



Classifying Climate Change Tweets

Joe Marx

Climate Tweets?



 **Katie Mack** @AstroKatie 15h
Honestly climate change scares the heck out of me and it makes me so sad to see what we're losing because of it.

  29  122 

 **Gary P Jackson (RAT)** @gary4205 7h
[@AstroKatie](#) Maybe you should learn some actual SCIENCE then, and stop listening to the criminals pushing the [#GlobalWarming](#) SCAM!

  7  

 **George Wallace**  @MrGeorgeWallace 
Global warming's for real y'all. Someday there won't be any rappers with ice in their names. Only water. Water-T. Water Cube. Vanilla Water.
12:37 PM · Apr 26, 2012 

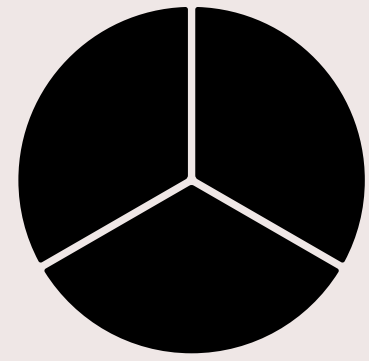
 174  251  Share this Tweet



Business Case

- Better understand the online conversation surrounding climate change
- See if fossil fuel emissions follow trend of sentiment
- Look at will of the people to design legislation to tackle climate change
- Target climate deniers with information that may persuade them to





Process



Data Collection

EDA

Modeling

Validation

Application



kaggle

Twint

Data Collection

Twarc



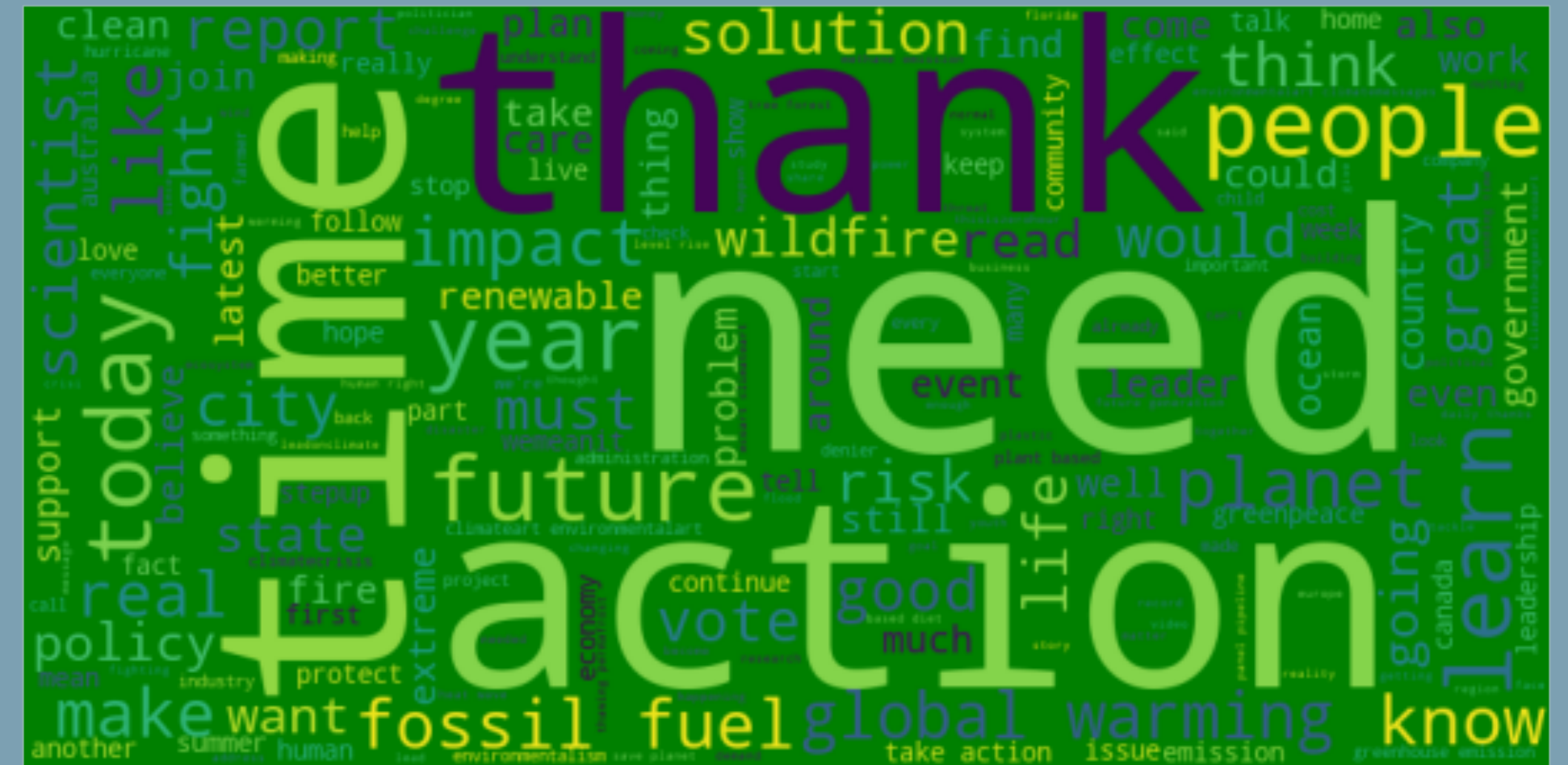
EDA

- Looked for classifying hashtags
- Class Distribution
- Class patterns

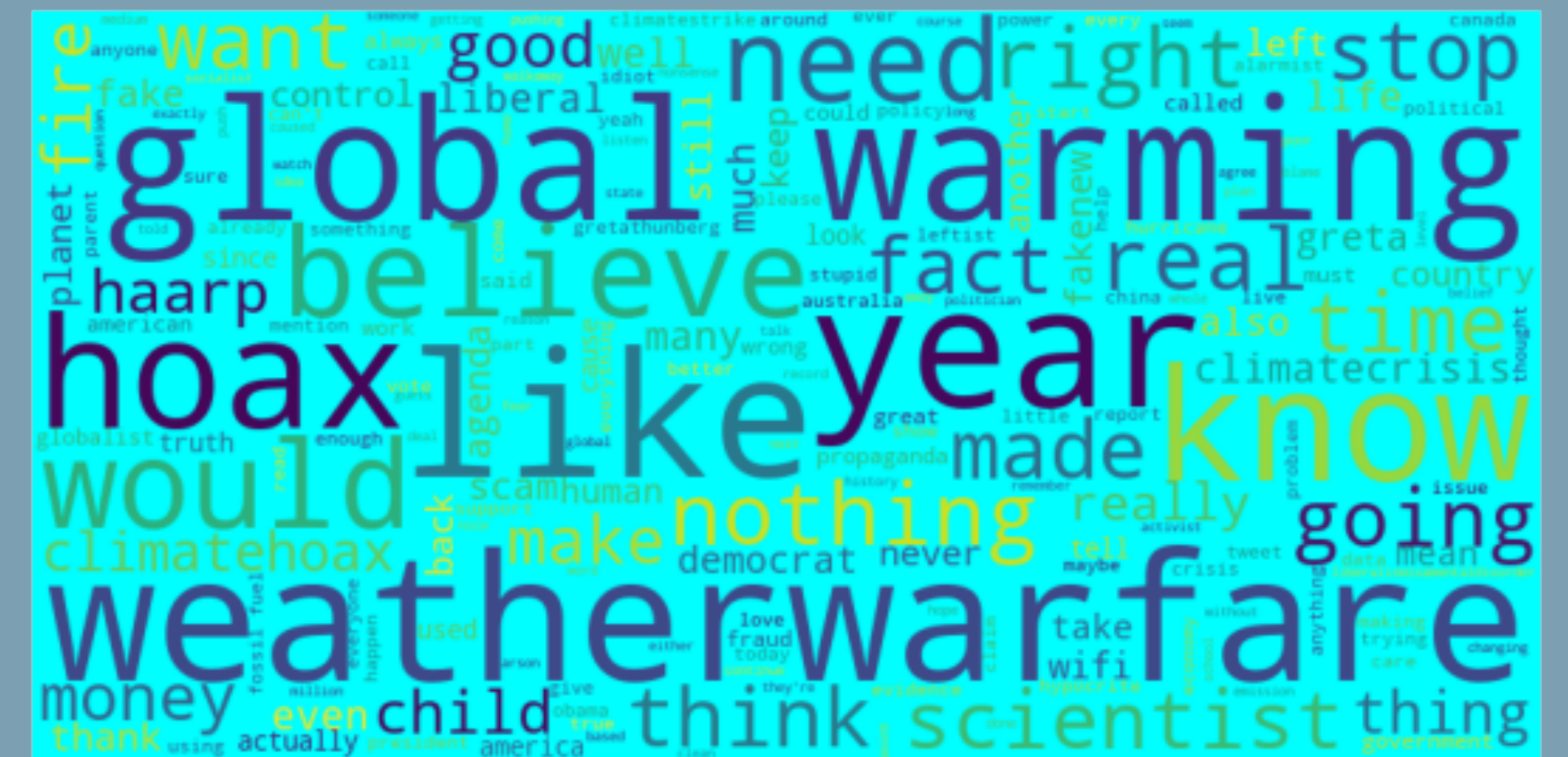
```
believer_tags = ['climatechangeisreal', 'actonclimate',
                 'extinctionrebellion', 'climateemergency',
                 'climateactionnow', 'capitalism', 'public_health',
                 'climateaction', 'humanityextinction',
                 'activism', 'noplanetb', 'savetheplanet']

denier_tags = ['climatechangeisfalse', 'climatechangenotreal',
               'climatechangehoax', 'globalwarminghoax',
               'tcot', 'ccot', 'tlot', 'pjnet', 'rednationrising',
               'voterred', 'libtard', 'libtards', 'maga']
```

Climate Believer Tweets



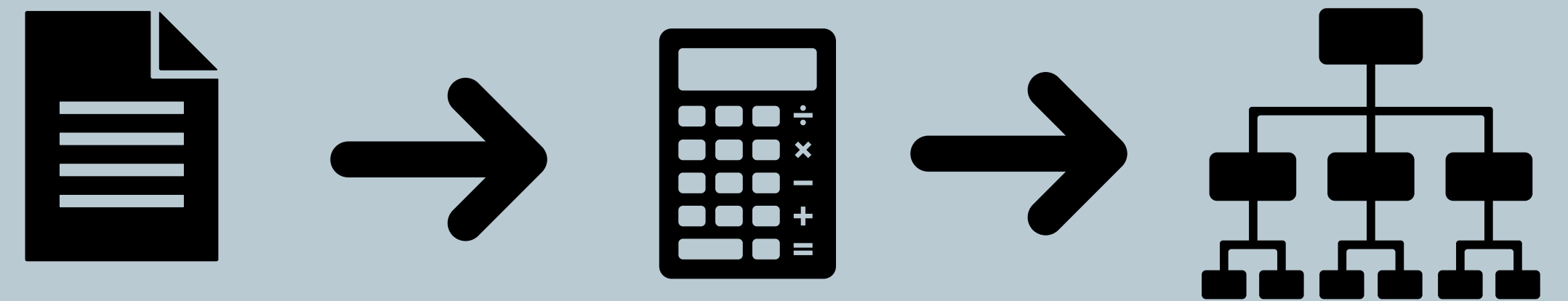
Climate Denier Tweets





Modeling

- Target variable engineered using hashtags
- Data prepared with regex, NLTK
- Vectorized with TF-IDF then passed into sci-kit learn classifiers using GridSearchCV
- Validated on separate pre-labeled dataset from Kaggle

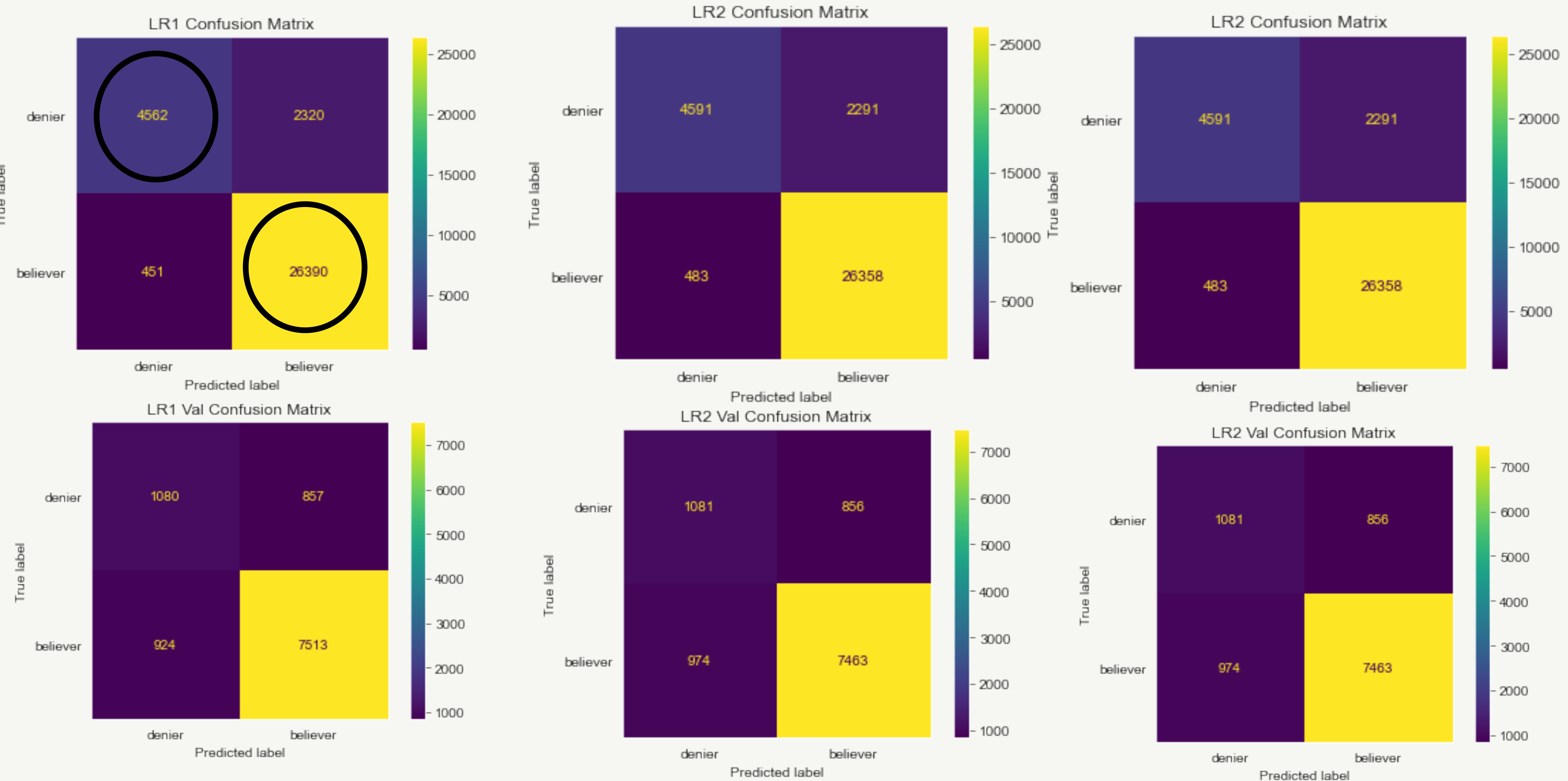




Results

Metrics	<i>Dataset</i>	Accuracy	F1 Score
LR1	Test	91.78%	95.01%
	Val	82.83%	89.4%
LR2	Test	91.77%	95.0%
	Val	82.36%	89.08%

Results



←

Thread

Dr. Jonathan Foley

@GlobalEcoGuy

Despite what some billionaires seem to think, climate change is a **time** problem, not a tech problem.

We already have the tech needed to solve the problem. We really do.

But what we don't have is time to wait or waste.

3:11 AM · Feb 18, 2021 · Twitter for iPhone

372 Retweets

23 Quote Tweets

1,311 Likes

Dr. Jonathan Foley

@GlobalEcoGuy · 2h

Replying to @GlobalEcoGuy

To clarify, I am glad that some technology R&D is happening for tools to scale existing solutions (making them cheaper, faster, better) and invent new ones.

But this should NOT become a reason to delay what we can do now, with the many effective solutions we have today.

1

4

43

Dr. Jonathan Foley

@GlobalEcoGuy · 2h

No amount of technology, short of a time machine, can make up for the time we have already wasted debating and delaying action on climate change. And any additional delay, no matter the reason, is very problematic.

1

7

47

Classify Climate Change Tweets

This crude web app uses a logistic regression model to classify a text as a beleiver or a denier of climate change

Enter the tweet text in the box below

Press Enter to apply

Source file changed.

Rerun

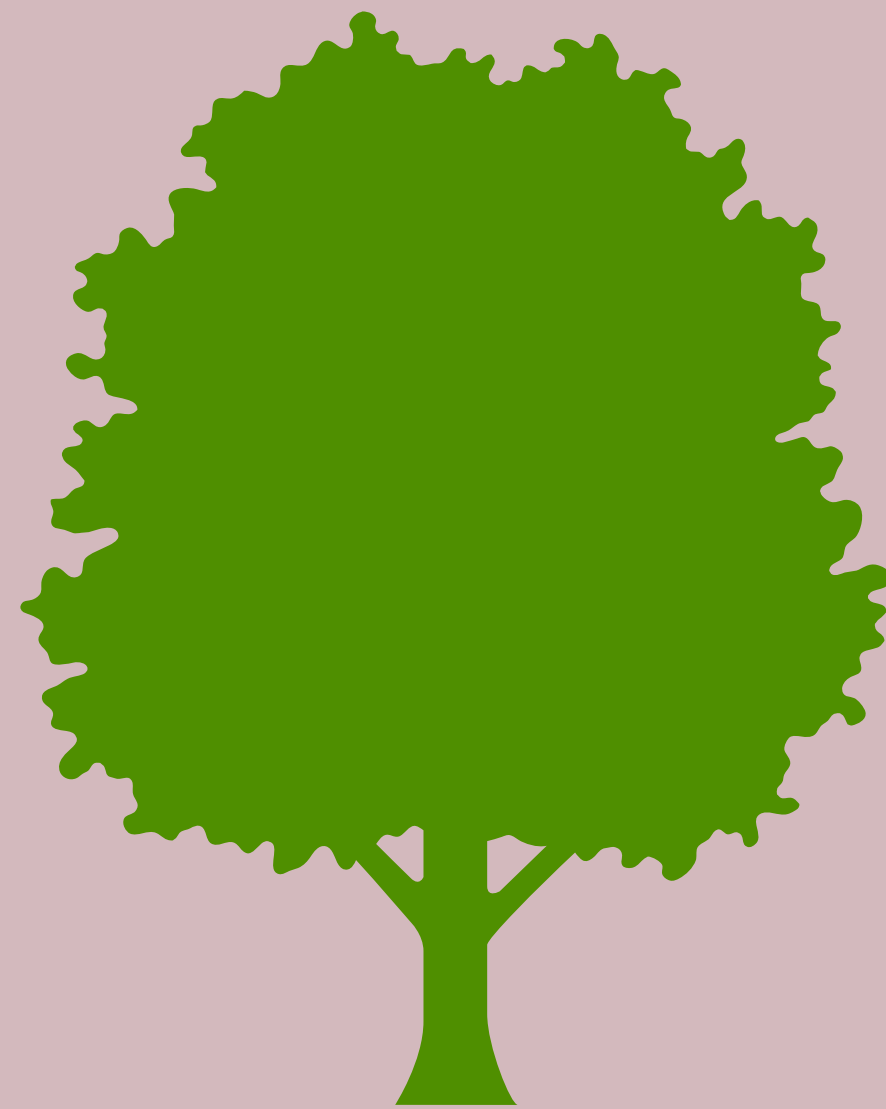
Always rerun

Next Steps

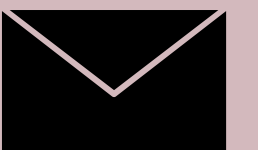
AND Shortcomings

- Robust Feature Engineering
- Analyze and work to reduce false positives
 - Utilize more of the GW dataset
 - Implement neural network model with transfer learning
- Use CO2 emissions API to look at emissions over time in relation to text sentiment
- Twitter no longer captures all right wing sentiment
- Hashtags are imperfect categorizers
- Twitter \neq Overall public sentiment

Thank You!



jmarx@hash.fyi



<http://linkedin.com/in/joe-marx-260a64102>



https://github.com/JoeBrowz/climate_prop

