I will be conducting a novel application of entity recognition, entity typing, and relation mining for the task of classifying moral rhetoric in a large corpus of annotated Tweets. In a recent paper, Lin et al. (2017) conduct "Entity linking" to supplement established document classification methods to predict a Tweet's five moral labels. These moral labels, which are theoretically derived from Moral Foundations Theory (see Graham et al., 2013, for a synopsis), are a set of binary labels that apply at the document level. Each label, when active, signifies that a given “moral concern" is used by the writer of the particular text. There are five labels, corresponding to the five moral foundations: fairness, care, authority, loyalty, and purity. In Lin et al. (2017), the authors link entities in Tweets and link them to an external Knowledge Base (Wikipedia) and append the text of the linked entities' abstracts to the document. They achieve superior performance using this method, in comparison to purely text-based methods.

In my project, I propose to improve this classification in two ways: (1) In terms of classification performance; (2) In terms of explainability. In this field, a key element of any predictive system is the ability to provide "reasoning" for each prediction, or a “trail of evidence" which influences the learner’s decision. Such information can facilitate collaboration between NLP researchers and domain experts, allowing the latter to design schema for future IE given the evidence provided to them in text classification.

The method I will be using for classification is an attention-based LSTM (e.g. Wang et al., 2016). To form a coherent set of features which offers explainability in the form of entities (and their types) rather than just words, I propose to conduct distantly-supervised entity recognition and typing in the style of Ren et al. (2015). This will supplement traditional features in text classification (word and character n-grams, word embeddings, syntactic feature mining).

In terms of the knowledge bases I will use to conduct distant supervision of the entity recognition and typing, I will use both established (i.e. Wikipedia) KB as well as less traditional, more specialized schema. This latter category of KB include actor- and event- ontologies in the political science community, notably CAMEO (Schrodt et al., 2008). Additionally, I have been working on developing a schema for structuring “political objects", including persons, groups, events, and concepts, which will be applied at later stages of the project, time allowing.

The data I am using for this project is a set of roughly 30,000 annotated Tweets. Each Tweet has multiple annotations for each of the 5 moral foundations. Evaluation will be done by comparing classification metrics (precision, recall, F-1, and accuracy) between several baselines and the proposed method. Those methods will be:

- Linear SVM with TF-IDF word feature weights
- LSTM with word embeddings
- LSTM with word embeddings (add attention)
- LSTM with embeddings & attention, segmented by entity recognition.

If the above modeling is finished early, I will experiment with: (1) New entity schema, with dictionaries of entities with corresponding hierarchical types; (2) Utilize large, unlabeled datasets of Tweets and political texts to learn embeddings of entities and entity types.

The big-picture outcome of the project, if successful, is an approach to content analysis for social scientists that offers a link between external knowledge and the text of the analyzed documents, such that each entity’s relation to the given construct (in this case, morality) is evident from the model.
References


Xiang Ren, Ahmed El-Kishky, Chi Wang, Fangbo Tao, Clare R Voss, and Jiawei Han. Clustype: Effective entity recognition and typing by relation phrase-based clustering. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 995–1004. ACM, 2015.