

Social media analytics : course projects

2021-2022

Below is a description of the different project possibilities for the course of Social Media Analytics. You can choose one of the topics and apply this to a company of your choice if applicable, as long as the language analysed is English.

Make sure to carefully read the documentation below. Importantly, don't stick to what we have seen in class but show that you understand or are able to combine the different elements.

Universal guidelines

1. Study the data
 - a. Manually on Twitter or the platform you are using
 - b. Using different structures and representations in R
 - c. Try different preprocessing/cleaning methods
2. Define and formalize the problem
 - a. Do we need a formal model or not? (not necessary for more descriptive insights)
 - b. What should the model do/predict?
3. Research and brainstorm about possible solutions/tutorials/packages
4. Prototype on smaller subsets
5. Refine and tune your model
6. Do not limit yourself to what we've done in class (!). If you simply apply a topic model or sentiment analysis, this is simply copy-pasting code. You must show your understanding of the materials and dare to go further.

Project planning

- Work on project during class
 - At the end of sessions if possible
 - Fourth session (Jan 20)
- Autonomous finalization of the project: until February 3, 2022. (two weeks after the last session)
- Submission: per group
 - Preferably a Shiny application and/or RMarkdown document with the results, interpretation and recommendations, and a description of the application functionalities, the challenges encountered and maybe what you would improve in the future
 - Code (well commented!)

Project evaluation

Project evaluation will be based on several items

- The effort put into collecting a relevant amount of data
- The analysis performed
 - Degree to which you go further/deeper than what is seen in class in terms of analysis
 - Correctness of the analysis
 - Perform analyses that explore elements we haven't discussed in class
- Nicely representing the analysis in terms of tables/graphics
- Making a good report/app that clearly shows your analysis and recommendations

Project Descriptions

1. Analyze a company's recent Twitter profile

The project involves a Twitter analysis of an American (or English-speaking) company (can be another profile as well, as long as it has sufficient data). The purpose is that you come up with a comprehensive analysis of the Twitter environment the company is operating in (both from the business itself and from customers). How and what are people tweeting and contacting the company? How does the company tweet and react to tweets (eg do they respond to positive and negative tweets? What is included in their responses?) Who are the followers or what do they tweet about? Use the elements you have seen in class. To get you started, here are some ideas and guidelines:

- First collect your data in a smart way. These can be company tweets, followers' tweets (e.g., collected using the timeline function). Look for how to create search queries that can provide a lot of tweets (you can use multiple searches). The search query goes 7 days back, so you can do this multiple times to get more data.
- Give an overall summary of the tweets (engagement, volume, directed/undirected,...), maybe over time
- Conduct a sentiment analysis and content analysis of the tweets. Make sure to do a detailed sentiment analysis (think of bigrams, reinforcement words).
- Who are the people who contact the company and what are they interested in (e.g., their list subscriptions)?
- ...

2. Analyse job descriptions

A lot of websites show job descriptions, or allow to upload CVs. Although these are more difficult to collect via an API publicly, companies that post their jobs on these websites can do so.

The goal of this analysis is to extract information from the text, so with a strong focus on entity recognition. Try to extract the job description, the location, the current company and skills.

You will get a dataset of ca 2000 job descriptions extracted from CVs. Take a closer look at research on named entity recognition and extraction. You might check the R package

openNLP, and use the Google natural language API. But you may also go deeper and try regular text packages like stringr combined with regex to get relevant info.

This project is somewhat further from the course, but I will of course take this into account in the grading. You can also use the document “Skill Finder: Automated Job-Resume Matching System” to get some ideas.

3. Perform advanced sentiment analysis

Try to implement some more informed sentiment analysis, based on the paper of Homburg et al. (2015) (<https://journals.sagepub.com/doi/abs/10.1509/jmr.11.0448>).

Compare their model with a more traditional customer sentiment model as seen in class. Which one is favorable and why?

Hints:

- The steps are clearly outlined in Homburg et al. (2015)
- Make sure you can compare the results
- Make your model as strong as possible

4. Predictions of online engagement for (company) tweets

Imagine you are in charge of the social media of a company. You want to evaluate what elements drive the engagement with your social media network (e.g., Twitter).

Hints:

- Use Twitter data, along with an engagement figure (eg number of retweets, shares, ...). Try to first get a good dataset to analyze.
- Make a set of independent variables to explain engagement. You can think of sentiment and topic variables, but also other variables related to the text can be included.
- Use a machine learning method or a traditional statistical model (e.g., regression) to model this.
- Make sure you include some evaluation measure
- There is already quite some academic literature on this topic.

5. Analyze a Vimeo/Youtube video

Analyze a video from Vimeo/Youtube. If you want to go advanced, you can extract the text mentioned in the video as well and analyze this with the Google Speech to text api.

- You can analyze the comments to the videos. What is the overall sentiment? What do the comments mainly talk about in terms of topics or concepts?
- You can analyze the description of the video. If you do this (or the transcribed text of the video) for multiple videos, try to find the concepts and topics mentioned in these videos.

6. Analyze Yelp reviews in terms of content and sentiment

Analyze a set of Yelp reviews (since the Yelp API is pretty limited, I will provide a dataset for this).

You can link the sentiment, and for instance try to find sentiment for specific elements mentioned (atmosphere, food quality, ...) in the review. This can be linked for instance to the helpfulness of the review. Moreover, other elements of the text (length, content, ...) can serve as independent variables to predict helpfulness.