RIGHTFLUENCER

Arin Ghosh
Karanjit Singh Tiwana
Manoj Karthick Selva Kumar

SFU

WHAT IS INFLUENCER MARKETING ?

Finding influential people in your product category to attract buyers by marketing your brand through them.

"Building relationships with people who can build relationships for you"

MOTIVATION

- Around 70% of marketers consider finding the right influencers to be the toughest part of their job.
- Exponential growth of social influencer marketing over the past few years; Estimated to be worth $1.07 billion.
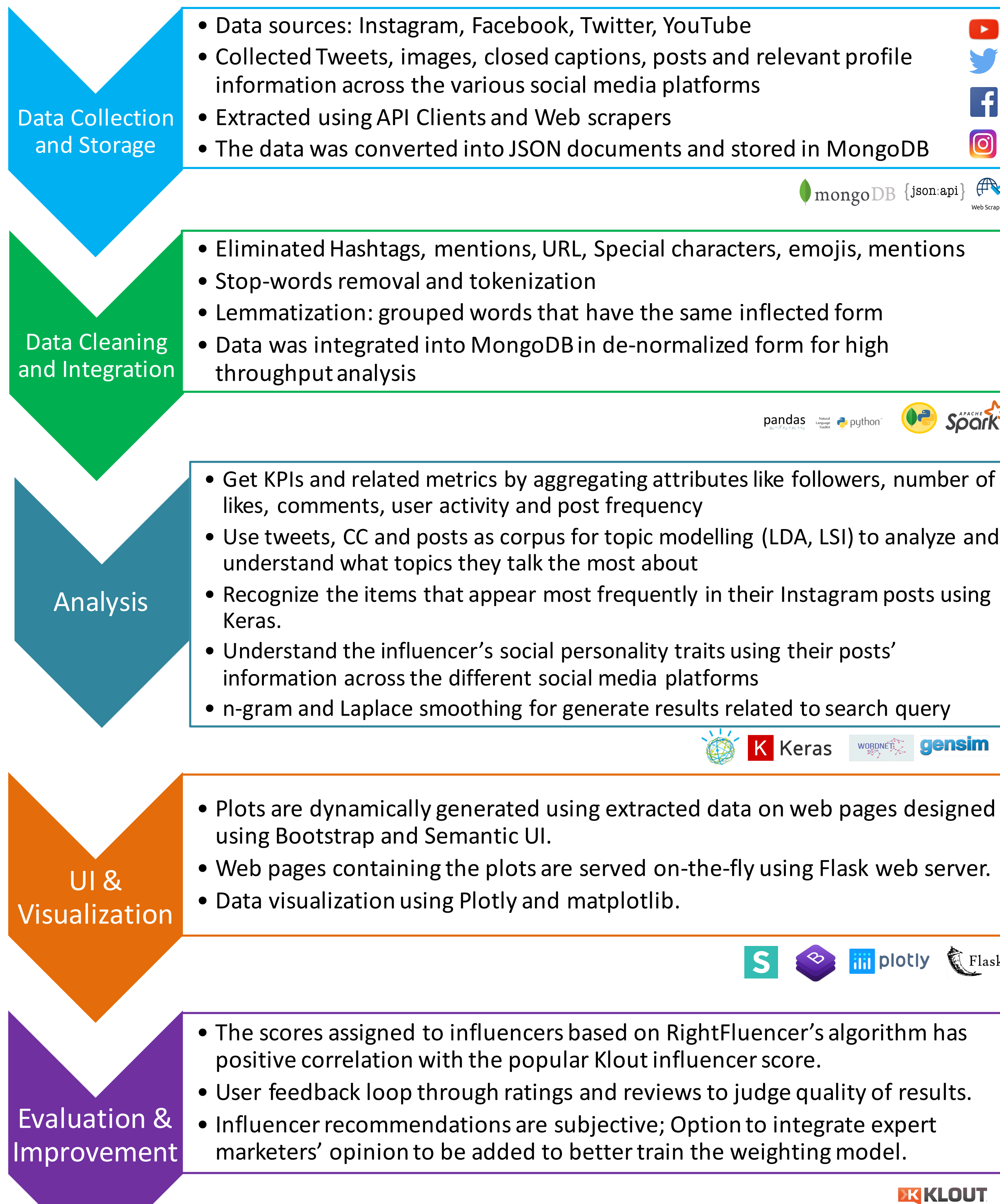- Narrow focus of current products on just metrics; fails to capitalize on user's niche and post content.

WHAT IS RIGHTFLUENCER ?

An interactive web application and dashboard that allows you to find the right influencers for your product and category by analyzing their posts, images and videos.
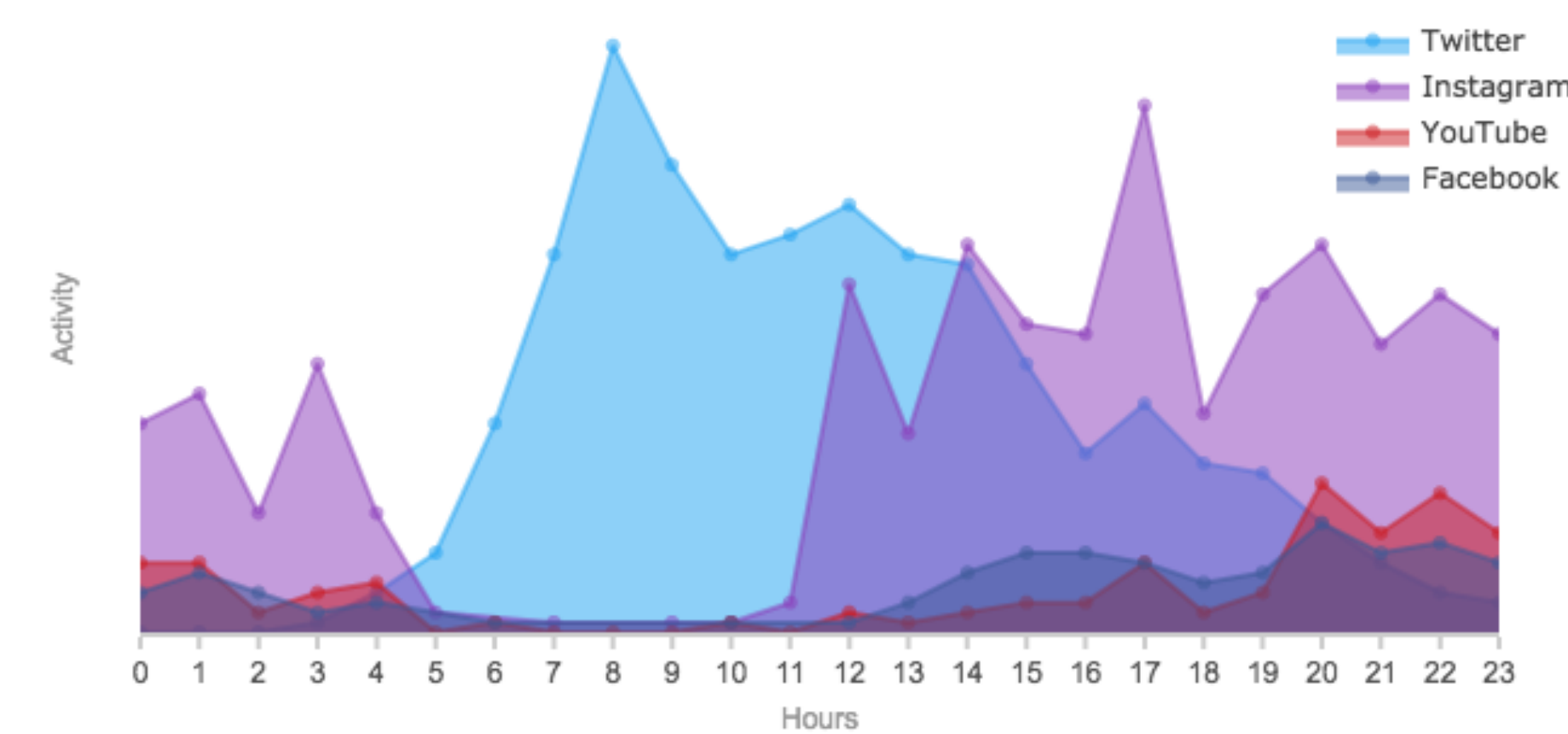
KEY DIFFERENTIATORS

- Collective analysis of multiple social media platforms in the same place using both metrics and content.
- Inspection of images to understand what products influencers are interested in to find their niche.
- Analyzing YouTube videos of influencers using closed captions to better interpret their expertise.
- Influencer score based on product and category.

DATA PROCESSING PIPELINE

Data Collection and Storage
- Data sources: Instagram, Facebook, Twitter, YouTube
- Collected Tweets, images, closed captions, posts and relevant profile information across the various social media platforms
- Extracted using API Clients and Web scrapers
- The data was converted into JSON documents and stored in MongoDB

Data Cleaning and Integration
- Eliminated Hashtags, mentions, URL, Special characters, emojis, mentions
- Stop-words removal and tokenization
- Lemmatization: grouped words that have the same inflected form
- Data was integrated into MongoDB in de-normalized form for high throughput analysis

Analysis
- Get KPIs and related metrics by aggregating attributes like followers, number of likes, comments, user activity and post frequency
- Use tweets, CC and posts as corpus for topic modelling (LDA, LSI) to analyze and understand what topics they talk the most about
- Recognize the items that appear most frequently in their Instagram posts using Keras.
- Understand the influencer's social personality traits using their posts' information across the different social media platforms
- n-gram and Laplace smoothing for generate results related to search query

UI & Visualization
- Plots are dynamically generated using extracted data on web pages designed using Bootstrap and Semantic UI.
- Web pages containing the plots are served on-the-fly using Flask web server.
- Data visualization using Plotly and matplotlib.

Evaluation & Improvement
- The scores assigned to influencers based on RightFluencer's algorithm has positive correlation with the popular Klout influencer score.
- User feedback loop through ratings and reviews to judge quality of results.
- Influencer recommendations are subjective; Option to integrate expert marketers' opinion to be added to better train the weighting model.

VISUALIZATIONS

Influencer activity by hour

Monthly user interaction quotient
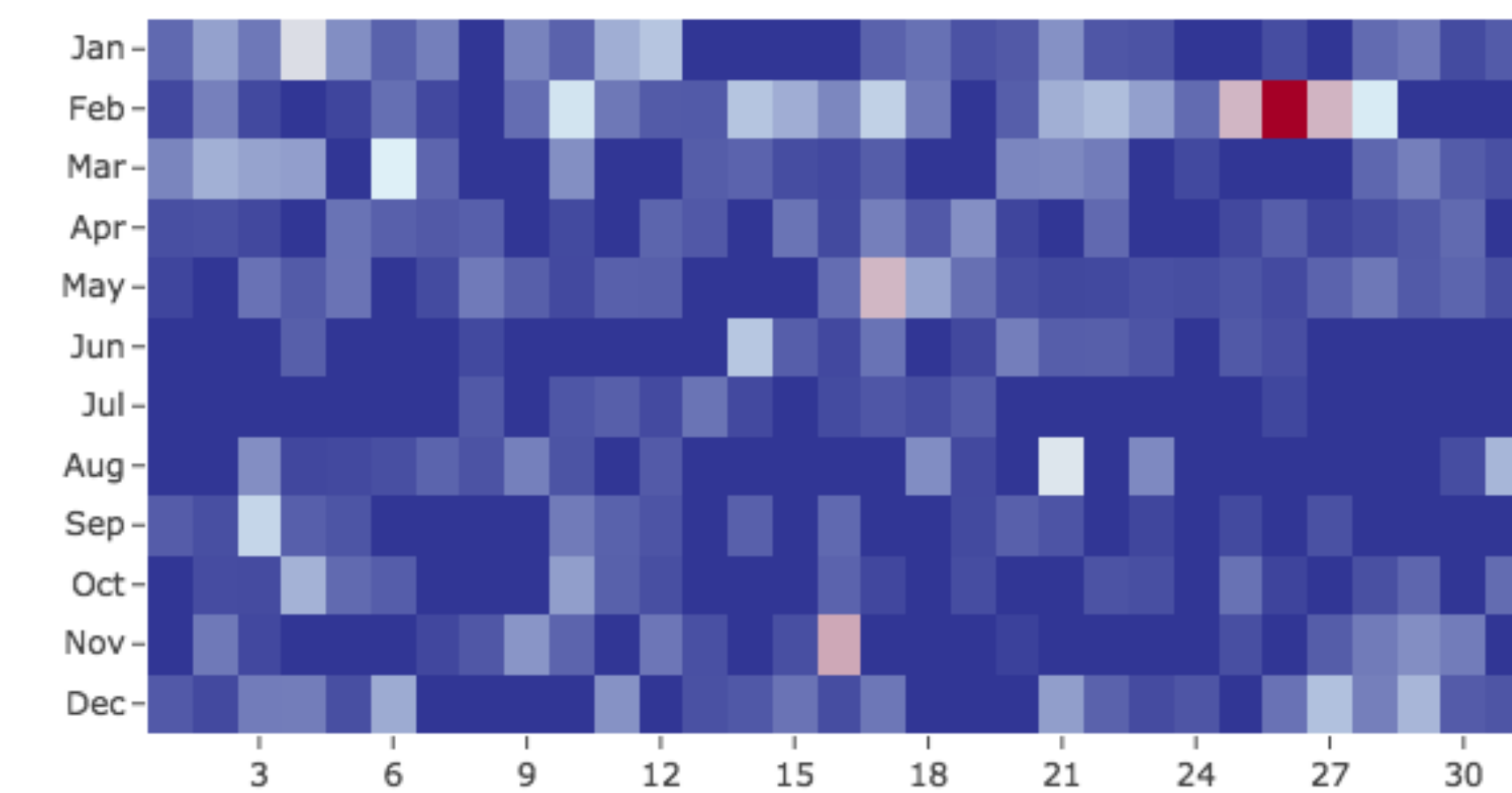
Get to know Logan Paul's personality

Izy Hossack talks the most about

baked · feta · bagels · basil · fancyporridgefriends · izyeatsfancyporridge · porridge · izyeatslondon · butter · aubergine · rye · guide · izybrunches · image · maple · morning · syrup · pinhead · mentioned · salted · tahini · looks · tasty · option · gonna
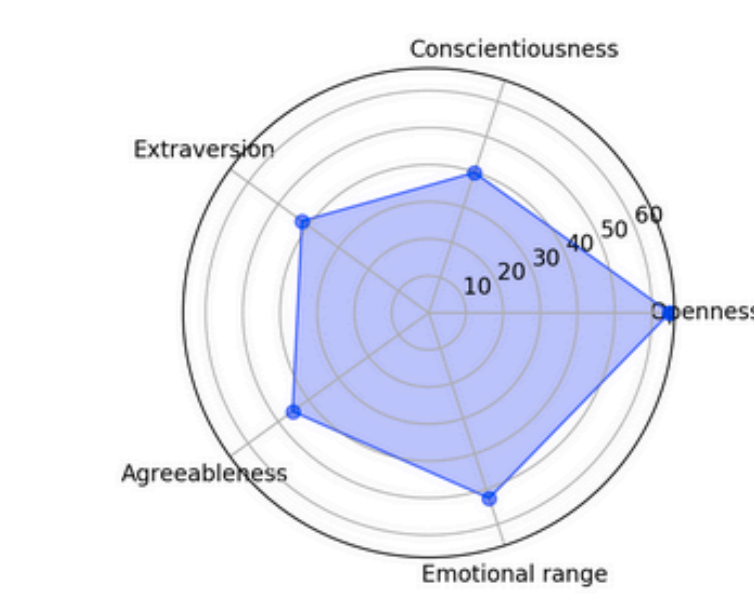
Topic modelling using LDA on YouTube data

Popular items in Izy Hossack's pics!

chocolate_sauce · wooden_spoon · bottlecap · spatula · bagel · tray · pizza · necklace · plate · ice_cream · pretzel · dough · bakery
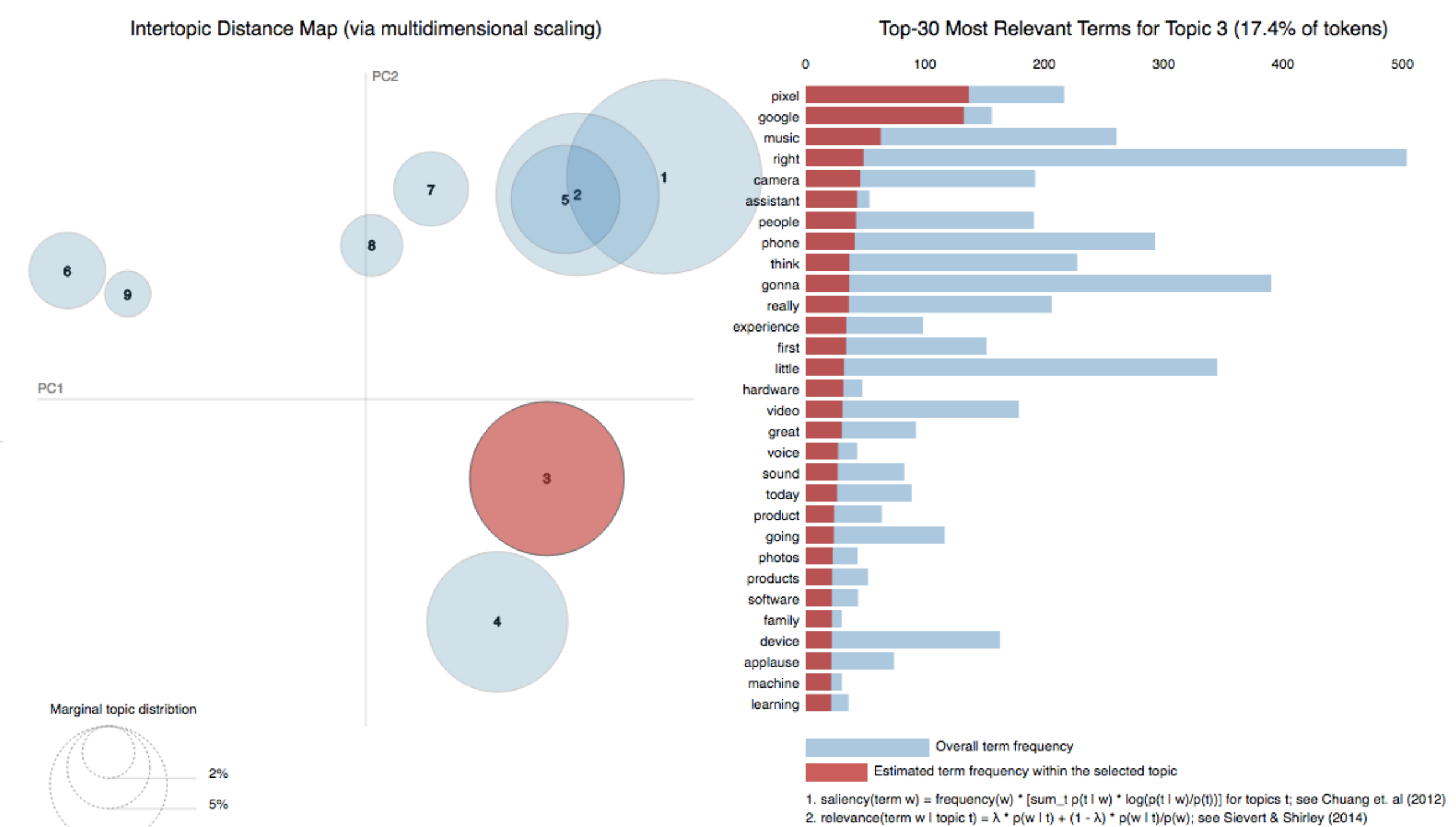
APPROACH

- ❑ Collect social media posts, images and videos by top influencers across multiple categories from various social media platforms.
- ❑ Compile absolute metrics like follower count, likes; Generate derived metrics like popularity, influencer & audience activity through aggregation.
- ❑ Discover the topics talked about in their posts through LDA / LSI and find objects in their posts using Computer Vision (Finding their niche).
- ❑ Generate weighted scores for the particular influencer-product pair through n-gram, topic and object relevancy along with absolute and derived metrics.
- ❑ Find the top influencers for the product using the weighted score.
- ❑ Visualize their metrics, the results of topic modelling and their personality to better understand the influencer and their niche.

LESSONS LEARNT

- ❑ Since open APIs are constantly evolving and introducing restrictions, enterprise access and/or partnerships might be required to maintain system functionality up to date.
- ❑ During data collection process, it is imperative to set an uniform timestamp format, to make timeseries calculations consistent.
- ❑ More custom tuning of the image classification models might be necessary to better understand images.