horizontal line

Sprint 1 Results

Sprint End: 10.20.2019

Sprint Captain: Mianchun Lu

Sprint Co-captain: Hongying Liu

**─**

# Sprint Goal

## BabbleLabble implementation first iteration

1. Run first iteration of BabbleLabble implementation **on track**

## Finalize the Pitch Presentation and description of how the additional feature engineering will be used

1. Provide draft (End of Monday) and make the pitch presentation Thursday **on track**

# Results

Main: We clarified the workflow of this project, and has developed a clear plan for our future steps, which includes feature engineering, labeling model optimization and neural network optimization.

* Feature Engineering:

We took nine samples from the training data and did a quick EDA(e.g. Term frequency, similarity matrix, sentiment analysis) on them, and we found that the term frequency and similarity coefficient should be added to our future models.

* Labeling model optimization:

We compared Snorkel and Babblelabble and figured out the difference between them. We also learnt the sample codes for Babblelabble on github.

* Neural network optimization:

We learnt the neural network pipeline that WRI applied, and figured out how to test deeper and more complicated neural networks. We also tested RoBERTa with same samples, and found some possible alternatives to improve the performance based on some readings.