

Reddit Project

Adrian Garrido

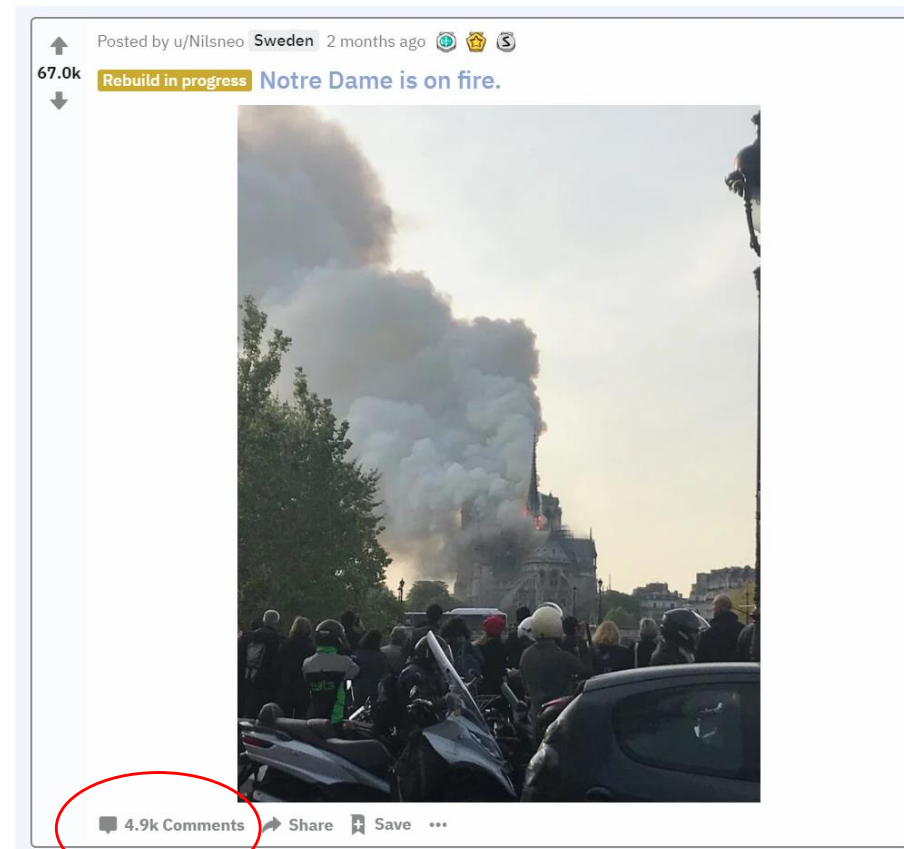
Objective

- Creating a function that takes the names of 2 subreddits and creates different models that predict whether a post is from one subreddit or the other.

Hot page



Top page



Data collection

Using top page instead of hot page.

Selecting Data

Title of the post

Text of the post

Primary comments

The screenshot shows a Reddit post on the subreddit r/space. The post title is "In the future if/when humans colonise other planets, what do you think will be the sociological differences between a Mars colony, a Moon colony and those who chose to stay behind on earth?". The post text is "Will we see people travelling between the colonies as 'immigrants'? Will people who stayed on Earth view the colonies as traitors who 'abandoned' Earth?". The primary comments are: "Watch The Expanse for a fairly reasonable look at that exact question.", "I love how in the Expanse it is not utopian but not totally dystopian either and everything seems as unplanned as our present", "Came here just to say this exact thing! It's such a good show.", "I don't think The Expanse is an accurate prediction of the future nor do I think it tries to be. I heavily doubt we'd colonize mars before figuring out how to make robots mine the asteroid belt. We also would never need to transport water like they do in the Expanse that is ridiculous. The ISS already has water reclaimers so efficient, they drain the water out of food before sending it up. I've only seen a few episodes but just these two things majorly affect the socioeconomic status of people in The Expanse.", "It is one of the best shows ever after The 100, I wanted to say the same", "Thanks to the Internet, I don't think there will be too much cultural drift. Sure, Mars has a 6 - 45 minute latency (is that even the right term?), but once a server farm is established, there will be regular, daily (even hourly) transfers of data, so Mars will be able to share memes with Earth and vice versa in very nearly real time. Sure, things will be different... some of the first native-born Martians might bawk at the idea of 150% gravity... but we live in the age of communication, and communication fosters a sharing of cultures and blending of diversity.", "To expand on that, barring a major breakthrough in propulsion tech, digital media will likely be Mars's main export after the initial founding of the colony.", and "It is still hard for me to grasp the economic incentive to build colonies on Mars. It is less hospitable than uninhabited places on earth".

reddit r/space Search r/space

42 In the future if/when humans colonise other planets, what do you think will be ... Discussion CLOSE

r/space · Posted by u/seawitches 12 hours ago

In the future if/when humans colonise other planets, what do you think will be the sociological differences between a Mars colony, a Moon colony and those who chose to stay behind on earth?

Discussion

Will we see people travelling between the colonies as "immigrants"? Will people who stayed on Earth view the colonies as traitors who "abandoned" Earth?

55 Comments Share Save Hide Report 76% Upvoted

What are your thoughts? Log in or Sign up LOG IN SIGN UP

SORT BY BEST

Birddawg65 40 points · 11 hours ago

Watch The Expanse for a fairly reasonable look at that exact question.

zig_anon 8 points · 6 hours ago

I love how in the Expanse it is not utopian but not totally dystopian either and everything seems as unplanned as our present

Mauimoves 8 points · 10 hours ago

Came here just to say this exact thing! It's such a good show.

Gymnopodies3 3 points · 4 hours ago · edited 4 hours ago

I don't think The Expanse is an accurate prediction of the future nor do I think it tries to be. I heavily doubt we'd colonize mars before figuring out how to make robots mine the asteroid belt. We also would never need to transport water like they do in the Expanse that is ridiculous. The ISS already has water reclaimers so efficient, they drain the water out of food before sending it up. I've only seen a few episodes but just these two things majorly affect the socioeconomic status of people in The Expanse.

AbandonedStarship -1 points · 6 hours ago

It is one of the best shows ever after The 100, I wanted to say the same

FallingStar7669 20 points · 12 hours ago

Thanks to the Internet, I don't think there will be too much cultural drift. Sure, Mars has a 6 - 45 minute latency (is that even the right term?), but once a server farm is established, there will be regular, daily (even hourly) transfers of data, so Mars will be able to share memes with Earth and vice versa in very nearly real time. Sure, things will be different... some of the first native-born Martians might bawk at the idea of 150% gravity... but we live in the age of communication, and communication fosters a sharing of cultures and blending of diversity.

carnold2414 5 points · 8 hours ago

To expand on that, barring a major breakthrough in propulsion tech, digital media will likely be Mars's main export after the initial founding of the colony.

zig_anon 1 point · 6 hours ago

It is still hard for me to grasp the economic incentive to build colonies on Mars. It is less hospitable than uninhabited places on earth

r/space

15.6m Members 8.5k Online

Share & discuss informative content on: * Astrophysics * Cosmology * Space Exploration * Planetary Science * Astrobiology

JOIN

About Careers Press Advertise Blog Help The Reddit App Reddit Coins Reddit Premium Reddit Gifts

Content Policy | Privacy Policy User Agreement | Mod Policy © 2019 Reddit, Inc. All rights reserved

Data Cleaning

Urls

Mentions of other subreddits

Numbers

Special characters

Models

➤ Logistic Regression

- Transforming with Count Vectorizer

➤ Random Forest Classifier

- Transforming with Tfidf Vectorizer

➤ Multinomial Naive Bayes

- Transforming with Count Vectorizer

Function Outputs

- Constant feedback
- Check to see if there is enough posts in total
- Warning for unbalanced classes
- 10 most important words

WARNING: Unbalanced classes
Do you wish to continue?

Next Steps

➤ Version 2.0:

- Additional argument in the function that determines the minimum number of posts that should be gathered
- Input the parameters you want for each model
- Adding replies
- Adding fancy animations

0.00% [0/10 00:00<00:00]

Questions?

