Apple Enhancement

Enhancing Customer Experience Through Text Mining

This Capstone Project is submitted in partial fulfillment of the requirements for the course Applied Business Analytics (MDA 720) during the Spring Semester of 2023.

While writing this Capstone Project, we have not witnessed any wrongdoing, nor have we violated any conditions of the LIU Honor Code.

Anthony Vincent

April 25, 2023

Table of Contents

Background

In today's business world, enhancing customer experience and understanding the customers is a crucial part of businesses and how they are able to grow. A company that understands the customer's needs and satisfies their requests is more susceptible to seeing a loyal customer following that will enhance business revenue.

Taking the company that I want to focus on is Apple. Monitoring and identifying certain products that include MacBooks, iPhones, Ipads, Apple Watches, or any other products will be a valuable step in recognizing products that customers value. Focusing on these products, and using sentiment analysis to extract relevant data, can provide valuable information on how customers view the products. That is why identifying crucial and relevant information will only enhance the process and company performance in the long run. It will also allow me to generate and understand customer opinions in order to generate new products or expand on existing ones.

Finally, with all of this information that I will gather, Apple will be able to make more informed decisions on how to enhance their products, where they need to place more priority on product development, and in turn, this will create a more enhanced customer experience. From there, the growth and success of the company can take continued positive strides.

Objective/Goals of the Project

My plan is to deliver a clearer message on common issues and create an approach that uses sentiment analysis to promote Apple products in a positive way. Sentiment analysis involves the process of analyzing text data to determine its tone. This will help generate insight into customer sentiment toward Apple products. I can even build classification models in order to

help in decision-making backed by data that will help enhance customer experience and in turn, hopefully, enhance the sales of Apple products.

Data Exploration

Apple is a multinational technology company that is known worldwide for its products that include iPhones, Ipads, Apple Watches, AirPods, and software including ios, macOS, and watchOS. They are known for their design-driven approaches that have formed a power-house technology business that is in the hands of millions and millions of people. So as I started my analysis, I first wanted to decipher an approach that I could use in order to have a specific starting point. In class, we explored topics that included web scraping, text mining, sentiment analysis, and Google trends to provide valuable insights into a company. What stood out to me was sentiment analysis.

Apple has such a strong brand image on social media, and that is one key reason why I decided to utilize sentiment analysis in order to explore this area more. For my data exploration, I found valuable information online, (data.world) where contributors were given a tweet and asked whether they were positive, negative, or neutral on Apple. However, this was not as easy as it sounds. I sifted through numerous datasets that involved Apple tweets from Twitter, but many lacked the necessary information for me to provide this sentiment analysis. Some lacked the sentiment analysis component, and others created headaches for me as I tried to run my analysis. Ultimately, this was the most accurate and reliable dataset that turned out to be the most helpful and helped enhance my goals of understanding the customer, and what product I could use to add benefits to Apple itself.

Data Visualization

Image 1. Calling Specific Columns

```
1]: data[['sentiment','text']]
```

1]:

	sentiment	text
0	0	#AAPL:The 10 best Steve Jobs emails ever...htt...
1	0	RT @JPDesloges: Why AAPL Stock Had Mini-Flash ...
2	0	My cat chews @apple cords. Such #AppleSnob.
3	0	I agree @jimcramer #IndividualInvestor trade #...
4	0	Nobody expects Spanish Inquisition #AAPL
...	...	...
3881	0	(Via FC) Apple Is Warming Up To Social Media -...
3882	0	RT @MMLXIV: avocado emoji may I ask @apple
3883	1	@marcbulandr I could agree more. Between @Appl...
3884	0	My iPhone 5's photos longer downloading automa...
3885	1	RT @SwiftKey: We're excited named @Apple's 'Ap...

3886 rows × 2 columns

Image 2. Sentiment Analysis Visualization (0 = negative, 1= positive)

```
In [32]: data['sentiment'].value_counts()

Out[32]: 0    3463
         1     423
         Name: sentiment, dtype: int64
```

Image 3. TF-IDF

```
]: consolidated=' '.join(word for word in data['text'][data['sentiment']==0].astype(str))
wordCloud=WordCloud(width=1600,height=800,random_state=21,max_font_size=110)
plt.figure(figsize=(15,10))
plt.imshow(wordCloud.generate(consolidated),interpolation='bilinear')
plt.axis('off')
plt.show()
```

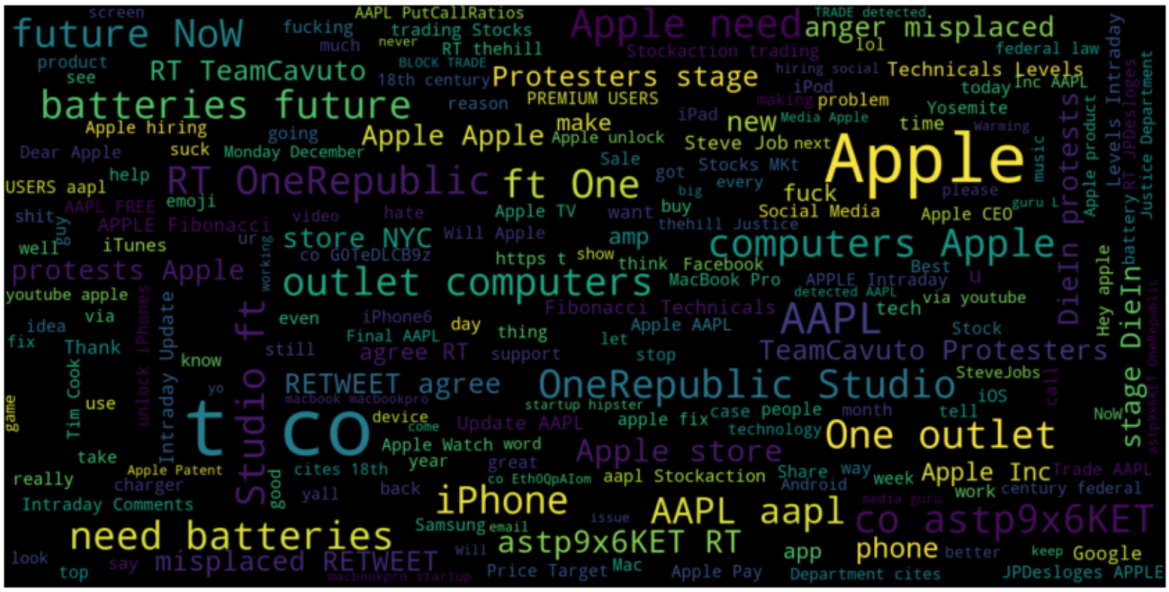

Image 3 shows TF-IDF or Term frequency-inverse document frequency. TF-IDF calculates how relevant or common a word is in a series to a text. The words increase proportionally to the number of times a text of a word appears in the data set. As you can see the extraction of Apple is very common as indicated by a larger visualization in image 3. Apple is the company that I am focusing on and it is clear that it appears several times in my first visual extraction technique. However the key focus is on extracting positive and negative words from Apple tweets in order to see how customers view the products, and from there I can make an informed decision on what product I might use or remove, and see what possible products I can implement or what possible additions I can add to make Apple a more successful company. I will explain this clearly in the next section of my data analysis/sentiment analysis extraction techniques.

Image 4. TF-IDF (negative words)

```
data_neg = data['text'][:800000]
plt.figure(figsize = (20,20))
wc = WordCloud(max_words = 1000 , width = 1600 , height = 800,
               collocations=False).generate(" ".join(data_neg))
plt.imshow(wc)
```

```
<matplotlib.image.AxesImage at 0x7fb9f674dd30>
```

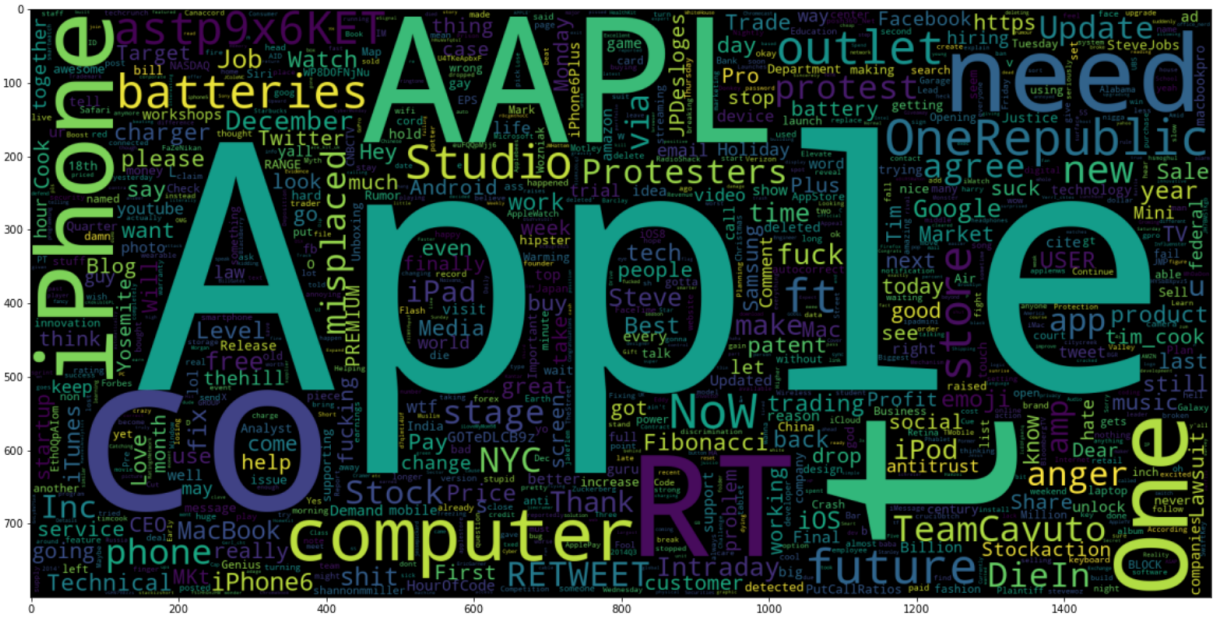

Image 5. TF-IDF (positive words)

```
data_pos = data['text'][:800000]
wc = WordCloud(max_words = 1000 , width = 1600 , height = 800,
               collocations=False).generate(" ".join(data_pos))
plt.figure(figsize = (20,20))
plt.imshow(wc)
```

```
<matplotlib.image.AxesImage at 0x7fb9e1e5d2b0>
```

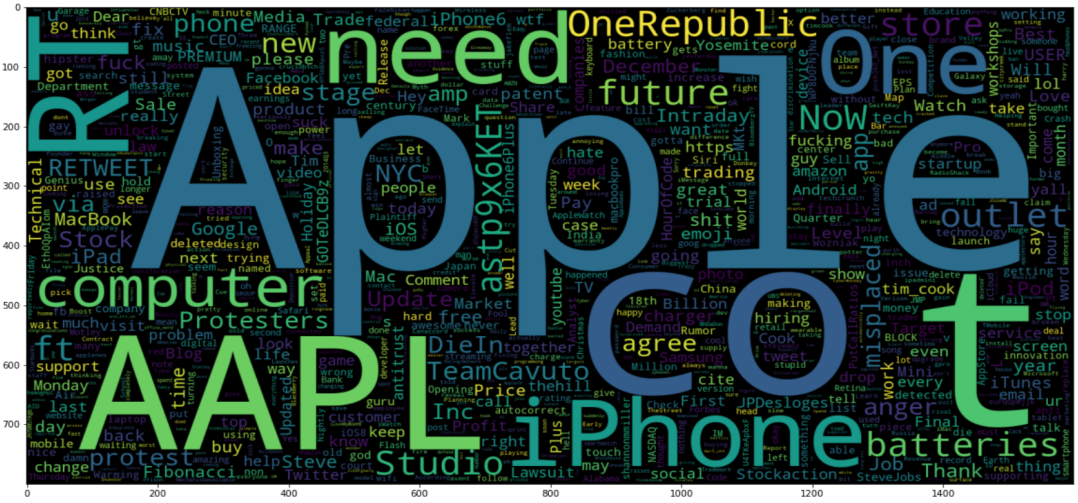

Image 6. Confusion Matrix

```
from sklearn import metrics
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test,pred)

cm_display = metrics.ConfusionMatrixDisplay(confusion_matrix = cm,
                                            display_labels = [False, True])

cm_display.plot()
plt.show()
```

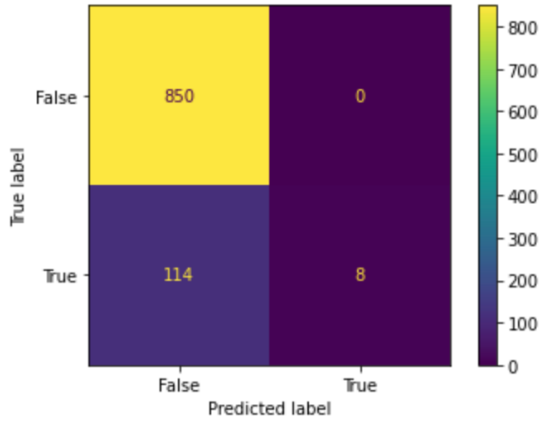

Image 6 builds off of Image 3 above. I had previously created a TF-IDF preprocessing technique to display words proportionally to the number of times they appeared in the dataset, and have now decided to use a confusion matrix model that is created to fit my training and testing data set. However, before I was able to create this model, I had to use a logistic regression model through the splitting of the testing and training data with the sklearn train/test/split function. Then a logistic regression model is created. This model becomes a way to make predictions on the testing data. I calculated the accuracy of the model and received 88.27%. Since my data does not need to be perfect, this model is a good indication that this model is useful to enhance my overall decision-making of the company Apple.

Data Analysis/Sentiment Analysis

Now that I have given you a general sense of the direction in which I am headed, my main goal in this analysis is to focus on the areas where Apple can improve the experience of the

customers, and in turn, what product needs to be introduced or added, to keep sales flourishing and the customers happy. As I explained previously, sentiment analysis uses algorithms to analyze the sentiment of what is expressed in the data. I decided to build off of the in-class examples through Twitter and extract positive or negative sentimental data to help me determine what product I wanted to introduce. However, I want to first walk you through the steps of my analysis.

First I collected the data. Then I cleaned the data and removed any irrelevant information that was going to disrupt my sentiment analysis. Next, I set my values to binary values between 0 and 1, and created a positive sentiment value prediction of any sentiment value greater than three to appear as a 1 on my analysis, and a negative sentiment value to show as 0 for any value less than or equal to 3. It was an easy way to convert a numeric sentiment value into a binary classification that would aid in my project. From there I continued my analysis by calling the function clean review and removing all stop words that will create confusion in my data if I were not to remove them. This text preprocessing helps prepare me for the ultimate goal, which is sentiment analysis on Apple tweets.

Sentiment analysis on Apple tweets involved the Word Cloud visualization techniques to display words that correlate to the positive sentiment and words that correlate to the negative sentiment parameters that I have created (Images 4 and 5). My key focus was on the larger words that involved technology because that is the area in which I wanted to expand my product line for Apple. Iphone, computer, Ipad, and studio seemed to pop out the most for me in the positive sentiment analysis. In the negative sentiment analysis visualization Iphone, computer, studio, and iPad reappeared, but what stood out to me was the word batteries. This is a clear indication that

in my Word Cloud visualization, and sentiment analysis, the word batteries is often mentioned a ton. This might be because customers are unhappy with the product's battery life, they are poorly made, or they are super difficult to replace. So what should I do to address this?

Conclusions/Recommendations

Can I improve the battery life, implement a faster charging system, make a device that is compatible with other systems that allow for the use of only one battery and charger over several, make a more environmentally friendly battery device, or create a device that allows for customers to share battery? All of these questions had been going through my mind as I spent numerous hours sifting through my data and trying to figure out what would best help my company, Apple. In the end, I decided that to create a better customer experience and improve the positive sentiment toward Apple products, creating a way to share battery life will be my course of action. Two areas that I will work on are a built-in sharing function and an app-based sharing function. For the built-in sharing function, I will be working with Apple to add it to all the new devices and create an update to add it to old devices. Essentially, all customers will be able to go into the settings of their phone, and through a built-in feature, with the tap of a button, any amount of battery that they set it to will be able to be shared with anyone through Bluetooth technology. No more relying on service or wifi, it will be a more reliable and sustainable way for customers to not have to worry about running out of battery when out and about with friends or even in a pinch when they are surrounded by strangers. However, if this is not good enough, my second implementation is the app-based sharing device that would enable anyone, on any device to share service. Say you forget a charger at an airport, coffee shop, or gas station, with the click of a button, anyone that is logged in to the app on the web will receive a notification that you

need a battery. It does not even have to be someone you know, but it will provide an endless supply of battery life whenever you need it.

Ultimately the goal is to make the customer's happy, and improve Apple's services and products to be the most competitive company out there. With all of this in mind, improving Apple devices to allow for battery-sharing capabilities is the way that I know will improve customer loyalty, satisfaction, and happiness with the products. Not only will this create a better experience for the customers, but this improvement will be another way that Apple can differentiate itself from its competitors. Through sentiment analysis, I was able to come up with a solution that I believe will remove a ton of negative connotations in tweets from customers about our products, enhancing the overall outlook of Apple products in a remarkable way.

Works Cited

https://data.world/crowdflower/apple-twitter-sentiment/workspace/file?filename=Apple-Twitter-Sentiment-DFE.csv