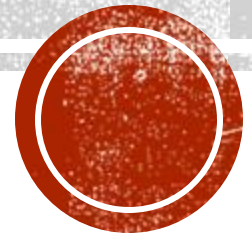


CYBERSECURITY

Youssef, Mendel Lion Sjin Tjoe, Boyoung Han

Jorge Edmundo Ortiz Durán, Jaekyu Lee, Petri Pellinen



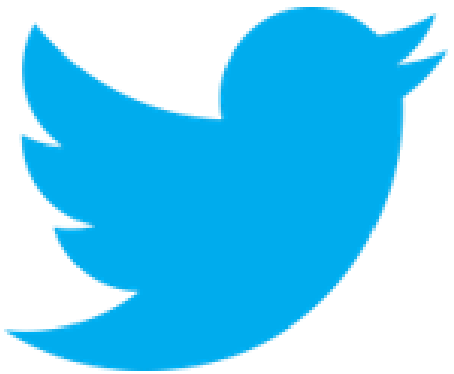
CONTENT

- Overview
- Approach
- Progress
- Next step



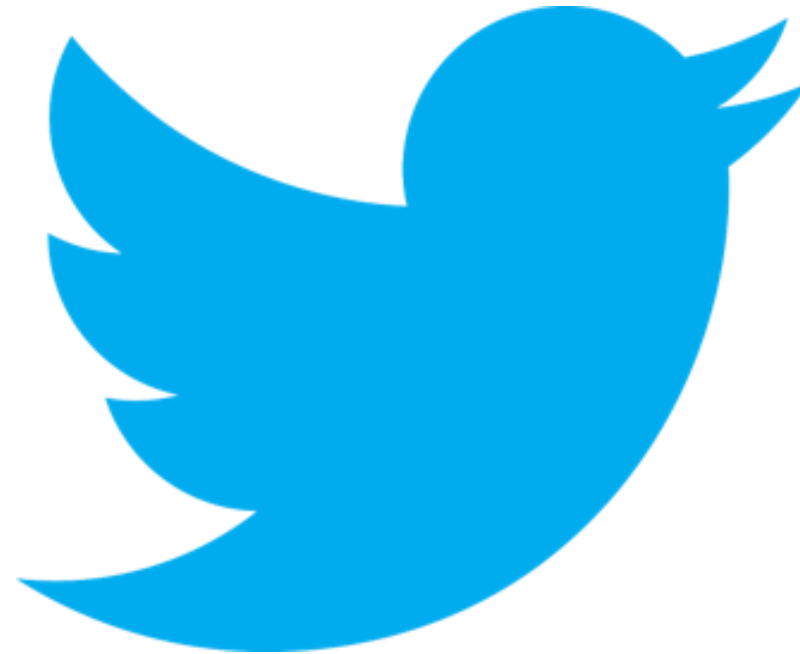
OVERVIEW

- “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



OUR WEEK

- Portfolio
- Research Report



OUR WEEK

	alpha	Train Accuracy	Test Accuracy	Test Recall	Test Precision
0	0.00001	0.963519	0.654241	0.654241	0.654241
1	0.11001	0.956234	0.709368	0.709368	0.709368
2	0.22001	0.948696	0.721550	0.721550	0.721550
3	0.33001	0.940889	0.726884	0.726884	0.726884
4	0.44001	0.932247	0.730188	0.730188	0.730188
5	0.55001	0.924306	0.732703	0.732703	0.732703
6	0.66001	0.916409	0.733855	0.733855	0.733855
7	0.77001	0.908826	0.733370	0.733370	0.733370
8	0.88001	0.901273	0.733309	0.733309	0.733309
9	0.99001	0.893810	0.731370	0.731370	0.731370

	alpha	Train Accuracy	Test Accuracy	Test Recall	Test Precision
0	0.00001	0.975874	0.907391	0.907391	0.907391
1	0.11001	0.967832	0.910359	0.910359	0.910359
2	0.22001	0.964322	0.910359	0.910359	0.910359
3	0.33001	0.963445	0.910359	0.910359	0.910359
4	0.44001	0.960813	0.909766	0.909766	0.909766
5	0.55001	0.955695	0.912140	0.912140	0.912140
6	0.66001	0.949700	0.911250	0.911250	0.911250
7	0.77001	0.942535	0.909172	0.909172	0.909172
8	0.88001	0.934493	0.906500	0.906500	0.906500
9	0.99001	0.928352	0.897596	0.897596	0.897596



OUR WEEK

alpha	0.660010
Train Accuracy	0.916409
Test Accuracy	0.733855

alpha	0.550010
Train Accuracy	0.955695
Test Accuracy	0.912140



NEXT STEP

- Prepare the 1.6 millions tweet dataset for testing
- Test the data with 1.6 millions tweet
- Work in some vizualitation of our results
- Prepare final presentation



Q&A

