

Molly McNamara

Springboard Capstone Project Proposal 2

Text Analysis of Tweets on Net Neutrality

Net neutrality is defined as “the principle that Internet service providers should enable access to all content and applications regardless of the source, and without favoring or blocking particular products or websites.” In the United States the Federal Communications Commission (FCC) has for several years classified the internet as a telecommunications service in order to protect net neutrality. Recently this order has been rescinded and a measure of disapproval is expected to be voted upon in the US Senate.

Twitter is an online news and social media platform with approximately 330 million active users. Given that users of Twitter are by its very nature users of internet services, sentiment analysis of tweets could provide a strong case for lawmakers trying to understand the public consensus on the subject or how their constituents wish the issue to be handled. The state of Montana, for example, has already enacted its own legal protections to force net neutrality from providers contracted with the state and the state of New York is expected to make similar provisions. By providing an analysis of current tweets, this information could be used to persuade state and national representatives to vote according to the position of the voters they represent.

The Data

Data will be collected from Twitter using Python and the Tweepy app to gather all tweets with the search term NetNeutrality over an adequate timeframe for a large dataset. This should include not only the text of the tweets but also user information such as username, location and any profile description.

Potential Analysis Strategy

The proposal for this project is to perform text/sentiment analysis of the tweets to identify trends or sentiment on the subject of net neutrality. Regional trends will be identified if possible, isolated by state or region as available, and these could be compared to the actual votes in the Senate if it takes place before conclusion of this project.

Deliverables

After analysis of the data, the code used to develop the models will be shared. A formal report will be developed detailing the results of the exploratory data analysis, the statistical analysis and the results of the modeling, and this information will also be captured at a high level in a slide deck. The end goal would be to have some recommendations by region or state on how voters want their representatives to act on the issue.