

Project: Practicum I for Batch II (Winter 2019)

Scope: Limited to courses taught in Terms I

Evaluators: Bhat Dittakavi

Maximum no. of members per team: Two

Weightage: 40% weightage (Remaining 60% for Practicum 2)

Evaluation Criteria and Weightage (Total 100 marks)

- 1) Data Story Telling (10 Marks)
- 2) Data Requirement & Collections (5 Marks)
- 3) Data Understanding (15 Marks)
- 4) Data Preparation (15 Marks)
- 5) Models, Evaluation & Feedback (20 Marks)
- 6) Business Recommendation (5 Marks)
- 7) Executive Summary (10 Marks)
- 8) Code Demo Video (20 Marks)

Deadline for Code, demo & Presentation: 28th Dec 2019 11.59pm

Date of Evaluation: 29th and 30th December 2019

Hope you have got good clarity about Practicum course and the benefits it brings to the table. Even though you may have a team member, please make a note that you are strongly advised to do hands-on work and both the team members are expected to code, visualize and analyse. Priority is for individual learning in a team setting. Formation of team is a matter of convenience only. You will submit one code as a team and do a team presentation on the day of valuation. If we come to know that only one of you have done the entire work, we may ask the non-contributing member to do a different project all alone.

Objective is to pick a problem in such a way that you try to apply concepts from as many courses from Term I as possible and put them to practice while solving a real-life problem you have hand-picked. **If you can't collect data say from twitter beyond a day or rate limit, remember that this is the same limitation for all the teams. Give your best to find alternate ways of extracting information or find deeper ways of deriving insights from the limited data.**

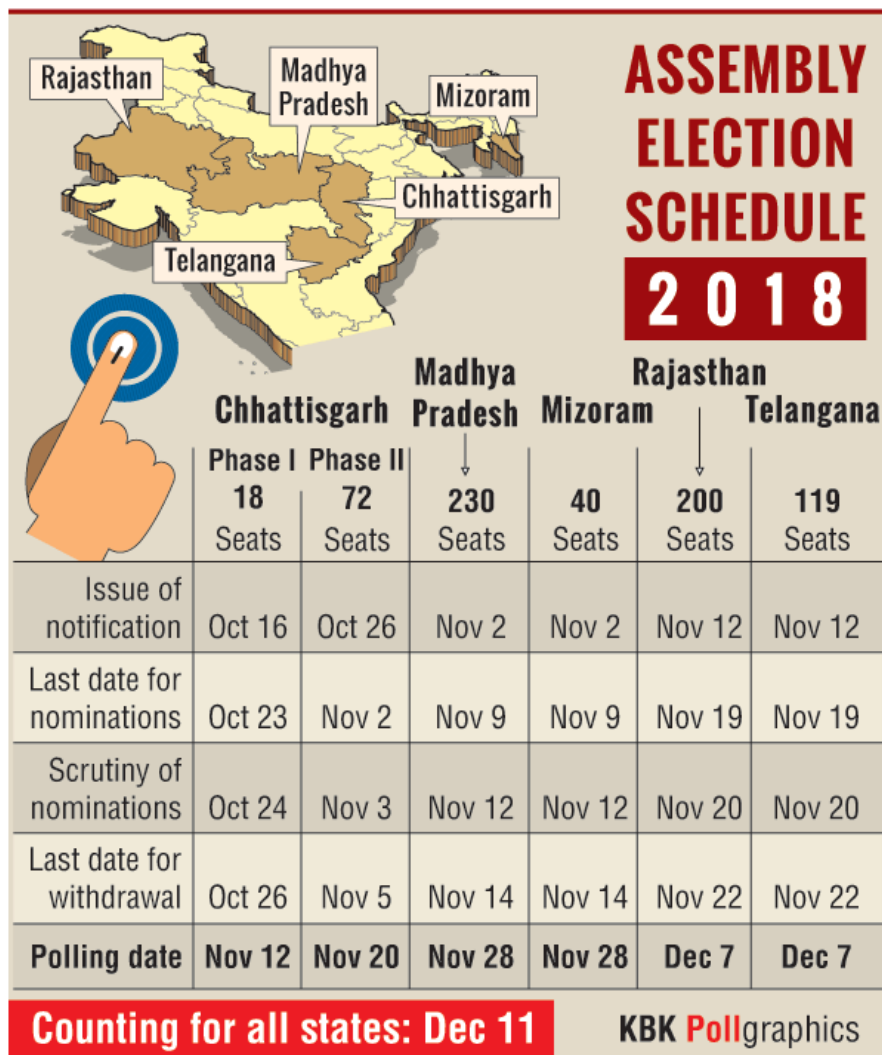
If you have any questions, please feel free to write to me at bhat.dittakavi@gmail.com. If it is important, call me at 8008804440.

You can choose any of the below projects for Practicum I.

Project I: Assembly Elections 2018

Background

Social media plays an important role in any elections across the world. Not long ago, companies like Cambridge Analytica influenced USA elections by doing “disinformation campaign” using Facebook. Digital media plays dominant role across the world in elections. In India, Chhattisgarh, Madhya Pradesh and Mizoram are done with voting few days ago and assembly elections for Rajasthan and Telangana are just around the corner.



Objective

We would like you to track and analyse the election chatter that happens in twitter, facebook and other social media channels, news sites and portals. Analyse trends and patterns and even predict the outcome of these elections. Choose a problem within the domain of 2018 assembly elections. For example, you may use different data collection methods as needed and collect different

opinions from influencers and key opinion leaders on social media and analyse the sentiment of the voters. Or, you may choose to check the veracity of the opinion poll and exit poll data done by popular news channels by applying statistical concepts learnt.

Whatever the problem you pick within the bounds of assembly elections, you are expected to leverage data visualization techniques learnt in the class room. Explore the data using visualization and do the first cut analysis and then deeper analysis. Apply text analytics to do various NLP tasks that help you derive election insights from social media and beyond. You can also run “Google Trends” to see the relevant trends on different elections for different time periods. Incorporate the trends in conjunction with the chatter from the media and do text analytics. Even you may do some big data analysis. You are welcome to choose any publicly available dataset of tweets, trends and posts. These questions are to generate curiosity in you.

Choose your own sub-problem within this domain.

Data Sources

- Social Media (Twitter, Facebook, Whatsapp, Instagram or others)
- Third party sites (where you can get historical tweets)
- Google Trends
- News Portals and Archives

Twitter provides archive search on tweets. For the sandbox version, access is free and includes 50 requests per month. Each request can return up to 100 tweets. So, you can extract 5,000 tweets per month. Between two of you, you may get 10,000 tweets. Click the link below for more details.

<https://developer.twitter.com/en/docs/tweets/search/overview/premium.html>

Project 2: Narendra Damodardas Modi

Background

On 16th May 2014, NDA got thumping majority and won Indian general election with 336 seats. Narendra Modi led the sweeping victory. It has been 4.5 years that our Prime Minister Narendra Modi has taken the office. Is PM Narendra Modi losing his popularity? Modi has brought in big initiatives such as Demonetization, GST, Swachh Bharat and many more.

Objective

What is the current sentiment of people about Narendra Modi as a leader? Are they happy? Are they sad? Are they frustrated? Have his initiatives worked? What do people think about each of these initiatives?

We would like you to track and analyse the popularity, perception and sentiment on Narendra Modi using twitter, facebook and other social media channels, news sites and portals. Analyse trends and patterns. You may also take one of his initiatives and do deep dive analysis on it. For example, you may take GST and see whether it has hurt small businesses or made common man suffer? Evaluate messages from influencers and key opinion leaders on social media and analyse the sentiment. You may take any specific topic about Narendra Modi and do your analysis on that topic.



Whatever the problem you pick within the bounds of Narendra Modi and his initiatives or the lack of, you are expected to leverage data visualization techniques learnt in the class room. Explore the data using visualization and do the first cut analysis and then deeper analysis. Apply text analytics to do various NLP tasks that help you derive insights from social media and beyond. You can also run “Google Trends” to see the relevant trends for different time periods. Incorporate the trends in conjunction with the chatter from the media and do text analytics. Even you may do some big data analysis. You are welcome to choose any publicly available dataset of tweets, trends and posts.

Choose your own sub-problem within this topic.

Data Sources

- Social Media (Twitter, Facebook, Whatsapp, Instagram or others)
- Third party sites (where you can get historical tweets)
- Google Trends
- News Portals and Archives

Twitter provides archive search on tweets. For the sandbox version, access is free and includes 50 requests per month. Each request can return up to 100 tweets. So, you can extract 5,000 tweets per month. Between two of you, you may get 10,000 tweets. Click the link below for more details.

<https://developer.twitter.com/en/docs/tweets/search/overview/premium.html>

Project 3: Apple Inc.

Background

Apple Inc. is an American multinational technology company headquartered in Cupertino, California, that designs, develops, and sells consumer electronics, computer software, and online services. The company's hardware products include the iPhone smartphone, the iPad tablet computer, the Mac personal computer, the iPod portable media player, the Apple Watch smartwatch, the Apple TV digital media player, and the HomePod smart speaker. Apple's software includes the MacOS and iOS operating systems, the iTunes media player, the Safari web browser, and the iLife and iWork creativity and productivity suites, as well as professional applications like Final Cut Pro, Logic Pro, and Xcode. Its online services include the iTunes Store, the iOS App Store and Mac App Store, Apple Music, and iCloud.



Apple became the world's first trillion dollar company in August 2018. Many things happened in the last couple of months and Apple has lost its market cap by 20% just within the last 30 days. Few weeks ago, Apple announced that it would end unit sales reporting for its hardware devices. On top of it, Trump has issued a threat of 10% tariff on iPhones just few days ago. Then there is a looming trade war between USA and China. After the departure of Steve Jobs, Apple hasn't brought any life changing products to the market except for incremental innovations on existing products.

Objective

Is Apple still the greatest tech company that it used to be? Investors are concerned. Is Apple's loyal fan base diminishing in size? US market hardly gives any growth to Apple. OnePlus is dominating Apple in key emerging markets. Is there a paradigm shift in the user sentiment of Apple? How about investor sentiment?

We would like you to track and analyse the popularity, perception and sentiment on Apple using twitter, facebook and other social media channels,

news sites and portals. Analyse trends and patterns. You may also take one of Apple's flagship products and do deep dive analysis on it. For example, you may take iPhone and see whether investor or consumer sentiment is changing? You don't have to limit your choice to Apple's products. You may choose any topic around Apple Inc. and do analysis on the same. Go beyond doing basic sentiment analysis.

Whatever the problem you pick within the bounds of Apple Inc., you are expected to leverage data visualization techniques learnt in the class room. Explore the data using visualization and do the first cut analysis and then deeper analysis. Apply text analytics to do various NLP tasks that help you derive insights from social media and beyond. You can also run "Google Trends" to see the relevant trends for different time periods. Incorporate the trends in conjunction with the chatter from the media and do text analytics. Even you may do some big data analysis. You are welcome to choose any publicly available dataset of tweets, trends and posts.

Choose your own sub-problem within this topic.

Data Sources

- Social Media (Twitter, Facebook, Whatsapp, Instagram or others)
- Third party sites (where you can get historical tweets)
- Google Trends
- News Portals and Archives

Twitter provides archive search on tweets. For the sandbox version, access is free and includes 50 requests per month. Each request can return up to 100 tweets. So, you can extract 5,000 tweets per month. Between two of you, you may get 10,000 tweets. Click the link below for more details.

<https://developer.twitter.com/en/docs/tweets/search/overview/premium.html>