1135-VN-2440 **Phylisicia Nicole Carter*** (pncarter@crimson.ua.edu), 345 Gordon Palmer Hall, Box 870350, Tuscaloosa, AL 35487. *Sparse Regression for Twitter Analysis.*

We consider sparse regression techniques as tools for classification of sentiment within Twitter posts. Analysis of Twitter usage suffers from several unique challenges. For example, the 140-character limit severely limits the amount of information contained in each post; this causes most tweets to contain an extremely small subset of the dictionary, presenting challenges for learning schemes based on dictionary usage. To remedy this undersampling issue, we propose usage of penalized regression. Here, we employ logistic regularization to avoid any degeneracy caused by the sparse usage of the dictionary in each tweet, while simultaneously learning which terms are most associated with each sentiment. As an illustrative example, we employ sparse logistic regression to classify tweets based on the users' perception of a connection between vaccination and autism. Our regression scheme provides a classification function as well as feature selection in the form of a list of words most associated with pro- and anti-vaccination sentiments. (Received September 26, 2017)