THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Online Community Monitoring

Tracking Changes in Subreddit Sentiments

Natural Language Processing

DATS 6312, Section 11

Team 3, Final Project

Ashish Telukunta, Jonathan Schild, Thomas Stanton

Project Outline

1. Conceptualization/description
2. Training model to predict emotion and sentiment of text
3. Scrape Reddit data
4. Apply model(s)
5. Plot sentiments over time

Model Approach

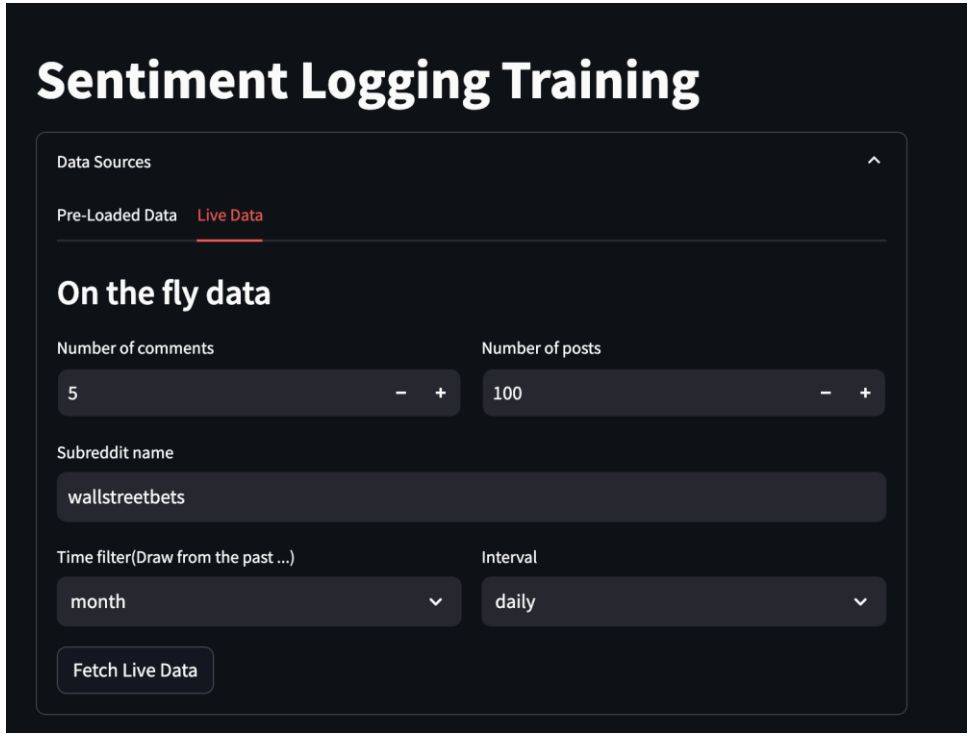
1. Data
2. Model identification
    a) Positive/Negative sentiment
        i. Hugging Face's pipeline
    b) Emotions
        i. BERT1, BERT2 (pseudo-labeling)
        ii. RoBERTa
        iii. ELECTRA
3. Tuning
    a) Hyperparameters
    b) Optimizer, Loss function

Model Performances

Model	Validation Accuracy	F1 Score
BERT1 (w/o Reddit data)	0.9520	0.9582
BERT2 (w/Reddit data)	0.9100	0.9140
Electra	0.9531	0.9530
Roberta	0.9502	0.9502

Reddit Scraper

1. Praw
2. Duration restrictions: day, week, month, year
    a) No ability to specify date ranges
3. Data retrieved
    a) Posts & comments; specify number of each
    b) Dates; enable grouping and averaging

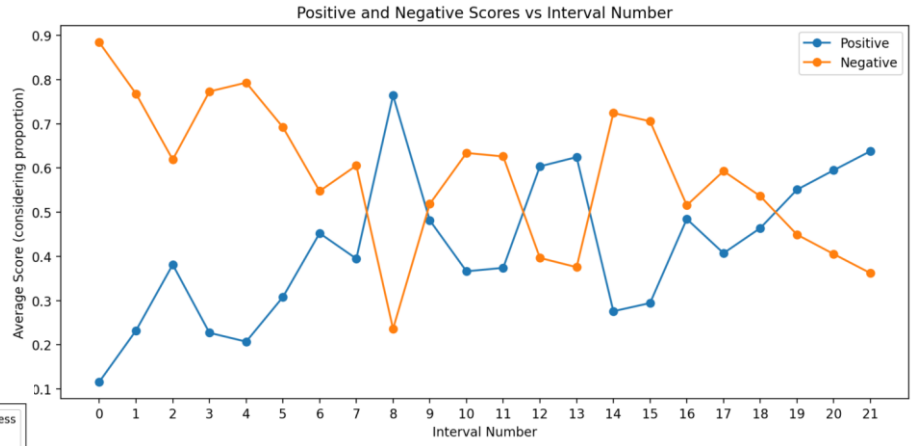


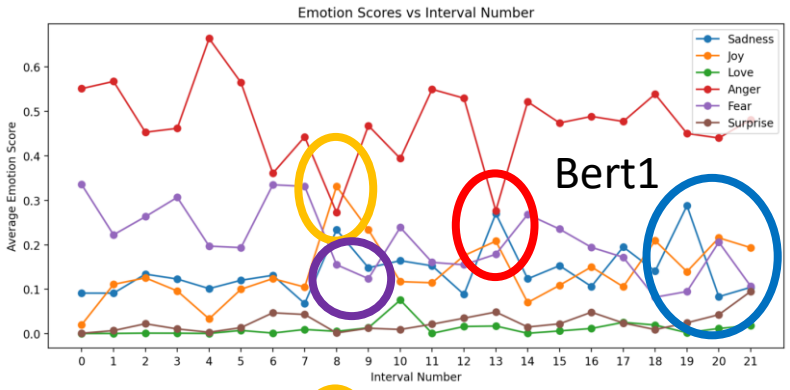
Model Application: Preview

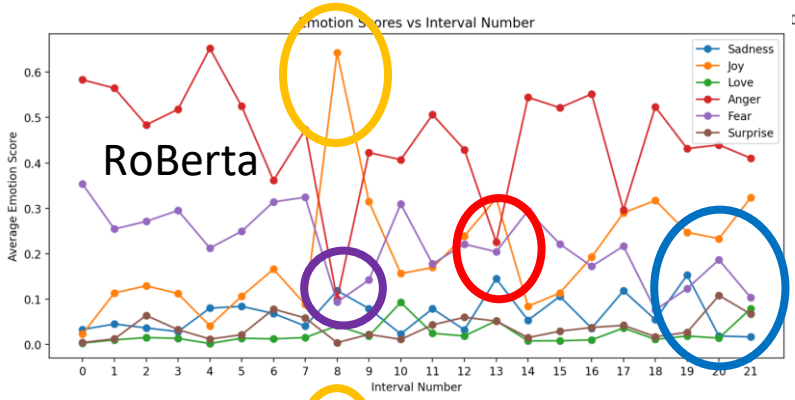
1. Scrape parameters
    a) Top 40 comments for the top 50 posts
    b) For past month, grouped to daily intervals

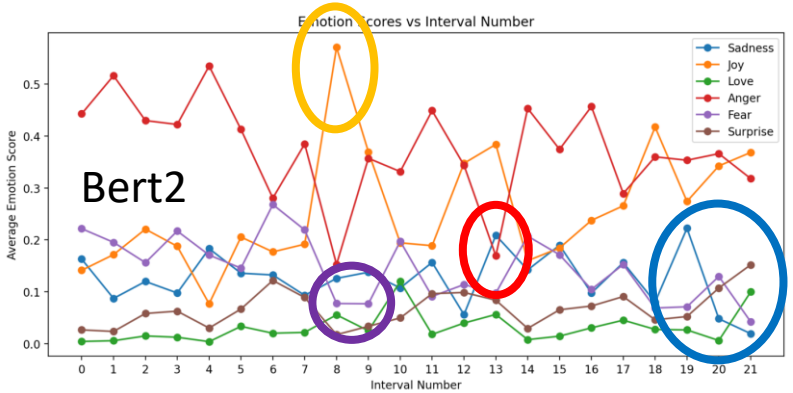
2. Examples…
    a) r/pics
    b) r/funny
    c) r/wallstreetbets
    d) r/gwu

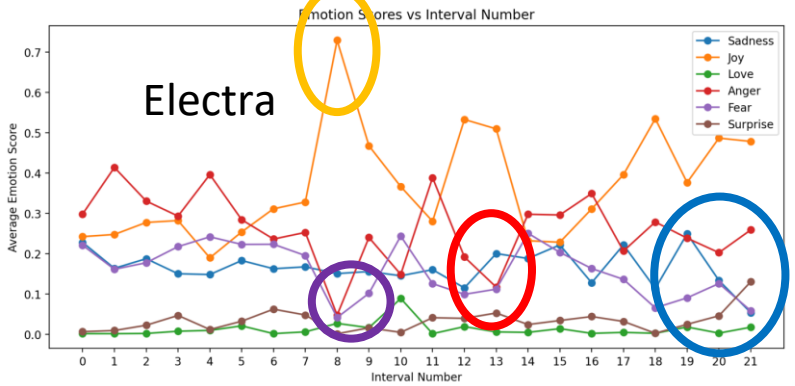
r/pics…


Positive and Negative Scores vs Interval Number
Positive
Negative
Average Score (considering proportion)
Interval Number

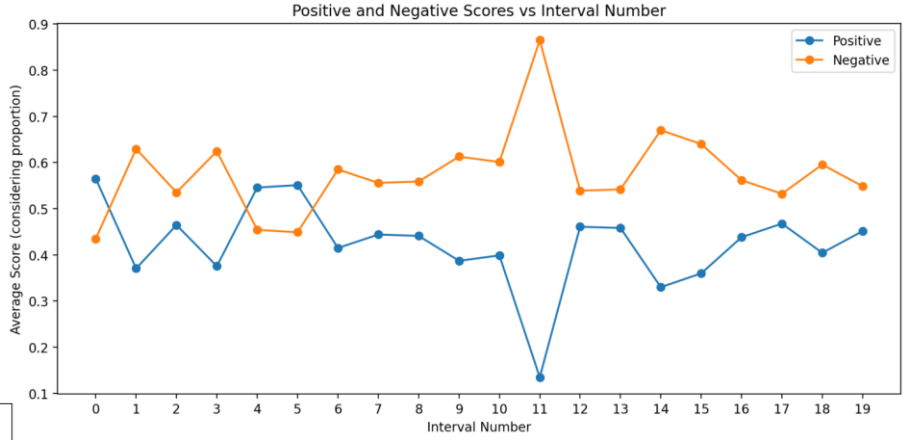

Emotion Scores vs Interval Number
Sadness
Joy
Love
Anger
Fear
Surprise
Average Emotion Score
Interval Number
Bert1

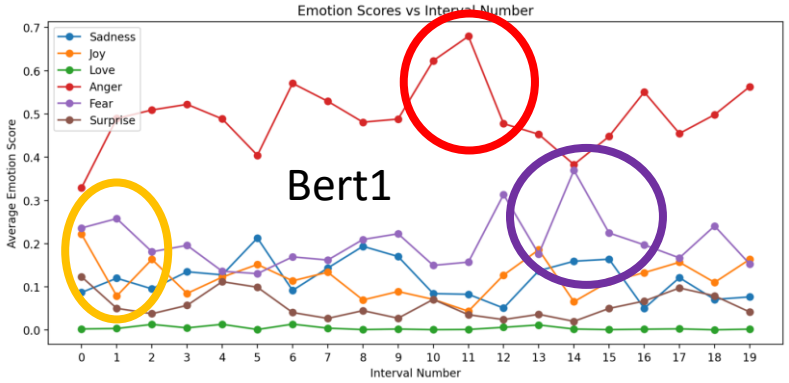

Sadness
Joy
Love
Anger
Fear
Surprise
Average Emotion Score
Interval Number
RoBerta

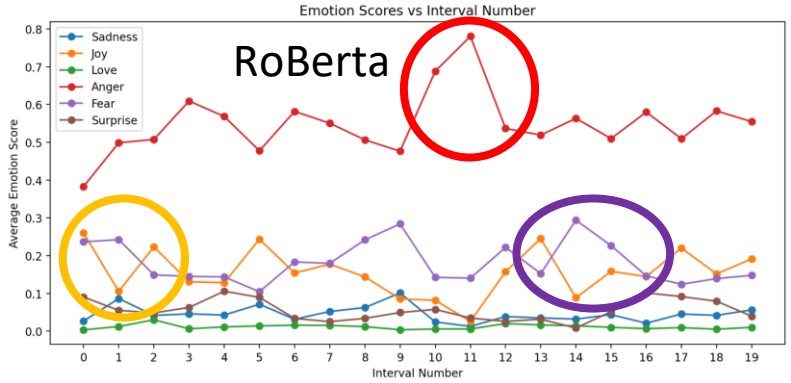

Sadness
Joy
Love
Anger
Fear
Surprise
Average Emotion Score
Interval Number
Bert2

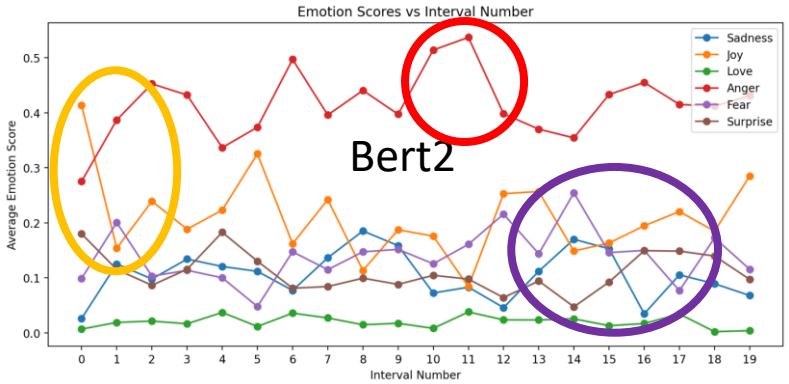

Sadness
Joy
Love
Anger
Fear
Surprise
Average Emotion Score
Interval Number
Electra

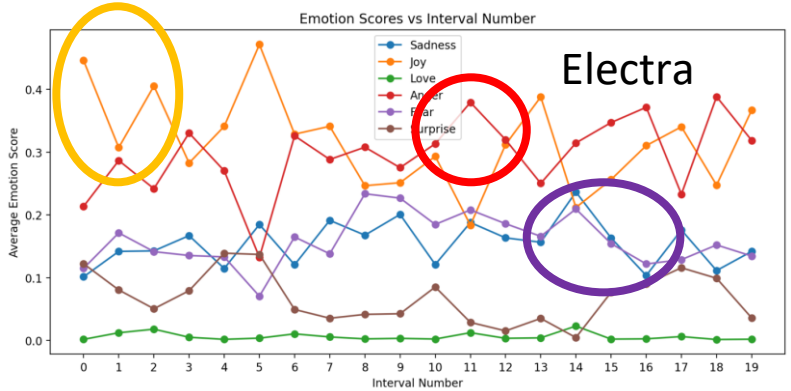
r/funny…


Positive and Negative Scores vs Interval Number
Positive
Negative
Average Score (considering proportion)
Interval Number

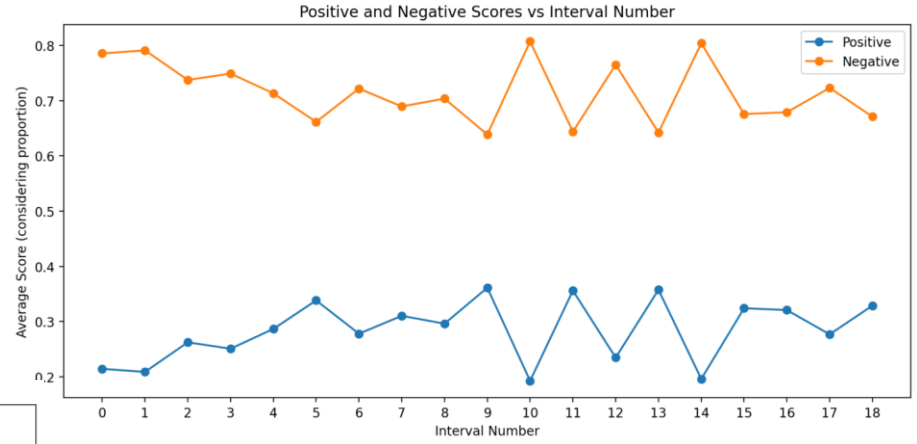

Emotion Scores vs Interval Number
Sadness
Joy
Love
Anger
Fear
Surprise
Bert1
Average Emotion Score
Interval Number

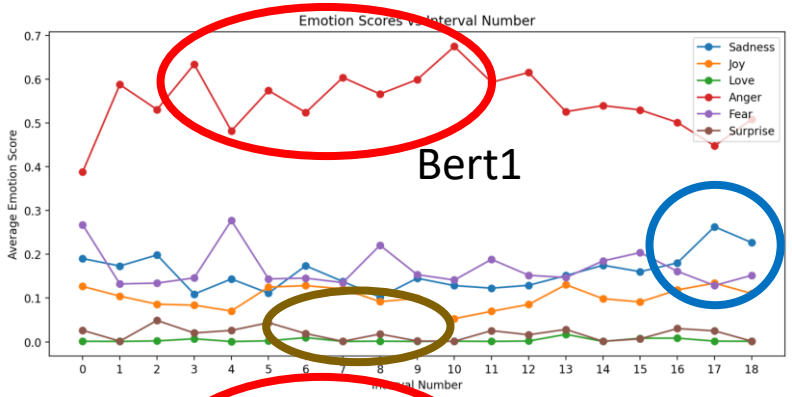

Emotion Scores vs Interval Number
Sadness
Joy
Love
Anger
Fear
Surprise
RoBerta
Average Emotion Score
Interval Number

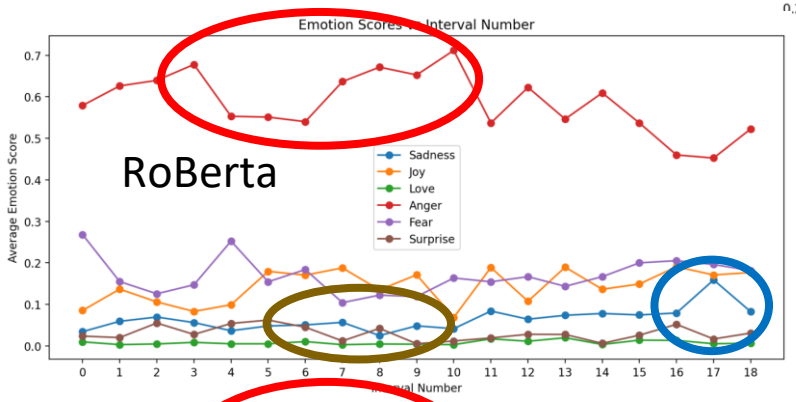

Emotion Scores vs Interval Number
Sadness
Joy
Love
Anger
Fear
Surprise
Bert2
Average Emotion Score
Interval Number

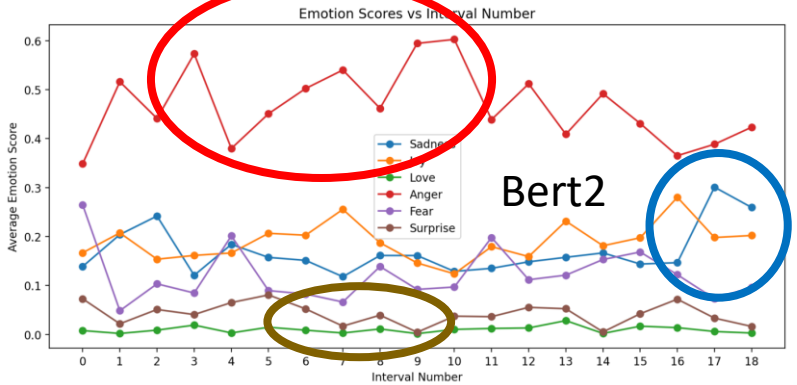

Emotion Scores vs Interval Number
Sadness
Joy
Love
Fear
Surprise
Electra
Average Emotion Score
Interval Number

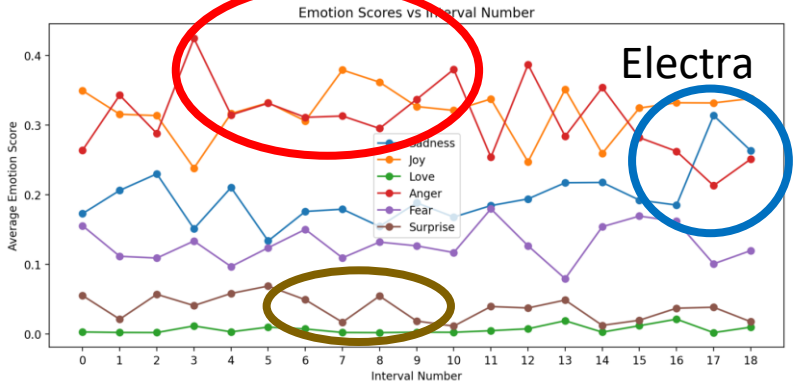
r/wallstreetbets...











Demonstration...