Civera SCRUM file 3

April 1, 2021

Meeting Notes and Next Steps

* What have we worked on?

Ammar and Sherry have done significant semantic analysis on the case descriptions using tfidf and glove embeddings to extract actors and actions. They have looked into different clustering methods like kmeans and random forests (and a few others?).

* What will we be working on next?

We are still working on identifying the best clustering strategy. Unsupervised learning methods are looking more promising than the supervised learning methods we have tried so far. We are also still working on converting the existing php code into python.

After we identify the best clusterings, we will do some post processing to fill in the rest of the actor/action columns.

* Have I run into any issues? Do I need help?

We are currently struggling to identify the best k value for kmeans clustering. The quantity of data is making it very difficult to run efficiently. We may need help with this.

* Have I talked to the client recently? When are we meeting with them next?

Our last meeting with the client was 3/25. Our next expected meeting with the client is April 8.