Politicians you can Trust: Classifying Legislator COVID-19 Tweets for Reliability

Adam Bryant, Ashton DeRousse, Ben Klein, Timothy Krenz

Problem and Setting

Can we trust information from our politicians about COVID-19?

Is misinformation being spread on twitter?

Can an ML model be trained to find out who's telling the truth?

Congressman Greg Steube ✔ @RepGregSteube · Mar 25
Symptoms of the Coronavirus include: fever, cough, and shortness of breath. If you are experiencing any of these symptoms, visit the CDC's website to learn more about next steps:

Rep. Jim Jordan ✔
@Jim_Jordan

Fake News.

9:26 AM · Sep 4, 2020 · Twitter for iPhone

Since February, members of Congress have tweeted about Covid over 112,000 times

Dataset

- Scraped over 500,000 tweets from members of Congress
- Filtered down to 100,000 based on keywords related to COVID-19
- Multiple features related to demographics
  - Identify correlation between feature and reliability
- Output label: reliability score ranging from 0 to 1

```
1   wordbank = [
2       'corona',
3       'coronavirus',
4       'covid',
5       'covid19',
6       'covid-19',
7       'virus',
8       'pandemic',
9       'epidemic',
10      'lockdown',
11      'testing',
12      'cases',
13      'hospitalization',
14      'related deaths',
15      'mask'
16  ]
```

1	0	Caring for a loved one with #dementia is
1	0	We must address shortages of personal
1	0	I joined my colleagues in urging the inclu
1	0	RT @OHdeptofhealth: Most people who
1	0	If nothing else, this pandemic should ser
0	1	Here is my recent e-newsletter with thou
0	1	People showing up to work next week: –
1	1	As healthcare staff work tirelessly to trea
1	1	Congress has approved over $3 trillion in
1	1	Research shows that homeowners facin
1	1	Wearing a mask is one of the best and e
1	1	Safe and smart steps, including continui
1	1	I'm thankful that UPMC Hanover was ab
1	1	Step 3: Access to Affordable Testing W
1	1	Iowa's egg industry has seen a drastic d
1	1	Each and every one of us has a role to p
1	1	REMINDER: Tomorrow at 8:15am, we w
1	1	RT @WhipClyburn: Wear your mask.
1	1	Yesterday @USDA began accepting app
1	1	President Trump was warned in January
1	1	RT @ElPasoTXGov: An all-time high of

lastName	firstName	state	party	chamber	gender	birthday	raceEthnicity	pre2018incumbent	citizenPopulation	medianHouseholdIncome	TweetCongress	reliability	tweets
Abraham	Ralph	LA	1	-0.5	-0.5	24172	1.0	1.0	744070.0	36718.0	repabraham	0.868421052631579	86.0
Adams	Alma	NC	-1	-0.5	0.5	27206	-0.5	1.0	727330.0	59268.0	repadams	0.7346938775510203	186.0
Aderholt	Robert	AL	1	-0.5	-0.5	20210	1.0	1.0	664720.0	41110.0	robert_aderholt	0.8247422680412371	114.0
Aguilar	Pete	CA	-1	-0.5	-0.5	15130	0.0	1.0	650185.0	58668.0	reppeteaguilar	0.7659574468085106	116.0
Alexander	Lamar	TN	1	0.5	-0.5	29360	1.0	-0.33	6349395.0	48708.0	senalexander	0.8618421052631579	346.0
Allen	Rick	GA	1	-0.5	-0.5	25216	1.0	1.0	692875.0	43933.0	reprickallen	0.8854961832061069	146.0
Allred	Colin	TX	-1	-0.5	-0.5	13734	-0.5	-1.0	646990.0	70119.0	repcolinallred	0.907563025210084	390.0
Amash	Justin	MI	0	-0.5	-0.5	14826	1.0	1.0	702545.0	57926.0	justinamash	0.5862068965517241	41.0

Approaches

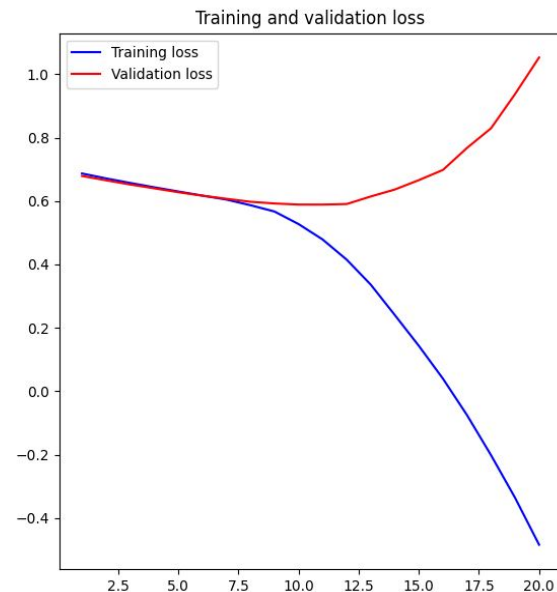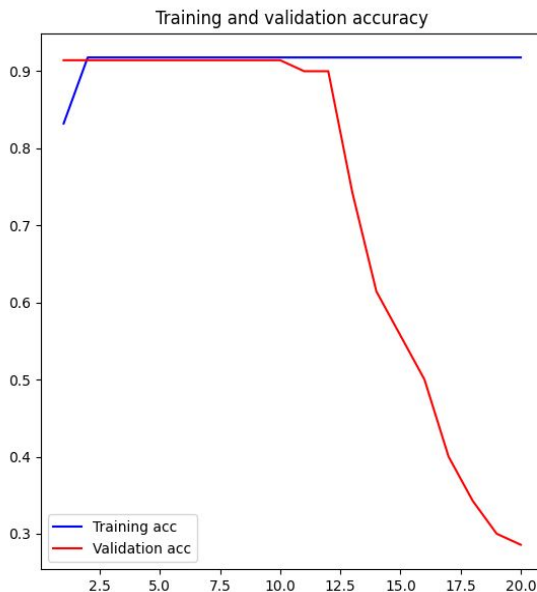Breaking down sentences: n-gram ranges (1,1) to (1,3)

Vectorizing: one-hot, tokenizer word embedding, counts

Reduce feature size:

Add a pooling layer

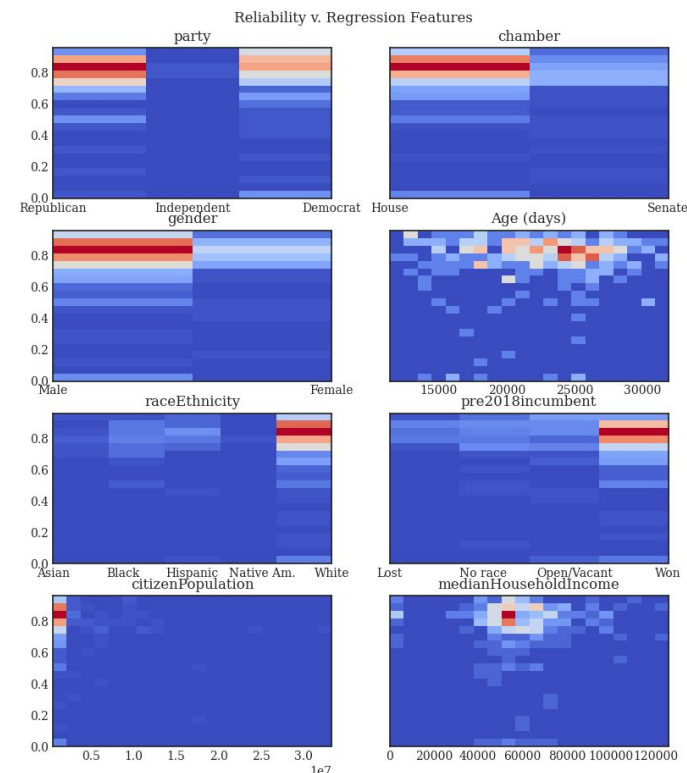Small dataset:

Split sets appropriately!

Findings and Analysis

Congress produced ~112,000 COVID-19 related tweets

Classification	Reliable	Unreliable
Number	~95,000	~17,000
Good Examples	"It's critical that we listen to experts about how to reopen schools safely. Classroom instruction is important, but our kids (and teachers) need to be safe in the classroom. That can only happen if and when we ramp up testing and actually *have a plan*. Some ideas https://t.co/7u2rPJkyKl" - @RepKatiePorter 2020-07-11 16:23:28	"@luke_mo_ Nothing racist about where the virus came from (China), who lied about it (China), or who covered it up (China). If you read the story, you'll be more informed about what they've done to your fellow Arkansans and Americans due to their actions." - @RepRickCrawford 2020-03-18 21:01:21

Plenty of bad examples too!

Linear Regression applied to demographics with a determination coefficient value of… $R^2 = -0.051$

Future Work

Sentiment does **not** equate to facts!

"US Capitol coronavirus COVID19 update by CDC FDA NIH
*90% of all USA drug content comes from China
*USA challenge is still low risk
*NO USA quarantine helpers have COVID19
*USA testing kits more available
*USA has PERFECT zero fatality record
*COVID19 NOT man-made https://t.co/Y4QZNM1GEr""

- @RepMoBrooks, 2020-02-28 13:36:01

Reliable

SenatorBaldwin	1255508267847081986	2020-04-29 14:44:31	1	-1	RT @ChrisMurphyCT: We're months into this crisis & w
RepVeasey	1306686340549935109	2020-09-17 20:07:55	1	-1	In the middle of a pandemic, the Trump administration raise
RepScottPerry	1253447417246158848	2020-04-23 22:15:26	1	-1	@SpeakerPelosi delayed add'l funding for PPP for over 2 w
repblumenauer	1313575329529819142	2020-10-06 20:22:18	1	-1	Because of Trump's short-sighted decision, your state and lo
NydiaVelazquez	1268249410410078211	2020-06-03 18:33:17	1	-1	These stories are abhorrent. Our #HealthcareHeroes are e
PattyMurray	1315770469094629377	2020-10-12 21:45:00	1	-1	Republicans are jamming through their anti-health care SCC
SenSherrodBrown	1250454579805896704	2020-04-15 16:02:58	1	-1	President Trump has: - fired the White House pandemic resp
RepMichaelWaltz	1311318914169737216	2020-09-30 14:56:06	1	-1	When China threatens to withhold ingredients for our pharm
RepValDemings	1288593549139808257	2020-07-29 21:53:37	1	-1	We passed critical #covid relief over nine weeks ago. The G
SenJohnBarrasso	1248256377589501957	2020-04-09 14:28:06	1	-1	.@SpeakerPelosi is once again playing politics with much-n
RepTedYoho	1243336182219657219	2020-03-27 0:37:00	1	-1	Thank you @RepMcCaul for championing this effort. The w
RepMarkGreen	1308815883951919111	2020-09-23 17:09:58	1	-1	As a physician, I know the science is definitive. The @CDCg
RepRickCrawford	1286349290944770048	2020-07-23 17:15:44	1	-1	.@FBI Director Wray has stated they open one new counter
RepAndyBiggsAZ	1306053935053012992	2020-09-16 2:14:58	1	-1	90% of Arizona's COVID-19 reported positive PCR results o
RepDebDingell	1252696705931378690	2020-04-21 20:32:23	0	0	My weekly telephone town hall will be tomorrow at 5:30pm.
WhipClyburn	1260300100334768138	2020-05-12 20:05:33	0	0	RT @COVIDOversight: Tomorrow, May 13, at 3PM ET, Chai
RepNewhouse	1257718809361866758	2020-05-05 17:08:26	0	0	The public meeting I had planned to get feedback from local
RepDelBene	1285617018289496066	2020-07-21 16:45:57	0	0	RT if you agree: Reducing child poverty shouldn't be partisar
senrobportman	1312534784124485632	2020-10-03 23:27:32	0	0	These figures are heartbreaking & illustrate why I've be
RepJackBergman	954546647448985600	2018-01-20 2:46:31	0	0	RT @HouseGOP: @SenSchumer, are you willing to shut th
RepJohnKatko	987424804759310339	2018-04-20 20:16:33	0	0	Proud to work with @graceanapolitano and the Mental Heal
TomColeOK04	974388255943790592	2018-03-15 20:53:58	0	0	Today, I had the opportunity to listen to Secretary of Health a
NydiaVelazquez	1301196894987849728	2020-09-02 16:34:49	0	0	RT @NYCHealthSystem: As the #YearOfTheNurse continue
RepSpanberger	1251510795298648070	2020-04-18 14:00:00	0	0	Just under a week left to submit to the congressional art con

Needs more data: 540 manually-classified tweets
More people too!

Importantly:
Don't adhere to this as an authority