**Meeting 09/10/20 Notes**

* Meetings now Thursdays at 9:30
* Do project on something academically novel – identify limitations and gaps in literature
* Look at datasets – be critical - check for class imbalance/size of dataset/if group underrepresented
* Research toxic language – look at the subject, why is the message toxic, figure out underlying reason/subject (e.g. everything against Hillary sexist)
* Look at misclassified instances – may discriminate against sub-group
* Users may tweak language to evade detection
* Come up with ideas for project – e.g. see what happens if mask offensive words from dataset – can we still accurately tell toxic language based on context? (might be innocent if between friends)
* Start GitHub repository – keep project diary, create table of datasets, summarize papers, list of ides, push changes regularly (later create slides for mini presentations)