Critique of Manik Insights (Function4) Progress Video

What we liked:

- The team made it extremely clear what the business problem was, and how it was converted into a machine learning problem
- It was explained clearly what useful attributes they would use from their initial data analysis.
- I liked how they explained what next steps they needed to take after the initial analysis of the data.
- NLP seems like a good approach since fine tuning could be easy and chance of interpretability depending on exact implementation.

What we didn't like:

- It was not explained how and why the other 15 attributes were not useful. It would have been nice to get at least an explanation, if not a brief showing of those 15 not useful attributes.
- They did not explain what NLP techniques they would use to aid in their binary classification of relevance. What features are to be generated from their NLP? What NLP techniques will they be using?
- They did not explain how to determine confidence level/determine relevance of event. Are they going to be using the prediction score that naturally comes with the model? Or is the team going to be using confidence intervals?
- The relevant dataset has a lot more instances than the irrelevant dataset. How will balancing be done?
- Presentation itself only had 3 or so slides.
- Sudden cut-off at the end of the video. We were waiting on a bit more information, actually.

Suggestions:

- If the specific instances are already tagged, will the model then be necessary? Maybe useful to have a separate naive algorithm.