

Group 21 - Homework 3 - Kafka Toxic Live Reddit Comment Filter

Kopal Bhatnagar, Tamanna Sharma, Tushar Yadav, Arjun Lamba, Rashmila Mitra

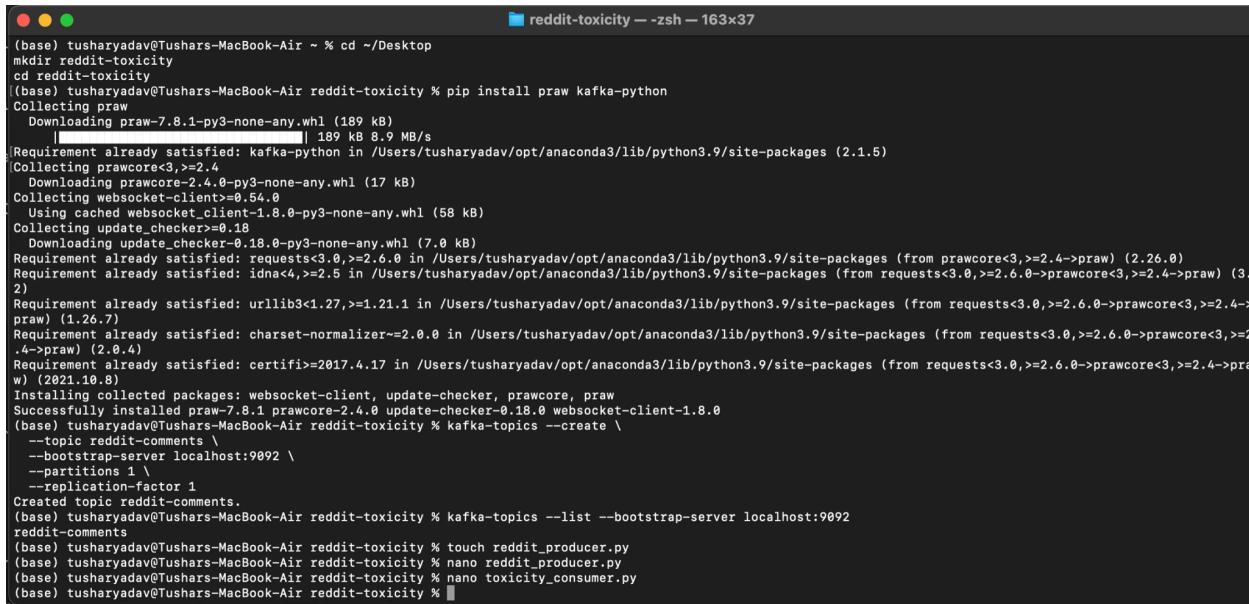
A) Steps Followed During Setup and Execution

1. Environment Setup

- Installed Kafka and Zookeeper using Homebrew on Mac.
- Started both services with brew services start kafka and brew services start zookeeper.
- Verified that Kafka was running on localhost:9092.

2. Folder Structure

- Created a dedicated folder named reddit-toxicity on the Desktop.
- Stored all script files: *reddit_producer.py*, *toxicity_consumer.py*, and *stream_data.csv*.



```
(base) tusharyadav@Tushars-MacBook-Air ~ % cd ~/Desktop
mkdir reddit-toxicity
cd reddit-toxicity
(base) tusharyadav@Tushars-MacBook-Air reddit-toxicity % pip install praw kafka-python
Collecting praw
  Downloading praw-7.8.1-py3-none-any.whl (189 kB)
    ━━━━━━━━━━━━━━━━ 189 kB 8.9 MB/s
Requirement already satisfied: kafka-python in /Users/tusharyadav/opt/anaconda3/lib/python3.9/site-packages (2.1.5)
Collecting prawcore<3,>=2.4
  Downloading prawcore-2.4.0-py3-none-any.whl (17 kB)
Requirement already satisfied: requests<3.0,>=2.6.0 in /Users/tusharyadav/opt/anaconda3/lib/python3.9/site-packages (from prawcore<3,>=2.4->praw) (2.26.0)
Requirement already satisfied: idna<4,>=2.5 in /Users/tusharyadav/opt/anaconda3/lib/python3.9/site-packages (from requests<3.0,>=2.6.0->prawcore<3,>=2.4->praw) (3.2)
Requirement already satisfied: urlib3<1.27,>=1.21.1 in /Users/tusharyadav/opt/anaconda3/lib/python3.9/site-packages (from prawcore<3,>=2.4->praw) (1.26.7)
Requirement already satisfied: charset-normalizer~=2.0.0 in /Users/tusharyadav/opt/anaconda3/lib/python3.9/site-packages (from requests<3.0,>=2.6.0->prawcore<3,>=2.4->praw) (2.0.4)
Requirement already satisfied: certifi>=2017.4.17 in /Users/tusharyadav/opt/anaconda3/lib/python3.9/site-packages (from requests<3.0,>=2.6.0->prawcore<3,>=2.4->praw) (2021.10.8)
Installing collected packages: websocket-client, update-checker, prawcore, praw
Successfully installed praw-7.8.1 prawcore-2.4.0 update-checker-0.18.0 websocket-client-1.8.0
(base) tusharyadav@Tushars-MacBook-Air reddit-toxicity % kafka-topics --create \
  --topic reddit-comments \
  --bootstrap-server localhost:9092 \
  --partitions 1 \
  --replication-factor 1
Created topic reddit-comments.
(base) tusharyadav@Tushars-MacBook-Air reddit-toxicity % kafka-topics --list --bootstrap-server localhost:9092
reddit-comments
(base) tusharyadav@Tushars-MacBook-Air reddit-toxicity % touch reddit_producer.py
(base) tusharyadav@Tushars-MacBook-Air reddit-toxicity % nano reddit_producer.py
(base) tusharyadav@Tushars-MacBook-Air reddit-toxicity % nano toxicity_consumer.py
(base) tusharyadav@Tushars-MacBook-Air reddit-toxicity %
```

3. Reddit App & Producer Script

- Registered a Reddit developer app and obtained the client ID, secret, and user agent.
- Used the *praw* library to stream live comments from subreddits like r/news, r/AskReddit, and r/politics.
- Created a Kafka producer (*reddit_producer.py*) to send these comments to the Kafka topic reddit-comments.

GET NEW REDDIT  MY SUBREDDITS  HOME - POPULAR - ALL - USERS | CHATGPT - ENTREPRENEUR - OPENAI - BUSINESS - ANNOUNCEMENTS | ASKREDDIT - PICS - FUNNY - MOVIE

reddit PREFERENCES options apps RSS feeds friends blocked password/email delete

developed applications

 **bd_hw3_g21**
personal use script
4A-eVGV9hn19AD9rQFerng

change icon

secret B7aj-dLWpI6y-3mb50O4pZL2GXXnNQ **developers** tush-ydv (that's you!) remove
add developer:

name

description

about url

redirect uri

update app **delete app**

create application

By creating an app, you agree to Reddit's Developer Terms and Data API Terms. You must also register to use the API.

name

web app A web based application
 installed app An app intended for installation, such as on a mobile phone
 script Script for personal use. Will only have access to the developers accounts

4. Execution Flow

- Ran *reddit_producer.py* in one terminal to stream Reddit comments.
- Ran *toxicity_consumer.py* in another terminal to classify comments and log results into *stream_data.csv*.

```
tusharyadav -- python3 reddit_producer.py -- 174x43
Last login: Thu May  1 14:53:36 on ttys000
(base) tusharyadav@Tushars-MacBook-Air ~ % python3 reddit_producer.py
python3: can't open file '/Users/tusharyadav/reddit_producer.py': [Errno 2] No such file or directory
(base) tusharyadav@Tushars-MacBook-Air ~ % cd ~/Desktop/reddit-toxicity
python3 reddit_producer.py
Streaming Reddit comments to Kafka...
[Sent: {'id': 'mq3hxtl', ' subreddit': 'politics', 'author': 'Gatonom', 'body': 'We need to make America something people want to love, not try to teach them to love it.', 'created_utc': 1746137526.0, 'score': 1}
[Sent: {'id': 'mq3hu16', ' subreddit': 'politics', 'author': 'veruca_seether', 'body': 'She is just as wishy-washy right-sliding. Her refusal to push back against trans attacks during the campaign was pretty Gavin like.', 'created_utc': 1746137527.0, 'score': 1}
[Sent: {'id': 'mq3huav', ' subreddit': 'news', 'author': 'Wolfram_And_Hart', 'body': 'I'll believe it when I see it. Big Lots is back again so...', 'created_utc': 1746137530.0, 'score': 1}
[Sent: {'id': 'mq3huc0', ' subreddit': 'AskReddit', 'author': 'patmurny', 'body': 'I dunno if I didn't why would they call me the ass whisperer lol', 'created_utc': 1746137530.0, 'score': 1}
[Sent: {'id': 'mq3huf3', ' subreddit': 'AskReddit', 'author': 'Armored-Elder', 'body': 'Unlimited ammo - GTA San Andreas for PS2', 'created_utc': 1746137531.0, 'score': 1}
[Sent: {'id': 'mq3hufo', ' subreddit': 'AskReddit', 'author': 'IlliropbissUoff', 'body': 'You punch them in the throat. I know it sounds harsh, but do you really have a chance in a fair fight between you and "bully"? Right? So you gotta do what you gotta do. Don't let any one walk on you. Watch a few YouTube videos on how to throw a punch.' , 'created_utc': 1746137531.0, 'score': 1}
[Sent: {'id': 'mq3hug6', ' subreddit': 'AskReddit', 'author': 'flaagan', 'body': 'Kind of interesting reading that. My grandmother on my dad's side had eight kids and lived a relatively modest living her entire life. At one point several years back, my dad had me help him take a screenshot of the stock / savings account he managed for her, because he wanted her to know she had over a million dollars. Still lived modest for the rest of her life. When she was reaching the end and her health was failing, it took my dad getting angry with her over the phone for her to accept that she was more than well off and could afford any care she needed, and that the money was hers to spend, not to suffer and 'save it for her kids'. Despite me generally disliking my father for a handful of other reasons, he made damn sure his mom was comfortable to the end, and shocked his siblings with how much inheritance they were all getting when grandma did pass on.", 'created_utc': 1746137531.0, 'score': 1}
[Sent: {'id': 'mq3huin', ' subreddit': 'AskReddit', 'author': 'AdThat328', 'body': 'Am I a kid again or still in my 30s? xD If I'm a kid again, rushing to see if Goosebumps is on TV while I organise my book collection...then probably putting on a Steps CD...\\nIf I'm an adult, immediately booking flights to Ibiza. Full travel agent experience.\n', 'created_utc': 1746137532.0, 'score': 1}
[Sent: {'id': 'mq3hujv', ' subreddit': 'AskReddit', 'author': 'dukeduke420', 'body': 'Go outside', 'created_utc': 1746137532.0, 'score': 1}
[Sent: {'id': 'mq3hulz', ' subreddit': 'AskReddit', 'author': 'DAETdatabase', 'body': '"the last time I sat/laid like this and angled my dick like this, I shot on my own face / directly into my mouth.... Oh well, maybe it won't happen this time" --that guy but for a dozen or more times', 'created_utc': 1746137533.0, 'score': 1}
[Sent: {'id': 'mq3humn', ' subreddit': 'AskReddit', 'author': 'Agile-Entry-5693', 'body': 'It feels good', 'created_utc': 1746137533.0, 'score': 1}
[Sent: {'id': 'mq3hurop', ' subreddit': 'politics', 'author': 'Initial-Toe-9512', 'body': 'Lots of words in this article, but whether or not it said anything, I don't know.\nIn short, if Trump is unpopular, it'll be easier for people to stand up to him. No Kidding, but that's not anything people can really do.\nOtherwise, they basically said to not freak out and overreact to what he does, some of what he is doing is technically legal, and in not overreacting, people can have a better conversation?\n\nIf someone got something else from this article, feel free to enlighten me.', 'created_utc': 1746137534.0, 'score': 1}
[Sent: {'id': 'mq3huse', ' subreddit': 'AskReddit', 'author': 'Cutiemyc', 'body': 'Go through my VHS Library and watch a movie.', 'created_utc': 1746137534.0, 'score': 1}
[Sent: {'id': 'mq3husg', ' subreddit': 'AskReddit', 'author': 'DeffJamiels', 'body': 'TRIANGLE, TRIANGLE, SQUARE, CIRCLE, X, L1, L1, DOWN, UP', 'created_utc': 1746137534.0, 'score': 1}
```

Reddit Producer Output

5. Basic Toxicity Detection

- Initially used a simple keyword match with 20 *common toxic words* (e.g., “idiot”, “dumb”, “trash”).

```
● ● ● tusharyadav — python3 toxicity_consumer.py — 188x52
Last login: Thu May 1 15:11:06 on ttys001
(base) tusharyadav@Tushars-MacBook-Air ~ % cd ~/Desktop/reddit-toxicity
cd: too many arguments
(base) tusharyadav@Tushars-MacBook-Air ~ % cd ~/Desktop/reddit-toxicity
python3 toxicity_consumer.py
Listening for Reddit comments...
✓ OK: we need to make america something people want to love, not try to teach them to love it.
✓ OK: she is just as wishy-washy right-sliding. her refusal to to push back against trans attacks during the campaign was pretty gavin like.
✓ OK: i'll believe it when i see it. big lots is back again so...
✓ OK: i dunno if i didn't why they called the ass whisperer lol
✓ OK: unlimited ammo vote an andrea for ps2
✓ OK: you punch them in the throat. i know it sounds harsh, but do you really have a chance in a fair fight between you and "bully?" right? so you gotta do what you gotta do. don't let any one walk on you. watch a few youtube videos on how to throw a punch.
✓ OK: kind of interesting reading that. my grandmother on my dad's side had eight kids and lived a relatively modest living her entire life. at one point several years back, my dad had me help him take a screenshot of the stock / savings account he managed for her, because he wanted her to know she had over a million dollars. still lived modest for the rest of her life. when she was reaching the end and her health was failing, it took my dad getting angry with her over the phone for her to accept that she was more than well off and could afford any care she needed, and that the money was hers to spend, not to suffer and 'save it for her kids'. despite me generally disliking my father for a handful of other reasons, he made damn sure his mom was comfortable to the end, and shocked his siblings with how much inheritance they were all getting when grandma did pass on.
✓ OK: am i a kid again or still in my 30s?
if i'm a kid again, rushing to see if goosebumps is on tv while i organise my book collection...then probably putting on a steps cd...
if i'm an adult, immediately booking flights to ibiza. full travel agent experience.
✓ OK: go outside
✓ OK: "the last time i sat/laid like this and angled my dick like this, i shot on my own face / directly into my mouth.... oh well, maybe it won't happen this time" --that guy but for a dozen or more times
✓ OK: it feels good
⚠️TOXIC: lots of words in this article, but whether or not it said anything, i don't know.
in short, if trump is unpopular, it'll be easier for people to stand up to him. no kidding, but that's not anything people can really do.
otherwise, they basically said to not freak out and overreact to what he does, some of what he is doing is technically legal, and in not overreacting, people can have a better conversation?
if someone got something else from this article, feel free to enlighten me.
✓ OK: go through my vhs library and watch a movie.
✓ OK: triangle, triangle, square, circle, x, 11, 11, down, up
✓ OK: ....don't cum where you sum loll11
✓ OK: i mean sure that, but also i got a 6 piece mcnugget meal yesterday and it was $16. so.. (in canada)
✓ OK: jumper - 3rd eye blind.
✓ OK: told him "not worth the effort" and walked away to join my friends, not for protection, just that being with friends was part of my normal day.

i never had any further trouble with him after declining his offer to fight. but then we never really hung around together at school, either.
✓ OK: texas democrats have consistently not shown up to vote, i'll believe it when i see it
✓ OK: amazon delivery where they use a code to leave your order inside
✓ OK: this! motherlode and rosebud are the first ones that came to mind
✓ OK: b, right, b, right, left, x,y, up - gta:sa
✓ OK: ok brother

and i am also a bit sad about this post, cuz a lot of my comments got some downvotes and it is bad for my account. i needed some so that i can post in the soccer community. but i talked to you about islam which made me feel happy and forgot about all the other stuff.
✓ OK: i cook using exact measurements and equipment. i never try to add my own twist to a recipe.
✓ OK: cheat code...
never knock when step sister is in the shower....
```

Note - we took only 20 random toxic words and ran the analysis based on it, so there might be some discrepancy for a comment/post to be toxic or ok!

```
that's how we tell people to fix these issues right?
✓ OK: you're damn right. i've been speeding like a bat out of hell lately. speeding limit my ass.
⚠️TOXIC: we definitely need to counter trump's marketing skills with better communication
✓ OK: i'm not gonna say i'm a fan of trump, but i think he's a good president especially during the height of covid. i def have pisd from something's i have seen.
✓ OK: olive - pearl jam.
✓ OK: lol that was too intense 😊😊
⚠️TOXIC: idk become a menace, overthrow the government, drive away from the police/military/whatever become a king, let myself get caught at 23h 55h and just die in the hands of the police
✓ OK: motherlode is so goated
⚠️TOXIC: i want to get my prostate milked. if i only planned my germany trip a little better i would have experienced it. but now i'm kinda in a committed relationship of sorts and my partner would definitely not be for shit like this.

and it would just feel weird to me to get my prostate milked while in a committed relationship.

even tho i'm paying for it and am not actively fucking a person i still can't bring myself to do it. which sucks cause i really want to try it
✓ OK: itt: horrible bowel movements. scattered with near-incest.
⚠️TOXIC: "lowkey" in reference to feelings towards something. sentences like "i lowkey hate him" makes me cringe, because "i don't like him much" would convey the same meaning.

also, "cook", "cooking" or "cooked" in reference to sportsmen's performances. such as "jackson was cooking the last game" or "he's a good manager, let him cook". so fucking annoying.
✓ OK: excrement ed.
✓ OK: "all research into brain worms will forthwith cease. no further questions will be answered", rfk jr. added with no further elaboration.
✓ OK: bro
✓ OK: $$$
✓ OK: yeah that's fair. part of the reason the connection it's hard is because just about anything of use to us is such a huge number of atoms so expecting to be able to simulate every atom is infeasible.

molecular interactions between proteins and a molecule of interest, lets say a potential drug, could be simulated now at this high level of accuracy (remember proteins can be a few hundred atoms to hundreds of millions of atoms and they exist in a solvent that affects how they fold). these interactions help screen whether a new potential drug is promising or not before it's made. but truly the use cases is vast. this simulation can be used to calculate mechanical properties of a material, electrical properties, optical properties and so on. for my research i focus on using it to study thermodynamic properties of polymers to try and discover new plastics that can be recycled from polymer back to monomer with only heat to allow for fully circular recycling at huge volumes
✓ OK: good question
✓ OK: the cycle must be discontinued.
```

Toxicity Consumer Output

- Upgraded to sentiment-based classification using **TextBlob**, labeling comments with polarity < -0.3 as toxic. -1.0 = very negative & +1.0 = very positive.

We treated anything below a threshold (like -0.3) as toxic. Re ran the producer file and consumer file after updating the toxic analysis code with text blob-based sentiment analysis code.

```
tusharyadav - python3 toxicity_consumer.py - 173x48
Last login: Thu May 1 15:33:37 on ttys003
(base) tusharyadav@Tushars-MacBook-Air ~ % cd ~/Desktop/reddit-toxicity
python3 toxicity_consumer.py
Listening with Textblob...
OK | Polarity: 0.16 | up down left right a + b + start together while sonic wags his finger :)
OK | Polarity: 0.30 | Smile
OK | Polarity: 0.00 | Get the band back together and sell out the Palace Hotel Ballroom in Chicago.
OK | Polarity: 0.59 | Lol that's good! That reminded me of the old video of the kid eating a plastic straw to ["prove that turtles are pussies"](https://youtu.be/OAwjs3NcbH4?si=6V3hpFie8ayd07v)
OK | Polarity: 0.13 | I've worked with some extremely wealthy people. They are often the most difficult to get to pay their bills - they're used to everything being given to them. Self made folks are usually nicer and have some real stories of resilience though some feel they're owed everything. Inherited wealth people are usually more trouble than they're worth.
OK | Polarity: 0.14 | I heard someone describe the felon-in-chief's dismantling of federal assistance programs and FEMA (while reappropriating FEMA funds for his golf trips) as the "Alabamafication" of the US and I can't stop thinking how true that is going to be. We're aiming for full third world status while the Kleptocrats steal everything not nailed down. Just like Russia.
OK | Polarity: 0.28 | idk honestly I can't really remember, I have really vague memories of her grip getting tighter whenever I was taking too long to answer but again I genuinely cannot remember lol this is like one of five memories I think happened from my childhood years lol, I don't know honestly, I was pretty young like 7 and younger I think. She never did anything other than that, I don't think. I don't know about ppd, maybe I think it's definitely a possibility and pretty likely but idk, her mom wasn't the greatest so I definitely know that generational trauma plays a role somewhere. She didn't do the same thing as your dad but she'd walk away and slam doors and say we were kissing her off and she needed to get away from me so she wouldn't get angrier and do something which isn't really the same thing and idk why I brought this up but I guess I'm going with it, anyway idk again I can't really remember like at all lol so this is going off of like idk what they even are lol feelings or something idk man.
OK | Polarity: 0.80 | Great answer. But, I'd watch it in theaters again and this time I'd be prepared to pee my pants.
OK | Polarity: 0.00 | playing outside till the lights turn off then its dinner time
OK | Polarity: 0.34 | Feels like grass
OK | Polarity: 0.34 | Fall in love don't do it unless you're ready.. These H_- ain't loyal.
TOXIC | Polarity: -0.61 | idk why, but everytime i end up liking someone, they are so mean to me... AND I STILL LIKE THEM!!!
OK | Polarity: 0.00 | That Martha's Vineyard was where Martha Stewart lived
OK | Polarity: 0.00 | Thank the Lord that d***** t**** isn't cosplaying as president!
OK | Polarity: 0.09 | I was never afraid of clowns, but nuns on the other hand...
OK | Polarity: 0.04 | Our home had a bedbug infestation when I was in high school, definitely the most traumatic experience I've been through. They are so difficult to get rid of.
OK | Polarity: 0.09 | ??? Using Back To The Future as my guiding light. Look for a newspaper bin/toll. First hint of year. From there, assess your skills, and aim for a construction job that won't ask too many questions. Hopefully Summer, but assess the season, and the weather that goes along with that. From there, assess on locations that you can hole up. Construction new home builds, underpass, etc. First thing is to NOT talk to family members as they will categorize you crazy and possible call the cops. From there, try to get by for 3 months. Save money when can. From there, any extra money, see about investing
TOXIC | Polarity: -0.48 | 500 to get cured from cancer, fuck yes.
OK | Polarity: 0.24 | In 2023, the year BEFORE Trump entered into his second term, 115.6 million fentanyl pills were seized. Trump wants you to believe that no one was working on this before him. How about asking how many were seized during any year of his first term? Yeah, I didn't think you would.
OK | Polarity: 0.62 | Love this! I also did this 😊
OK | Polarity: 0.11 | I'm not sad, maybe a little nostalgic today but it's part of letting go of things and healing, you gotta feel to heal. I'm actually grateful and feel at peace with everything happening and whatever may come.
OK | Polarity: 0.15 | be a bad bishhhhhhhhhhhh do what make u happy and you will find someone when the time is right...trust!!!!
OK | Polarity: -0.17 | That it's not true
OK | Polarity: 0.00 | Metroid for NES. Justin Bailey
OK | Polarity: 0.20 | Chicken wings every day for two weeks. I loved wings I don't understand why we never had more of those.
OK | Polarity: 0.05 | What is one thing you NEED but cannot afford or have resources for right now?
```

The accuracy of toxic comments/posts are still low as many comments/posts do have a toxic sentiment but the model predicts a positive polarity.

```
ut it's very hard for two people to have a long-term relationship with no sex whatsoever.
OK | Polarity: 0.08 | I'm not up with the play on lab-grown meat, but would they be able to grow different cuts of steaks and different types of meat like beef, pork, chicken etc? or would it be more just meat patties?
OK | Polarity: 0.58 | Can't view without a Twitter account, can someone elaborate?
OK | Polarity: 0.00 | I
OK | Polarity: 0.00 | I guess the oral was too vigorous😊
TOXIC | Polarity: -0.35 | Suddenly the upcoming Windows 10 End of Support ain't so bad after all. /s
TOXIC | Polarity: -0.40 | People like you give the rest of us Democrat voters the sort of reputation that Republican voters hate us for.

Pull your head out of your ass and act like a human being with empathy, or else just shut the fuck up.
OK | Polarity: 0.09 | Step 1 tell the director that you've enjoyed your time on set, but you're not coming anymore

Step 2 clarify that you mean both literally and figuratively
OK | Polarity: 0.28 | If you haven't already, time to work on this with a professional. Last year I underwent EMDR for trauma from decades ago. With the support of a wonderful counsellor, my life has vastly improved.
OK | Polarity: 0.20 | Intergenerational trauma and/or being raised in a family with lots of skeletons in the closet that no one really acknowledges so you just figure out through tid bits and throw away comments over the years.
OK | Polarity: 0.00 | Never make eye contact with a clown
```

TextBlob Toxicity Consumer Output

6. Hugging Face Transformer Integration

- Integrated the *unitary/toxic-bert transformer model* using the Hugging Face pipeline().

```

tusharyadav -> python3 toxicity_consumer.py - 188x41
Last login: Thu May  1 15:49:16 on ttys001
(base) tusharyadav@Tushars-MacBook-Air ~ % cd ~/Desktop/reddit-toxicity
python3 toxicity_consumer.py
config.json: 100%|██████████| 811/811 [00:00<00:00, 283kB/s]
model.safetensors: 100%|██████████| 438M/438M [00:33<00:00, 13.0MB/s]
tokens.txt: config.json: 100%|██████████| 174/174 [00:00<00:00, 1.93kB/s]
vocab.txt: 100%|██████████| 232K/232K [00:00<00:00, 3.18MB/s]
special_tokens_map.json: 100%|██████████| 112/112 [00:00<00:00, 60.2kB/s]
Device set to use mps:0
listening and classifying with HuggingFace model...
▲TOXIC | Score: 0.001 | We need to make America something people want to love, not try to teach them to love it.
▲TOXIC | Score: 0.026 | She is just as wishy-washy right-sliding. Her refusal to push back against trans attacks during the campaign was pretty Gavin like.
▲TOXIC | Score: 0.002 | I'll believe it when I see it. Big Lots is back again so...
▲TOXIC | Score: 0.737 | I dunno if I didn't why would they call me the ass whisperer lol
▲TOXIC | Score: 0.001 | Unlimited amms - GTA San Andreas for PS2
▲TOXIC | Score: 0.526 | You punch them in the throat. I know it sounds harsh, but do you really have a chance in a fair fight between you and "bully?" Right? So you gotta do what you gotta do. Don't let any one walk on you. Watch a few YouTube videos on how to throw a punch.
▲TOXIC | Score: 0.001 | Kind of interesting reading that. My grandmother on my dad's side had eight kids and lived a relatively modest living her entire life. At one point several years back, my dad had me help him take a screenshot of the stock / savings account he managed for her, because he wanted her to know she had over a million dollars. Still lived modest for the rest of her life. When she was reaching the end and her health was failing, it took my dad getting angry with her over the phone for her to accept that she was more than well off and could afford any care she needed, and that the money was hers to spend, not to suffer and 'save it for her kids'. Despite me generally disliking my father for a handful of other reasons, he made damn sure his wife and grandparents to the end, and looked his siblings with how much inheritance they were all getting when grandma did pass on.
▲TOXIC | Score: 0.001 | Am I a kid again or still in my 30s?
If I'm a kid again, rushing to see if Goosebumps is on TV while I organise my book collection...then probably putting on a Steps CD...
If I'm an adult, immediately booking flights to Ibiza. Full travel agent experience.
▲TOXIC | Score: 0.147 | Go outside
▲TOXIC | Score: 0.963 | "the last time I sat/laid like this and angled my dick like this, I shot on my own face / directly into my mouth.... Oh well, maybe it won't happen this time" --th at guy but for a dozen or more times
▲TOXIC | Score: 0.001 | It feels good
▲TOXIC | Score: 0.001 | Lots of words in this article, but whether or not it said anything, I don't know.
In short, if Trump is unpopular, it'll be easier for people to stand up to him. No kidding, but that's not anything people can really do.
Otherwise, they basically said to not freak out and overreact to what he does, some of what he is doing is technically legal, and in not overreacting, people can have a better conversation?
If someone got something else from this article, feel free to enlighten me.
▲TOXIC | Score: 0.005 | Go through my VHS library and watch a movie.
▲TOXIC | Score: 0.001 | TRIANGLE, TRIANGLE, SQUARE, CIRCLE, X, L1, L1, DOWN, UP
▲TOXIC | Score: 0.987 | I mean don't compare to SUM lolllll
▲TOXIC | Score: 0.049 | I mean sure that, but also I got a 6 piece McNugget meal yesterday and it was $16. So... (in Canada)
▲TOXIC | Score: 0.003 | Jumper - 3rd eye blind.
▲TOXIC | Score: 0.001 | Told him "not worth the effort" and walked away to join my friends, not for protection, just that being with friends was part of my normal day.

```

Faced an issue: similar to Textblob in the Hugging Face model, but here all the comments/posts are being classified as Toxic even if it is a neutral sentiment because it's a binary classifier trained only on toxic vs not toxic Reddit data.

```

▲TOXIC | Score: 0.001 | You Oughta Know - Alanis Morissette
▲TOXIC | Score: 0.022 | A very unpopular girl at my school developed a crush on me because I was nice to her. Then she started telling people we were dating...
▲TOXIC | Score: 0.001 | Triple Play Baseball 2000 automatic home run - triangle, square, triangle, circle, cross, square, L1, R1
▲TOXIC | Score: 0.055 | What's your fav ice cream
▲TOXIC | Score: 0.001 | The plane is about $9000/hr to operate and it was just under 11 hours of flight each way. The airport we landed at cost us $2000/day to stay at and we put the plane in a hangar for another $2500/night. We were there for two weeks. So we're looking at ~270K in airplane costs alone.
I've no idea what they spent the rest on. I was just told by the book keepers what it all cost. I would assume ultra high end lodging, dining, ski trips. We brought a ton of stuff home so I also assume shopping. There was at least one painting all boxed up in padding.
Who even knows. The owner is very generous but also spends money like the world is ending tomorrow.
▲TOXIC | Score: 0.013 | klapaucius
✓OK | Score: 0.584 | Me too! I have fond memories in 4th grade when everyone had to do typing tutor for **weeks if not months** and I got to play web games because I had already beaten typing tutor twice in a single day or two. I was always just doing whatever I wanted on the computer during computer lab because I was more proficient with computers than my teachers in 2006. I attribute that mainly to StarCraft being hard as fuck and without cheat codes I stood no chance. Plus having to learn *133t* speak to join battle.net servers in brood war. Good memories.
▲TOXIC | Score: 0.001 | Fusion power.

```

Hugging Face(Unitary/toxic bert model) Toxicity Consumer Output

7. Model Replacement for Balanced Classification

- Switched to *facebook/roberta-hate-speech-dynabench-r1-target*, which provides better gradation: LABEL_0, LABEL_1, and LABEL_2.

```
tusharyadav — python3 toxicity_consumer.py — 175x34
Last login: Thu May 1 16:13:48 on ttys001
(base) tusharyadav@Tushars-MacBook-Air ~ % cd ~/Desktop/reddit-toxicity
python3 toxicity_consumer.py
Device set to use mps:  

Listening and classifying with HuggingFace updated model...
✓ OK | Score: 1.0 | Where am I???
✓ OK | Score: 0.593 | Lock him up.
✓ OK | Score: 0.855 | Having worked in DC, the true fire is set to one day disappearing and is usually just shit like Ron Swanson throwing his computer in the dumpster.
✓ OK | Score: 0.999 | Remember when DOOM multiplayer and you thought those codes (that you thought only you knew!) would work, only to end up finding out it was for single player only. 😂
✓ OK | Score: 0.984 | i was about to say this. in active addiction i'm pretty sure almost everyone i knew saw me naked
✓ OK | Score: 0.534 | My source is I made it the fuck up!
✓ OK | Score: 1.0 | People from the US are not only super easy to spot but also to become casual acquaintances with. We are usually very friendly and easily engage in "small talk" such as the weather, traffic, best restaurants or museums in town. We also smile in public and will say hello to almost anyone. We are socialized this way.
Visiting the US is an easy way to meet Americans, but if you can't do that yet if you live in a place friendly to US tourists we are pretty easy to find.
✓ OK | Score: 1.0 | Your post has been removed as it violated [Rule 3](http://www.reddit.com/r/AskReddit/wiki/index#wiki_rule_3):
* Questions seeking a correct answer, that can be researched elsewhere or provide a limited scope for discussion are not allowed in r/AskReddit.

Not sure where to post? Check out r/AskReddit's [list of related subreddits](/r/AskReddit/wiki/sidebarsubs), or get recommendations in r/findareddit. Please read the rules before posting in other communities as they have different rules than r/AskReddit.
✓ OK | Score: 1.0 | A couple of quotes about common sense I listened somewhere some years ago.

"If you have to say you are using common sense, you are not using common sense"
"common sense is in fact very uncommon"
✓ OK | Score: 0.999 | Heard a song about a case in the Philippines.

[https://www.youtube.com/watch?v=DJp3n9P2NbU](https://www.youtube.com/watch?v=DJp3n9P2NbU)
✓ OK | Score: 1.0 | Never take love for granted it's an emotion it and will be reforged to something greater. Just be glad it's going through.
✓ OK | Score: 0.999 | Definitely stop on the way and enjoy some affordable food in an economy that isn't trashed yet
✓ OK | Score: 0.999 | Breaking the Law - Judas Priest.
```

Mapped these labels to "non-toxic", "neutral", and "toxic" and used the confidence score(0 to 1) to display severity.

```
OK | Score: 0.981 | I usually bring my own sandwiches and beverages to the host of something that I know they'll like. I feel bad if I show up empty handed.
✓ OK | Score: 0.981 | Stephen Miller only got into politics because he lacks the upper body strength to strangle prostitutes.
✓ OK | Score: 1.0 | Three Days Grace - Pain
✓ OK | Score: 0.995 | I just don't have the patience to take care of a kid. I also don't want my free time taken from me
✓ OK | Score: 1.0 | No reason for small donors to spend here. They will get enough funding, find small local elections where every dollar counts. State legislation wins are almost as valuable as federal ones and your money goes a ton further
✓ OK | Score: 1.0 | No bacon on naked days! If I do then just an apron would help.
✓ OK | Score: 1.0 | This is probably the best one. Everyone else is picking either stereotypically early 90's or stereotypically late 90's songs, and anyone who lived then knows that the early and late 90's were incredibly different.

Bittersweet symphony is probably the best inflection point for both.
✓ OK | Score: 1.0 | If i'm my current age I'd go find my mom and dad who would also be in their 20s and not have any kids yet and hang out with them :(
✓ OK | Score: 0.996 | To most Americans racism means you actively hate and show disdain for a minority.
✓ OK | Score: 0.998 | I like this sentiment so much. Every time I have a weird thought about sleeping with a coworker or some rando at the food store, I try to remind myself it's hormones and this has what has kept the human line going.
```

Hugging Face(facebook/hate-speech model) Toxicity Consumer Output

B) Brief Explanation of Producer & Consumer

1. **Producer (*reddit_producer.py*):**
 - a. Connect to Reddit API using *praw*.
 - b. Streams new comments from multiple subreddits.
 - c. Sends each comment (ID, author, body, etc.) as a JSON object to Kafka.

2. **Consumer (*toxicity_consumer.py*):**
 - a. Connect to the Kafka topic reddit-comments.
 - b. Classifies incoming comments using either:
 - i. **TextBlob** for sentiment polarity,
 - ii. or **Hugging Face transformer models** for toxicity detection.
 - c. Prints results and appends them to a CSV for further analysis.

C) Real-World Application & Use Case

This Kafka-based setup simulates a real-world real-time content moderation system, such as those used by:

1. Social media platforms (Reddit, Twitter, YouTube)
2. Online communities & gaming chat systems
3. Customer support platforms and comment moderation tools

Example Use Case: Live Toxicity Flagging Dashboard for Reddit Moderators

In a real-world application:

1. The Reddit producer could continuously stream comments from monitored subreddits.
2. The consumer could apply fine-tuned transformer models to classify each message in real-time.
3. A frontend dashboard could alert moderators when toxic comments exceed a certain confidence threshold (e.g., score > 0.75).
4. All flagged content could be logged for further review or automated takedown.

This system is scalable and could be enhanced with:

1. Multi-label classification (threat, insult, obscene, etc.)
2. Real-time dashboards (e.g., using Dash or Grafana)
3. Integration with Slack or email alerts