SMU
SINGAPORE MANAGEMENT UNIVERSITY

Team 7, G1, IS434, AY2018-19 T1
Social Analytics and Applications
Project Proposal
28 September 2018

THESMARTLOCAL
SINGAPORE

Vera Low
Hong Yang
Lau Zi Quan
Agnes Lim Jia Hui
Gokarn Mallika Nitin
Deborah Sim Hsien Minn

Client

The Smart Local is a media publishing company focusing on travel and lifestyle features located both in and out of Singapore. The company has accounts on various social media platforms, including, their website, Instagram, Facebook and YouTube.

Platform	Outreach
Website	1.1 million monthly unique visitors
Facebook	364,788 followers, 358,000 likes
YouTube	215,809 subscribers
Instagram	92,300 followers

* All results are updated as of 25 September 2018 from the respective platforms

Instagram is an up and coming social media platform and the outreach TSL has on their Instagram account is the lowest amongst all their social media platforms. Thus, our group has decided to run our analyses on their account to improve their outreach on Instagram.

TSL's Instagram account currently boasts of about 4K posts and 92.3K followers. With each post having roughly 1K likes and 100 comments, the ratio of likes to followers can be considered quite lacking.

Some recent efforts made to increase engagement on TSL's Instagram include promoting their newest series, 'The $100 Nomad' on their feed. Along with advertising, hashtags like #TSLMakan and #TSLTravel were generated to encourage users to use these hashtags when writing and posting recommendations and increasing awareness of the series and presence on the platform.

Business Problem

The Smart Local is not effectively targeting the appropriate audience for the wide array of content they post. The company currently wishes to recommend posts to consumers based on what they think the consumers may like, instead of sharing posts and hoping to gain more views from the general public. There have also been talks about the current social network link of TSL to build more brand awareness for TSL and more awareness of TSL's posts.

Data Collection

The data that we plan to crawl is as follows:

1. The number of Instagram likes and comments on pictures sorted by category
2. The followers of the users that liked and commented on TSL's posts up to 3 degrees
3. The followers of the user's followers that follow TSL up to 3 degrees.

Analysis

Market Basket Analysis

Market basket analysis is used to uncover associations. Based on what TSL's users usually like, share or watch on TSL's pages, we will identify posts and videos that these users may possibly also be interested in.

Example

- * Assume there are 100 users
- * 10 of them watch travel videos, 8 watch food videos and 6 of them watch both travel and food videos
- * watch travel videos => watch food videos
- * support = P(Travel & Food) = 6/100 = 0.06
- * confidence = support/P(Food) = 0.06/0.08 = 0.75
- * lift = confidence/P(Travel) = 0.75/0.10 = 7.5

Social Network Model (Target Marketing and Community Building)

Our objective is to be able to identify the centrality with high power of influence. By building a Directed Network Graph, we are hoping to be able to identify the agents with high power of influence. The data that we are looking at would be based on what we can crawl from Instagram. We are able to access information such as who are the current followers, who liked that particular post, who commented and if they also tagged any other users. Although there are limitations like not being able to view private accounts and which of those users (ones with private accounts) liked posted videos, there is sufficient information to allow us to experiment with.

To build a Network Graph, we would need to determine what contributes to the following factors:

Directed Network Graph	This represents the user who is following TSL and likes TSL's post(s). It also represents the connection of how users are related to each other by following or being followed.
Node	Each node represents an Instagram user who is connected to TSL
Edge	Each Edge represents the user that is connected to another user by following the user or liking the user.
Node Size	The size can be determined by the centrality of the node. The Centrality is represented by the degree. The degree is calculated by the number of interactions the user has with TSL by liking, following and commenting.
Edge Distance	This represents the number of hops the user has from TSL. The distance would be determined by the number of degrees of freedom of how the user can reach TSL. The closeness centrality can be calculated by the inverse of the shortest path away from TSL.

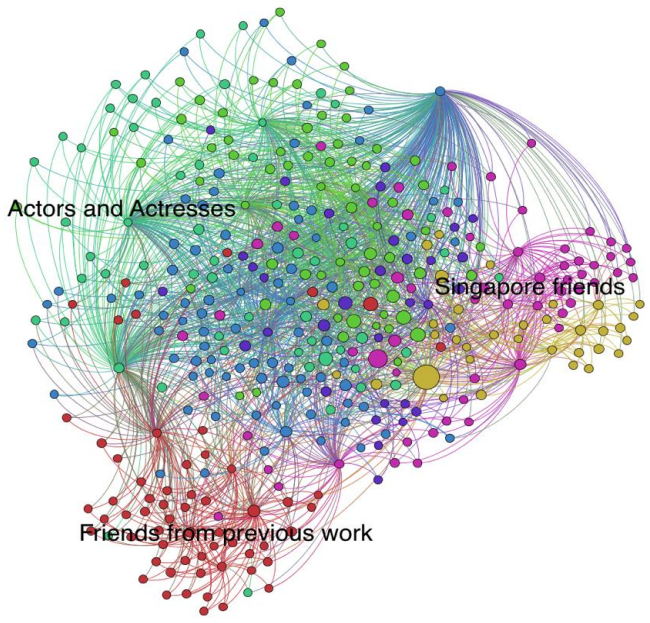
Edge Weight (Thickness)	This represents the strength of interactions, which can be derived from the number of likes, number of comments, and number of mentions. To determine the level of "closeness" between followers, recency, frequency and intensity of interaction would be evaluated.
Node Colour	Topic modelling could be applied to identify if users have higher influence toward food, travel or product related posts. Another colour selection can be done for the number of in-degree centralities of the Node.

Clustering and Classification Model for Topic Identification

Validation

1. Create Ground Truth
2. Create a recommendation model
    a. Create recommendations, get a number of people to see if they would see the video recommended

Expected Outcome



Value that we can add to the company

1. Determine the type of posts and videos TSL can post depending on the time of the day to generate more likes and comments. When there is a certain amount of likes or views on a post, it would show up on the explore page of users who do not follow them. This would potentially allow them to gain more awareness just by targeting the type of post at different timings
2. Determine the type of content TSL can focus on more. Some types of content would definitely be more interesting to the general public and TSL can determine what those

types of posts are through our analysis and generate more posts based on those interesting contents

Potential Limitations

1. Some accounts are set to private because of which we are not allowed to crawl their details
2. We cannot view who has viewed the video but only see who liked the video on Instagram
3. Facebook does not let anyone inspect their pages (web crawling on Facebook is prevented)
4. How can we determine if users would finish watching a video or the videos that was generated using the recommendation model

References

- Elisa, E. (2018). Market Basket Analysis Pada Mini Market Ayu Dengan Algoritma Apriori. Journal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 2(2), 472-478. doi:10.29207/resti.v2i2.280
- Noam Segev, Noam Avigdor, Eytan Avigdor. (n.d). Measuring Influence on Instagram. Retrieve from: https://arxiv.org/pdf/1806.00881.pdf