

# COVID19 Kaggle Challenge

## Week 2

# Tasks for the next week

1)

Read (skim) 5 papers to understand the general structure of peer reviewed articles.

# Tasks for the next week

2)

Make a list of objectives you are most interested in (ranking from 1 to 10). A survey will be sent to gather your responses.

# Tasks for the next week

3)

## Data processing step

- Tokenization
- Remove punctuation and special characters

# Tasks for the next week

4)

- The objective is to create a Search engine to find papers based on key words.
- Use Regular expressions to find one word
- A combination of words

# Tasks for the next week

5)

- Cluster papers

# Tasks for the next week

6)

- Topic Extraction (opposite of Clusters, similar to PCA)

# Tasks for the next week

7)

- Can clustering be done as the second step, after narrowing the scope using step 4?



# Tasks for the next week

8)

- How can Word2Vec be used?

9)

- How can Sen2Vec be used?

10)

- How can Doc2Vec be used?