

Executive Summary

Business Insider writes articles about different financial, political, and global topics. In this project, we're interested in extracting the names of all CEOs, all companies, and all percentages from articles spanning 2013-2014. I built a NER that is based on logistic regression to extract those features, and in total I found 989 CEO names, 663 companies, and 14351 percentages. These lists are by no means exhaustive, but they consider a wide scope of candidates. Ultimately, the models portray an exploration in building an NER, but are limited by their simplicity.