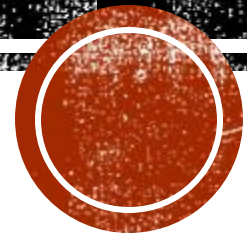




## SPRINT 4

Youssef, Mendel Lion Sjin Tjoe, Boyoung Han

Jorge Edmundo Ortiz Durán, Jaekyu Lee, Petri Pellinen



# CONTENT

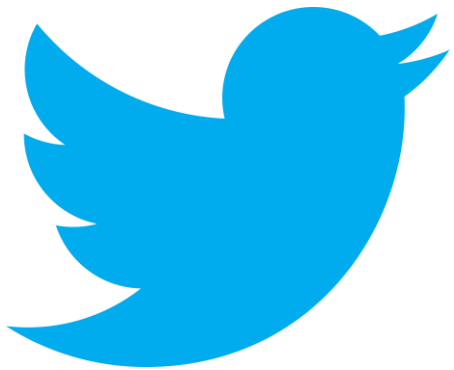
- Overview
- Approach
- Progress
- Next step



# OVERVIEW

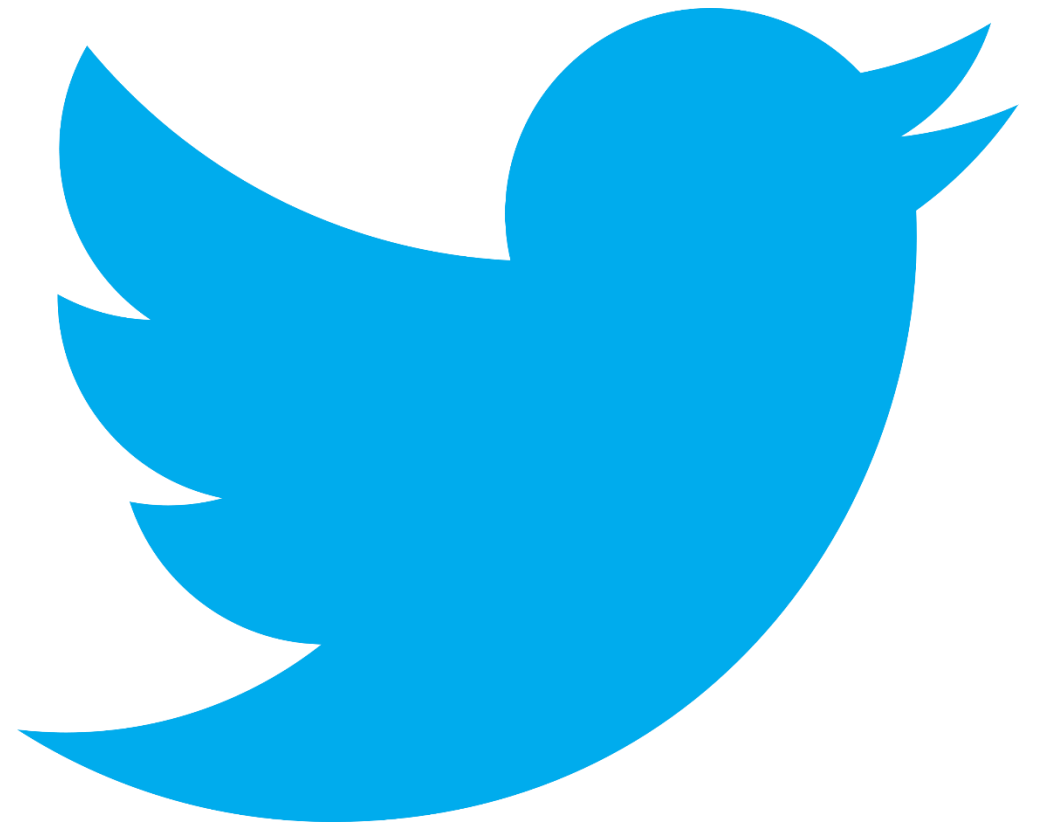
- Objective

“Collecting twitter feeds and  
making Classifier & Sentiment analysis  
Specified for CyberSecurity”



# APPROACH

1. Mining twitter data
  2. Text pre-processing
  3. Terms frequencies
  4. Data analysis
  5. Data visualization
  6. Sentiment analysis
- } Machine learning



# OUR WEEK

- Made classifier with keywords

- using word count function
- from 70 Hacktivist list
- Sum frequency of keywords
- (best 20 list of keywords)

```
: #Counter  
  
counted = pd.Series(np.concatenate([x.split() for x in data]))  
  
print(counted[0 : 30])
```

anonymous	135
opgabon	79
mywhiteninja	72
scode404	72
xpektusanon	67
anonghost	66
cyberwarriortim	55
pursuance	53
youranonrevolt	53



# OUR WEEK

- Made classifier with keywords
  - find average & range about frequency of keywords from 70 hacktivists
  - use this range for distinguishing whether new user is a hacktivist or not
  - train it with machine learning again to make it perfect



# OUR WEEK

- Make new sentiment analysis for new dataset
  - Using 14000 tweets (70user \* 200tweets)
  - Review & label tweets with our hand
  - Negative = 0 , Neutral = 2 , Positive = 4
  - Train it with machine learning again



# OUR WEEK

	Tweets	SA
0	@dburkes I wouldn't be surprised if he would suggest everyone to turn the power off as much as possible when they aren't using it :-p	2
1	@dburkes No computer, no cybersecurity breach. Brilliant move.	0
2	@miyagawa Listen to "rain sounds" now. It immediately fixes the problem and really soothing (until leaving the office).	2
3	@miyagawa I've been listening to "rain sounds" all day today, and been thinking this rain should help the situation a lot.	2
4	@niw I've already decided that I should not use it to mean "sorry" or "thanks" when writing in English.	2
5	@larrychen We can't wait to see new MacBook Pro with ARM, Face ID, Escape, and no goddamn Touch Bar. #takemymoney @pschiller	2
6	@RoKhanna short term: free college because investing our next generation is the most efficient way to make our coun... <a href="https://t.co/6xL6GBNI2h">https://t.co/6xL6GBNI2h</a>	4
7	@MetroManTO @gruber Unfortunatly, he's not the one for _them_ because he doesn't take PAC money. #Beto2020	2
8	For the first time, I just downloaded my twitter data for no particular reason, and learned that the oldest tweet i... <a href="https://t.co/GQMWELeLOh">https://t.co/GQMWELeLOh</a>	0
9	@hak As much as I'm so impressed and excited by the technology, this photo isn't as impressive as it gets, tbh. The... <a href="https://t.co/QDobR237of">https://t.co/QDobR237of</a>	0
10	@kenn Yup. 🙄	0
11	@kenn Huh, typical pg. So many great insights doesn't necessarily mean he's always right.	0
12	@kenn iCloud Photos works great for me too but I also want to automatically post and share random photos to public.... <a href="https://t.co/L54uXwoDIY">https://t.co/L54uXwoDIY</a>	4
13	@kenn Although this doesn't apply to me as a Pro user, I seriously need an alternative to move all my photos.	0
14	The girl came home with a lot of candies and a raccoon doll. "That's not a candy. Where did you get that?" "Don't k... <a href="https://t.co/OtQomLFI2E">https://t.co/OtQomLFI2E</a>	0
15	USB-C to iPad didn't happen because everything becomes USB-C. It's because it makes more sense to replace laptops w... <a href="https://t.co/2gcZX8MA9">https://t.co/2gcZX8MA9</a>	0
16	@miyagawa I hate micro USB so much so that I won't buy any gadgets with it. Hint, hint, @amazon.	0
17	@larrychen @Apple When Dashboard on macOS is obsolete, something like widget on main screen may not appeal as much,... <a href="https://t.co/iXLH8ZEXy3">https://t.co/iXLH8ZEXy3</a>	0
18	@Psychs @tapbot_paul It's definitely not for Xcode. It's not "Pro" in short. I'm kinda considering iPad Pro over Air, now..	0
19	@larrychen @Apple I don't think it's a good idea. 4-column is perfect. Maybe 2X bigger icons may work good, tho.	0
20	Hey @Flickr are you kidding me? It's ios 12 and why do I still have to type my email and password? Please. <a href="https://t.co/1YW8UoEIUI">https://t.co/1YW8UoEIUI</a>	0
21	@moiombo Because the new shallow keyboard is way better. Only problem is that the super sensitive too big touchpad fr... <a href="https://t.co/7NPEzN98vv">https://t.co/7NPEzN98vv</a>	0





# NEXT STEP

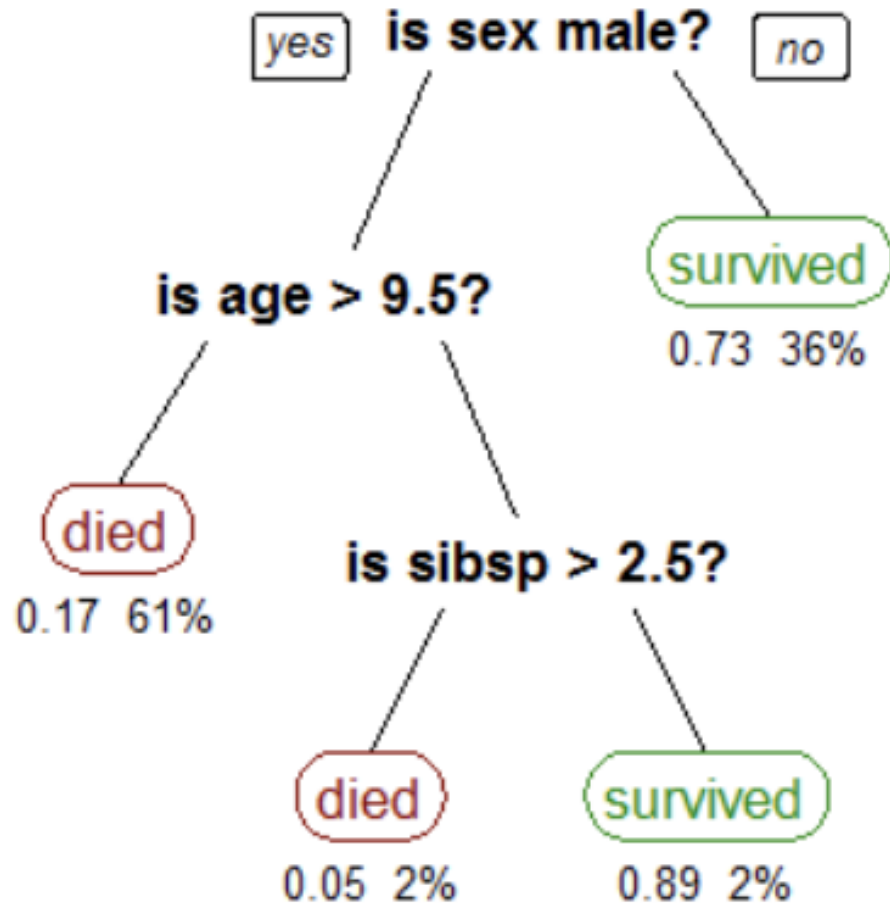
- Review & score tweets with our hand
- Making classifier again



# Q&A



# DECISION TREE



- Data recovery algorithm that creates prediction models,
- What characteristics do these users belong to?
- The figures below the leaf of the tree show the probability of survival and percent of the observations in each leaf.e

