Project_info

March 9, 2024

1 The Final Project

Xiaoyi_Wang, Qitian_LI, Yolanda_Meng

1.0.1 Objective:

Develop a bot that identifies and shares the top 10 trending posts from specified Reddit tags or subreddits with users.Our objective is to identify and engage with active users within this subreddit by offering them a personalized daily digest of top posts.

1.0.2 Introduction

Our bot aims to initiate a personalized interaction by sending a private message to users interested in specific tags, inquiring if they would like to receive daily updates on the hottest posts under these tags for the next ten days. Users expressing affirmative interest will be subscribed to a custom push notification service, receiving a daily digest of the most popular content. Conversely, those who decline or express disinterest will not receive further communications, respecting their preference and privacy.

In our beta version,we plan to use a bot to identify the top 10 hottest posts within a specific subreddit and directly message those users. The message will inquire if they are interested in subscribing to daily updates related to that subreddit. If they respond affirmatively, we will proceed to send them a daily digest for a duration of either one week or three days, depending on their preference. Should they decline, no further messages will be sent.

We need to be able to get at least the publishers of the top 10 most popular posts for a specific post under a specific tag. Then we could send a specific subscription message to the users in the users list, expressing our intention to have them subscribe.There also needs to be the ability to be able to get the top 10 hottest posts, which is called daily digest, everyday after their consent to our subscription, and these require specific LOOPs to do so.

In our idea version, we are more inclined to user customization options that enable a personalized content discovery journey. This feature periodically suggests new content sections or tags of potential interest through private messages. Users can directly respond to these suggestions to subscribe or decline. If they opt in, they have the flexibility to choose updates for either a week or three days. Additionally, implement a simple rating system allowing users to rate the suggested content via private message replies. For those unsure about subscribing, offer a glimpse into the current top 10 trending hashtags as an alternative. This approach ensures users are engaged and can tailor their content experience to their preferences.We will also analyze this post through the sentiment analysis function. If the score is too low, we will consider pushing other posts. It would be better

if we also include commenters of those popular posts to increase the diversity and collect as many potential users we want as possible. These functions can make the user's browsing process efficient, and can also push some optimistic posts through sentiment analysis to have a positive effect on the user's psychology.

Implementing the dream version of a dynamic content discovery mechanism presents significant challenges due to its high technical complexity, which requires extensive data analysis and machine learning capabilities to accurately understand user interests. Additionally, designing an intuitive and user-friendly interface that allows for effective feedback on preferences poses a challenge, especially considering the varying levels of technical proficiency among users. Moreover, as the number of users increases, maintaining the system's performance and scalability becomes increasingly difficult, further complicating the realization of this vision.

1.0.3 Load praw code

Tips: make sure praw library is installed

[3]:
```
# Import the praw library to work with the Reddit API
import praw
```

1.0.4 Load csv file

[4]:
```
# Import the time library for time-related functions
import time

# Import the nltk library for natural language processing tasks
import nltk

# Download the VADER lexicon, which is used for sentiment analysis
nltk.download(["vader_lexicon"])

# Import the SentimentIntensityAnalyzer class from nltk's sentiment module
from nltk.sentiment import SentimentIntensityAnalyzer

# Instantiate an object of SentimentIntensityAnalyzer
sia = SentimentIntensityAnalyzer()
```

```
[nltk_data] Downloading package vader_lexicon to
[nltk_data]     /home/jovyan/nltk_data…
[nltk_data]   Package vader_lexicon is already up-to-date!
```

1.0.5 Run the key.

Keep it fake when testing to prevent my account being blocked.

[5]:
```
# load your reddit keys
%run reddit_keys.py
```

[6]:
```
# Give the praw code your reddit account info so
# it can perform reddit actions
reddit = praw.Reddit(
    username=username, password=password,
    client_id=client_id, client_secret=client_secret,
    user_agent="a custom python script for user /" + str(username)
)
```

```
Version 7.7.0 of praw is outdated. Version 7.7.1 was released Tuesday July 11,
2023.
```

Get the top 10 trending posts from tags.
It takes time when executing. Wait a few seconds.

[15]:
```
# get the top 10 posts
# Set the subreddit to "science" and retrieve the top 10 hot posts
subreddit_name = "science"
posts = reddit.subreddit(subreddit_name).hot(limit=10)
```

Part1: Get the users' data We are going to separate the extracting data and requesting subscription parts to prevent being shadow-banned.

[16]:
```
# Create a set to store unique users
users = set()
print("Let's Start Collecting Commenters' information")

################## Part 1 #################

# Loop through the posts and add the authors and the authors of the top 3␣
 ↪comments to the set
for post in posts:
    # Add the post's author
    users.add(str(post.author))
    # Look at the first three comments
    for comment in post.comments[:3]:
        # Check if the comment has an author attribute
        if hasattr(comment, 'author'):
            # Add the comment's author
            users.add(str(comment.author))

# Output the number of unique authors and commenters collected
print(f"Number of authors and commenters collected: {len(users)}")

# Remark:
#    - Since this bot is very likely to be blocked by reddit, we design␣
 ↪separate sections
```

```
#    to collect users and interact with the users. In Part 1, we are able to␣
 ↪view the users
#    information prior to the interactions.
print("Those are the users we collected: ")
print(str(users))
```

```
Let's Start Collecting Commenters' information
Number of authors and commenters collected: 17
Those are the users we collected:
{'ACR96', 'chrisdh79', 'mvea', 'Osiris62', 'DresdenFormerCypher', 'Klomenko',
'None', 'smurfyjenkins', 'Wagamaga', 'RickyWinterborn-1080', 'Irish_Whiskey',
'giuliomagnifico', 'AutoModerator', 'Sciencebang', 'synth003', 'Sryzon',
'JustDirection18'}
```

Part2: We don't want to really send the users' messages to prevent being shadow-banned. Since we've already get the real users' data from Part1, we are going to mimic the process of sending messages to those users, rather than really sending messages. Let's call the simulated Part2 "Fake Part 2" as below, and the real DM process simply "Part 2".

[17]:
```
################## Fake Part 2 #################
# We use fake part 2 to mimic sending messages to those users.
for user in users:
    print("------- DM -------")
    print("Daily Digest Subscription")
    print("Hi " + user + ", would you like to receive a daily digest of the top␣
 ↪posts from " + subreddit_name + " over the next 10 days? Please reply with␣
 ↪'yes' or 'no'.")
    print("------- END DM -------")
```

```
------- DM -------
Daily Digest Subscription
Hi ACR96, would you like to receive a daily digest of the top posts from science
over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi chrisdh79, would you like to receive a daily digest of the top posts from
science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi mvea, would you like to receive a daily digest of the top posts from science
over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi Osiris62, would you like to receive a daily digest of the top posts from
```

science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi DresdenFormerCypher, would you like to receive a daily digest of the top posts from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi Klomenko, would you like to receive a daily digest of the top posts from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi None, would you like to receive a daily digest of the top posts from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi smurfyjenkins, would you like to receive a daily digest of the top posts from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi Wagamaga, would you like to receive a daily digest of the top posts from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi RickyWinterborn-1080, would you like to receive a daily digest of the top posts from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi Irish_Whiskey, would you like to receive a daily digest of the top posts from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi giuliomagnifico, would you like to receive a daily digest of the top posts from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi AutoModerator, would you like to receive a daily digest of the top posts from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------

```
Daily Digest Subscription
Hi Sciencebang, would you like to receive a daily digest of the top posts from
science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi synth003, would you like to receive a daily digest of the top posts from
science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi Sryzon, would you like to receive a daily digest of the top posts from
science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
------- DM -------
Daily Digest Subscription
Hi JustDirection18, would you like to receive a daily digest of the top posts
from science over the next 10 days? Please reply with 'yes' or 'no'.
------- END DM -------
```

The code below shows the real Part2. However, we can only safely use the fake part2 because Reddit doesn't allow us to send private message to those users, and we won't be able to get the replies from them. However, the functionality works well, and the fake part 2 already shows how the DM works in our code.

[ ]:
```python
################## Part 2 #################

# don't run this part of the code when testing to prevent being blocked.
# Send a message to each user in the set asking if they want to subscribe to␣
 ↪the daily digest
# Note: This block is commented out to prevent accidental execution

for user in users:
    message = f"Hi {user}, would you like to receive a daily digest of the top␣
 ↪posts from r/{subreddit_name} over the next 10 days? Please reply with 'yes'␣
 ↪or 'no'."
    reddit.redditor(user).message('Daily Digest Subscription', message)

# Wait for a day before proceeding (to give users time to respond)
time.sleep(24 * 3600)


# manage the subscription
subscribted_users = set()
# Check the inbox messages for responses to the subscription message
for message in reddit.inbox.messages():
    # Check if the message is related to our subscription
    if message.subject == 'Daily Digest Subscription':
```

```
        # User wants to subscribe
        if message.body.lower() == 'yes':
            # Add user to the subscribed set
            subscribted_users.add(message.author)
            # Subscribe the user to the subreddit if they haven't already
            if not reddit.redditor(message.author).
 ↪has_subscribed(subreddit_name):
                reddit.redditor(message.author).subscribe(subreddit_name)
        elif message.body.lower() == 'no':  # User does not want to subscribe
            # Unsubscribe the user from the subreddit if they have previously␣
 ↪subscribed
            if reddit.redditor(message.author).has_subscribed(subreddit_name):
                reddit.redditor(message.author).unsubscribe(subreddit_name)

################## Part 3 #################
 # Get the top 10 hot posts from the subreddit
top_posts = reddit.subreddit(subreddit_name).hot(limit=10)
 # Send the daily digest to subscribed users
for post in top_posts:
    # Get sentiment score for the post title
    sentiment_score = sia.polarity_scores(post.title)["compound"]
    # If the sentiment is neutral or positive
    if sentiment_score >= 0:
        # Loop through subscribed users
        for user in subscribted_users:
             # Create a message with the post title and URL
            message = f"Here's a trending post from r/{subreddit_name}:
 ↪\n\n{post.title}\n\n{post.url}"
            reddit.redditor(user).message('Daily Digest', message) # Send the␣
 ↪message to the user

# Print a confirmation that the daily digest has been sent
print("Daily Digest Sent!")
```

Part3: We don't want to really send the users' messages to prevent being shadow-banned. Since we've already get the real users' data from Part1, we are going to mimic the process of sending messages to those users, rather than really sending messages. Let's call the simulated Part2 "Fake Part 3" as below, and the real DM process simply "Part 3".

[19]:
```
for user in users:
    print("------- DM -------")
    print( user + "Reply: Yes" )
    print( f"Here's a trending post from r/{subreddit_name}:\n\n{post.
 ↪title}\n\n{post.url}" )
    print("------- END DM -------")
```

------- DM -------
ACR96Reply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
chrisdh79Reply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
mveaReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
Osiris62Reply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
DresdenFormerCypherReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
KlomenkoReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
NoneReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
smurfyjenkinsReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
WagamagaReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

```
https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
RickyWinterborn-1080Reply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood.
The findings showed that adults who had amblyopia or lazy eye as a child, had a
29 per cent higher chance of developing diabetes, 25 per cent higher odds of
high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
Irish_WhiskeyReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood.
The findings showed that adults who had amblyopia or lazy eye as a child, had a
29 per cent higher chance of developing diabetes, 25 per cent higher odds of
high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
giuliomagnificoReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood.
The findings showed that adults who had amblyopia or lazy eye as a child, had a
29 per cent higher chance of developing diabetes, 25 per cent higher odds of
high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
AutoModeratorReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood.
The findings showed that adults who had amblyopia or lazy eye as a child, had a
29 per cent higher chance of developing diabetes, 25 per cent higher odds of
high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
```

SciencebangReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
synth003Reply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
SryzonReply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------
------- DM -------
JustDirection18Reply: Yes
Here's a trending post from r/science:

Children with ‘lazy eye’ are at increased risk of serious disease in adulthood. The findings showed that adults who had amblyopia or lazy eye as a child, had a 29 per cent higher chance of developing diabetes, 25 per cent higher odds of high blood pressure and a 16 per cent higher risk of obesity

https://www.eurekalert.org/news-releases/1036761
------- END DM -------

1.1 Reflective Questions:

1.1.1 Q1: What steps can we take to prevent bots from perpetuating bias in content moderation?

Utilitarianism: Under this frame, preventing bias in content moderation is essential for improving users fulfillment and confidence in the platform. Biased bots may coincidentally advance particular sorts of content over others, prompting a skewed opinion of data. Moreover, incorporating different training datasets and guaranteeing transparency in the bot's decision-making process can assist with alleviating the bias. By focusing on user welfare and fairness in content moderation, our bot can add to a more comprehensive and adjusted web-based local area, in this manner lining up with utilitarian standards.

Deontological Ethics: According to a deontological point of view, the bot has an ethical obligation to respect users' freedoms and independence in satisfied control. Biased bots that sustain separation abuse users' rights to access various opinions and information. To maintain moral standards, our bot ought to be planned with severe rules for fairness and transparency. This incorporates transparent correspondence about its balance models and provides users with choices to pursue content decisions. By focusing on moral standards of fairness and respect for users' autonomy, our bot can satisfy its obligation to advance a comprehensive and impartial online environment.

1.1.2 Q2: This question explores the ethical implications of automated systems that might narrow the scope of information provided to users?

Utilitarianism: Automated systems that narrow the scope of information should be evaluated based on their impact on user satisfaction and information relevance. While narrowing the scope may improve the quality of information for some users, it may also limit access to diverse perspectives and essential information for others. To address this issue, the bot ought to focus on user-driven design by permitting users to redo their data inclinations and giving straightforward clarifications to data sifting. By generally expanding user satisfaction and guaranteeing transparency, the bot can line up with utilitarian standards.

Deontological Ethics: According to Deontology, our bot holds an honest conviction to respect users' autonomy and freedom to access different perspectives. Restricting the scope of information without direct justification disregards users' autonomy and may lead to the dissemination of biased or incomplete information. To uphold moral standards, our bot grants users control over their information preferences and provides transparent explanations of its filtering algorithms. By focusing on users' autonomy and respecting the moral principles, our bot ensures that its actions align with principles of fairness and respect for users' rights. In our code, users have consultation functionality, so they already have control. We have already explained to users how information is filtered, and if they choose to subscribe, we will provide the most popular posts. That shows we are aligned with this frame.

1.1.3 Q3: How to prevent trolling and does it affect the functionality of the bots?

Utilitarianism: Preventing trolling is essential for enhancing user satisfaction and maintaining platform functionality. Trolling conduct can upset online communities, decrease user commitment, and sabotage the adequacy of the bot's usefulness. To address trolling, our bot carry out robust moderation policies, for example, we have content filtering calculations and user reporting systems by users choosing whether to subscribe or not. By focusing on user welfare and platform usefulness, our

bot can add to a positive and favorable online environment, consequently lining up with utilitarian standards.

Deontological Ethics: According to a deontological viewpoint, preventing trolling is an ethical basic to maintain users' rights to safety and secure and respectful engagement. Trolling behavior violates users' privileges to participate in online communities free from harassment and misuse. Our bot has an obligation to implement local area norms, safeguard users from unsafe way of behaving, and cultivate a comprehensive web-based environment. By carrying out proactive measures to prevent trolling and respecting ethical norms of safety, our bot can satisfy its obligation to advance a positive and respectful online platform while maintaining its functionality through our subscription management. But in order to align with this frame, we can do better on designing the part of filtering information and preventing misuse of our bot.

[ ]: