

Progress Report for Project Chameleon (Jacob Huss, Chaitanya Jujjavarapu, and Edward Park)

1. We have successfully scraped the needed IEM data from the IEM website into a CSV. We then wrote IEMcsvTransformer.py, which took this data and calculated the normalized price for each date. These normalized prices were stored in data/IEMPrices.txt. We then wrote a function in DataReader.py to read these prices into a dictionary. For the 2nd part of our scraping process, we had to create a program that fetches the data from all 47000 xml files between the dates of May 2000 and October 2000, and then parses the text of each article so that only our matching query terms were pulled out and outputted and cleaned up into a text file through a 'bag of words' approach.
2. Now that we have our data completely scraped and prepared, we need to implement the algorithm detailed in the paper. In order to do this, we will first need to break down our algorithm into smaller pieces that can be tackled by each team member. Doing this will require us to first gain a better understanding of the algorithm itself.
3. The biggest challenge that we are facing is understanding how the algorithm for our function works.