***Project 3: “Fake News!” Detector***

******

**Team Members:**

|  |  |  |
| --- | --- | --- |
| **Name** | **Email** | **Phone Number** |
| Paola Lopez | paola.lopez1@yahoo.com | 408-763-7216 |
| Alicia Ly | aliciaseavly@gmail.com | 510-708-4079 |
| Medha Aravind | medha.795.a@gmail.com | 669-215-9234 |
| Miguel Eamiguel | mgeamiguel@gmail.com | 562-668-1992 |

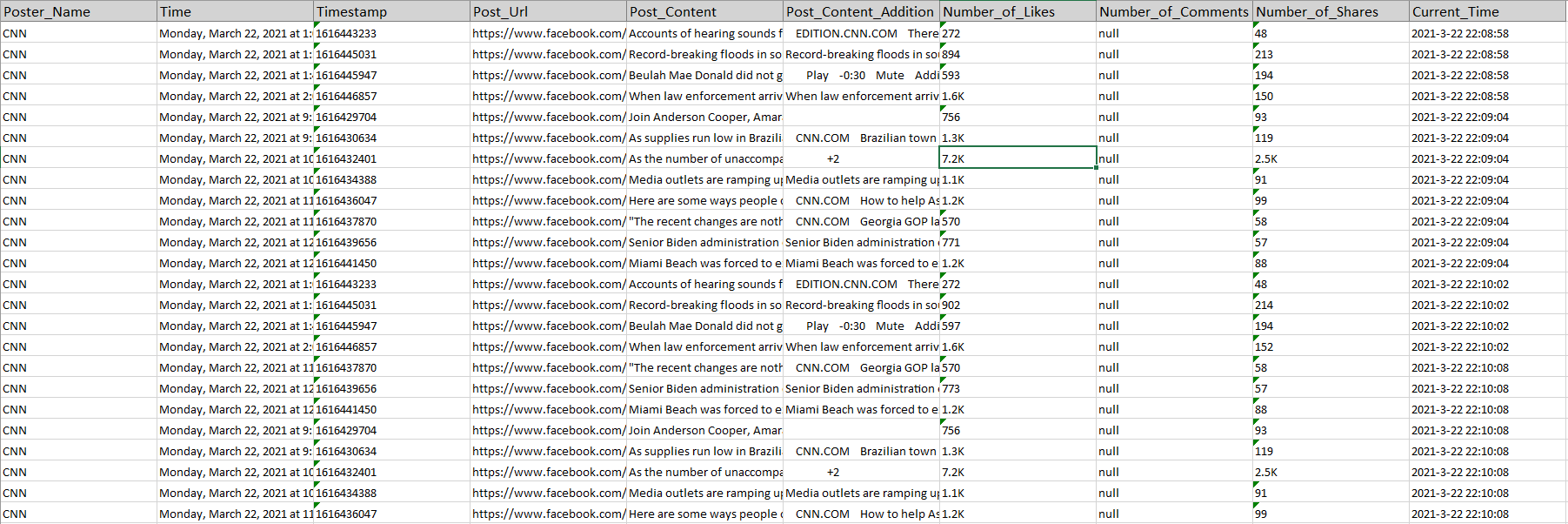
**Background and Motivation:** In the age of social media, it has become more important than ever to be wary of what we read on all platforms. Social media has made it easier to connect with the world and has great influence on the masses as anyone can share anything. What we wanted to do was to find a way to weed out the real news from the fake news. And, more importantly, we wanted to find a way to measure the impact that fake news has on society as a whole.

**Questions to Answer:**

1. Which social media sites contribute the most “Fake News.”
2. Which news sources or individuals create the most “Fake News.”
3. Which is shared and liked more, “Fake” or Real News? Why?

**Data Sources (Data to be extracted using Octoparse):**

* Facebook
* Twitter
* Reddit



**Requirements (Highlighted are the tools and modules that we plan to use):**

* Use scikit-learn and/or another machine learning library
* Use 2+:
  + python pandas
  + html/css/bootstrap
  + Matplotlib
  + plotly/D3/leaflet
  + SQL or mongoDB
  + Google cloud SQL or Amazon AWS
  + Tableau
* Host app using Heroku

**Checklist/Responsibilities:**

|  |  |  |
| --- | --- | --- |
| ***TASK*** | ***NAME*** | ***DATE OF COMPLETION*** |
| Extract and Clean data | Miggy, (Alicia, Paola providing additional sources) | 3/24/2020 |
| Use ML to create/fit/predict if post is considered to be “fake” or “real” | Paola | 3/26 |
| Launch on Heroku | Miggy | 3/29 |
| Tableau and Visualizations | Medha and Alicia |  |

**Tasks and timeline**

|  |  |  |
| --- | --- | --- |
| **Date** | **Task** | **Notes** |
| **3/22** | Project proposal |  |
| **3/23** | Code development | Miggy: Extracting and Cleaning Data |
| **3/24** | Code development | Miggy: Extracting and Cleaning Data |
| **3/25** | Code development | Paola: ML Model development |
| **3/26** | Code development | Paola: ML Model development |
| **3/27** | Create Visualizations | Medha, Alicia |
| **3/28** | Code development |  |
| **3/29** | Create Visualizations | Paola: Launch on Heroku |
| **3/30** | Presentation prep |  |
| **3/31** | Mock Presentation |  |
| **4/1** | Finalize the project |  |
| **4/3** | PROJECT PRESENTATION |  |

**Presentation**

Divide your presentation steps to tasks and assign it to members.