COVID-19: Live dashboard

Steffy Dias
Binghamton University
Binghamton, NY, USA
sdias1@binghamton.edu

Sagar Sahani
Binghamton University
Binghamton, NY, USA
ssahani1@binghamton.edu

Arun Reddy Gummala
Binghamton University
Binghamton, NY, USA
agummal1@binghamton.edu

Figure 1: COVID19-Live Dashboard

ABSTRACT

The sufficient amount of data collected through online social networks can better lead to evaluation of the pandemic situations and how quantifiable analysis from this data can be carried out to answer some relevant research questions.With a live data collection system running as well as a set of analyses performed on the data, it's time to provide some interactivity. In this project, we will thus be developing a simple dashboard that allows a user to vary some parameter of some of the analyses you performed and be shown a result. To put an interactive dashboard in place we propose to use Streamlit, an open source python library to build custom web pages. The user will be able to put in some parameters in place and accordingly the analysis graph will pop up.

KEYWORDS

Streamlit, analysis, plotting, parametrised

ACM Reference Format:
Steffy Dias, Sagar Sahani, and Arun Reddy Gummala. 2020. COVID-19: Live dashboard. In *Binghamton '20: ACM Symposium on COVID-19: Live dashboard, November 03–05, 2020, Binghamton, NY*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/1122445.1122456

Binghamton '20, November 03–05, 2020, Binghamton, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

The COVID-19 pandemic has led to many efforts being taken to analyze the data collected in many countries and cities to study and predict its growth so that it can help plan for healthcare resources and socioeconomic decision making. The pandemic has caused huge loss in public health, affecting millions of people world wide throughout countries.
The hour of the pandemic in 2020 has brought intensified efforts and experimental analyses by researchers and scientists alike to

analyze the dynamics of COVID-19. This has led to great importance to the related data that is readily available in huge amounts these days on social networking sites. The pouring amount of data available every week, every day and nearly every hour can be of great importance once brought to its correct use.Collecting this data to make meaningful insights can help a great deal.
The sufficient amount of data collected through online social networks can better lead to evaluation of the pandemic situations and how quantifiable analysis from this data can be carried out to answer some relevant research questions.
With all this in place and well developed and analysed from a researcher's end how well can this be portrayed and put forth in front of a lay man. How well can this be seen and be pictorially represented to easily understand the demographics of the whole situation. To well implement a web page and to pictorially represent all the research that has gone we propose to develop a live dashboard that will showcase all the analysis done with the help of the data collected from Reddit and Twitter with mere button pushing that will help users to apply filters, that is his own parameters for the to the analysis to show results.

2 MOTIVATION

To put forth relevant analyses of the behavioural change that is observed in mankind and how he has been affected by the peaks and declines in the cases this pandemic has brought that can be pictorially represented to better suit the understanding that can be well understood by a lay man.

3 PROBLEM STATEMENT

To pictorially represent insightful and impactful analysis of COVID19 data, collected from popular social media websites to that can .

4 ANALYSES TOOL

4.1 Streamlit

Streamlit [1] is an open-source platform for data science teams to create data applications with python libraries that uses python scripting, APIs, etc. to help data scientists to create python- based applications. This open-source Python library which is blazingly fast makes it easy to build beautiful custom web-apps for machine learning and data science. [2]It is an awesome tool that allows you to create highly interactive dashboards just with some knowledge of python. We will start by installing the Streamlit and see how it works.

4.2 Plotly

Plotly's [3] Python graphing library makes interactive, publication-quality graphs. Built on top of the Plotly JavaScript library (plotly.js), plotly enables Python users to create beautiful interactive web-based visualizations that can be displayed in Jupyter notebooks, saved to standalone HTML files, or served as part of pure Python-built web applications using Dash. The plotly Python library is sometimes referred to as "plotly.py" to differentiate it from the JavaScript library. This tool will be used to plot graphs from the data collected from Reddit api[4] and Twitter developer api[5].

5 ANALYSIS FROM PROJECT 2

As we have collected data from two online social networks which could best describe the thoughts, views and behaviour of users through their social media posts and tweets and which was used as our basis for analysis, the user will have two options to select from. It would Reddit or Twitter. After the social media is selected, the next part would be to select the date from the calendar that the user intends to have pictorial analysis of. Keeping in mind the data collected through that particular period, he will be given options to choose from those dates. After the options are selected from we propose to give our user following options to choose from.

5.1 Twitter Data Analysis Plot

- **Sentimental Analysis:**
  Sentimental Analysis of tweets count throughout the time period the data was collected. This can further be divided with respect to whether they want the positive sentimental count, neutral sentimental count or the negative sentimental count. This analysis describes the positive, neutral or negative thoughts or behavior of the user through that particular period of time.
- **Sentimental Analysis of Tweet Count every day:** This plot will pictorially represent the amount of tweets with their sentimental analysis whether positive, neutral or negative through that period.

5.2 Reddit Data Analysis Plot

- **Sentimental Analysis everyday:** Sentimental Analysis tweet count throughout the time period the data was collected. This can further be divided with respect to whether they want the positive sentimental count, neutral sentimental count or the negative sentimental count. This visualisation is derived from previous analysis done in project 2 and similar insights are represented here in an interactive way.
- **Sentimental Analysis at particular hour everyday:** Similar insights are drawn from the plot below showing the number of posts every hour for all the 7 days, though this sample includes posts for 6 hours of each day. The plot that shows data for all the days separately consists of an average of 140 negative, 607 neutral and 3 positive posts making it identical.

6 ACCESS DASHBOARD

StreamLit makes it easier to create dashboards like this for Data Visualization.
The dashboard consists of multiple features, the first filter allows you to select between 'Reddit''Twitter', after the selection has been made on social media.
If we select Twitter we will get 2 options one for selecting the date and one for selecting the plot type . If we check the filters options while selecting twitter and selecting the date and plot type it displays the plot for given date with count of negative polarities with red, neutral polarities with blue and positive polarities with green. If we uncheck the filter with twitter and select a date and plot; if we select positive it displays only the count for positive tweets on that date with count interval as for 2 hrs. If we select

negative it displays the count for negative tweets on that date with count interval as for 2 hrs similarly for neutral. If we select the wordCloud option this displays the word cloud related to specific positive or negative or neutral polarities depending on whichever is selected.

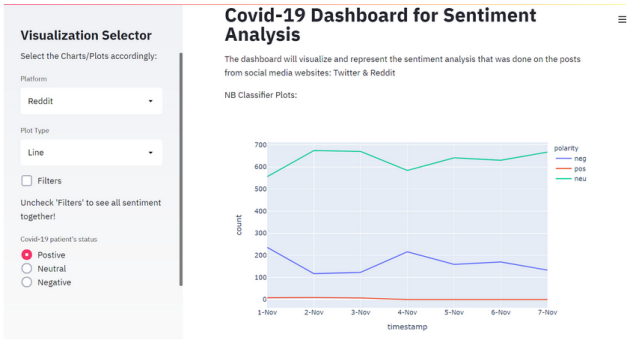

Figure 2: Reddit Line Plot of Positive Prediction

For Reddit all 7 days data is displayed and this could be further categorized by using the radio buttons for Positive, negative neutral. All in all the data needs to be viewed at once just uncheck the filter option.

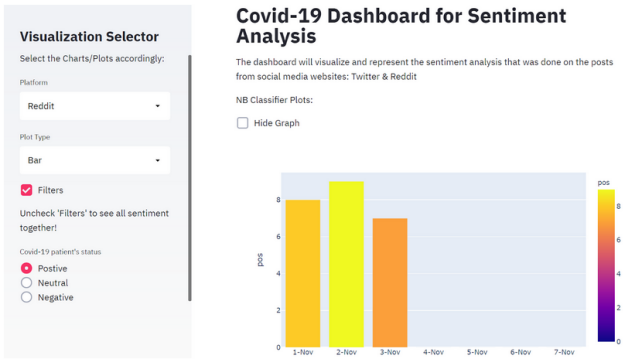

Figure 3: Reddit Bar Plot

For Twiiter, the date option makes it easier to get in depth analysis of peak hours for everyday, similar plotting options could be found here for all polarities.

WordCloud basically gets the top sentiments in the data and provides insights on how it maps with the plots.

7 DASHBOARD SUPPOSED TO DO

Dashboard has 2 options for Reddit and Twitter. If we select Reddit we will get Plot-Type options for Bar, Line, WordCloud selecting any of these options with the filers option checked gives us the plots related to Reddit displaying the information of count of neg, new, pos polarity tweet counts for all the dates . If we select reddit with filters option unchecked and select any kind of plot type from above plots we get graphs displaying the count for all polarities specific to date.

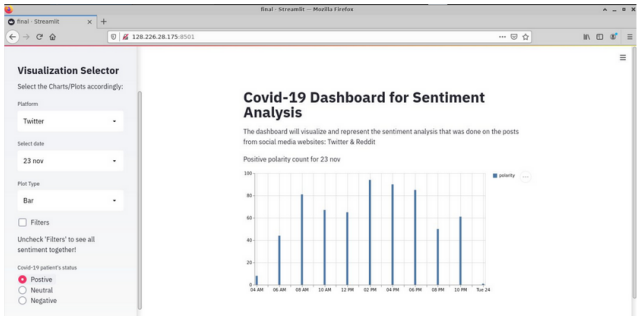

Figure 4: Twitter BAr Plot

Figure 5: Reddit Word Cloud

Figure 6: Twitter Word Cloud

8 HOW TO LAUNCH IT

Make sure you have Python StreamLit installed if not then:

- *pip install streamlit*

In addition plotly, wordcloud and matplotlib too:

- *pip install plotly*
- *pip install matplotlib*
- *pip install wordcloud*

StreamLit provides an IP address itself so as to view the dashboard and it will run with

- *streamlit run sar.py*

Here is a snap shot of the file running on the VM with X2GO client:

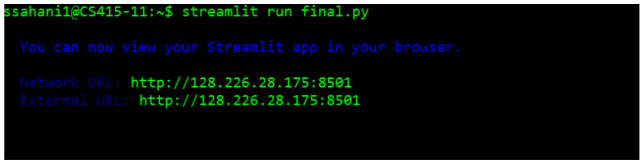

Figure 7: File running on the VM with X2GO client

Make sure you have all CSV's in the dir After running the pyfile: You will get the local and network URL, initially when you run it should open in the browser automatiocally. If not you could use the URL's.

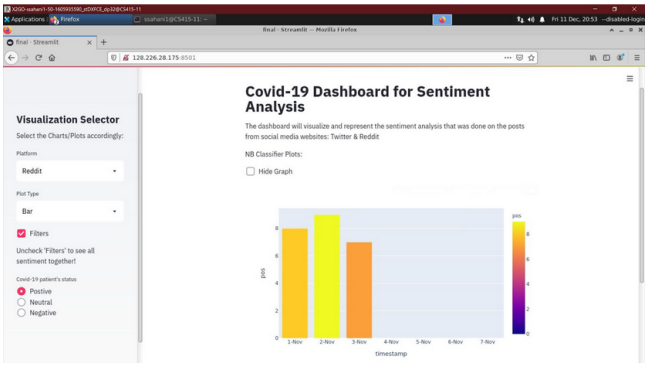

Figure 8: VM Launch Page

9 CONCLUSION

Finally, after seeing the analysis done on both the sides in different ways,pictorially; with the help of StreamLit, it seems that both the social media websites correspond to the same results. This interactive dashboard paired with live stream processing could be well extended.

10 ACKNOWLEDGMENTS

We would like to thank Prof. Jeremy Blackburn for guiding and helping us come up with this project and providing his insights on the various steps that could help us plan this project.

REFERENCES

[1] [n.d.]. https://www.streamlit.io/
[2] [n.d.]. https://towardsdatascience.com/creating-streamlit-dashboard-from-scratch-59316a74fa1
[3] [n.d.]. https://plotly.com/python/
[4] [n.d.]. https://www.reddit.com/dev/api/
[5] [n.d.]. https://developer.twitter.com/en/docs/twitter-api