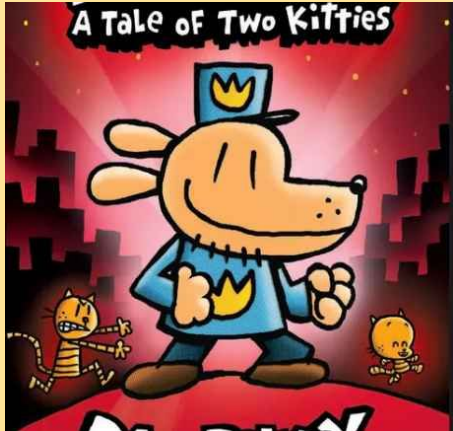# How do you feel about the UCLA?

: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}

- ## Post => Comments

{comment_id, post_id, body_score, created_utc, parant_id}

- **Most / Least used words**

Most: {UCLA, would, know, like,......}
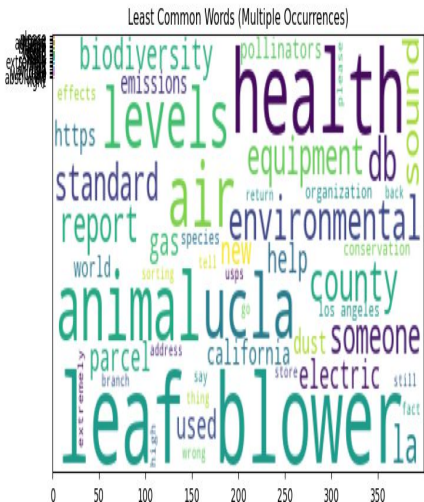
Least: {weekly, creating, feed, prices, mini……}

EDA - 2
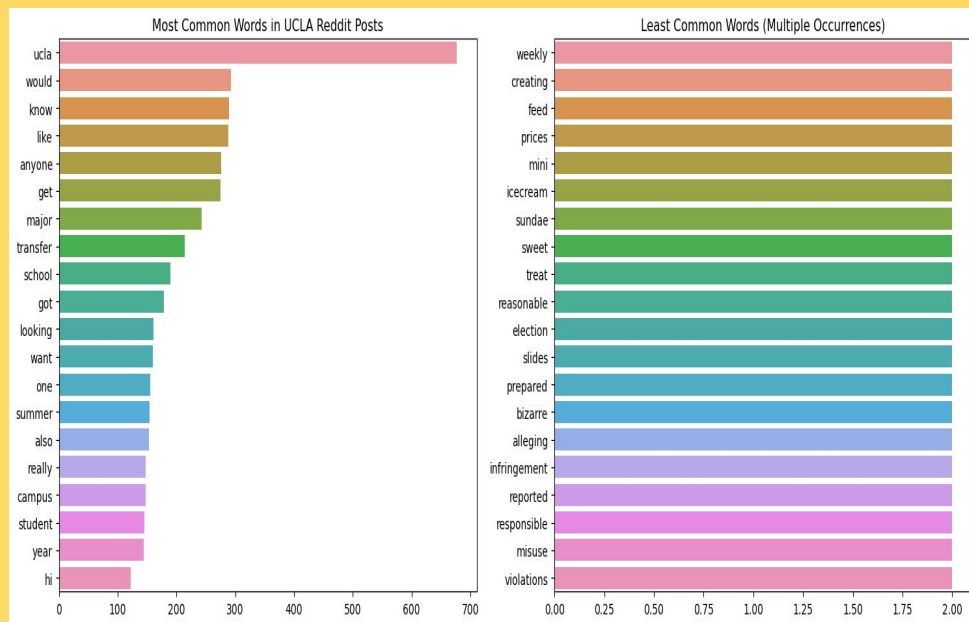
- **Subreddit Post** (e.g. r/LosAngeles)   - **Subreddit Post** (e.g. r/USC)

Next Procedure

- Classify the sentiment based on the context keywords {like, love, bad, aweful….}
- 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)