Stock Chatter: Using stock sentiment to predict price direction

This paper examines a popular stock message board for eleven stocks and finds that only 17.3% of posts include a user-provided sentiment of “strong buy”, “buy”, “hold”, “sell” or “strong sell”, whereas the remaining include sentiment only within the context of the post. We use several supervised machine learning algorithms to infer the sentiment from these post's text and find nearly double contain sentiment regarding the stock. This predicted sentiment is found to be more predictive of the underlying stock's price directional change than the user-provided sentiment. Additionally, the profit potential in trading stocks and the anonymity of the boards creates an environment prone for abuse. We found groups of users that display greater likelihoods of writing spam and provide higher quantities of sentiment. The term for users that entice others to buy and then sell at artificially inflated prices are “pump and dumpers.” A study by Frieder and Zittrain in 2007 found that investors that bought stocks where “pumpers” previously sent large quantities of touts, lost, on average, 5.25% . We eliminate these participants from the dataset to determine if the market is more predictable, but find no discernible difference.