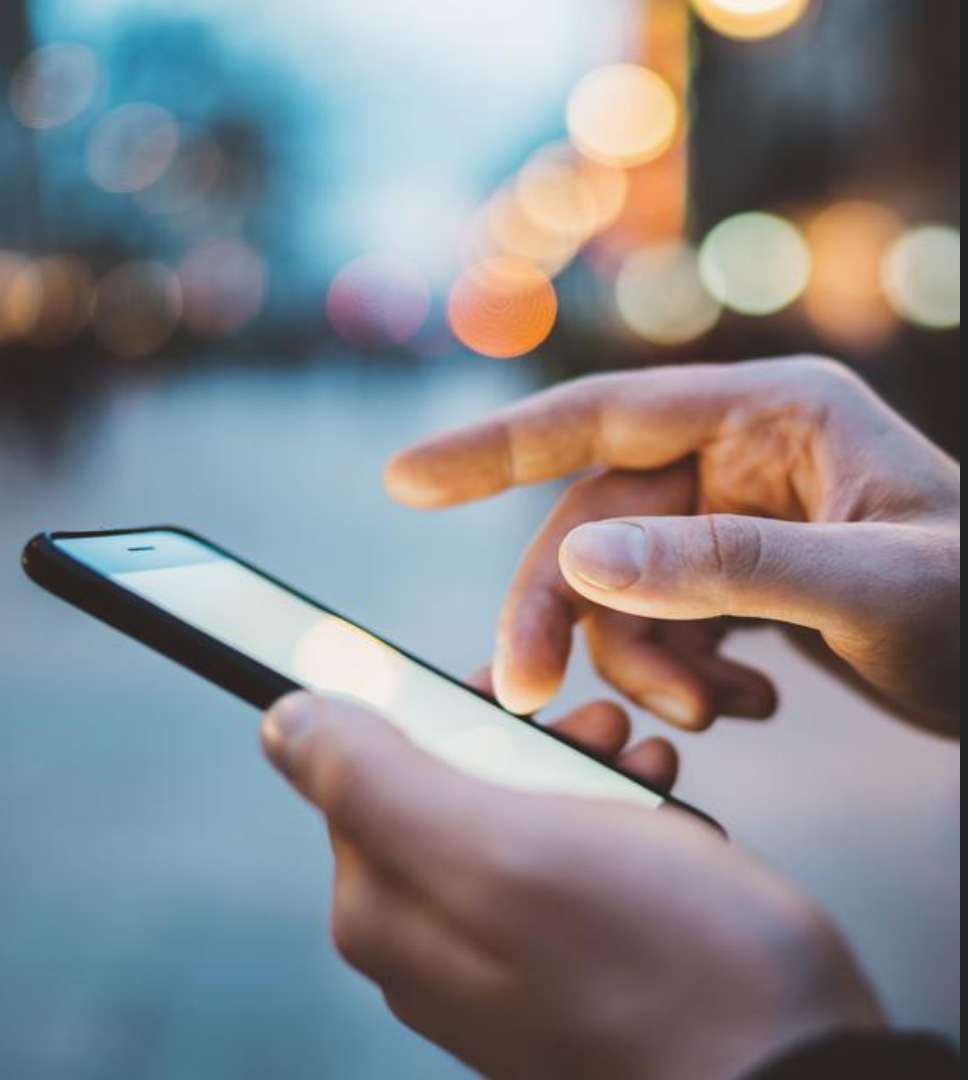
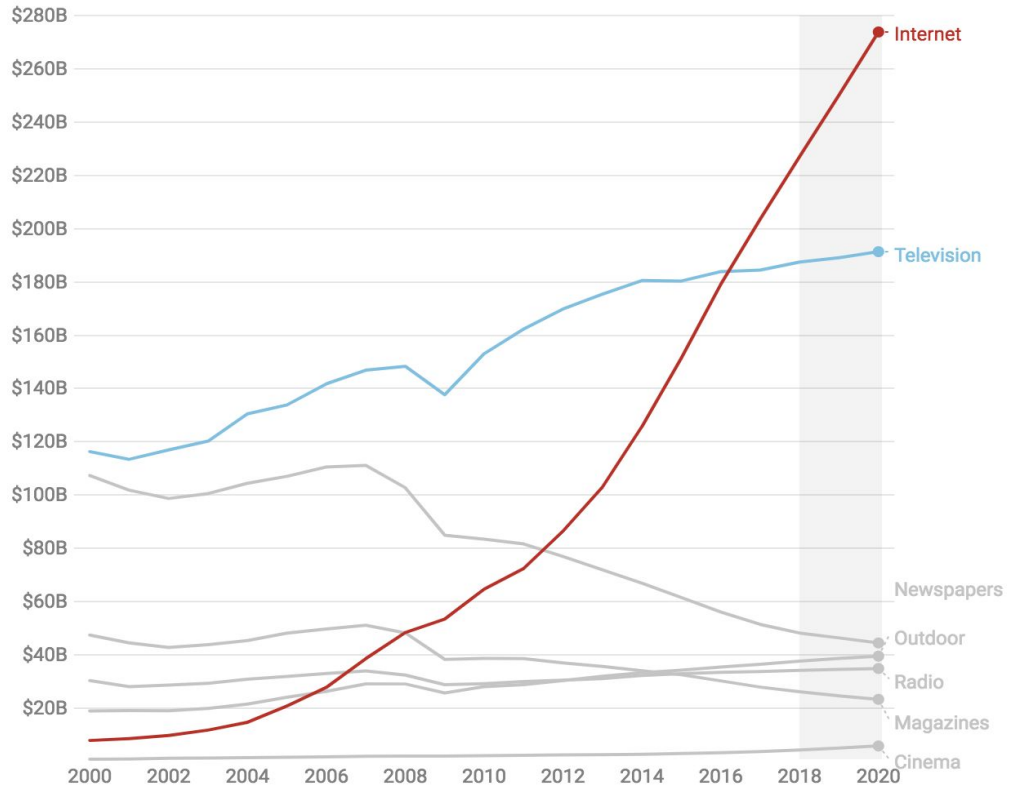


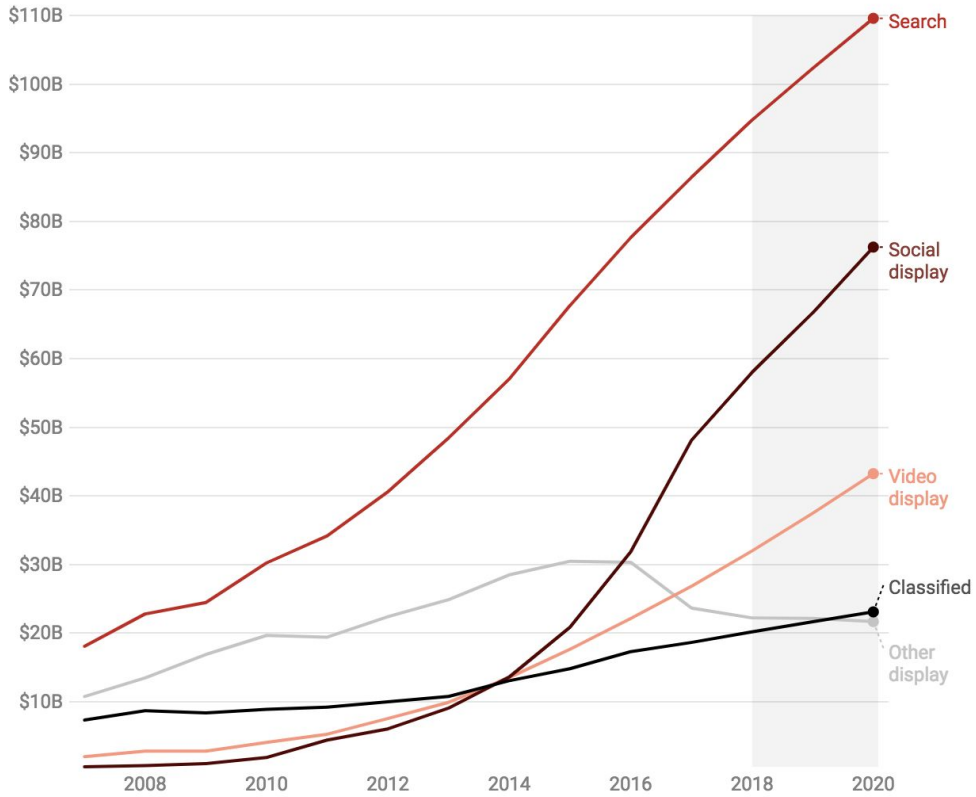
InstaPromote

By Arjun Sarup, Rajandeep Singh, Sajel Shah, VIneet Yellu, Yash Bhate

Problem we are solving



Problem we are solving



Problem we are solving

Importance of Social Media Advertising:

- Fastest growing form of digital advertising
- Higher level of user engagement
- Higher relevance of social media ads => larger ROI (target)

Problem we are solving

Why Instagram:

- 1B+ MAU
- Fastest growing social network
- Visual nature => extremely effective for certain industries

98% Of fashion brands

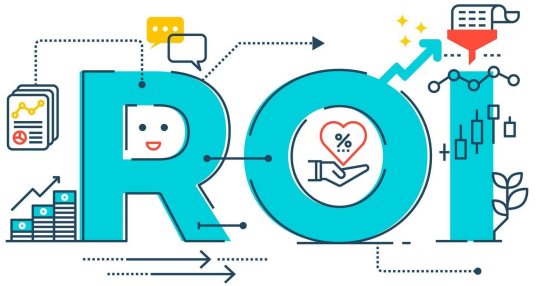
Problem we are solving

Suggest optimal brand influencers for each use case/industry

Suggest the best photo subjects to maximize likes

Predicts the expected number of likes

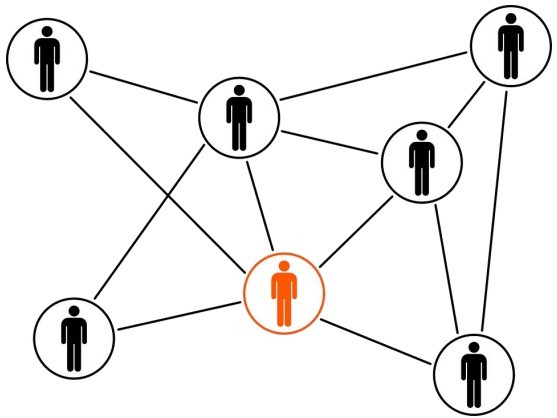
=> Increased ROI and lower advertising costs



Our Approach (Part 1)

Problem: Choose an optimal brand influencer

Primary Question: How many likes notable brand influencers would get for a particular post. What features in a post would correlate to a higher number of likes?



Our Approach (Part 2)

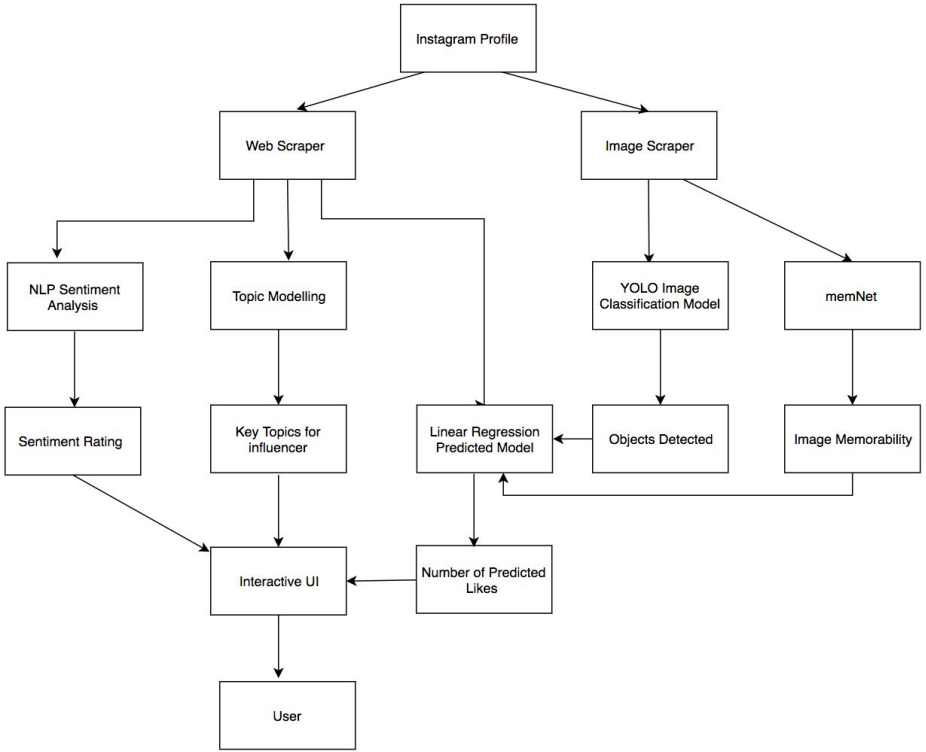
Solution: Creating a like prediction model for the influencers - to improve accuracy we use historical, categorical data.

Our Approach (Part 2)

Concern during Iteration: Present an overall "portfolio" of the brand influencer besides predicting the number of likes they would receive.

Solution: memnet classification, sentiment analysis, topic modelling.

Architecture of Solution



Web Scraper

- Tool integrates instagram get requests, BeautifulSoup, Selenium and various HTTP protocols
- Collected the media metadata - comments, user post information, followers count, etc

Model

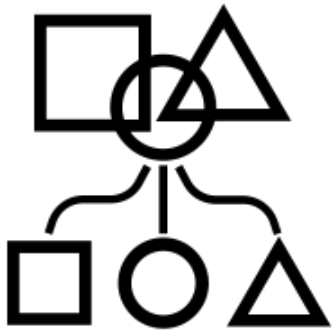
- Built a Data pipeline to include this data into a working environment
- Use these features as inputs for the like predictor and perform sentiment analysis on comments

Image Scraper

- Modified a tool to collect images and videos and story from any profile
- Stores images in the correct format in chronological order to perform further analysis

Image Classification

- Apply YOLO framework to parse the image and extract features
- Use memNet CNN structure to decipher memorability of image to ultimately improve accuracy of like predictor
- Combine with the results of Web Scraper to improve the overall accuracy



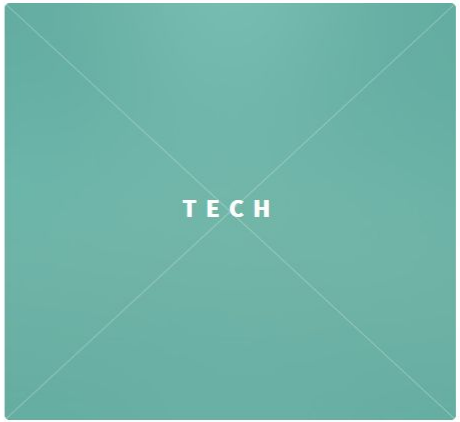
Intended User Interface

INSTAPROMOTE

This is InstaPromote. The best way for you to find the right influencer for YOU.

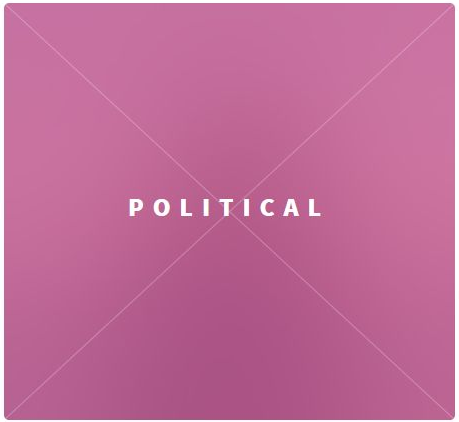
InstaPromote uses Machine Learning algorithms to help you understand Instagram influencers better, so that you can find the perfect match to promote your brand.

We provided services ranging from sentiment analysis to like predictors. Go ahead and choose your industry area or enter the influencer you want to know more about.



Choose Influencer

Select a Fasion Influencer you want to know more about.



OR

ENTER INFLUENCER'S HANDLE

aliaabhatt

SUBMIT

INSTAPROMOTE

Tell us about the post you plan to make

FORM

Check out these really cool shoes!!! Thanks Nike!!!

https://i.imgur.com/sZPOFHk.png

Date

05 / 09 / 2019

Time

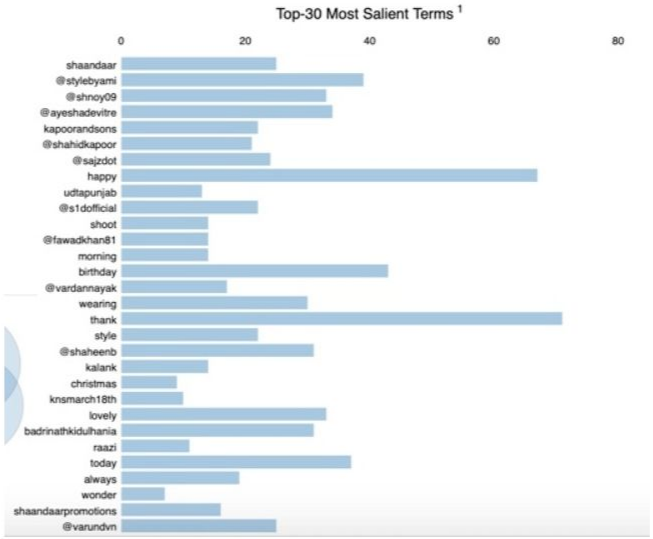
12 : 00 PM

SUBMIT

@aliaabhatt

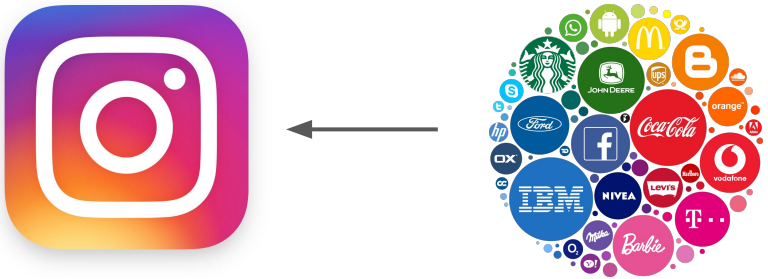
PREDICTED LIKES: 1156650

AVERAGE SENTIMENT OF FOLLOWERS:
81.2% POSITIVE
10.3% NEUTRAL
8.5% NEGATIVE



Summary of learning path

- Initial idea: A broader-themed social media optimizer which was going to scrape data off of LinkedIn, Insta, Tumblr etc. to make useful predictions for industry.
- Refinement needed: Too generic. Needed a specific industry-domain to play around with. Chose optimal brand influencers on Instagram



Summary of learning path

- Going beyond qualitative and immediate-location based data for like predictor
  - Outcome: incorporated historical/categorical data as features for a higher accuracy through a previous_likes metric, epoch and a YOLO object detection model

Summary of learning path

- Suggesting subjects to include in the photo
  - Outcome: because of YOLO detection, we are going to release the most important items/people to include in the post image; based on the trained weights of the categories/objects for each influencer

Summary of learning path

- Modifying the scraper we used to get the top 100 comments for each post so that we could use the sentiment of each post in our feature set
    - Outcome: successfully modified the web scraper, couldn't carry out sentiment analysis before the presentation because we just got the web scraper to work properly
- Using topic modelling to analyze the correlation between topics and likes
    - Outcome: incorporated into portfolio

Summary of learning path

- Constructing a hashtag popularity metric to make a hashtag_used score as a feature
  - Outcome: made a heuristic to track hashtag popularity by searching for #of posts related to that hashtag in the timeframe that the post was made
  - Ultimately rejected the idea because of amount of scraping needed on the fly, could use suggestions

Summary of learning path

- Using the memorability score of each image (calculated using Memnet) as a feature in our predictive model and as a scoring/ranking system for the image that the advertiser wants the influencer to promote
  - Outcome: Memnet is a caffe-based model --> need help translating it to Keras

Questions

- Using memnet to estimate visual memorability - the architecture involves CNNs but we cannot seem to implement the Caffe library to build the neural network
- For the NLP Classification - how do we integrate VADER (Valence Aware Dictionary and Sentiment Reasoner) a tool built specially for social media analysis into the prediction model

Github Link

https://github.com/arjunsarup1998/team_30_data-x