Opportunity and Significance
The purpose of this version of Gekko is to implement a sentiment analysis module capable of determining the market sentiment for a user-defined company.

Technical Objectives
- Create a dictionary of keywords exclusively used in advertising
- Create regular expressions to recognize grammatical patterns used in ads
- Implement a neural network classifier to recognize subjective/objective statements
- Collect training data in order to train bayesian classifier to score tweets as positive or negative sentiment.
- Fine tune classifier through natural language processing for improved results

Related Work and State of Practice
Currently, the calculation of sentiment analysis through natural language processing is a popular technique for financial analysts. Machine learning techniques are often applied to pick up on patterns that may be imperceptible to humans. With this in mind, Gekko utilizes several neural networks for the purposes of filtration of irrelevant data and the calculation of sentiment scores. One unique approach is Gekko’s use of keywords and regular expressions to detect sentence structures most commonly found in ads or objective statements.

Technical Approach, Accomplishments and Results
Gekko utilizes the Twitter Streaming API to extract Twitter data to a MongoDB database on a remote server. Tweets consisting of advertisements or objective statements are filtered out using a dictionary approach, common grammatical patterns, and a Multinomial Naive-Bayes classifier. Sentiment analysis is performed on the remaining tweets and the associated sentiment scores are stored in the database for analysis.

Achieved Sentiment Accuracy: 73% (with slight variation)
Achieved Filtration Accuracy: 86% (with slight variation)

Next Steps for Development and Test
The module is currently deployable in its present state. It was determined that a Multilayer Perceptron yields higher accuracy than Naive-Bayes, but at a severe cost to performance. It may be worth it to examine this route further in the future.

Commercialization Plan & Partners
Team members worked in collaboration with Jameson Ruble of WSU’s Department of Linguistics. General requirements were determined by the client company, Locksley Technologies.

Regarding potential commercialization, Gekko’s sentiment analysis process was designed in a modular fashion, to best allow it to be incorporated into future projects.

The main requirement for commercialization would consist of implementing a unique strategy for stock prediction, based on Gekko’s determined market sentiment. The main hurdle would be in finding a truly unique strategy, as similar tools for financial analysts are already popular in the industry.

References
http://www.nltk.org/
https://api.mongodb.com/python/current/
http://docs.tweepy.org/en/v3.5.0/