Name: Relation Extraction with Deep Learning
Members: Ankur Bapna (ankurbpn), Melvin Jose (melvinj), Mikhail Sushkov (msushkov)

Our project aims to apply deep learning methods to the problem of relation extraction. In particular, we plan to use the dataset and evaluations from the slot filling track of the TAC KBP annual competition held by NIST. The competition defines a set of 41 relations and provides a publicly-released training corpus of 6 million documents. We plan to start with a simple RNN and then extend our approach using existing related work. In particular, we hope to apply the techniques described in [1]. Related work in [2] shows benefits from matrix factorization style algorithms which look at the global context as well. This indicates that a deep learning-based approach could be beneficial.

Since the KBP competition provides an evaluation set each year, we will use the 2010 evaluation set as development data and use the 2013 evaluation set as test data to evaluate our models. Additionally, the NLP group has developed a pipeline to test the performance of relation extraction systems that are part of the KBP framework. We will use this pipeline for evaluating our models and comparing their performance against existing results.

[1] Socher et. al., Semantic Compositionality through Recursive Matrix-Vector Spaces. EMNLP 2012.
[2] Riedel et. al., Relation Extraction with Matrix Factorization and Universal Schemas. NAACL 2013.