

ProtectMe

Detect malicious use of video/photographic content online

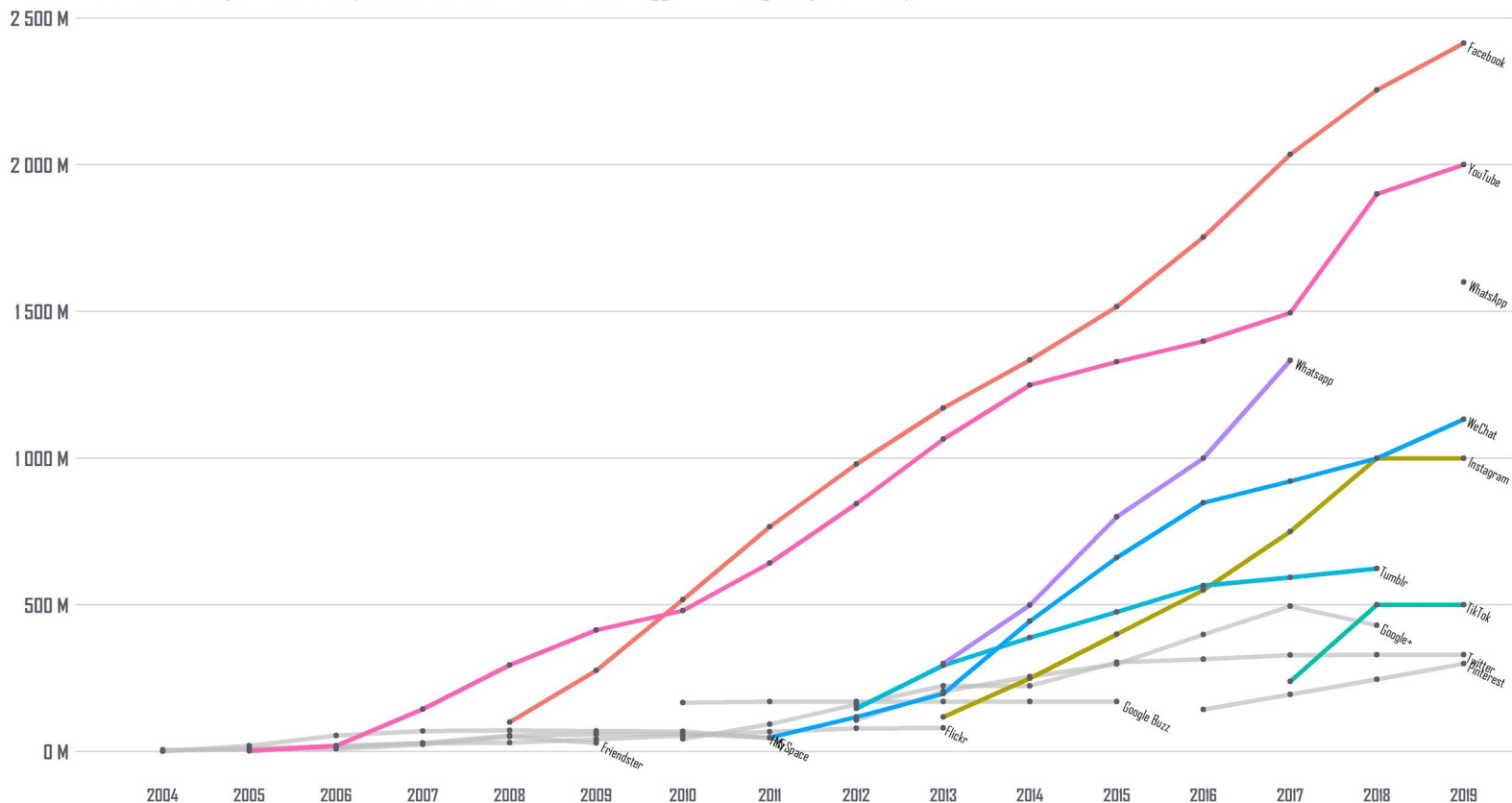
Intro

In the current days, the various forms of social media that exist have become a very impactful part of the common person's life, and day by day, we can see a rapidly rising user base for every social media platform.

And with this fast growth in usage, there is also a growth in users who try to take advantage of these social media platforms to spread malicious content and deceitful information.

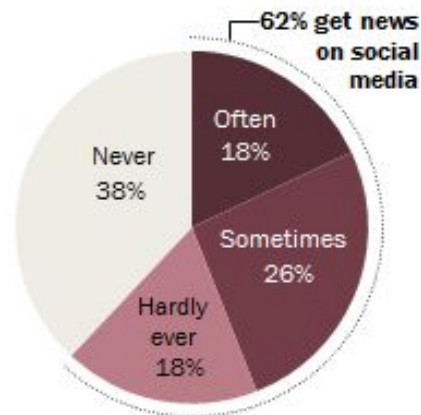
Number of people using social media platforms (2004-2019)

Estimates correspond to monthly active users (MAUs) that have logged in during the past 30 days.



About 6-in-10 Americans get news from social media

% of U.S. adults who get news on a social networking site ...

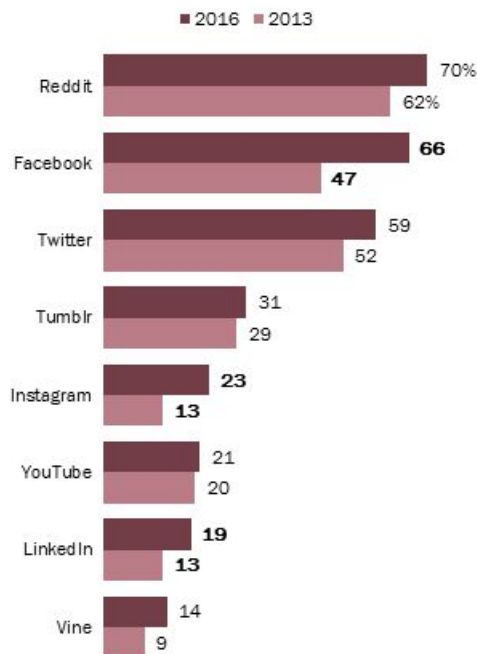


Source: Survey conducted Jan. 12-Feb. 8, 2016.
 "News Use Across Social Media Platforms 2016"

PEW RESEARCH CENTER

Growth in use of social media for news

% of users of each social networking site who get news there

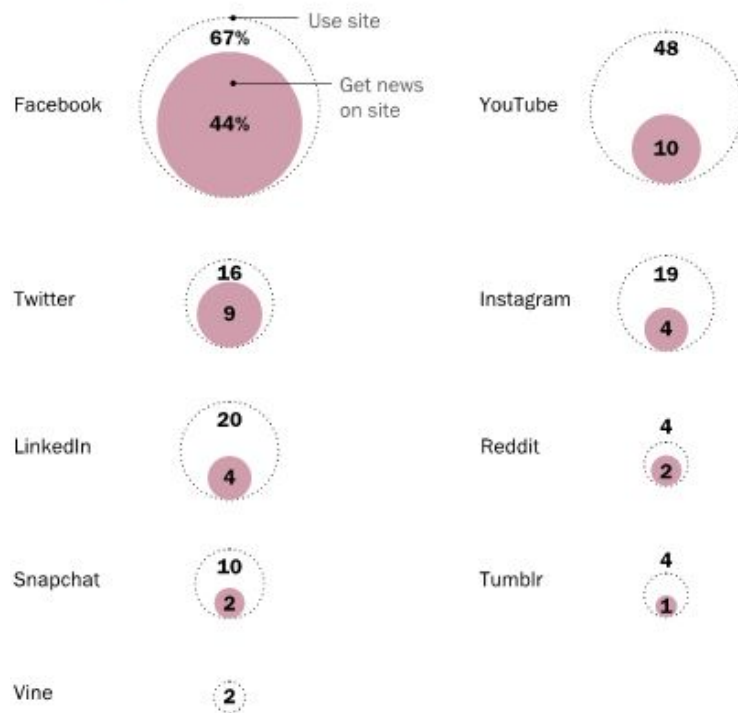


Note: Statistically significant differences in **bold**.
 Source: Survey conducted Jan. 12-Feb. 8, 2016.
 "News Use Across Social Media Platforms 2016"

PEW RESEARCH CENTER

Social media news use: Facebook leads the pack

% of U.S. adults who ...



Source: Survey conducted Jan. 12-Feb. 8, 2016.
 "News Use Across Social Media Platforms 2016"

PEW RESEARCH CENTER

A senator is demanding to know how Facebook will stop misinformation from spreading online

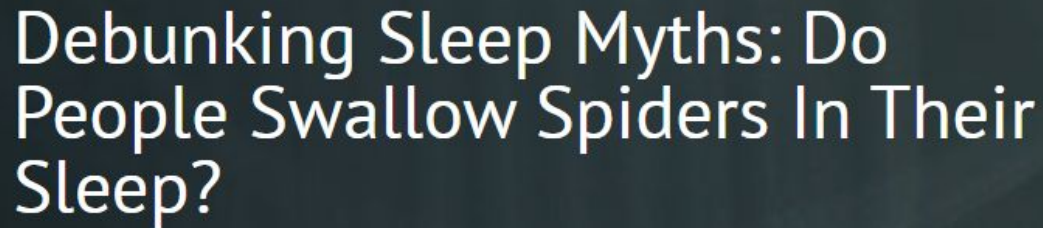
US Sen. Michael Bennet sent a letter to Mark Zuckerberg asking how he's preventing the company from helping undermine democratic elections.

<https://www.vox.com/recode/2020/2/24/21147428/facebook-2020-elections-misinformation-senator-michael-bennet-letter-zuckerberg>



José Figueiras said to be with professional contracts at risk for being accused of being the author of the attacks on the Twin Towers

<https://www.topinfluences.com/jose-figueiras-anima-poder-estar-com-contratos-profissionais-em-risco-por-ser-acusado-de-ter-sido-o-a-torres-gemeas/utor-do-ataques-as>



Debunking Sleep Myths: Do People Swallow Spiders In Their Sleep?

<https://www.sleep.org/articles/debunking-sleep-myths-people-swallow-spiders-sleep/>

Goal

Our project aims to monitor the rate at which this kind of content is spread.

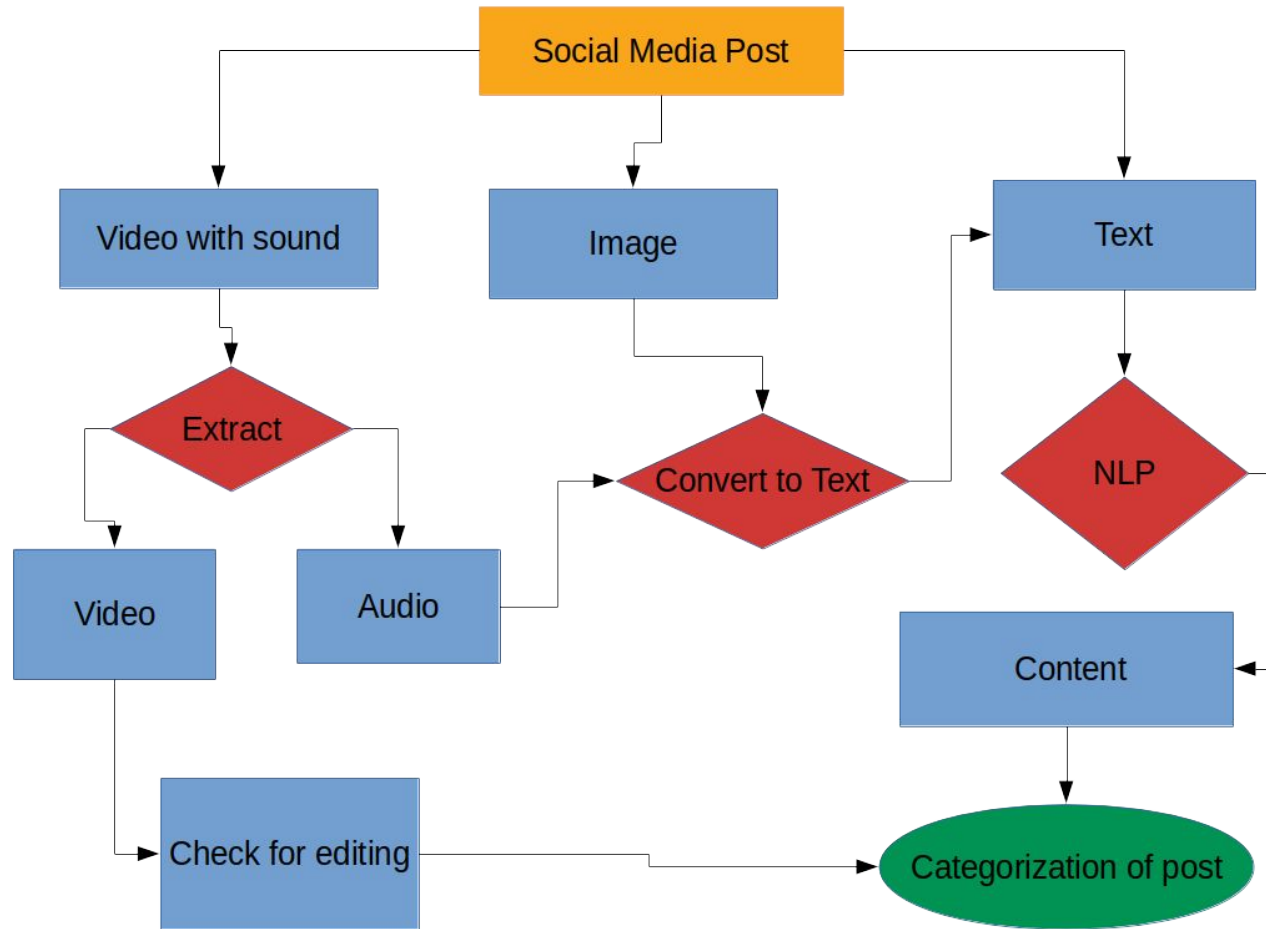
In short, we pretend to create an application which, for a given social media post, will try to extract as much information from it as possible and making use of that information, try to categorize the post as either malicious or harmless.

Tasks

In order to accomplish our goal, we will have to start by researching what is the current State of The Art for Natural Language Processors (NLPs) and softwares that detect manipulation in images or audio files.

After learning and choosing the technologies we deem more appropriate for this job, we will start by defining a data set of social media posts (not only text posts, but also images, video and audio included in the post itself) in order to fine tune everything in order to detect content that was intentionally manipulated from its original source.

Having an already fine tuned core for the application, it will be necessary to find a way to categorize a post as potentially malicious or harmless (using all the relevant data acquired) .



Expected Results

By the time we reach the last milestone of this project (in the context of the PEI class 2019/20), we expect to have a working application that will serve as a proof of concept for what can be done with the technologies we used in order to monitor the spreading of harmful content in a social media environment.