sentiment analysis iphone vs galaxy   
report   
for helio corporation

Alert Analytics

Overview:

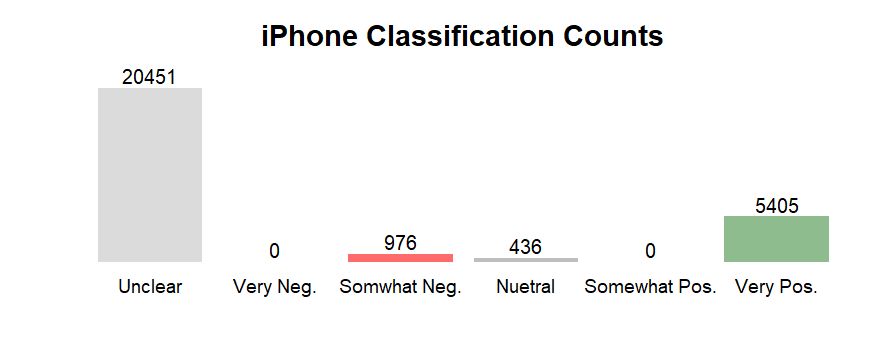
This is a report on the findings of a sentiment analysis conducted on the iphone and Galaxy smart phones. The analysis is done for Helio Corporation whom is developing a suite of smart phone apps to be used by the government health agency in developing countries. The health agency needs the app to be used on only one phone due to limited support and resources. As such, Helio needs to understand which smart phone has the most positive sentiment world-wide to facilitate the apps use. A collection of over 27000 recent cell phone reviews was acquired from the web via Common Crawl on Amazon’s Public Data Sets (aws). Machine learning models were then then applied to the data to determine the sentiment of each smartphone review.

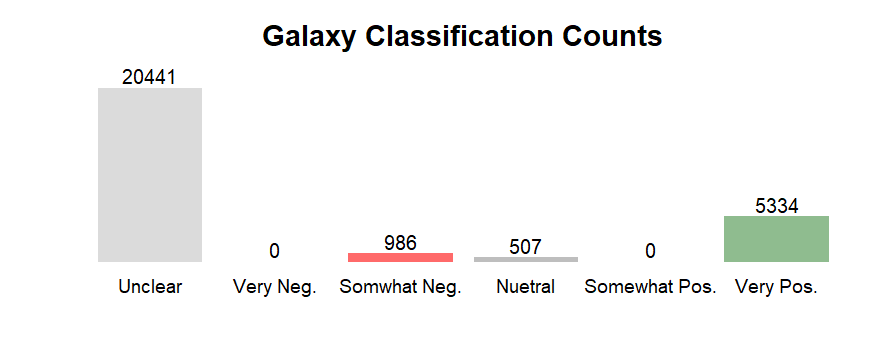
Findings:

For each model, the sentiment of each smart phone review is classified into one of six categories:

* 0 => sentiment unclear
* 1 => very negative
* 2 => somewhat negative
* 3 => neutral
* 4 => somewhat positive
* 5 => very positive

Results from the two models are shown below:





Both models are very close in sentiment and given the confidence of our models there is no clear favorite between the two phones. Sentiments are very similar for each phone.

Note, not all records are specific to the iPhone or Galaxy, there are 3 other phone models included in the data set. This explains the large number of “unclear” counts in each model results.

Confidence:

The performance of both models is very similar as shown in the table below.



At these performance levels, we can say that we have good confidence in each model’s result.

Implications:

Since both phones are equal in sentiment, other factors should be considered for selecting which phone to develop the apps for. Some possible factors may be: cost of app development, cost of support, and market share of each phone.

Methodology:

The data for this analysis was taken from a scrape of the Common Crawl data base that resides on Amazon Web Services servers. Common Crawl is a monthly capture (snap shot) of all data on the Web. The scrape was performed by a text analysis program specifically designed and written for this exercise. In the end, a data set with 59 columns (features) and 27,273 records was captured.

For each phone, smaller data sets were taken from this data and each record was manually read and given a sentiment score by the reader in line with the 6 classification levels. These two data sets were then used to train the models for the iPhone and Galaxy. Once models were optimized, they were then applied to the entire data set.