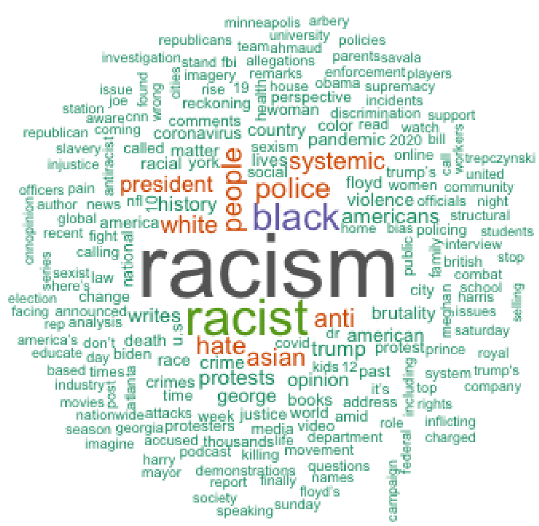

Figure 5: Top 200 Most Frequently Used Words in Tweets predicted as related to reporting of AAPI hate crimes

B. Result Limitations

Although there is a discrepancy between the predicted results and the content of the tweets, the random forest classifier performs well in terms of answering our research question and preliminary hypothesis. Regardless, the analysis has limitations. First, because it's time consuming to manually label all the tweets in the testing set, we used the top 100 tweets with the highest probability of them having anti-Asian hate related content to select the better performing of the two models. As such, we only have a slight inclination on how well the random forest model performs on the testing set. If we have enough time, we would label all the tweets in the testing dataset and also increase the number of tweets in the training set. Moreover, we can use an advanced language processing model to build the desired tweet classifier. As such, there are still rooms for improvement in terms of the tweet dataset and classification model.

VI. CONCLUSION

In this study, we used Twitter data to analyze the top 6 U.S. news media coverage on anti-Asian hate crimes by applying a random forest classifier, which was selected out of the two ML models based on the prediction accuracy on a subset of the testing set. From the model result, we found that between March 2020 and March 2021, these media in total have less than 1% of their tweets that potentially talked about anti-Asian hate crimes. This emphasized the minimal role news media have played in combating the surge in discriminations and violence against Asian Americans.

From our findings, we noted that there was a slight increase in media coverage of AAPI hate crimes after the Atlanta shooting event. Prior to this incident, the news media was silent in reporting hate crimes against Asian Americans despite the alarming rate of incidents, some of which resulted in death. This silence in return explains why most of us don't hear about anti-Asian hate crimes on the mainstream news media as often as the rate of which it has increased over the past year. By underreporting a prominent issue that had been affecting the livelihood of Asian Americans proved that overall media's flawed coverage of racialized communities, which remains invisible in the media.

Following this study, we suggest further research around hate speech against minorities on social media. Through these studies, we hope to raise more awareness and resources to combat the racism targeting minority communities including Asian Americans.

REFERENCES

Badaoui, S. (2020) *Black Lives Matter: A New Perspective from Twitter Data Mining*

Kumar, S., Pranesh, R. R., Pandey, S. C. (2020) *TweetBLM: A Hate Speech Dataset and Analysis of Black Lives Matter-related Microblogs on Twitter*

Random Forests. (n.d.). UC Business Analytics R Programming Guide. Retrieved May 1, 2021, from https://uc-r.github.io/random_forests

APPENDIX

Table 1

Variable Name	Coefficient
(Intercept)	-3.72
racist	10.70
aapi	0.77
report	0.01
stopaapihate	-1.92
community	1.95
incident.s	8.48
racism	1.50
asian.s	2.29
china	1.17
hate	4.12
asian_american.s	-0.91
aapi_hate	6.33
anti_asian.s	0.00
hate_crime.s	-0.50
anti_aapi	0.00
china_virus	8.53
asian_hate	5.56
asian_racism	6.51
hate_incident.s	0.00
aapi_community	-3.33

Table 2

mtry	Node Size	OOB RMSE
6	7	0.2347
6	9	0.2348
6	5	0.2349
6	3	0.2350
8	9	0.2357
8	7	0.2359
8	5	0.2362
8	3	0.2364
10	9	0.2367
10	7	0.2371