

Sentiment analysis for starwars rogue one movie

Prepared by:

Fares Hasan 1141123390

Mundher Al-Shabi 1131120942

Contents

[CHAPTER 1 INTRODUCTION 2](#_Toc474752287)

[Overview 2](#_Toc474752288)

[Sentiment Analysis 4](#_Toc474752289)

[Project objectives 4](#_Toc474752290)

[CHAPTER 2 METHODOLOGY 5](#_Toc474752291)

[Overview 5](#_Toc474752292)

[Dataset and Data collection 6](#_Toc474752293)

[Dataset 6](#_Toc474752294)

[Data Collection 6](#_Toc474752295)

[Pre-processing & Classification 8](#_Toc474752296)

[Term Frequency –inverse document frequency Method 10](#_Toc474752297)

[Extracting Sentimental Terms 10](#_Toc474752298)

[Visualizations 12](#_Toc474752299)

[CHAPTER 3 IMPLEMENTATION 12](#_Toc474752300)

[Training SVM linear Classifier 12](#_Toc474752301)

[The Graphical User Interface 13](#_Toc474752302)

[Data Analysis 14](#_Toc474752303)

[Data Acquisition 14](#_Toc474752304)

[Data Wrangling 16](#_Toc474752305)

[Visualizations 17](#_Toc474752306)

[CHAPTER 4 CONCLUSION 22](#_Toc474752307)

[CHAPTER 5 REFERENCES 23](#_Toc474752308)

[Index 1 JSON tree 24](#_Toc474752309)

# INTRODUCTION

## Overview

Sentiment Analysis is the process of categorizing, classifying, and identifying the opinions expressed in text format. It's referred to as opinion mining, inferring the conclusion or disposition of a speaker. For human beings, it’s an easy task to read between the lines and understand the writer’s sentiments and feelings. The computers have no concept of natural languages and cannot know what the words can mean in certain contexts, so performing an analysis require the understanding and analysis of natural language. A word cannot be simply understood if it means joy, anger or neutrality in certain context. The problem has to be solved in another way that overcomes the limitations of the computers, and therefore this problem should be reduced to mathematics (Rush, 2016). Sentiment Analysis looks to take care of this issue by natural language to perceive keywords inside an archive of data and subsequently classify the emotional narrative of the text.

The need for enhanced natural language processing techniques increase day by day. As machines gets involved more into our life, the need for machines that can process natural language increased. The rapid development and increase in need reached the peak with the revolution of social media. Several business needs and applications emerged and as our information expands every minute, the desperate need for new ways to process languages and provide response increased. The amount of conversations takes place online became a source of information and valuable data for business owners. To read new trajectory or change in the business landscape, have fast and handy feedback about various products and services. Though all the needs may take different contexts, variations and levels of complexity. One constant factor is, the crucial need for powerful models and accurate algorithms to analyze natural languages (lexalytics, 2017).

As advertisers for example, having the capacity to catch the many-sided quality of passionate reactions helps us figure out whether our social and substance showcasing activities are driving the activities that we made arrangements for, while additionally giving us hard signs for adjusting our procedure, if our touch focuses are not resounding with our clients.

The advantages of reading people’s sentiments are many. It can give an organization the upper hand to perceive new trends in fashion, lifestyle, political and personal preferences. Marketers can benefit from this to derive insights and feedback about products and services they provide. Governments can estimate public opinions regarding certain issues, elections, and new policies. In a digital era the power of machine learning and natural language processing have merged with a catalyst of huge data publically available online social media, into a rich ground for exploration. The applications are unprecedented and can be applied to many fields (iProspect, 2016).

For instance, if you would like to know if people in the UK are with or oppose the BREXIT, if people in Malaysia are enjoying or disturbed by the rainy season, or if the American election does or doesn’t matter of people from across the globe. All the previous examples are hypothetical, however we have read such analytics in the news. Shaping the public perception is another topic that could be in the other end of this end, but we aren’t in the interest of discussing it here. Service providers can measure customer’s satisfaction based on the customer’s feedback online and in social media. Ups or downs in sentiment results hints for beginning of new moves to improve products, training and enhance customer care agents or a new trajectory for marketing strategy (What is Sentiment Analysis?, 2017).

Recently sentiment analysis had been applied in politics, marketing and even in stock market. According to the Brand-watch, a correlation between changes in stock market and public sentiments in social media. It has been affirmed that the administration of the president of the United States, Barak Obama exploited the power of sentiment analysis to measure public sentiments to policy announcements and campaigns ahead of his second term in office (Bannister, 2015). Expedia Canada have aired a television advert that have started receiving a negative feedback about its music in most of the social media conversations. Expedia have dealt of the incident in a smart way, rather than abandoning the campaign, they have recognized the negative opinions and aired a modified advert where the violin is being smashed. Big magazines like the New York Times conduct their own search on the feedback, comments and conversation in the social media (Pena, 2016).The project was built by summer intern Isaac Pena, its objective was to go through every said article published by the Times to return categorize, topics and sentiments. The output of the system can be integrated to the Times search to enhance customer’s experience.

## Sentiment Analysis

In the basics natural language process is used to classify documents – texts – into positive, negative or neutral. Software engineers and data scientists develop systems that feed data into the sentiment analysis algorithm and present the results in a visual mediums. Among the most basic technique used to classify data is, to spot the keywords. Words like, love, interesting, good, like, fun, joy are obviously positive words, whereas, bad, hate, ugly, sucks, worst are obvious negative keywords. A week sentiment model can easily detect the phrase “good movie” as positive sentiment, and the phrase “bad story” as negative sentiment. The problem comes when we have the sentence “good movie but bad story”, the model will consider it neutral. This poses the need for developing better algorithms than spotting the keywords.

Algorithms differ in the method of scoring the data to determine whether a certain phrase indicate positive, negative, or neutral sentiment. The short coming in the previous example can be solved by separating the sentence by the word “but”, which will result in two score that can be combined. The accuracy and precision of the sentiment analysis model relays on the variety of the training set and the machine learning model (Algorithmia, 2017).

## Project objectives

This project is outlined to go beyond the construction of a sentiment analysis model. The objectives of this project is as below:

1. Build a sentiment Analysis model and comparing the performances of two methods.
2. Crawl twitter to collect tweets and store them into a database.
3. Clean and wrangle the data.
4. Gauge the sentiments in from the collected data.
5. Derive exploratory analytic and visualize findings.
6. Build simple GUI for sentiment analysis

# METHODOLOGY

## Overview

In this section the framework of the system will be explained and all the details elaborated. There will be two phases for this section. First phase will go on the sentiment analysis phase, constructing and training the model. The second phase will explain the data analytics part and system flow. In general the system is built to demo the power of the said model by applying it in a real life data collected from twitter. The diagram below shows the flow of the system.

As shown in the diagram. The output of the system are visualizations that represent the key findings from the analysis on the social media data collected. Embedding the all the components together showing a real life demo of what can be derived from conversations that are taking place in not so long time in the past.

## Dataset and Data collection

The term “dataset” in this context refers to the dataset that will be used in the training of the sentiment analysis model. Whereas the data collection indicates the data we intend to collect from social media for the purpose of performing data analysis.

### Dataset

The training dataset for the sentiment analysis model is from kaggle.com platform and called “*IMDB 5000 Movie Dataset”*. It has the reviews of 5000 movies with 50000 reviews portioned between 25 thousand positive and 25 thousand negative reviews. The data has been collected using the Python library called SCRAPY. The dataset is in TSV file (tab separated values). In the entire collection, no more than 30 reviews are allowed for any given movie because reviews for the same movie tend to have correlated ratings. Further, the train and test sets contain a disjoint set of movies, so no significant performance is obtained by memorizing movie-unique terms and their associated with observed labels (chuansun, 2016).

### Data Collection

The data collection phase was engineered to use the official API provided by the social media platform – in this case twitter – to collect data related to our topic. There exist two types of APIs from Twitter. The REST APIs and the Streaming APIs. The streaming APIs privileges developers with access to a global Strems of Tweet data. Twitter Streaming APIs have three distinct streaming endpoints (Dev Twitter, 2017). Each endpoint serves a specific purpose as follows:

|  |  |
| --- | --- |
| Endpoint | Use case |
| Public Streams | Streams of the public data flowing through Twitter. Suitable for following specific users or topics, and data mining. |
| User Strems | Single-user streams, containing roughly all of the data corresponding with a single user’s view of Twitter. |
| Site Strems | The multi-user version of user streams. Site streams are intended for servers which must connect to Twitter on behalf of many users. |

In order to access Twitter Streaming API, one need to have 4 pieces of information that will be used in the oath. API key, API secret, Access token, and Access token secret. All can be generated after signing up and create Twitter application. The four credentials are unique and confidential information for the individual that has to be protected against abuse.

The Twitter Streaming APIs will retrieve data in JSON format. The preliminary step we take right on receiving the data chunks is to parse them to extract the relevant details. The JSON tree will have all the information about the tweet. Index 1 shows a full tree of one tweet. The diagram below describe how each JSON tree is parsed to extract the details we are interested in.

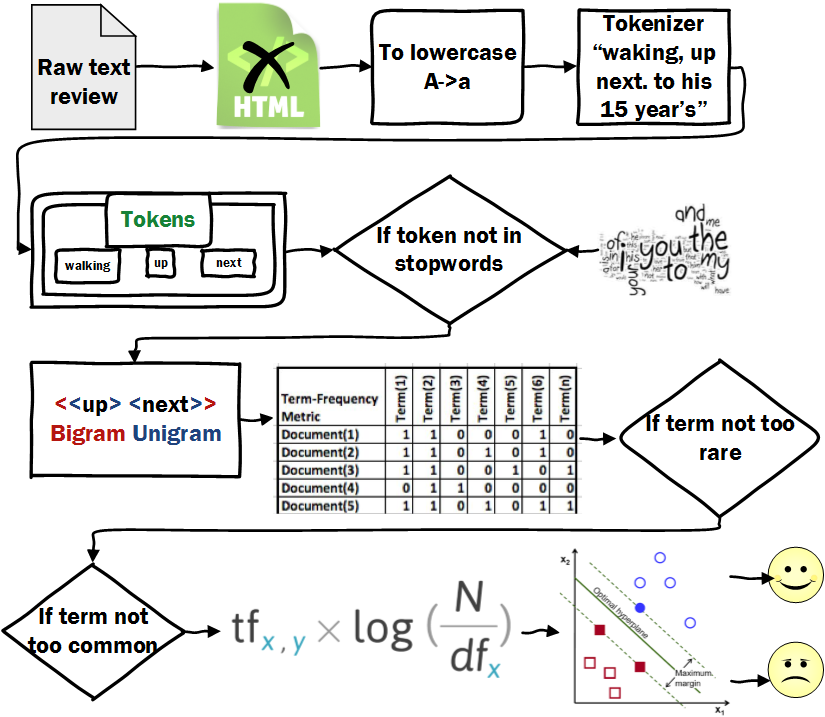
The saved tweet will have only the location, language, time and data, and the tweet body. Here we have used the SQLite as our database management system. Due to its ease of use, smoothness and lightness. The SQLite database will generate the identification numbers as primary key for each record. Later on the analysis can be initiated by reading on the data stored in the database.

## Pre-processing & Classification

The dataset has 50000 reviews about 5000 movies from global producers. Variations in length aren’t take into account perhaps due to the fact they may have no relevance. In the pre-processing phase we tried to prepare the reviews and process them before we extract the features and feed them into the model in order to perform training as shown in the figure below. Several steps are involved in this phase as described in an ordered manner below:

1. Remove HTML tags. Since some of the reviews contain HTML tags or referrals to external website, removing such noise is important to maintain the quality of the data.
2. Convert all words into lower case. The sensitivity of the words case could be interpreted as feature, which shouldn’t be the case, thus converting to lower case is performed to avoid that.
3. Extract words that doesn’t contain any numbers or symbols within them and that have more than one character. Up holding a standard natural language where numbers and symbols aren’t part of the English language vocabularies.
4. Removing stop words. Stop words like “a”, “and”, “or” and etc. as a matter of fact these could be noise if to be taken as features for the data.
5. Grouping reviews in two, a unigram set and bigram set. This step is for the purpose of testing the performance of the model in each method.
6. Consolidate unigram data and another set of unigram and bigram consolidated.
7. Removing words that appeared rarely, in one document.
8. Removing the words that exists in 80% or more of the reviews. The redundancy of the vocabularies is not a useful feature to be included.
9. Feed the extracted features to a linear Support Vector Machine Classifier.

The construction of features after the pre-processing phase is conducted using the two methods. The term frequency and inverse term frequency. In the next section the two methods will be introduced and explained in further details.



## Term Frequency –inverse document frequency Method

The Term Frequency –inverse document frequency (tf-idf) is a statistical method that gauges the significance of a word of document in the overall corpus collection. It is often used as a weighting factor in the topics of opinion mining and information retrieval. This method increases significantly the tf-idf value to the appearance frequency of a word in a document, but is offset by the frequency of a word in a corpus. Tf-idf is among the most popular term-weighting schemes and used in the majority of the text-based recommendation systems. In this project tf-idf has been used.

Assume we have an arrangement of English content archives and wish to figure out which report is most applicable to the question "the sweet milk". A basic approach to begin is by disposing of reports that don't contain each of the three words "the", "sweet", and "milk", however this still leaves many records. To additionally recognize them, we may check the quantity of times each term happens in each archive and sum them all together; the quantity of times a term happens in a record is called its term frequency. Since the expression "the" is so normal, term frequency will tend to inaccurately stress records which happen to utilize "the" more often, without sufficiently giving weight to the more significant terms "sweet" and "milk". The expression "the" is not a decent catchphrase to recognize important and non-significant archives and terms, not at all like the less regular words "cocoa" and "cow". Henceforth a reverse document frequency variable is joined which lessens the heaviness of terms that happen regularly in the record set and expands the heaviness of terms that happen once in a while.

## Extracting Sentimental Terms

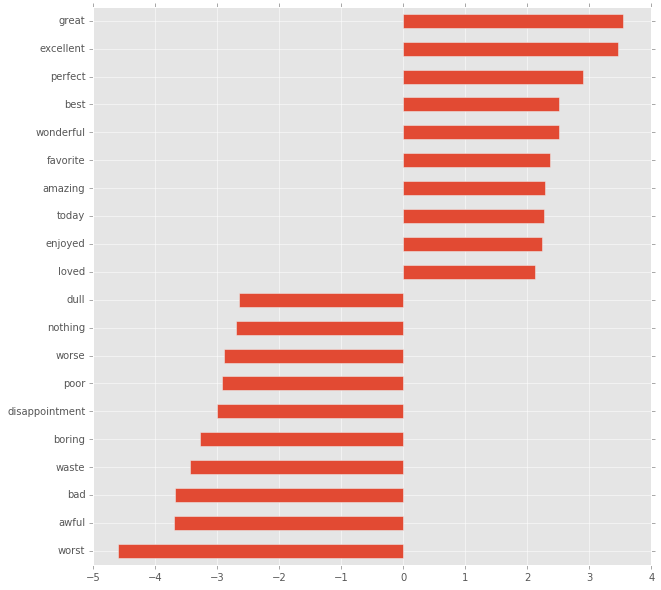
There is a big question of how the SVM classifies the reviews?

Is SVM has detected the positive and negative words?

To answer these questions we need first to at the SVM decision function which defined as follow

So the equation similar to the linear classifier where each input x has weight w. The function returns 1 if the total sum of weighted input is positive (1+) otherwise 0. In our case, we have two type of output 1 for positive sentiments and 0 for negative sentiments. The SVM training algorithm will try to assign high positive weights to the positive features and negative weights to the negative features. The not important words will have values close to zero.

After the training phase, we extracted the weights of the SVM features and analyzed. The figure below shows the most positive and negative that detected by the model automatically.



## Visualizations

The visualizations generated using two distinct libraries. The first in the JavaScript visualization library D3. D3 stand for Data Driven Documents and produces dynamic, interactive data visualizations in web-bowsers. The other tools are the python libraries ggplot and matplotlib. Visualizations are reported in images in PNG format.

# IMPLEMENTATION

## Training SVM linear Classifier

The linear SVM is trained in two feature sets that represent the textual data in two different representational form. The unigram and bigram features will be fed to the classifier and based on the performance we will utilize the classifier to handle the rest of the sentiment classifications in the project. The output of the training program will yield the information below:

* Accuracy of the model.
* Most positive sentiment words.
* Most negative sentiment words.
* Most un-useful words.

When running the model results are generated as stipulated at the table below.

|  |  |  |
| --- | --- | --- |
| Details | Feature set | |
| **Unigram** | **Bigram** |
| Accuracy | 89.6% | 90.72% |
| Most Positive | Wonderfully, enjoyed, best | Loved, amazing, favorite, best |
| Most Negative | Worst, waste, awful, poorly | Boring, poor, awful, waste |
| Most Un-useful | Jarhead, athon, inextricably | Terrorize community, leaves bit, |

As observed from the table above, there is similarity to some extent in the sentiment words detected. This similarity correlate to the accuracy result that is slightly close. Based on this performance, the unigram will be considered to perform the classification in the coming phases of the project.

## The Graphical User Interface

After successfully trained our model, we built lightweight API to handle the classification. This interface has been used by the graphical user interface (GUI) and to analyze the tweeter data. The interface takes a review as input and returns the sentiment. The caller function should not be worried about the complexity of the implementation.

The GUI showed below designed using C#. The C# part only gets input from the user and pass it to the python sediment interface.

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Data Analysis

### Data Acquisition

The acquisition of data is performed via the Streaming APIs provided by Twitter. Related details on the type of end points are as detailed in the methodology chapter. The pieces of information we are seeking from the data chunks are as below:

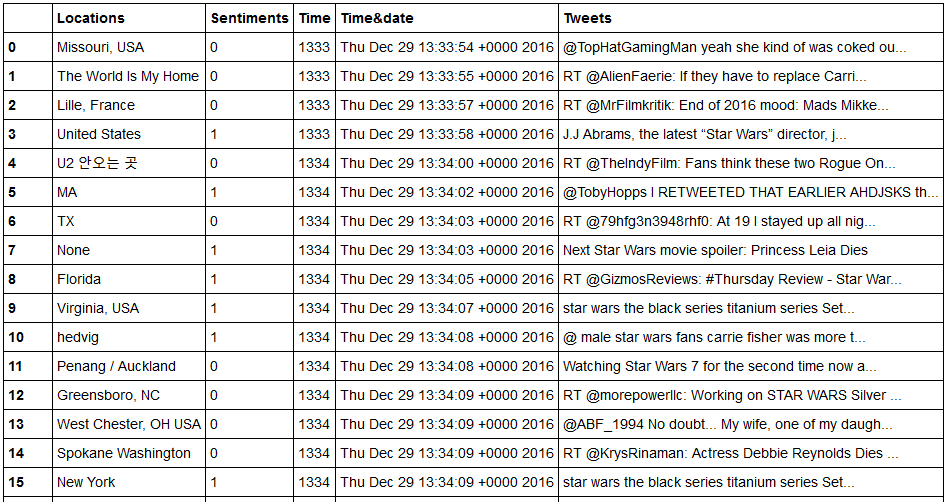
* Tweets text.
* User name.
* Language.
* Location.
* Data & time.

The parsed data will only have the five pieces mentioned above in addition to the primary key identification number that will be generated in the database management system all saved into SQLite database. The figure below shows how data looks like in the database.



Languages are abbreviated into two litters like “en” for English, “ja” for Japanese. The locations are set by the users, so there are some locations with none values due to the user setting. Each record in the database is represent one tweet.

### Data Wrangling

Data wrangling is the phase where we model and clean the data. Due to the fact that we have run the crawler for a number of hours that doesn’t exceed one day, the time and date will have a very limited variations. Thus, we have combined the hour and minute values and then we referred to them as the time. Though it is an unconventional method, yet it serves the purpose later on when we come to visualize the data in a chronological order. We have collected 32668 tweets in 5 SQLite databases. The reason of having 5 SQLite databases is due to the time out error and exceeding the number of requests allowed by twitter. This had caused us to regenerate a new database file to store new coming tweets each time we re-run the crawler. The tweets are all consolidated into one pandas data frame added to it the sentiment classification resulted from classifying the tweets text using the model. The figure below present the data-frame contains all the data.

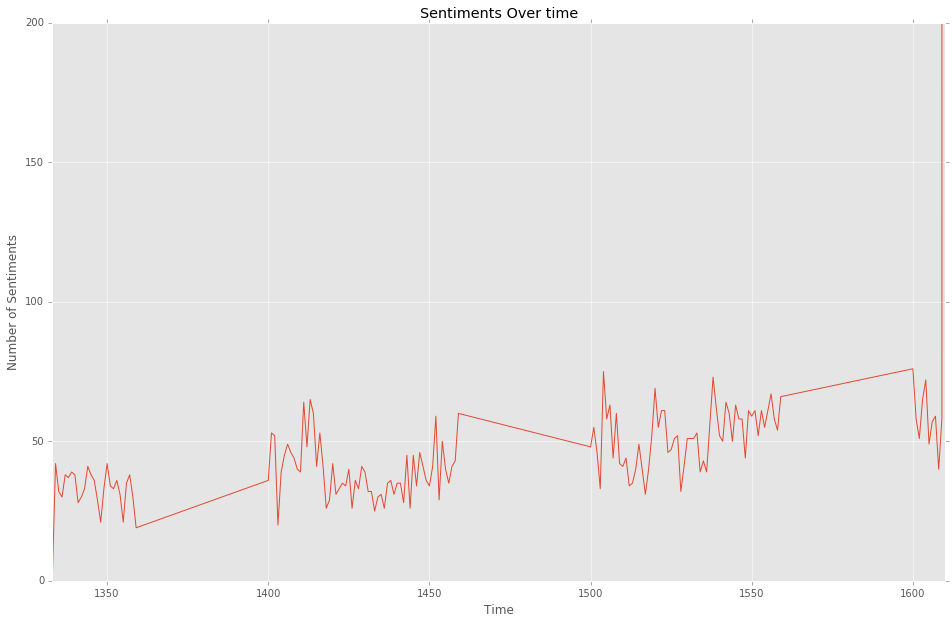
The data have about 5123 unique locations (with redundancy) and about 158 unique time values which reflects 158 minutes that is the running time for the crawler. The table below demonstrate the grouping of the tweets by the sentiment class in correspondent to the other variables.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sentiment class | Locations | Time | Time & date | Tweets |
| 0 | 4015 | 5797 | 5797 | 5797 |
| 1 | 24949 | 26871 | 26871 | 26871 |

The number of tweets per sentiment class is similar in time, time & data, and tweets because it is reflect upon the same thing. The user name was ignored in the data frame because it is an irrelevant factor to be considered in this analysis.

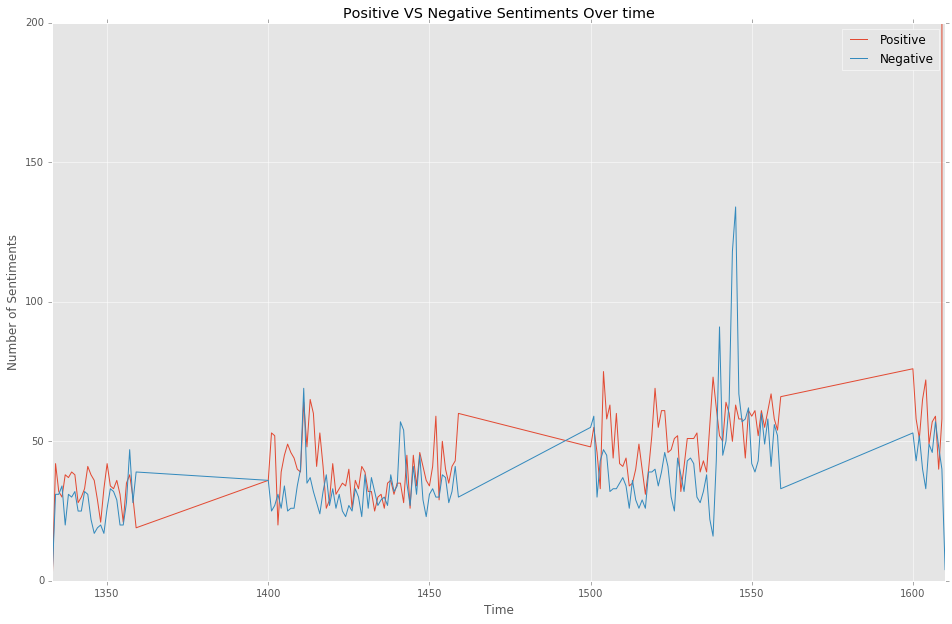
### Visualizations

#### Tweets Over time

The graph below shows the overall numbers of tweets over time. Ups and downs correspond to increase and decrease of the conversations taking place in that specific time.

It is noticeable that far many tweets were generated in the last minute compared to the overall time frame. This could be a co-incidence or it has some unforeseen indication.

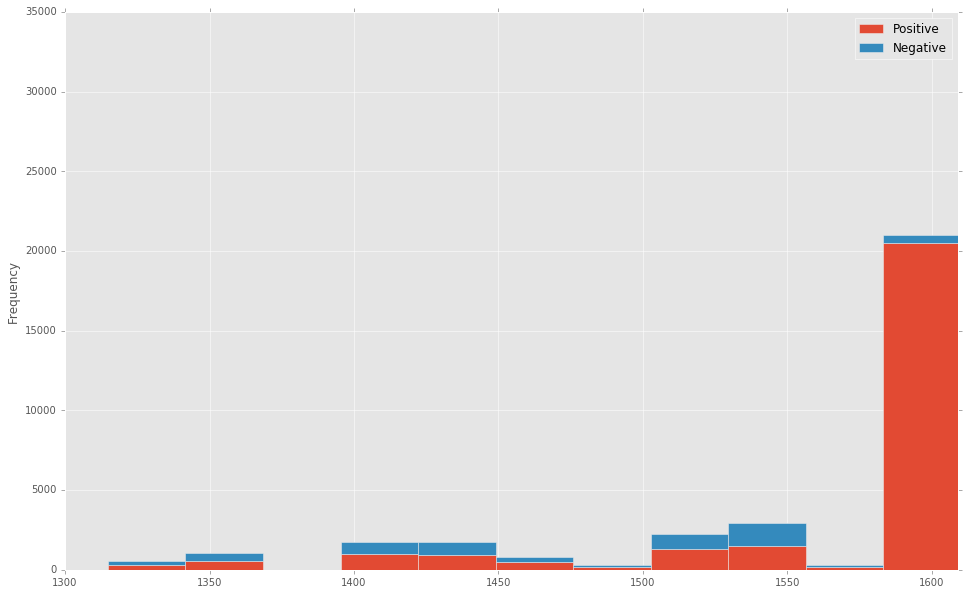
#### Positive vs Negative tweets

Looking at the graph below, we can see the variation in the negative vs positive tweet over the time line.

We can observe the peak of negative tweets near time 15:50pm followed with slight increase for the positive tweet till time 16:00pm. Yet a huge number of positive tweets peak in the last minute.

#### Sentiment Distribution

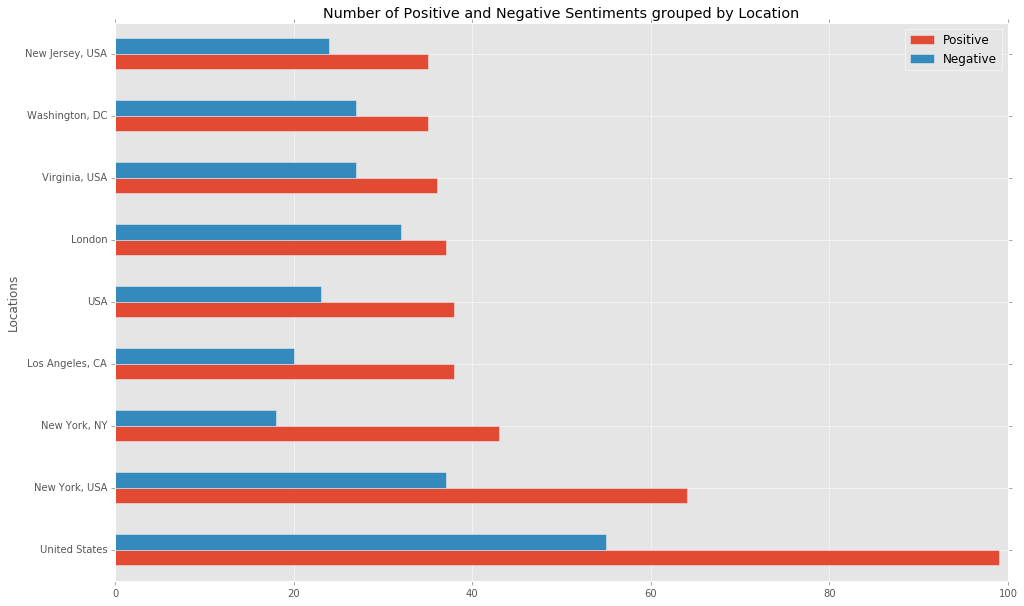
Looking into the histogram below shows the distribution of sentiments by the time. This representation is pretty much similar to the previous graph, but here with the histogram bins more details can be emphasized visually. From observing the time line of the histogram we can easily notice the disappearance of one of the bins between time 13:50 to time 14:00. This might not be visible in the previous graph. The justification of this absent bin is that the crawler wasn’t collecting tweets in that time. Again in the last minute we can size the relatively high bin of the histogram indicating the peak in the number of tweets in that time.



#### Sentiment per location

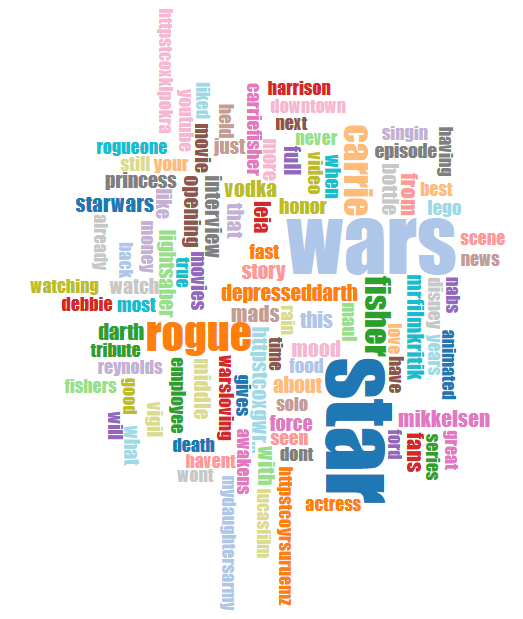
The locations are an interesting way to demonstrate opinions in a consideration of geographical distributions. It is crucial to have clean irredundant data in order to generate meaningful analysis with geographical distributions. Here we present our findings of the locations without standardization. Though it might not give a clear distinction between what is unique location and what is redundant, but we brought it here to show the purpose of using such details what is the significant value it can add. For example, New York, NYC City, and NYC USA are all the same place. Yet, it will be considered as three different places. The standardization is very useful when it comes to location sensitive or targeted analysis because without it, inaccurate, misleading findings could be found. The table below shows the highest 10 location in terms of the number of tweets. The histogram bins below the table shows the number of sentiment per location class grouped and presented. Clearly seen, the United States has the highest number of tweets, though other locations like USA and New Orleans, LA refer to the same country too.

|  |  |  |
| --- | --- | --- |
|  | **Sentiments** | **Neg-Sent** |
| **Locations** |  |  |
| **New Orleans, LA** | 19885 | 4 |
| **United States** | 99 | 55 |
| **New York, USA** | 64 | 37 |
| **New York, NY** | 43 | 18 |
| **Los Angeles, CA** | 38 | 20 |
| **USA** | 38 | 23 |
| **London** | 37 | 32 |
| **Virginia, USA** | 36 | 27 |
| **Washington, DC** | 35 | 27 |
| **New Jersey, USA** | 35 | 24 |
| **United Kingdom** | 28 | 41 |
| **UK** | 27 | 23 |
| **New York** | 27 | 20 |
| **California, USA** | 26 | 17 |
| **Beverly Hills** | 26 | 1 |
| **England, United Kingdom** | 26 | 18 |
| **London, England** | 25 | 18 |
| **Nebraska, USA** | 22 | 9 |
| **Canada** | 21 | 10 |
| **Chicago, IL** | 17 | 16 |



#### Words Map

The words map is a technique where we look for the highest frequency words in the data we have. It is a very useful technique to understand the topics and vocabularies used in the conversations. It gives a hit of what could be the topics or expressions talked in the social media. The size of the words correspond to the frequency of that word. The bigger the size of the word the higher frequency they have. The word map below is from our twitter data. Visually noticeable the words wars, rogue, and star. But many more can give us more hints like great series, fast, story, honor, vodka, and darth. These simple word indicate simple hints but not a full analysis, they are beautiful and neat in their presentation. With cleaner, standardized data this tool would be very useful and significant in provide a peak view of what the data hints.



# CONCLUSION

In the course of this project we have built a sentiment analysis model using the linear Support Vector Machine classifier. Experimenting with the unigram and bigram have lead us to adopt the bigram features due to its performance with accuracy 90.72%. The model has been utilized to classify the sentiments of 32668 of tweets about the movie Star Wars Rogue One. The classification of the tweets have then been analyzed with the use of location and time frame to generate visualizations and exploratory analytics.

This work have integrated two distinct domains of science, the machine learning and data science with the use of real data. It can be seen clearly how a powerful Natural Language processing technique can enhance the performance of the classifier. The analysis of the data about the latest movie is a real example of what can be done for many applications across the business, politics and social sectors of our societies.

Further development would to standardize the locations so the insights be more useful. Cleaner data is always a crucial part of any analysis project and in this work we have tried to clean and prepare the data to generate insights of interest. However, the findings we have presented in the visualizations give clear cues of what was the people talk about in social media in regard to the star wars rogue one movie. In the time of releasing the movie everyone was talking about what could the generation’s long movie has for us. The difference of show times from a country to another could be a slight drawback.

The project has been published online at Github: <https://github.com/mundher/sentimentNLP>

# REFERENCES

Algorithmia. (2017). *Determine positive or negative sentiment from text*. Retrieved from Algorithmia: https://algorithmia.com/algorithms/nlp/SentimentAnalysis

Bannister, K. (2015, January 26). *Understanding Sentiment Analysis: What It Is & Why It’s Used*. Retrieved from Brandwatch Blog: https://www.brandwatch.com/blog/understanding-sentiment-analysis/

chuansun. (2016, Septmber 10). *IMDB 5000 Movie Dataset*. Retrieved from kaggle: https://www.kaggle.com/deepmatrix/imdb-5000-movie-dataset

Dev Twitter. (2017). *Streaming APIs*. Retrieved from Twitter Developer Documentation: https://dev.twitter.com/streaming/overview

iProspect. (2016, November 6). *10 Sentiment Analysis Tools to Track Social Marketing Success*. Retrieved from social media: https://www.iprospect.com/en/ca/blog/10-sentiment-analysis-tools-track-social-marketing-success/

lexalytics. (2017). *What is Sentiment Analysis?* Retrieved from Sentiment Analysis: https://www.lexalytics.com/technology/sentiment

Pena, I. (2016, August 10). *Searching for Feelings — An Intern Works on Topic and Sentiment Analysis*. Retrieved from The New York Times: https://open.blogs.nytimes.com/2016/08/10/searching-for-feelings-an-intern-works-on-topic-and-sentiment-analysis/?\_r=0

Rush, i. (2016). *5 Algorithms Every Web Developer Can Use and Understand.* gitbooks. Retrieved from algorithms for webdves: https://lizrush.gitbooks.io/algorithms-for-webdevs-ebook/content/chapters/sentiment-analysis.html

*What is Sentiment Analysis?* (2017). Retrieved from Darabridge: http://www.clarabridge.com/sentiment-analysis/

# Index 1 JSON tree

{

"created\_at": "Sat Jan 23 14:52:22 +0000 2016",

"id": 690909948289486800,

"id\_str": "690909948289486848",

"text": "RT @Wennietrue: I love you being you\nBeautiful \n#amazing",

"source": "<a href=\"http://twitter.com\" rel=\"nofollow\">Twitter Web Client</a>",

"truncated": false,

"in\_reply\_to\_status\_id": null,

"in\_reply\_to\_status\_id\_str": null,

"in\_reply\_to\_user\_id": null,

"in\_reply\_to\_user\_id\_str": null,

"in\_reply\_to\_screen\_name": null,

"user": {

"id": 1148215213,

"id\_str": "1148215213",

"name": "Blue Lotus",

"screen\_name": "James\_P\_Lotus",

"location": "British Columbia Canada",

"url": null,

"description": "Poet and Digital Graphic Artist\r\nin Persuit of the Inheritance of Divine Reality. Sharing Words of Truth Inspired by The Source of Faith.",

"protected": false,

"verified": false,

"followers\_count": 2166,

"friends\_count": 2150,

"listed\_count": 41,

"favourites\_count": 3672,

"statuses\_count": 6824,

"created\_at": "Mon Feb 04 14:50:05 +0000 2013",

"utc\_offset": null,

"time\_zone": null,

"geo\_enabled": false,

"lang": "en",

"contributors\_enabled": false,

"is\_translator": false,

"profile\_background\_color": "C0DEED",

"profile\_background\_image\_url": "http://pbs.twimg.com/profile\_background\_images/783878076/17910423974652a079478b85cb4afbec.jpeg",

"profile\_background\_image\_url\_https": "https://pbs.twimg.com/profile\_background\_images/783878076/17910423974652a079478b85cb4afbec.jpeg",

"profile\_background\_tile": false,

"profile\_link\_color": "0084B4",

"profile\_sidebar\_border\_color": "000000",

"profile\_sidebar\_fill\_color": "DDEEF6",

"profile\_text\_color": "333333",

"profile\_use\_background\_image": true,

"profile\_image\_url": "http://pbs.twimg.com/profile\_images/3207205004/8df4769438fb2da7edab62aca24d6fa6\_normal.jpeg",

"profile\_image\_url\_https": "https://pbs.twimg.com/profile\_images/3207205004/8df4769438fb2da7edab62aca24d6fa6\_normal.jpeg",

"profile\_banner\_url": "https://pbs.twimg.com/profile\_banners/1148215213/1401394344",

"default\_profile": false,

"default\_profile\_image": false,

"following": null,

"follow\_request\_sent": null,

"notifications": null

},

"geo": null,

"coordinates": null,

"place": null,

"contributors": null,

"retweeted\_status": {

"created\_at": "Sat Jan 23 14:49:23 +0000 2016",

"id": 690909197433573400,

"id\_str": "690909197433573376",

"text": "I love you being you\nBeautiful \n#amazing",

"source": "<a href=\"https://mobile.twitter.com\" rel=\"nofollow\">Mobile Web (M5)</a>",

"truncated": false,

"in\_reply\_to\_status\_id": null,

"in\_reply\_to\_status\_id\_str": null,

"in\_reply\_to\_user\_id": null,

"in\_reply\_to\_user\_id\_str": null,

"in\_reply\_to\_screen\_name": null,

"user": {

"id": 1336970263,

"id\_str": "1336970263",

"name": "Wendy Joy",

"screen\_name": "Wennietrue",

"location": "BrisVegas",

"url": "http://favstar.fm/confirm\_email/eWAk2d6BZpfusP2QYTjS",

"description": "Happy-go-lucky, always feeling alive, excited to make new friends, number one saying love more it can't hurt, hug until you can hug no more, be you, be special",

"protected": false,

"verified": false,

"followers\_count": 3724,

"friends\_count": 4982,

"listed\_count": 121,

"favourites\_count": 34463,

"statuses\_count": 30674,

"created\_at": "Mon Apr 08 15:44:49 +0000 2013",

"utc\_offset": null,

"time\_zone": null,

"geo\_enabled": false,

"lang": "en",

"contributors\_enabled": false,

"is\_translator": false,

"profile\_background\_color": "C0DEED",

"profile\_background\_image\_url": "http://abs.twimg.com/images/themes/theme1/bg.png",

"profile\_background\_image\_url\_https": "https://abs.twimg.com/images/themes/theme1/bg.png",

"profile\_background\_tile": false,

"profile\_link\_color": "0084B4",

"profile\_sidebar\_border\_color": "C0DEED",

"profile\_sidebar\_fill\_color": "DDEEF6",

"profile\_text\_color": "333333",

"profile\_use\_background\_image": true,

"profile\_image\_url": "http://pbs.twimg.com/profile\_images/458461025902157824/YQAyk2jt\_normal.jpeg",

"profile\_image\_url\_https": "https://pbs.twimg.com/profile\_images/458461025902157824/YQAyk2jt\_normal.jpeg",

"default\_profile": true,

"default\_profile\_image": false,

"following": null,

"follow\_request\_sent": null,

"notifications": null

},

"geo": null,

"coordinates": null,

"place": null,

"contributors": null,

"is\_quote\_status": false,

"retweet\_count": 2,

"favorite\_count": 2,

"entities": {

"hashtags": [

{

"text": "amazing",

"indices": [

32,

40

]

}

],

"urls": [],

"user\_mentions": [],

"symbols": []

},

"favorited": false,

"retweeted": false,

"filter\_level": "low",

"lang": "en"

},

"is\_quote\_status": false,

"retweet\_count": 0,

"favorite\_count": 0,

"entities": {

"hashtags": [

{

"text": "amazing",

"indices": [

48,

56

]

}

],

"urls": [],

"user\_mentions": [

{

"screen\_name": "Wennietrue",

"name": "Wendy Joy",

"id": 1336970263,

"id\_str": "1336970263",

"indices": [

3,

14

]

}

],

"symbols": []

},

"favorited": false,

"retweeted": false,

"filter\_level": "low",

"lang": "en",

"timestamp\_ms": "1453560742815"

}