## **Internal Meeting Minutes 1**

|  |  |
| --- | --- |
| Date/Time | 12th January  |
| Venue |  |
| Attendees | Eric, Ivan, Shing Hei |
| Agenda | 1. Check with Prof with current data if we can proceed
2. Check with Prof on our proposed outcomes/problem statement
 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | Task/Description | Person in Charge | Due Date |
| 1 |  Data cleaning - Pulling category  | All | Nil |
| 2 |  | All  | Nil  |
| 3 |  | All  | Nil  |

Do we have to look at the demographics/post time/article type/content/date?

* Not possible to compare demographics across platforms

Engagement as the common basis of comparison

Posting title as feature engineering

Shing: Gantt chart template

Eric: Wiki

Ivan: Meta data for the report

Blog as the primary key (since every post starts off from the blog??? )

**Prof Kam:**

Chicken and egg problem? Catering to diff group and generation

By and large, people who started social media will be using FaceBook (2012); Instagram afterwards

There are differences in medium (Insta can be cross-linked to FB; Prof might not be able to find the same post across the diff platforms)

Ivan: This category of post performing better in this social media, why is it performing on another platform? It should all be correlated?

Prof: Should not jump to conclusion. If familiar with social media, anything can go viral, doesn’t mean if in IG, means go viral then if go FB, doesn’t go viral. We can proof him wrong with the data. \*Points to take note\*

1. Look at the senior work. Same term previous year, at least two teams (SGAG and competitors of SGAG). There will be diff in each of them and see how they solve them.
2. Data: What kind of data they are going to give u? FB -> FB insights (export). TSL extracted for us. One year data

Ivan: FB changed algo in Dec; people wanted to comment and share to win prices (push down the rank for FB -> to analyse the effect of change) Investigative problem

**Prof:**

1. Get the data from them first. Extract from them? Or they extract for us? It would be to our advantage to extract for ourselves. Not useful if extract from an overview level. Check if it is granular enough
2. Social media: If I post any content out, I might get immediate response, some post might drag on and still get a lot of response. Some on the next day, no more responses received. Time interval (for definition of metric) is very important. **Survival time** of post is different.
	1. How should we normalise this? Can I come out with a metric and method? If a particular post continuously receive 20/50 comments, I will continue to calculate the metrics. If go below 20, I will stop. Need to look through all the data and determine the best cut off line -> for normalisation.

**Ivan:**

Aggregated data, with some granularity. One post (ID link, time post, organic/paid impression, average time of video). Take from diff platforms and belong to the same campaign, title might be different.

* Master campaign file: FB, Instagram (serve as the primary key to integrate the data tgt)

**Prof:**

Check it out first. Sign the NDA and fill it all the particulars, and commit themselves to the project. Print a copy out and let Prof sign.

2 week default is a yes clause. Before presentation, they have to vet through without delaying. Confirm the date. We CANNOT change the clause, the 14 days are to ensure we graduate in time.