# Problem Statement

- U.S.A. Department of Health and Human Services would like to *identify COVID-19 vaccine online misinformation* and reduce its spread
- Creating algorithm to classify post as belonging to one of two subreddits: *pro-vaccine* or *vaccine debate*
- In vaccinated circles, what words/phrases are most popular?
  - Can *incorporate them* into future campaigns
- In unvaccinated circles, what words/phrases are most popular?
  - Can *debunk them* in future campaigns

The Race for a COVID-19 Vaccine

- In the U.S.A.:
  - **81,349,065** confirmed cases
  - **993,733** deaths
  - **66.62%** of population fully vaccinated
  - **Pfizer BioNTech, Moderna, and Janssen (J&J)** vaccines available
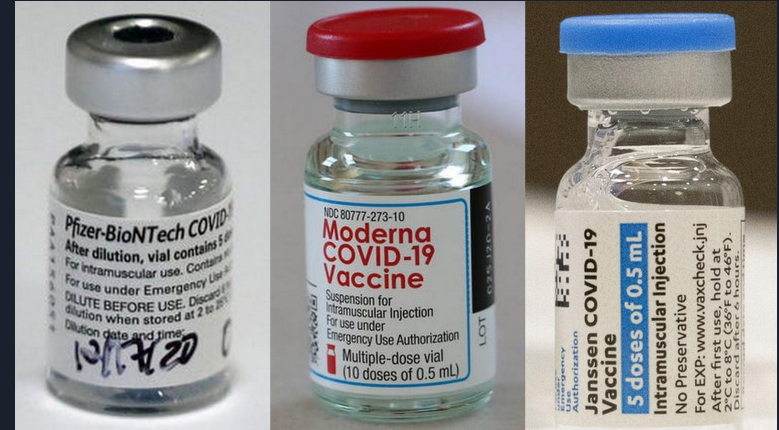
Debate Vaccines
r/DebateVaccines

[Join]

Posts

🔥 Hot | ✳️ New | 📊 Top | ⋯

⬆️ Vote ⬇️ | 📄 Let's review the countries response to Covid anecdotally.
Posted by u/dureshdalit 16 minutes ago | 💬 0 | ⋯

⬆️ Vote ⬇️ | `</>` Unheard concerns: WHO investigating possible link between COVID vaccine and hearing issues   youtube.com/watch?... ↗
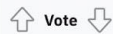**COVID-19 Vaccines**
Posted by u/pmabraham 26 minutes ago | 💬 1 | ⋯

⬆️ Vote ⬇️ | 📄 Heads Up: 80,000 NEW PFIZER PAGES coming tomorrow, MAY 1ST
Posted by u/jorlev 1 hour ago | 💬 3 | ⋯

⬆️ Vote ⬇️ | ↗️ SARS-CoV-2 vaccination can elicit a CD8 T-cell dominant hepatitis   ncbi.nlm.nih.gov/pmc/ar... ↗
Posted by u/jorlev 1 hour ago | 💬 0 | ⋯

⬆️ 0 ⬇️ | ☰ A 28 year old basketball player dies suddenly on the court. He received the COVID vaccine 6 months ago. Family history reveals sudden cardiac death in his father (1990's) and uncle (early 2000's). Biopsy reveals hypertrophic cardiomyopathy, confirmed by genetic testing. What was the cause of death? 36 votes 2 days 21 hours left | 💬 10 | ⋯
Posted by u/AlarmTablet 2 hours ago

⬆️ 3 ⬇️ | ↗️ Bad vaccines never die, they just get recycled   precisionvaccinations.com/vaccin... ↗
Posted by u/SftwEngr 2 hours ago | 💬 1 | ⋯

⬆️ 2 ⬇️ | ↗️ EXPLAINER: How China is using metal barriers to fight COVID   abcnews.go.com/Health... ↗
Posted by u/SftwEngr 3 hours ago | 💬 4 | ⋯

⬆️ 8 ⬇️ | ↗️ The Rise Of The Rockefeller Pharma Industry   globalpossibilities.org/the-ri... ↗
Posted by u/Tough_Perfect 4 hours ago | 💬 2 | ⋯

⬆️ 0 ⬇️ | ↗️ Covid deaths no longer overwhelmingly among unvaccinated as toll on elderly grows   news.yahoo.com/covid-... ↗
Posted by u/SftwEngr 5 hours ago | 💬 4 | ⋯

About Community

Debate and discuss issues surrounding vaccinations.

17.5k Members | **165** Online

🗓️ Created Apr 6, 2017

Filter by flair

COVID-19 Vaccines | **Question**

r/DebateVaccines Rules

1. Post requirements ⌄

2. Be civil, no personal attacks

3. No ad hominems ⌄

4. No trolling or spam

vaccine	1727
covid	1407
vaccines	1092
people	546
19	487
vaccinated	453
anti	357
vaxxers	340
pfizer	316
vaccination	300

covid 19	476
covid vaccine	204
anti vaxxers	157
19 vaccine	144
pro vaxxers	120
covid vaccines	113
19 vaccines	72
vaccines cause	67
spike protein	58
sars cov	58

covid 19 vaccine	144
covid 19 vaccines	71
question pro vaxxers	40
vaccines cause autism	39
covid 19 vaccination	25
pfizer covid 19	15
vaccines safe effective	14
pro vaxxers believe	13
covid 19 deaths	13
dr robert malone	12

pfizer covid 19 vaccine	12
athlete cardiac arrests issues	10
question pro vaxxers believe	8
believe vaccines cause autism	7
covid shots european database	6
pro vaxxers believe vaccines	6
following covid shots european	6
pfizer biontech covid 19	5
sars cov delta variant	5
biontech covid 19 vaccine	5

Most common words/phrases in DebateVaccines subreddit

CovidVaccinated

r/CovidVaccinated

Join

Posts

🔥 Hot   ✳️ New   📊 Top   •••   ☰ ⌄

⬆️ 0 ⬇️	📄	considering getting a second j&j...still waiting for novavax `J&J Booster` 💬 4 •••
		Posted by u/katzenjammerr 7 hours ago

⬆️ 6 ⬇️	▶️	Bringing a baby or child into your life can be an exciting (and sometimes overwhelming!) time. It starts with getting prepared—emotionally and financially. We're here to help you through your journey with guidance, information, and resources from people who have been through it themselves. 💬 6 •••
		u/fidelityinvestments · Promoted

Fidelity.com   Learn More

⬆️ 0 ⬇️	📄	What to expect for booster shot `Pfizer Booster` 💬 1 •••
		Posted by u/FarAcanthopterygii83 12 hours ago
⬆️ 4 ⬇️	📄	Sinovac booster after 2 Pfizer shots `Question` 💬 4 •••
		Posted by u/MitchM007 15 hours ago
⬆️ 25 ⬇️	📄	Side Effects, Long Covid, Covid Vax, Covid Unvax and Getting Support instead of Criticism `General Info` 💬 4 •••
		Posted by u/XulaSLP07 23 hours ago
⬆️ 3 ⬇️	📄	important question about booster side effects `Question` 💬 0 •••
		Posted by u/xj4y 2 days ago
⬆️ 1 ⬇️	📄	If you've had covid before being vaccinated, what was your second dose like? Getting my second shot tomorrow. Trying to plan for some effects. `Question` 💬 12 •••
		Posted by u/Crafty-Information42 2 days ago

About Community

Vaccines are here! Many types, from many manufacturers, at different times, for different people and in different places. Lets talk about vaccination and vaccines

46.7k Members   **45** Online

🎂 Created Nov 24, 2020

Filter by flair

`J&J Booster` `Pfizer Booster`

`Question` `General Info` `Pfizer`

`J&J`

booster	1500
pfizer	1486
vaccine	1393
shot	821
covid	785
dose	781
moderna	696
got	511
amp	491
effects	490

pfizer booster	297
moderna booster	263
booster shot	189
covid vaccine	168
second dose	165
pfizer vaccine	165
2nd dose	137
dose pfizer	130
pfizer shot	127
pfizer dose	93

second dose pfizer	32
pfizer booster shot	30
second pfizer shot	25
2nd pfizer shot	25
got moderna booster	24
swollen lymph nodes	24
covid 19 vaccine	22
pfizer booster experience	20
2nd pfizer dose	19
2nd dose pfizer	19

just got moderna booster	10
moderna booster flu shot	6
second dose pfizer vaccine	6
just got pfizer booster	5
got second pfizer shot	5
moderna booster half dose	5
just got pfizer shot	4
booster flu shot time	4
swollen lymph nodes neck	4
pfizer booster flu shot	4

Most common words/phrases in CovidVaccinated subreddit

Modeling

Count Vectorizer

- Logistic Regression
- KNeighbors Classifier
- SVC
- *Multinomial Naive Bayes Classifier*
- Decision Tree Classifier
- Extra Trees Classifier
- Random Forest Classifier
- Gradient Boosting Classifier

TF-IDF Vectorizer

- Logistic Regression
- KNeighbors Classifier
- SVC
- *Gaussian Naive Bayes Classifier*
- Decision Tree Classifier
- Extra Trees Classifier
- Random Forest Classifier
- Gradient Boosting Classifier

TFIDF-Vectorized Text
Model Performance

Models	Logistic Regression	KNeighbors Classifier	SVC	Gaussian NB	Decision Tree Classifier	Extra Trees Classifier	Random Forest Classifier	Gradient Boosting Classifier
Accuracy	.6017	.6421	.6701	.7164	.6695	.7217	.6666	.6203
Precision	.5814	.7418	.8844	.7969	.8820	.7253	.7373	.6009

Count-Vectorized Text
Model Performance

Models	Logistic Regression	KNeighbors Classifier	SVC	Multinomial NB	Decision Tree Classifier	Extra Trees Classifier	Random Forest Classifier	Gradient Boosting Classifier
Accuracy	.8585	.7161	.8482	.8670	.7783	.8098	.7898	.8458
Precision	.8609	.6614	.8615	.8757	.7675	.8101	.8631	.8499

Best Model: Multinomial Naive Bayes Classifier

- Alpha: 0.4
- Tokenizer: Lemma
- Stop Words: None
- Ngram Range: (1, 1)
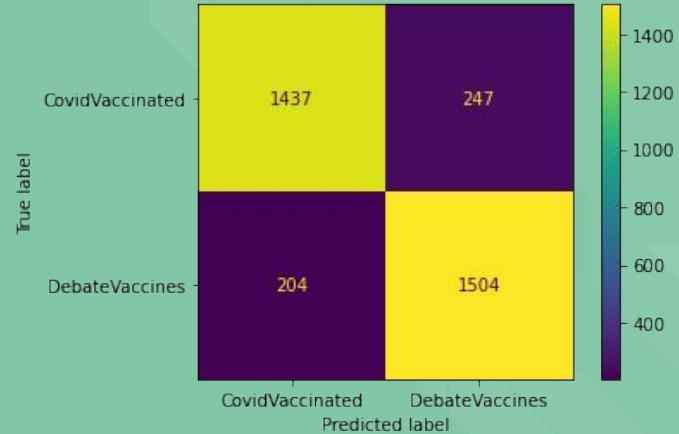
Correctly Classified Reddit Posts

⬆ 3 ⬇ 🔵 **r/CovidVaccinated** · Posted by u/cv2706 4 months ago 🔔

6 months late but got my second Pfizer dose today

`Pfizer`

Needles and scared of side effects, but finally did it today

💬 3 Comments ↗ Share 🔖 Save 👁 Hide 🚩 Report   80% Upvoted

⬆ 12 ⬇ 🔵 **r/DebateVaccines** · Posted by u/LumpyGravy21 5 months ago 🔔

Fully vaccinated English adults under the age of 60 are now dying at twice the rate of their unvaccinated counterparts.

Official Data: Vaccinated Adults Below 60 Are Dying at TWICE the Rate of Unvaccinated (humansarefree.com)

💬 4 Comments ↗ Share 🔖 Save 👁 Hide 🚩 Report   87% Upvoted
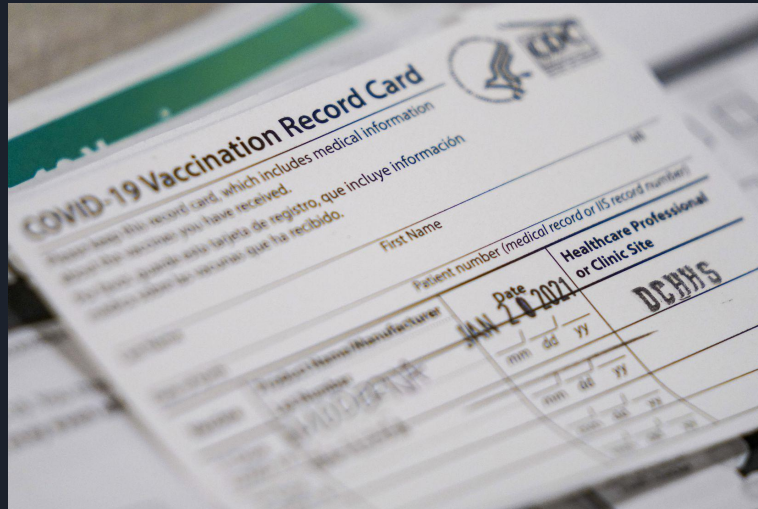
Misclassified Reddit Posts

Future Analysis

- COVID-19 vaccine misinformation hub r/NoNewNormal banned last summer
- No stop words used
  - Could fabricate stop words list of shared terms
- Sentiment Analysis integration
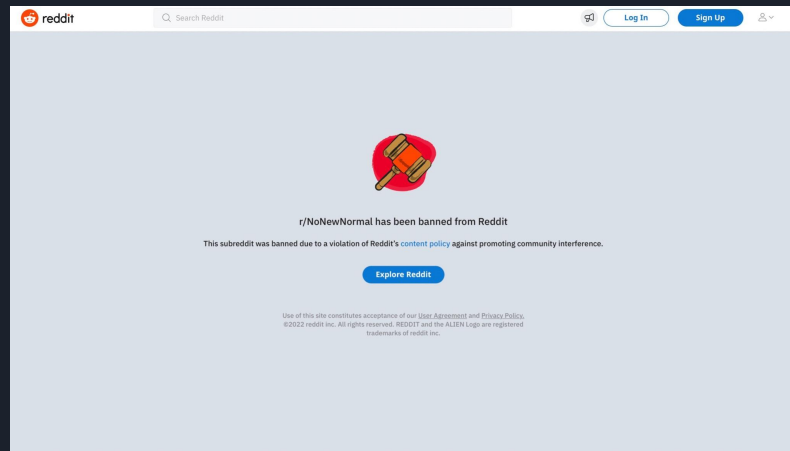
Conclusion

- Important for HHS to understand misinformation currently going viral
- Future campaigns can focus on *debunking myths popular in unvaccinated circles*
- Future campaigns can include *most persuasive arguments for vaccination*
  - Some are more convinced after hearing from non-public health officials (*"real Americans"*)

Questions?

Thank you!