# #feels: Detecting and Visualizing Regional Sentiment from Cross-lingual Tweets for Specific Hashtags Using SVM and Naive Bayes Classifiers

Alhamdan, Abdul Rahman

There is a wealth of available information on sites like Twitter that is not being utilized to its full productive potential. Many barriers exist between users, such as mistranslation or misrepresentation by the media. This project aims to create a tool to compare and analyze the opinions of people from different regions. A program was developed that collects tweets in a user-specified hashtag, classifies each tweet into positive/neutral/negative in terms of sentiment, and visualizes the sentiment in each country by assigning it a representative color. Multiple algorithms were tested, including Support Vector Machines and Naïve Bayes Classifiers. After evaluation, Naïve Bayes was found to work best, achieving an accuracy of 74.% when trained and tested on the Semeval 2013 English tweet dataset. An efficient method of reverse geo-coding that runs on a local client was developed, as opposed to submitting a request to an API. While other projects target the United States, the focus of this study had a global scope. Different multi-language support methods were examined in a comparative analysis and the best cross-language classification approach was chosen. Support for multiple languages will eliminate the bias of only analyzing English tweets. The program can be used for social purposes, such as analyzing the change in sentiment with the difference in geographical factors, such as location or culture. It can also be used for consumer behavior or business analytics, such as detecting customer dissatisfaction in a certain region or recommending systems to aid in decision-making.

**Awards Won:**