**Analysis of Overall Sentiment Toward iPhone and Samsung Galaxy**

Analysis and predictions for this assignment were performed using Large Matrix Files containing the following information:

* positie, negatie, and uncertain review counts of the following characteristics:
* display
* camera
* performance
* for the following devices:
* iPhone
* Samsung Galaxy
* Sony Xperia
* Nokia Lumina
* HTC Phone
* iOS
* Google Android
* Sentiment Score Value (for iPhone and Galaxy)

The Sentiment Score information was grouped into the following meaningful categories:

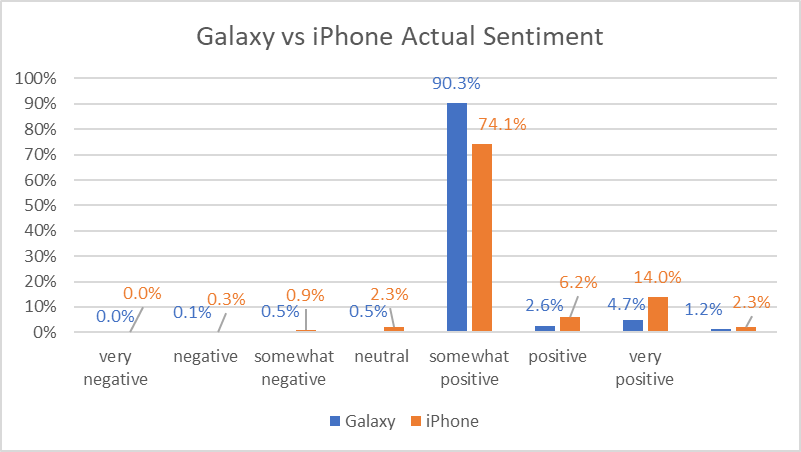
|  |  |
| --- | --- |
| **Score Range** | **Score Label** |
| -Inf - (-50) | very negative |
| (-50) - (-10) | negative |
| (-10) - (-1) | somewhat negative |
| (-1) - 1 | neutral |
| 1 - 10 | somewhat positive |
| 10 - 50 | positive |
| 50 - inf | very positive |

The same amount of data was analyzed for iPhone as it was for Galaxy. iPhone file contained 13,741 records and Galaxy file contained 13,741 records as well.

Galaxy and iPhone both had most of the sentiment rated as neutral. iPhone had lower rate for neutral than Galaxy, however iPhone had much higher positive sentiment scores.

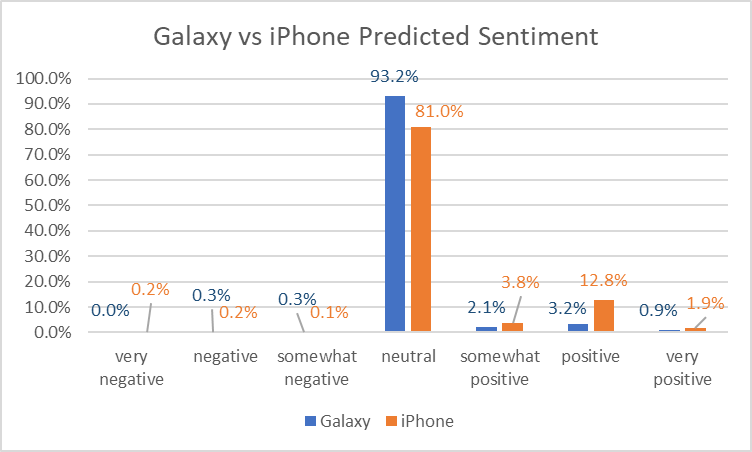
Even though negative sentiments were very small for both devices, Galaxy’s negative scores were even smaller than iPhone’s.

|  |  |  |
| --- | --- | --- |
| **Actual** | **Galaxy Sentiment** | **iPhone Sentiment** |
| very negative | 0.1% | 0.3% |
| negative | 0.5% | 0.9% |
| somewhat negative | 0.5% | 2.3% |
| neutral | 90.3% | 74.1% |
| somewhat positive | 2.6% | 6.2% |
| positive | 4.7% | 14.0% |
| very positive | 1.2% | 2.3% |



Based on the results of predictive analysis, I expect with 95% of accuracy the following sentiment results for Galaxy and 90% of accuracy for iPhone:

|  |  |  |
| --- | --- | --- |
|  | Galaxy | iPhone |
| very negative | 0.0% | 0.2% |
| negative | 0.3% | 0.2% |
| somewhat negative | 0.3% | 0.1% |
| neutral | 93.2% | 81.0% |
| somewhat positive | 2.1% | 3.8% |
| positive | 3.2% | 12.8% |
| very positive | 0.9% | 1.9% |



As with the actual sentiment scores, most of the predicted scores have neutral sentiment toward both devices -- Galaxy 93% and iPhone 81%. Neutral sentiments can potentially be changed in either positive or negative score direction. While negative scores were relatively very small for both iPhone and Galaxy, in the positive group category iPhone had an obvious lead.

Some of the attributes were highly correlated amongst themselves and therefore needed to be removed from the analysis, the final number of the attributes used in the analysis was 29 out of 59.

The following attributes were the most influential in determining sentiment score for iPhone:

* iPhone display positive
* iPhone display uncertain
* iPhone camera positive
* iPhone performance uncertain
* iPhone

The following attributes were the most influential in determining sentiment score for Galalxy:

* Samsung camera positive
* Samsung display negative
* Samsung performance uncertain
* Samsung Galaxy
* Google performance negative

The best classifier algorithm for iPhone data was Random Forest, while C50 performed the best on Galaxy dataset.

The predicted results of the sentiments toward iPhone and Galaxy were close, however positive sentiment group was larger for iPhone. I recommend iPhone as a preferred choice.