

---



# Predicting Engagement: Social Network Analysis and Deep Learning Techniques

Raffi Mesrobian





# Introduction

---

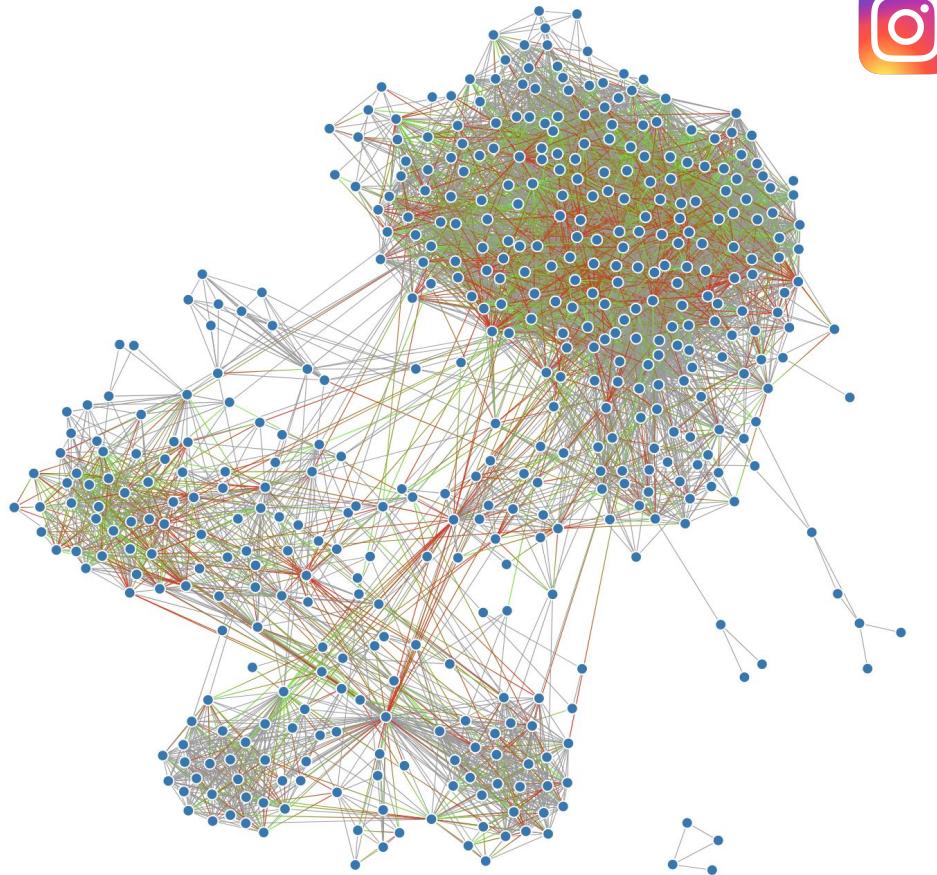
In recent years, social networks have played an increasing role in our daily lives. The social network has become a priority for brand marketing, with Instagram being a platform with one of the highest engagement rates. The data from these networks can be useful for other things, too. [This study](#) (1) uses publicly available data from Reddit to examine the effectiveness of different methods that can provide an early detection of major depressive disorders (MDDs).

The objectives of this project are to visualize my social network, applying graph theory to better understand it and to create a model that will predict the engagement level of a post.



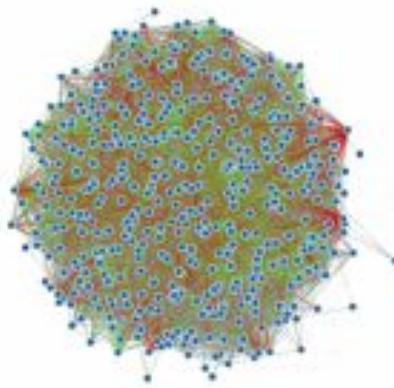
# Visualization

- Process:
  - Two web scraping bots
  - txt files to json (d3.js library)
- Network properties:
  - 463 nodes → 11,988 edges
  - Density: 5.6%
  - Avg. Degrees: 25.9
  - Avg. Shortest Path Length: 2.9
- This [gist](#) and [medium](#) post are great references for those who are interested in visualizing their network and learning about global and local network analysis



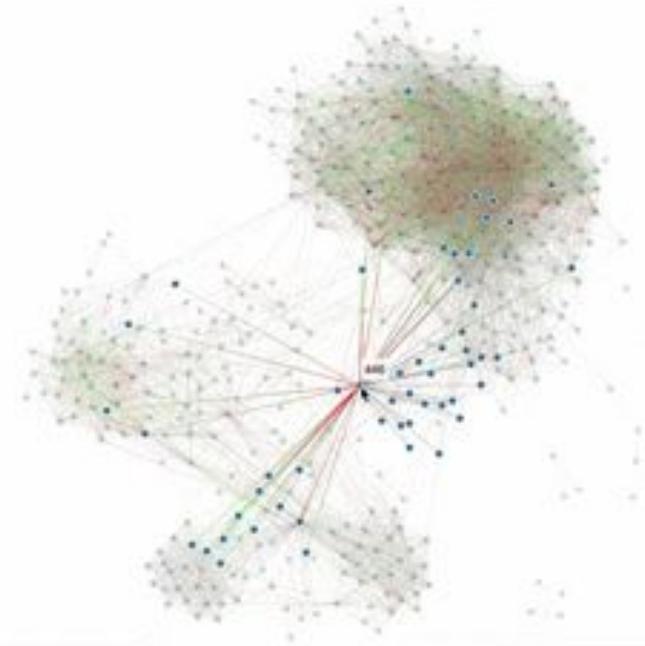
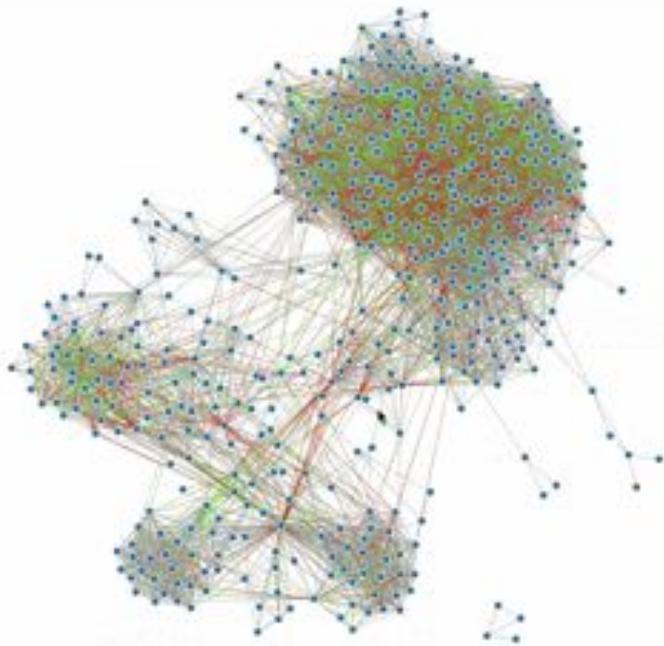


## Force Directed Graph



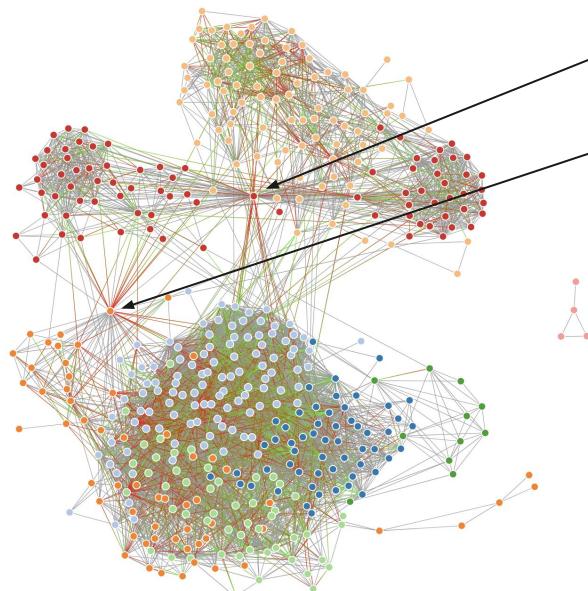


## Force Directed Graph (cont.)

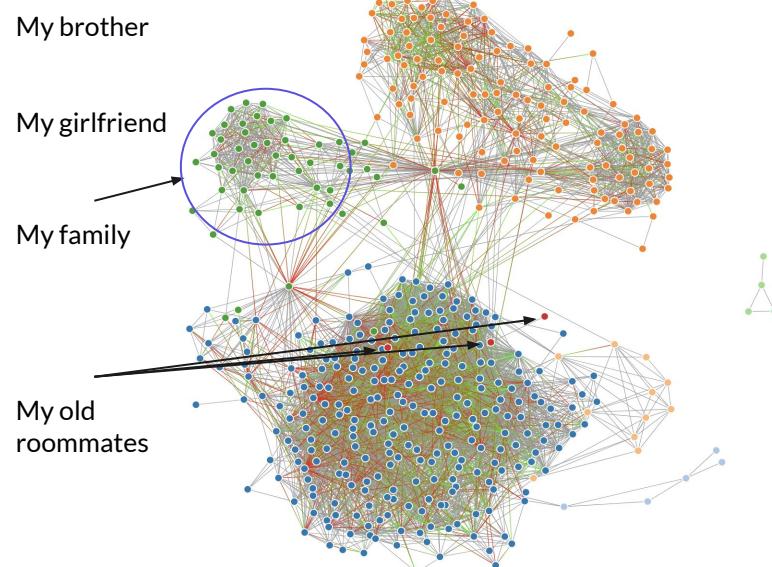




# Clustering



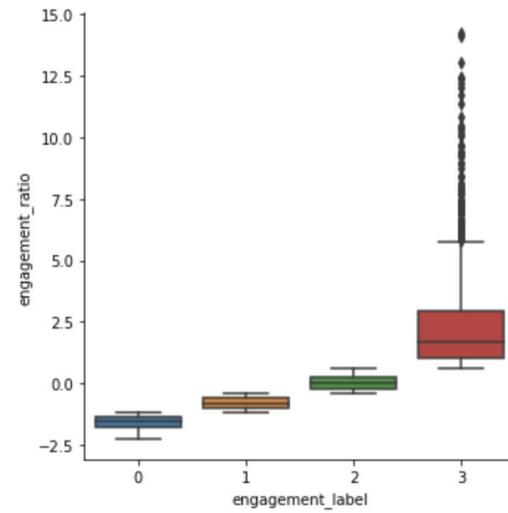
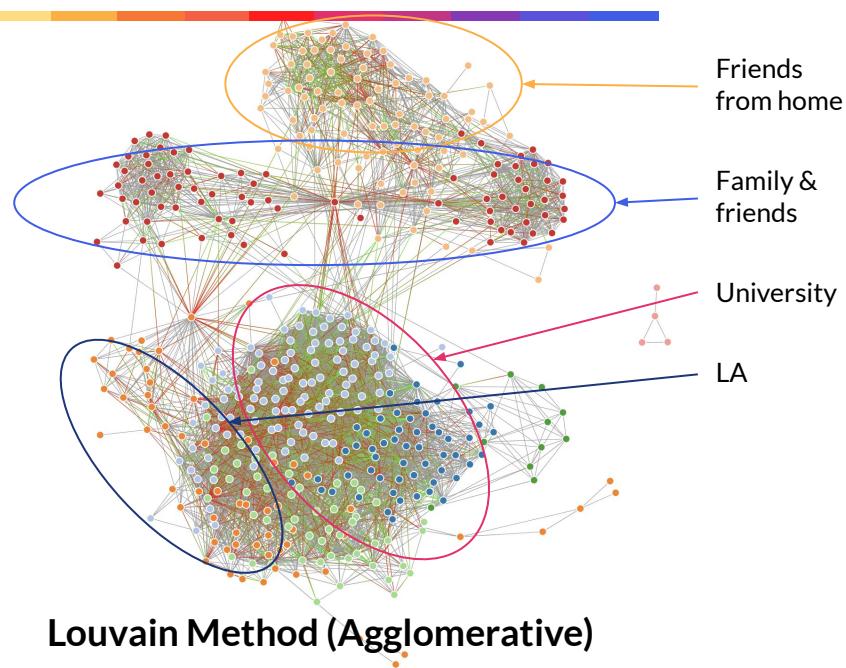
Louvain Method (Agglomerative)



Girvan-Newman Method (Decisive)



# Classification



Note: LR = likes / followers; CR = comments / followers; Mean Norm LR (MNL) =  $(LR - \text{avg. } LR) / \text{std. } LR$ ; Mean Norm CR (MNC) =  $(CR - \text{avg. } CR) / \text{std. } CR$ ; Engagement Ratio = MNL + MNC



# Post Data

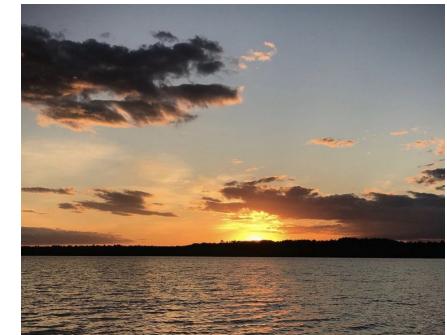
The image displays a composite view of an Instagram post. On the left is a photograph of a colorful mural on a wall behind a chain-link fence. On the right is the Instagram post interface with various data points labeled by arrows:

- Username:** Following (top left of the post area)
- Location:** Brooklyn, New York (top center of the post area)
- Caption text:** Missing that thicc, palpable, sweet august heat and the sweaty friends who made it so highly tolerable (center of the post area)
- Comment count:** 3w (time ago)
- Comment text:** 1 like, Reply (comment 1)
- Comment text:** 1 like, Reply (comment 2)
- Comment text:** I want that grilled (comment 3)
- Image URL:** The main image itself (the mural)
- Like count:** Liked by [redacted] and 75 others (bottom left of the post area)
- Type of image:** A photograph (indicated by the camera icon)
- Date and time of post:** DECEMBER 17, 2020 (bottom center of the post area)



# Classification Challenges

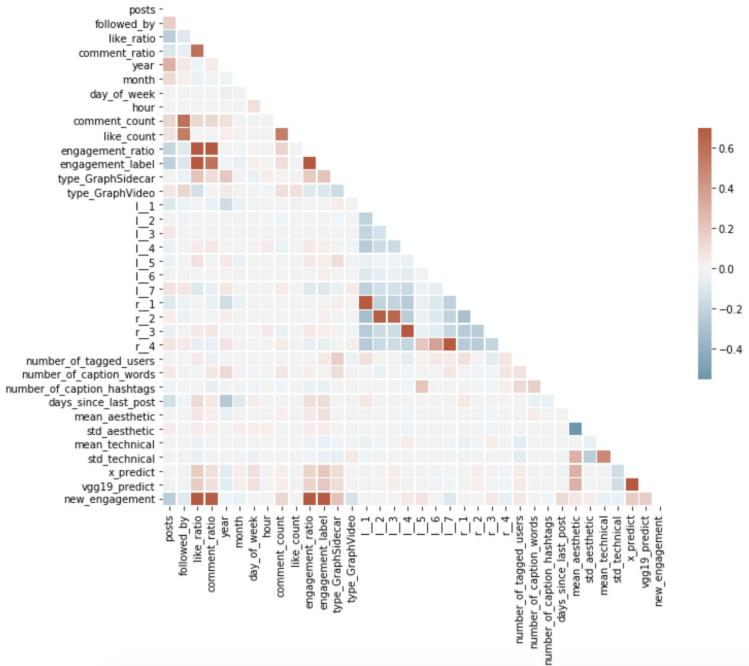
- Images:
  - Xception & VGG-19
  - Accuracy: ~33%
  - 😞
- Natural Language Processing:
  - Multinomial NB, SGD, Logistic Regression
  - Accuracy: ~45%
  - 🤔



Like Ratio:	<b>0.14</b>	<b>0.24</b>
Comment Ratio:	<b>0.002</b>	<b>0.007</b>
Caption:	Spent a week in the mountains wondering what month it was. Turns out it's Juntober. Looking forward to Febgust.	null
Hashtags:	#undersocialized, #oversaturated	null
Community:	<b>1</b>	<b>0</b>
Engagement Group:	<b>1</b>	<b>2</b>



# Regression

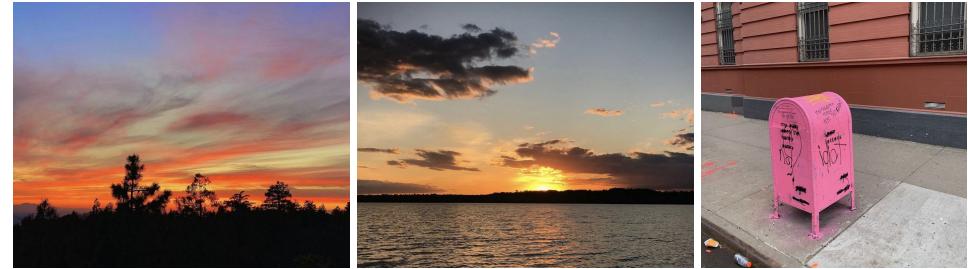
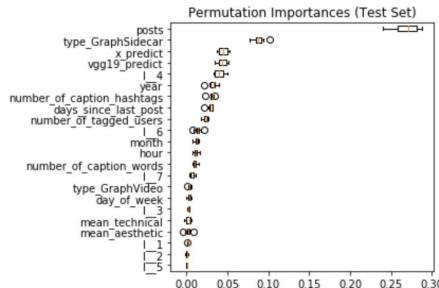


Engagement!



# Regression (cont.)

- **Images:**
  - Xception MSE: 3.11
  - VGG-19 MSE: 3.11
- **Tabular Data:**
  - LR MSE: 2.66
  - SVD MSE: 3.20
  - GB RF MSE: **1.89**



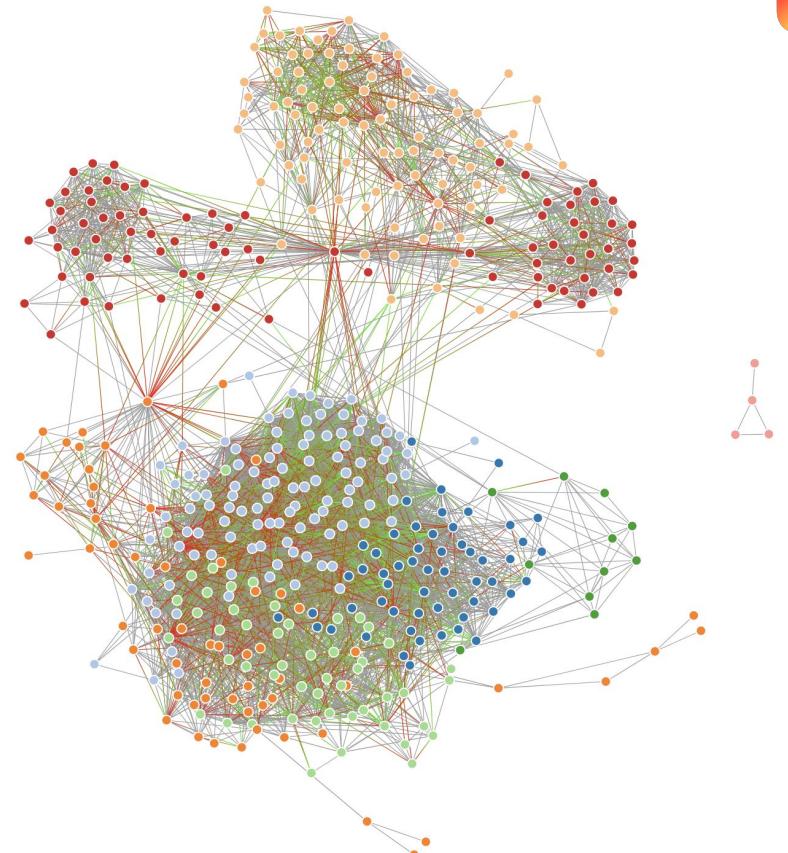
	<b>Engagement Ratio:</b>	<b>-0.775</b>	<b>0.484</b>	<b>-0.275</b>
Xception:	-0.009	-0.006	-0.106	
VGG-19:	-0.009	-0.006	-0.106	
Linear Regression:	-0.086	0.521	1.455	
Singular Value Decomposition:	0.012	0.440	1.392	
Random Forest:	-0.540	0.640	0.392	



## Next Steps

---

- Scrape more data / add more features
- OpenCV + Tesseract → brand logos
- Become an influencer





# Thank you!

Xception Image Labels for top 100 most engaging posts:

**Siamese\_cat**, marimba, maypole, Band\_Aid, **sunglasses**, Windsor\_tie, theater\_curtain, stole, candle, Brittany\_spaniel, restaurant, Sunglass, suit, paddle, cliff, pajama, cowboy\_hat, bikini, coho, studio\_couch, ice\_lolly, alp, jack-o-lantern, coil, academic\_gown, web\_site, altar, swing, sunscreen, stage, **hot\_pot**, palace, swimming\_trunks, seashore, sarong, jean, **burrito**, stingray, ski, **wig**, shower\_cap, Pembroke, sombrero, Maltese\_dog, breakwater, book\_jacket, velvet, **giant\_schnauzer**, cannon, seat\_belt, rugby\_ball, lakeside, valley, bib, pop\_bottle, **diaper**, comic\_book, cradle, tub, wok, cliff\_dwelling, Chesapeake\_Bay\_retriever, pay-phone, hoopskirt, racer, Australian\_terrier, **mobile\_home**, slot, croquet\_ball, vase, **lab\_coat**, maillot, motor\_scooter, water\_jug, gown, balance\_beam, abaya, canoe

Hashtags in top 100 most engaging posts:

#WhenToomsBecomeOne, #backstreetboys, #dnatour2019, #bsbmil, #cincodemacha, #Cindyisababe, #SHS, #randomactsofkindness, #motherslove, #motherdaughter, #loveoflife, #loveofmother, #godsmasterpiece, #actsofkindness, #kindness, #kindnessisfree, #kindnessrocks, #mothers, #cousinlove, #sisterslove, #payitforward, #mariokart, #mariokart8deluxe, #mariokartpumpkin, #mariopumpkin, #nintendo, #nintendopumpkin, #tanooki, #tanookimario, #spazzadallas, #tanookipumpkin, #mayorann, #tanookimariopumpkin, #900lumens, #varmit, #racingslicks, #marieloseszinnyc, #mustardinmyhair, #hotdogger, #hotdogsog, #birthdaydog, #nightmaresquad, #fallofsaigon, #Bozscaggs, #midnighttrain, #secondbank, #fazolis, #alpastor, #1917, #elchapo, #apocalypto, #minne, #gogops, #bigcat, #ChallengeAccepted, #pgachampionship, #womensupportingwomen, #sacrécoeur#sacrecoeur#frame, #RollTide, #nfl, #oakley, #oakleysunglasses, #theGrandestCanyon, #LTD, #liv michelleconnolly, #baltimore Ravens, #ravens, #lamarjackson, #truzz, #bigtruss, #womenempowerment, #acrylicpainting, #acrylic, #acrylicpaint, #football, #portrait, #portraitpainting, #painting, #fantasyfootball, #baltimore, #team43, #worldracingleague, #baltimoremaryland, #contemporaryart, #fineart, #art, #artwork, #latheater, #cantzophis, #nflmemes, #mvp, #usmc, #wrl, #enduranceracing, #miata, #prettyyoungthing, #wishesluck, #thisisbeyondnnndme, #felixjohnstone, #munclegiarekracing, #gp4, #womenempoweringwomen, #oscarwilde, #devilpup, #teufelswelpe, #hydrate, #gerberMRE, #america, #glendale, #brandpark, #missprism, #MostFunWins, #MarriedMurphy, #TeamSpecialSauce, #Shamrocks, #ScoochButts, #ThankYouEveryone, #adelinafahey

Raffi Mesrobian

[rmesrobi](#)

[raffimesrobian](#)

[rpmesrobian](#) 13



BeautifulJoup



Keras



NetworkX  
Network Analysis in Python



NumPy



pandas



SciPy



Selenium



seaborn

TensorFlow