Assignment 3

- By Aayushi Vyas

The following proposal is for the problem in Question 3 of the Assignment.

Introduction

While Social Media has been booming these days, it is getting more and more tough to understand the trends, the workings and the analysis of products and posts. Twitter has been a great platform for discussions, marketing and spread of word and it has become important to understand the logistics behind every tweet, every picture posted. To start with, one of the ways that we can analyse and come to a conclusion is by working with the data we already have like the tweets that have been posted so far, analysing the likes, comments and sentiments of those comments. This is the first layer of data which can be observed by simply gathering data.

Challenges and Analysis

To go to the next level, Twitter has a time and a date stamp on every tweet using which one can predict where the audience is from(geographically) and what is the best time of engagement. This is one of the biggest factors to know if the audience that follows the handle are looking at the tweet at the right time or not. For example, the best time to post tweets is between 12pm and 1:30pm EST due

to the corporate lunch breaks in the US where most of the people engage on twitter. By looking at the demographics provided by twitter itself, information like gender, age, and location can be extracted. The content of tweet is very crucial, in order to determine if the tweet will work as meme or not, or how much the people would like it. This forms another problem. What works for one, might not work for another, this is where we do sentiment analysis of what people post about, what they like and retweet things. The gathered data can show the trends of the topics that people are discussing about. The post type that people like the most, eg: Videos, memes, Pictures, GIFs, or texts. Every category has different trends based on the type of product the company is marketing and on what context the tweet is about. Such as a case study that I have worked on was about Netflix, whose most popular posting content are its seven second clips/videos of series and movies of its original content.

Solutions for further analysis

To recognise these trends, one can simply use Excel tools such as Pivot tables or to have a better analysis can use an already existing software while analyzes content for you or create an algorithm for it. There are many data analysing software like Tableau, SAS, Alteryx and many more, which can help gathering the data at one place and analysing through each segment. It is more logical to use the business analyst tools which are already in the market if the analysis doesn't require to be very segmented and spread out. But for big companies with lots of data and user following on social media, it is important to do the analysis precisely with the focus of reaching out to as many users as possible even by one tweet.

Ethics and Issues

When it comes to Social Media, there are very few rules that can be broken online but it gets equally risky to offend someone due to its large audience and widespread. It is important that while posting a meme or tweet, copyrights issues are avoided, data is gathered legally and nothing offensive is posted. It is essential to note that the strategy used for other companies might not work for this particular one. Memes especially tend to get a lot of attention these days, but it also can be relatable to only few people sometimes and might create a negative impact on the users. Even a word on social media gone wrong can destabilize the reputation of the entire company. So while tweeting things, one needs to be careful about the way they put things out, not create a negative sentiment which can have backlash from the audience.

Conclusion

For Social Media analysis, it is important to cover consumer analysis to sentiment analysis. Usually these social media platforms such as Twitter, Instagram and Youtube provide a statistical demographic and geographic analysis about the followers. So to get a rough idea about it is feasible to use these statistics, but when it comes to marketing companies launching campaigns, then they need to go further into analysis step by step in order to reach the maximum audience. For these node-to-node analysis, one can use decision trees to target specific audience. Which tweets are reposted the most, what type of consumers like which kind of posts, how the users react to certain

posts and what time do we get the largest group of people online. The communication through these channels is easy but to understand the behind the scenes logic and patterns that form is challenging and that is why we still need better and accurate algorithms to work on the data set we have and the focus on the nodes and spread of the data in the direction we need.