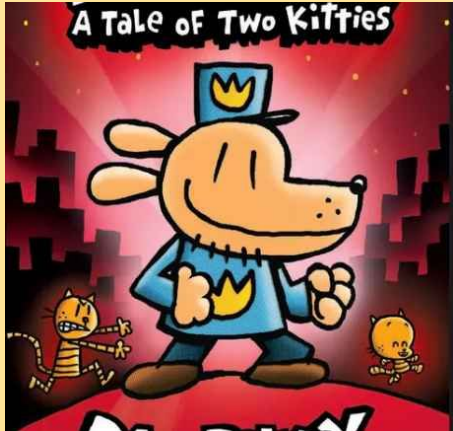


UCLA, MASDS, STATS-418-S25

How do you feel about the UCLA?

:Real Time Sentiment Analysis via Reddit data



Hochan Son

Objective

- Building Real time Sentiment Analysis system
- What sentiments? {Like , Hate, Frustrated, Super or emojis}
- Scrape data from Reddit / Subreddit channels
- Rate Limits (only 1000 post from the current datetime)
- Scoring system
- Keyword to any social sentiment (i.e. USC, LA Olympic 2028, or Tariff)

EDA

- Subreddit Post (e.g. r/UCLA)

{id, title, body, created_utc, score, num_commend, rul}

```
1 posts_df.head()
```

	id	title	body	created_utc	score	num_comments	url
0	1kbzww6	In the professional world, do UCLA and USC peo...		2025-04-30 21:06:24	43	48	https://www.reddit.com/r/USC/comments/1kbzww6/...
1	1k9gsoa	Has anyone chosen USC Viterbi over UCLA's engi...	If yes, why so? Do you think you made the righ...	2025-04-27 15:49:58	32	35	https://www.reddit.com/r/USC/comments/1k9gsoa/...
2	1k50vyu	Trojan Network is Insane! Okay, does it help I...	I was admitted undergrad aero but am switching...	2025-04-22 00:58:00	7	8	https://www.reddit.com/r/USC/comments/1k50vyu/...
3	1k3djev	Will I get jumped if i wear a UCLA sweater here?	Like in gangs if you wear a certain color u wi...	2025-04-19 20:29:40	0	15	https://www.reddit.com/r/USC/comments/1k3djev/...
4	1k1x9j8	USC vs Cal vs UCLA: Civil Engineering	Deciding between UC Berkeley, UCLA, and USC fo...	2025-04-17 22:20:36	20	30	https://www.reddit.com/r/USC/comments/1k1x9j8/...

- Post => Comments

{comment_id, post_id, body_score, created_utc, parant_id}

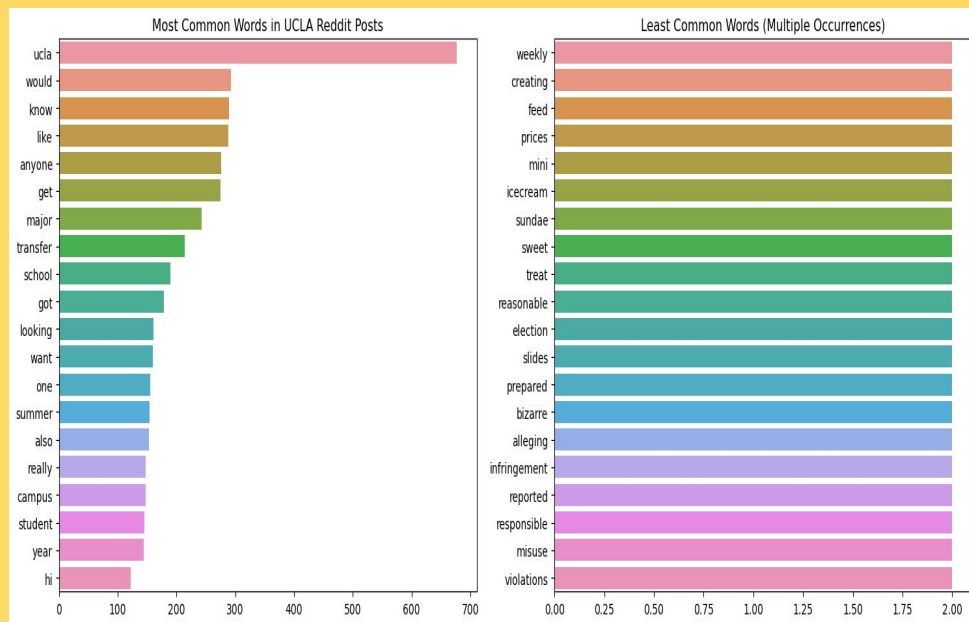
```
1 comments_df.head()
```

	comment_id	post_id	body	score	created_utc	parent_id
0	mpyrfffi	1kbzww6	nah dude it's on sight	260	2025-04-30 21:11:36	t3_1kbzww6
1	mpyr787	1kbzww6	No, they don't tell you about this but some of...	238	2025-04-30 21:09:59	t3_1kbzww6
2	mpyr7x3	1kbzww6	[Removed by Reddit]	74	2025-04-30 21:10:07	t3_1kbzww6
3	mpyrtzs	1kbzww6	In the professional world, it's actually quite...	127	2025-04-30 21:14:35	t3_1kbzww6
4	mpyt0i7	1kbzww6	Yes, they do get along, but many UCLA people t...	71	2025-04-30 21:23:29	t3_1kbzww6

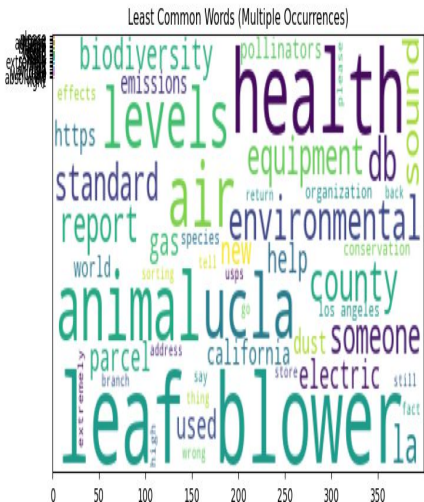
- Most / Least used words

Most: {UCLA, would, know, like,.....}

Least: {weekly, creating, feed, prices, mini.....}



- Subreddit Post (e.g. r/LosAngeles)

[illegible]

Next Procedure

- Classify the sentiment based on the context keywords {like, love, bad, awful....}
- Final result will be based on scoring of the sentiments
- Use a lightweight Local LLM (i.e. Gemma3-2b, Bitnet-2b, or Tiny LM)
- Fine tune model to increase accuracy (optional)
- Implement in docker on Kubernetes (k3s) on Raspberry PI 4 (optional)
- Dashboard with interactive graph (optional)