**Analyzing Reddit Behaviour**  
User Manual



|  |  |
| --- | --- |
| **Name:** | Rachel Solomon |
| **Student No:** | 14345381 |
| **Completed**: | 19/05/2018 |

Table of Contents

[**1. Beginning a Session**](#_9ddqcarqdnlp) **3**

[**2. Reddit Analysis Web Application**](#_h3fkd39vevsg) **4**

[2.1. Home](#_x874k8t3h70m) 4

[2.2. Viewing overall dataset results](#_72m81swm2ri) 4

[2.2.1. View Statistical Results](#_aryd8d6aeku7) 4

[2.2.2 View charts](#_c32r4w2a5fsj) 6

[2.2.3 View Word-Clouds](#_9ujy951jo9dp) 8

[2.2.4 View machine learning](#_jlji9bo2indt) 11

[2.2.5 Interactive machine learning](#_ar68e24pr0da) 11

[2.3 Subreddit User Query](#_8jez9ri2dui) 13

[2.3.1 Subreddit Results](#_s1mg33w2tzls) 13

[2.4 About](#_xe4dk5stftdp) 15

[2.5. Journey](#_xb0b9rj1r6ff) 15

[**3. Ending a Session**](#_g8ikvborsnbx) **15**

## 1. Beginning a Session

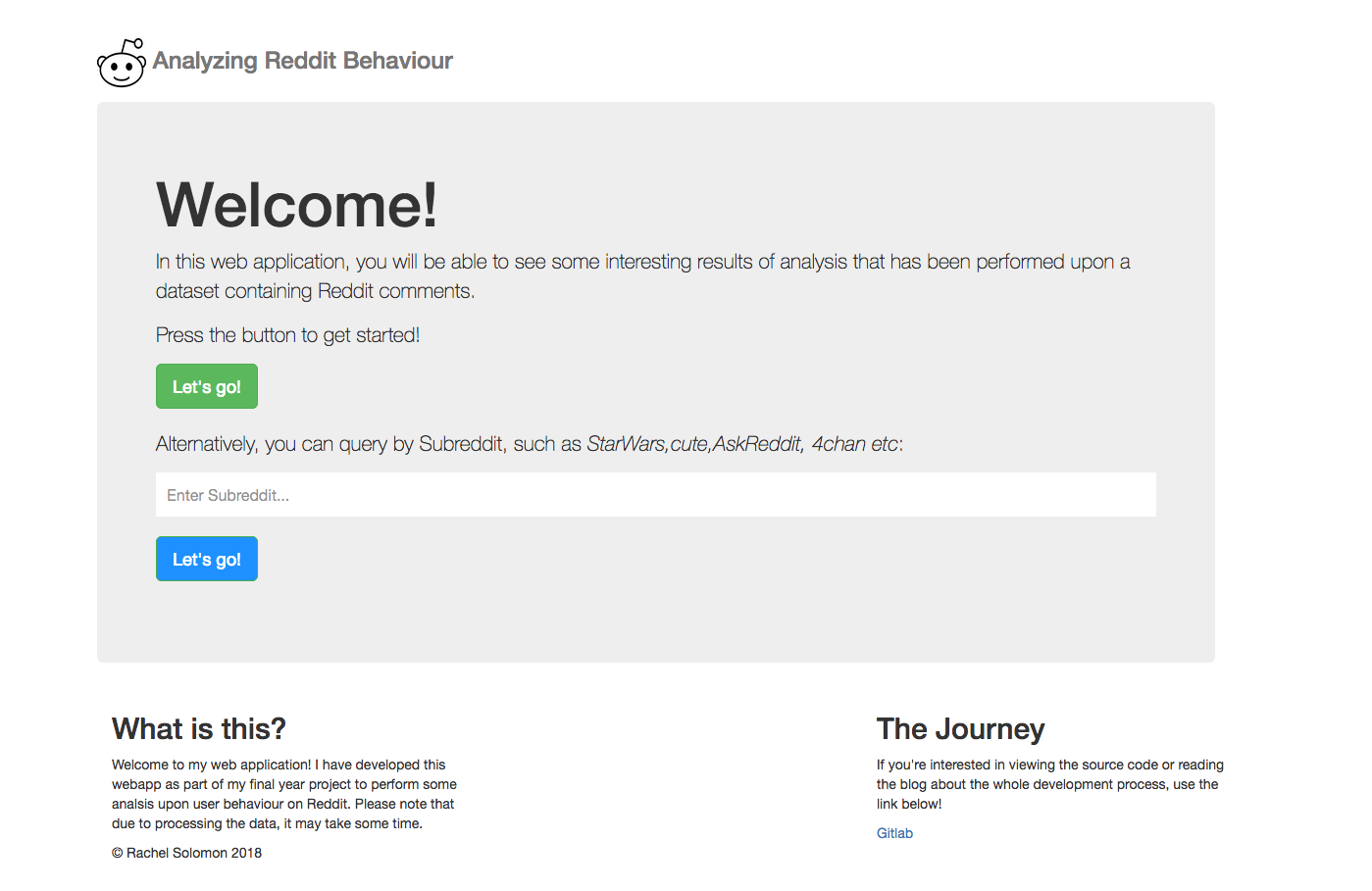
To begin using the Reddit Behaviour Analysis tool, go to your desktop and open your favourite web browser such as Google Chrome, Mozilla Firefox, Safari etc.

Once you have loaded your browser, click on the URL search bar and type in this URL:

<http://www.analyzingredditbehaviour.com/> .

  
*Fig 1.1 Entering the URL*  
  
Once that is clicked, let the web page load and you will be redirected to the homepage of the

web application.

*Fig 1.2. Analyzing Reddit Behaviour homepage*

## 2. Reddit Analysis Web Application

Once you have reached the homepage, your attention is immediately directed towards the two different paths that you can take from here. There will be more detail about these paths in the following subsections.

### *2.1. Home*

On the homepage, you can see that there is a brief overview of the purpose of the application and also the option to see the overall results of analysis upon the complete dataset or you, the user, have the option to query by Subreddit in order to extract some information regarding the newest posts within your chosen Subreddit.

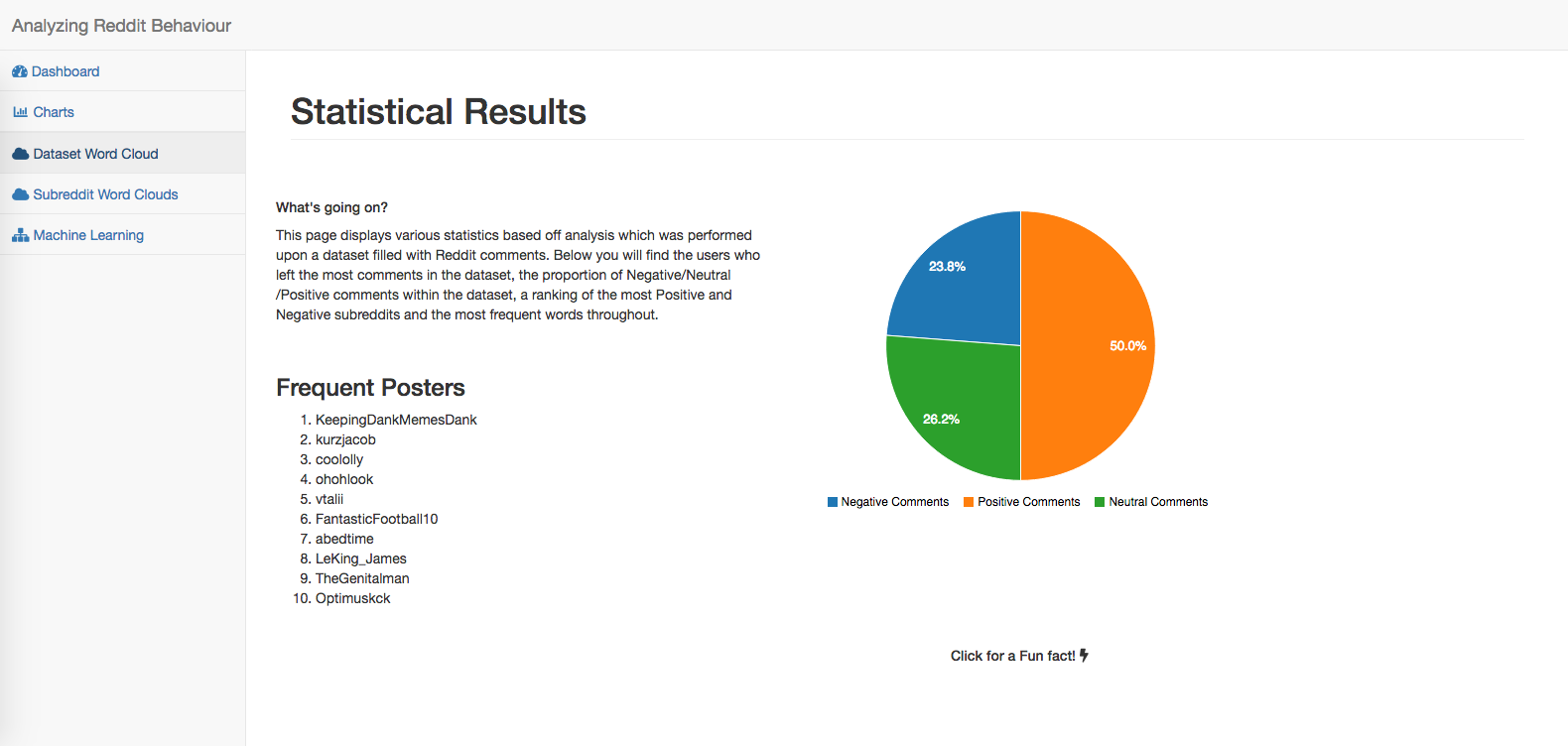
To return back to the homepage at any time during the process, you can click the logo and title in the top left corner.

### *2.2. Viewing overall dataset results*

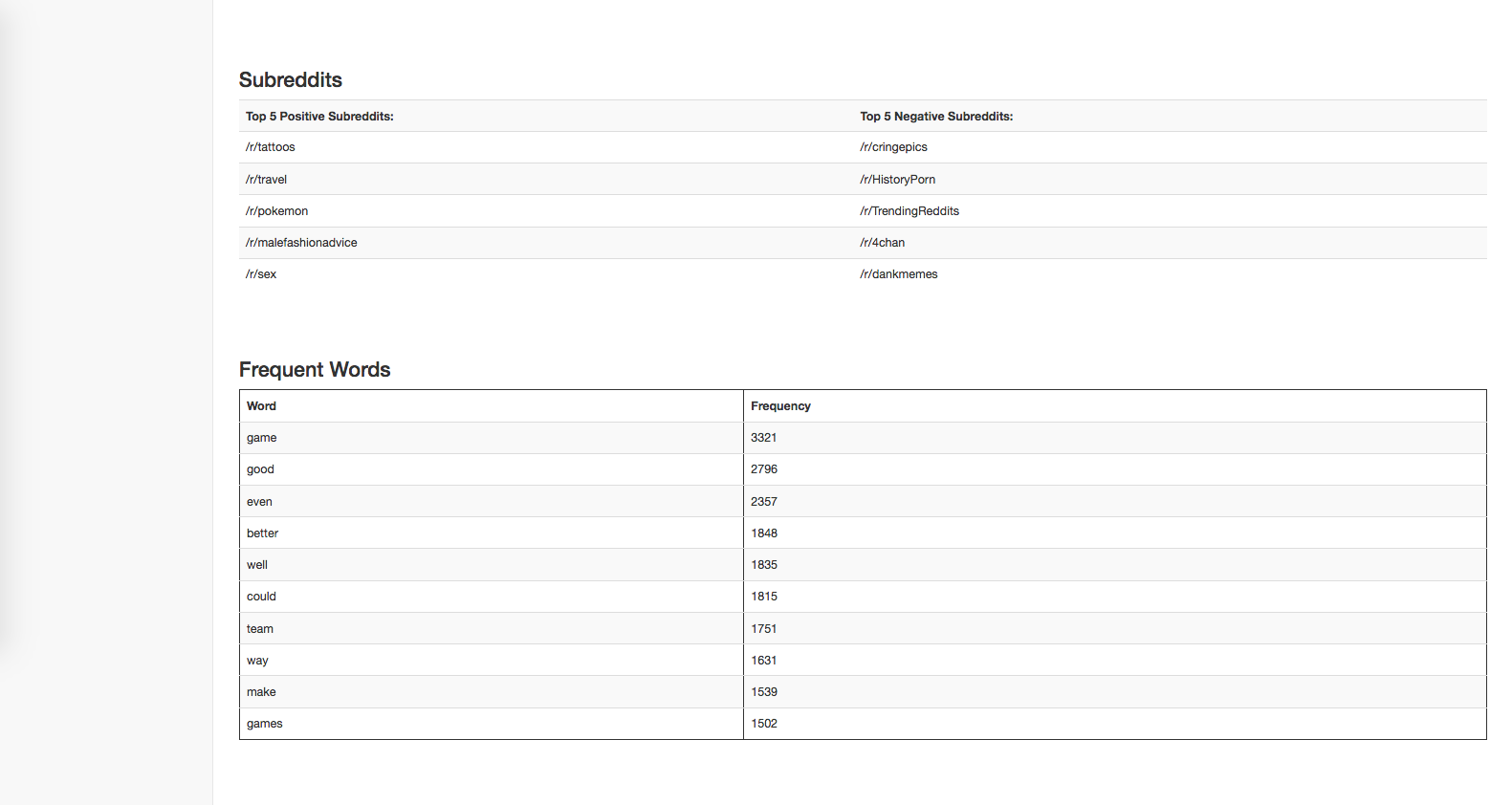
If you wish to view the results of the analysis of the entire dataset, you are introduced to a web page displaying various statistical analysis. You are able to navigate in between pages with the menu on the left-hand side of the window.

### *2.2.1. View Statistical Results*

Within the statistical results tab, you are initially presented with an overview as to what the page is, the distribution of positive, negative and neutral comments, the users who have commented the most within the dataset.  
There is also the option where you can click for a fun-fact which tells you about the self-declared sarcasm percentage within the dataset.

*Fig 2.2.1. Statistical Results dashboard, part 1/2*

You can then scroll down to view the top 5 most positive and negative subreddits within the dataset as well as the most frequent words and their wordcount.

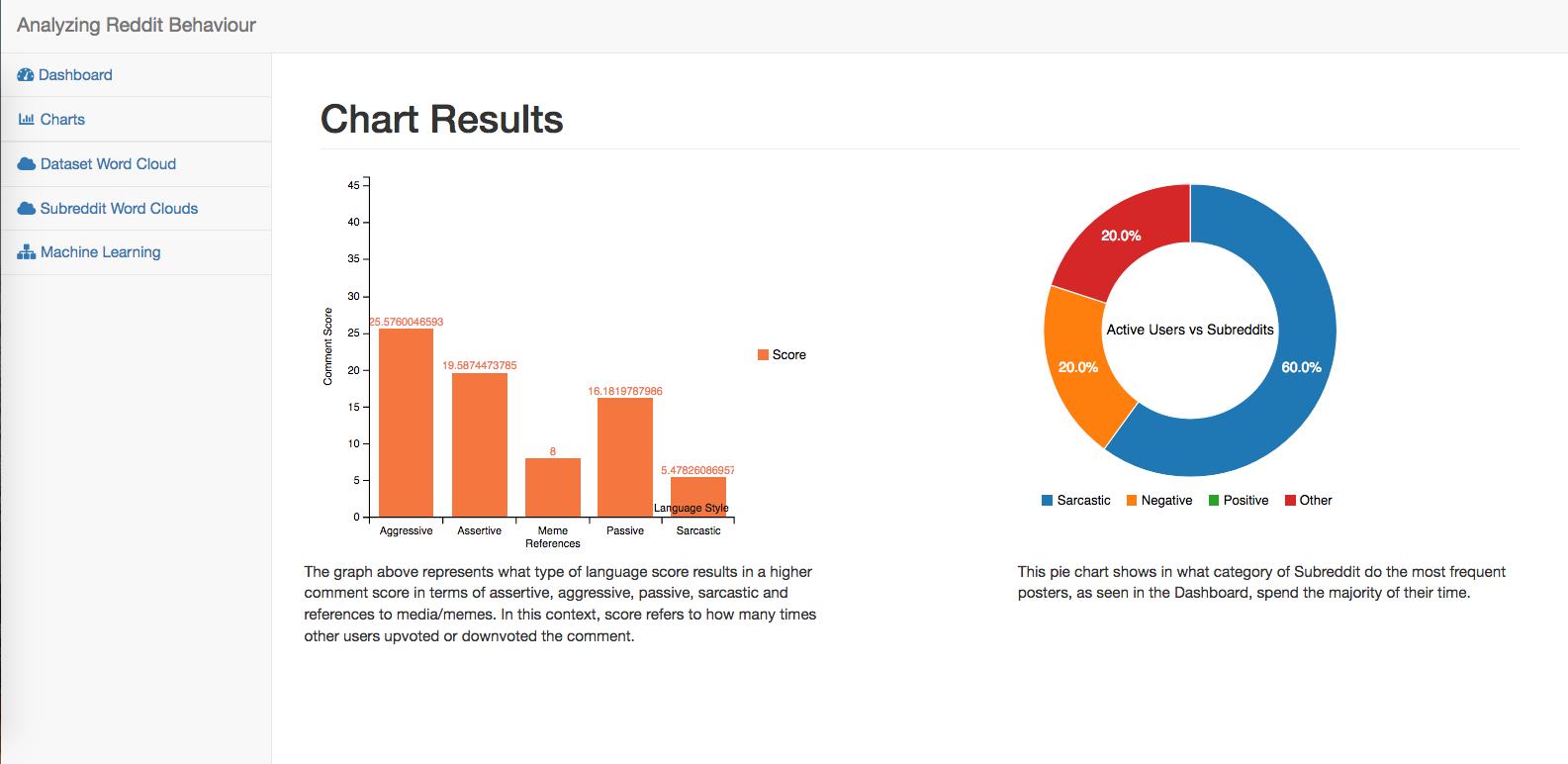
*Fig 2.2.1. Statistical Results dashboard, part 2/2*

**

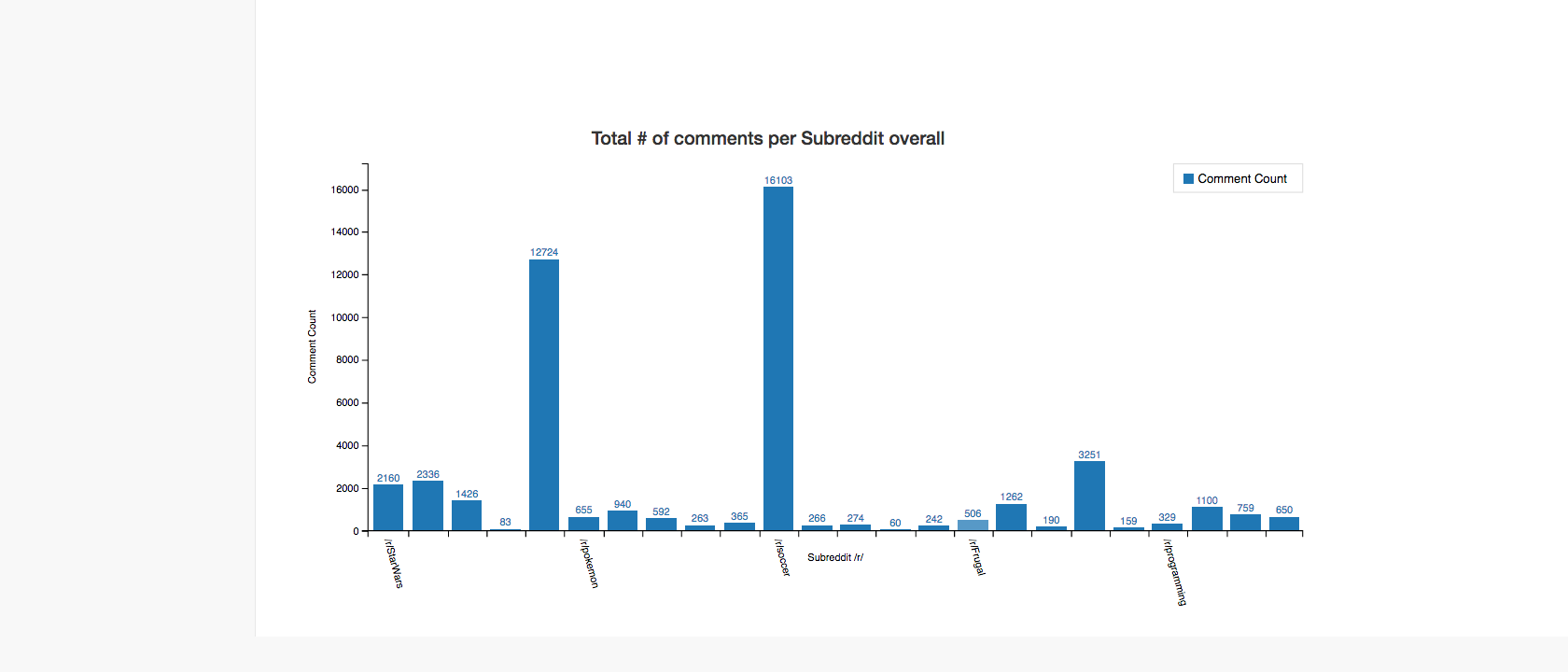
*Fig 2.2.2 Sarcasm fun-fact*

### *2.2.2 View charts*

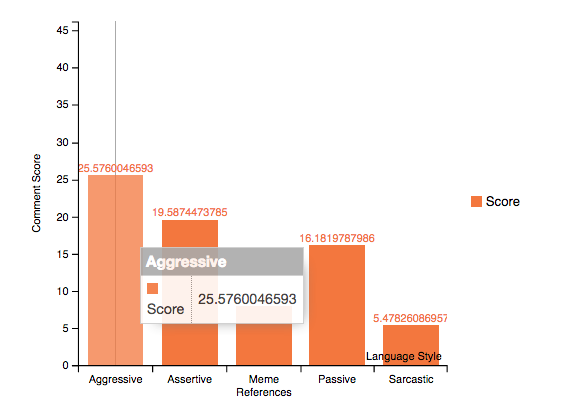
When you navigate to the Charts tab, you are presented with some graphical information in the form of two bar charts and a donut chat. There are some bits of information that explain what each chart represent and the user can hover over each chart to gather some more precise information.

**

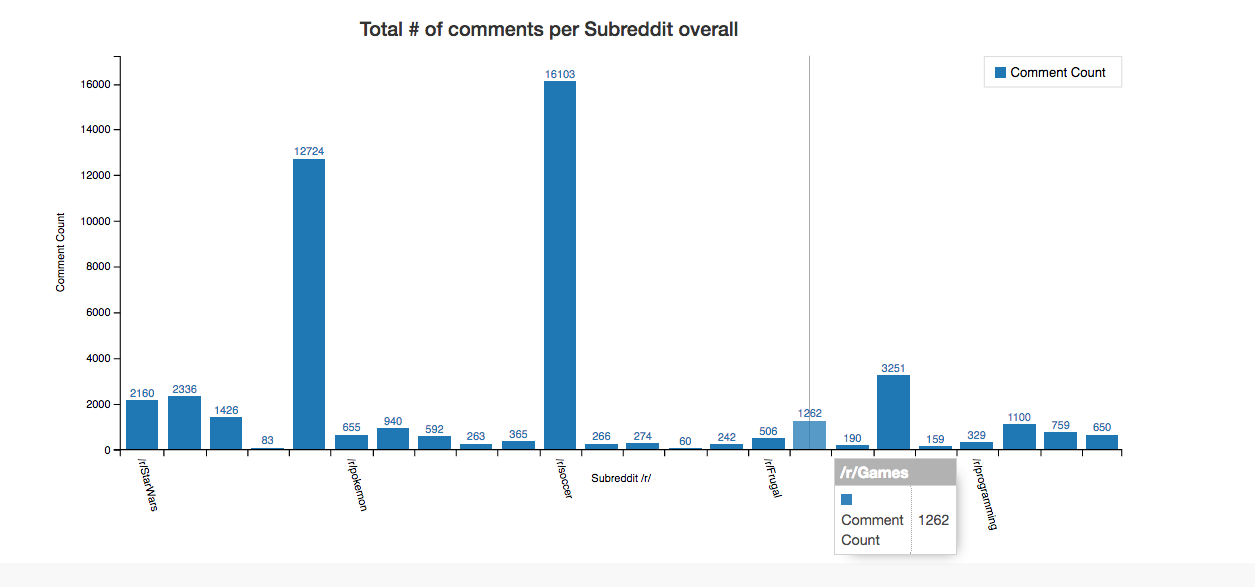
*Fig 2.3 Chart Results dashboard, part 1/2*

**

*Fig 2.3 Chart Results dashboard, part 2/2*

**

*Fig 2.4.1 Hover Chart Result*

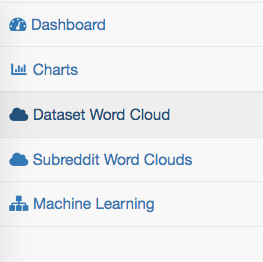
**

*Fig 2.4.2 Hover Chart Result*

### 

### *2.2.3 View Word-Clouds*

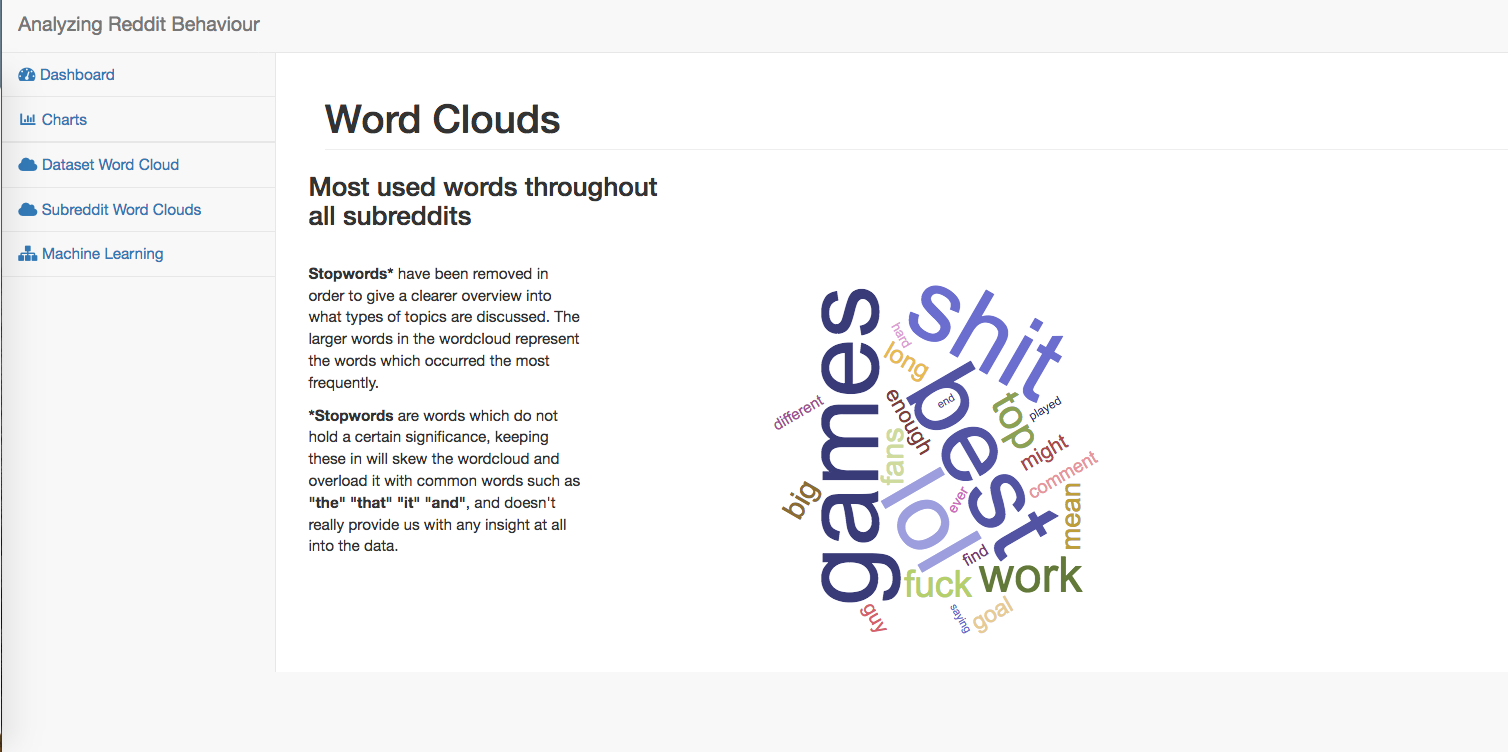
There is the option to view either a word-cloud that represents the overall dataset or to view the word-cloud corresponding to a subreddit of your choice. To view the word-cloud for the overall dataset, simply select the option *Dataset Word Cloud* from the menu on the left-hand side as demonstrated in the image below.



If you wish to view the word-cloud for a subreddit of your choice, simply select the option *Subreddit Word Clouds.*Upon clicking either of these options, you will be redirected to a new page.

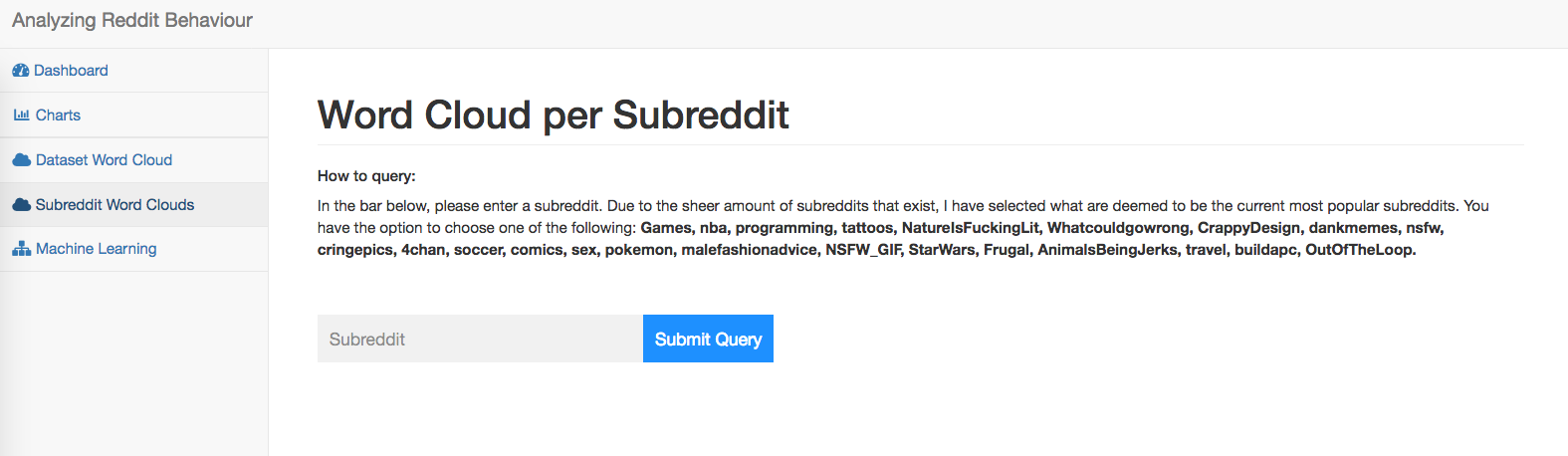
*Fig 2.5 Left-hand Menu*

Upon being presented with the results for the word-cloud for the overall dataset, there is a brief description giving an outline and more description about the cloud. The word-cloud represents the frequency of words within the dataset. The larger the word, the more frequently it was used.

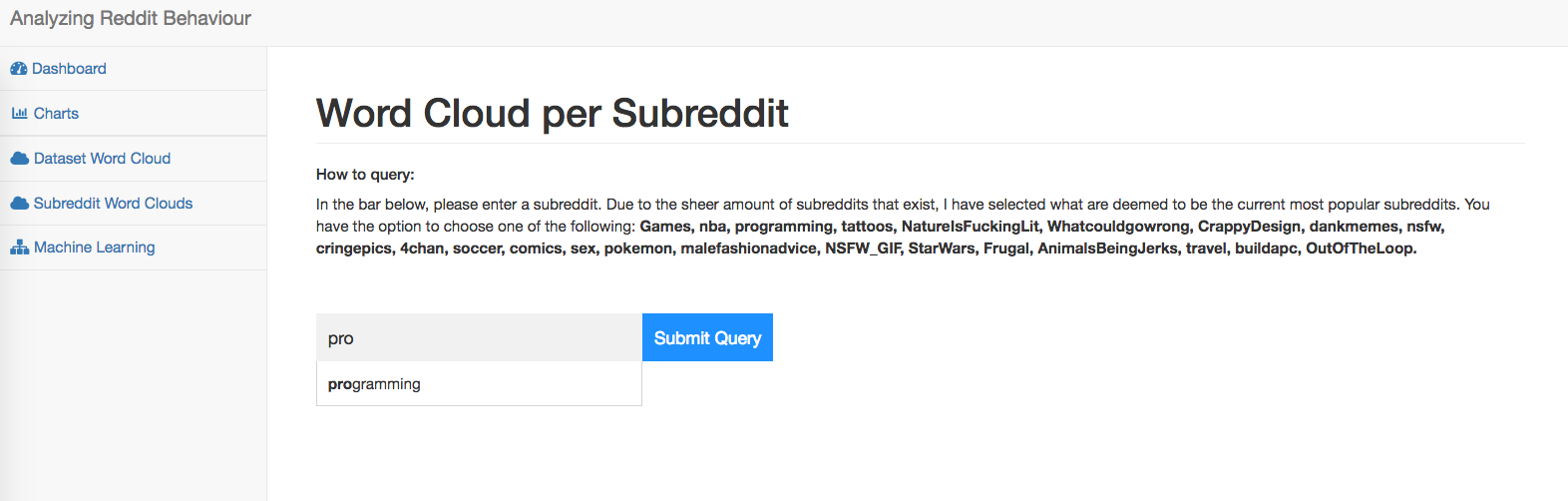
**

*Fig 2.6 Dataset Word Cloud Result*

Within the *Subreddit Word Clouds* tab, you can query by subreddit according to the list that the page offers to you. You can start this by inputting a letter and the search bar should provide you with a list of suggestions to pick from to autocomplete the query, seen below.

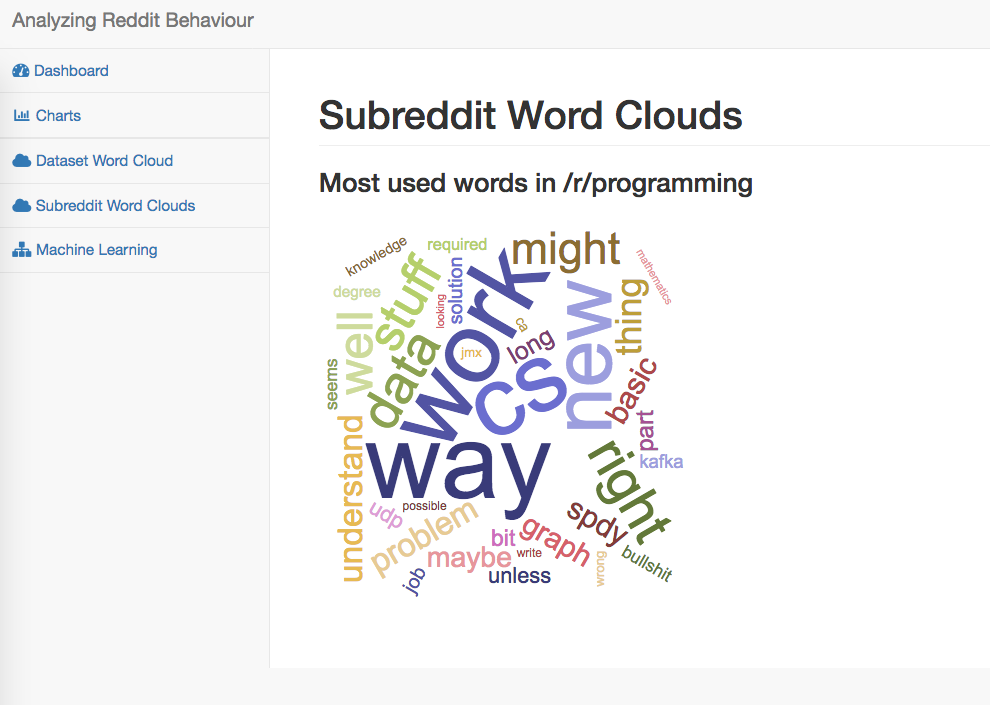
**

*Fig 2.7 Subreddit Word Cloud Query 1/2*

**

*Fig 2.7 Subreddit Word Cloud Query 2/2*

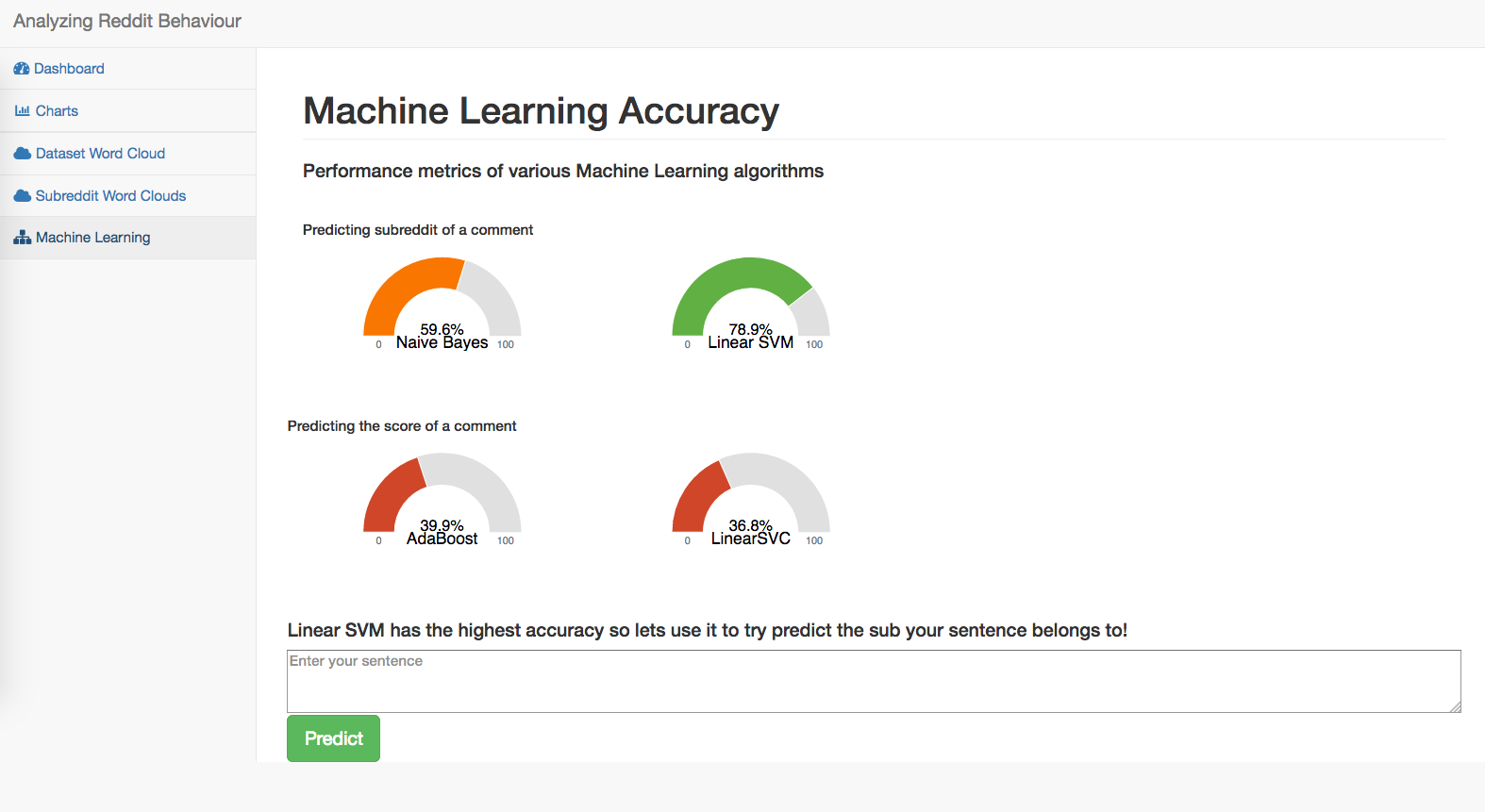
When you have selected a subreddit, you can click Submit Query. You will then be presented with your results outlining the most frequent words.

**

*Fig 2.8 Subreddit Word Cloud Result*

### *2.2.4 View machine learning*

Within the Machine Learning tab, you are presented with four Gauge charts. These charts represent the prediction accuracy across four different Machine Learning algorithms that the application achieved when trying to predict the Subreddit of a comment and the Comment Score of a comment within the dataset.

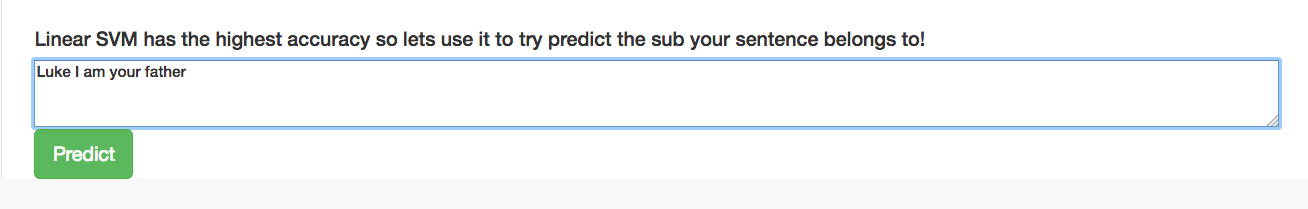


*Fig 2.9 Machine Learning Results*

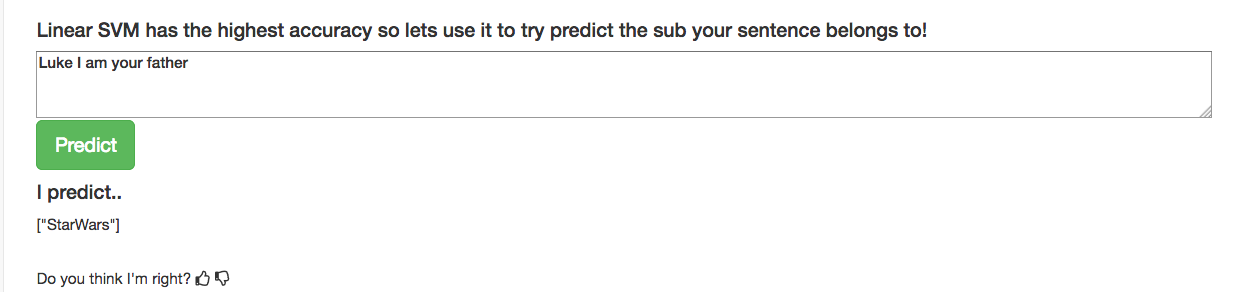
### *2.2.5 Interactive machine learning*

Underneath these gauge charts, you, the user, have the option to input a sentence and to interact with the web application to see if it can correctly predict the expected Subreddit of the users input using the Linear SVM algorithm.   
  
The Linear SVM algorithm was chosen due to the fact that it performed better than Naive Bayes. The Linear SVM algorithm combines Linear Classifiers with SGD training. Training Linear classifiers with SGD is idea for this type of data as it can be successfully applied to approach and rectify the problem of sparse data which is often encountered within text classification and natural language processing.

To use, simply enter your phrase into the textbox and click predict.



*Fig 2.10 Entering input for Interactive Machine Learning*

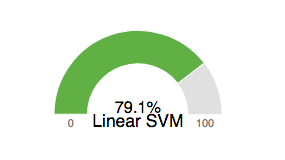


*Fig 2.11 Interactive Machine Learning Results*

There is the option for the user to give feedback if they agree or disagree with the result of the prediction, they can do so by clicking either the thumbs-up or the thumbs-down button. The feedback from the user alters the gauge chart representing the result of the Linear SVM algorithm. It alters it by either increasing or decreasing it depending on which button the user has clicked.



*Fig 2.12 User feedback for predictive accuracy*



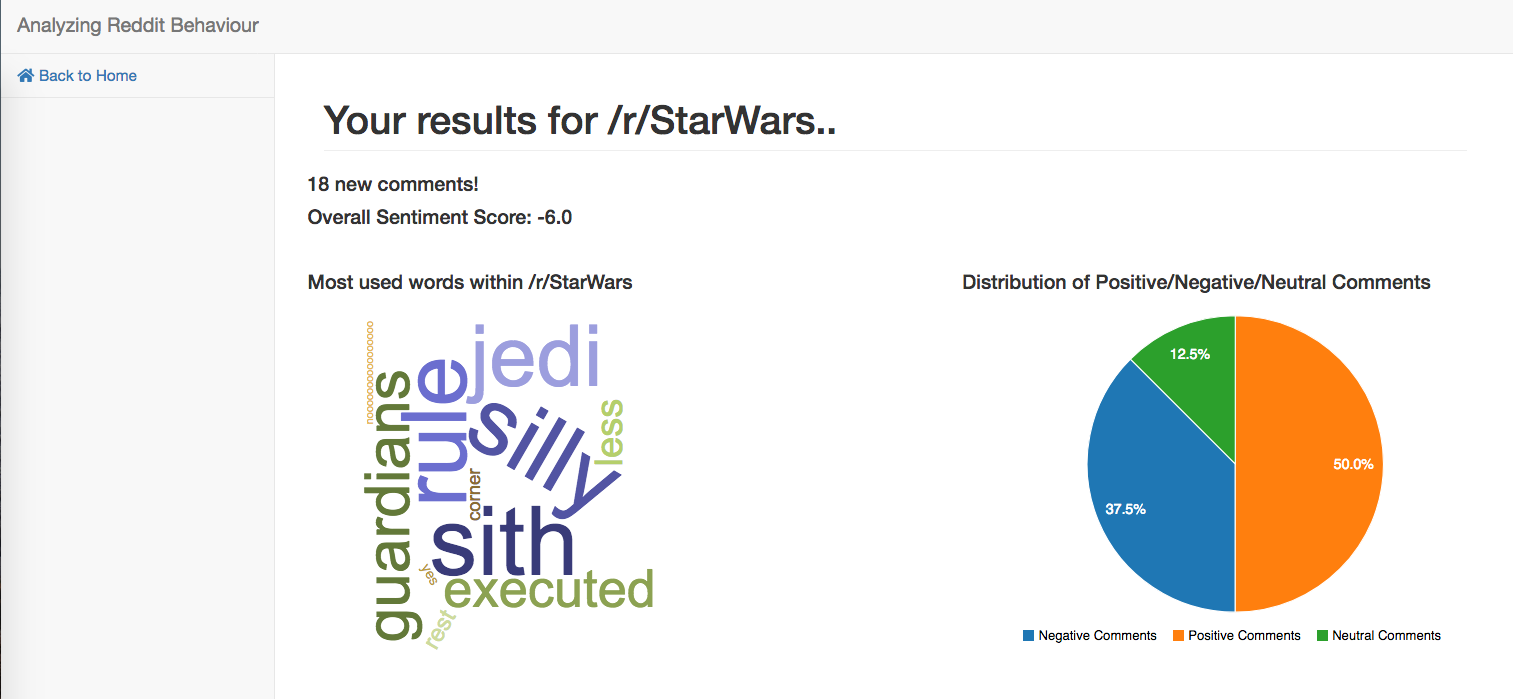
*Fig 2.13 Updated gauge chart from user feedback*

### *2.3 Subreddit User Query*

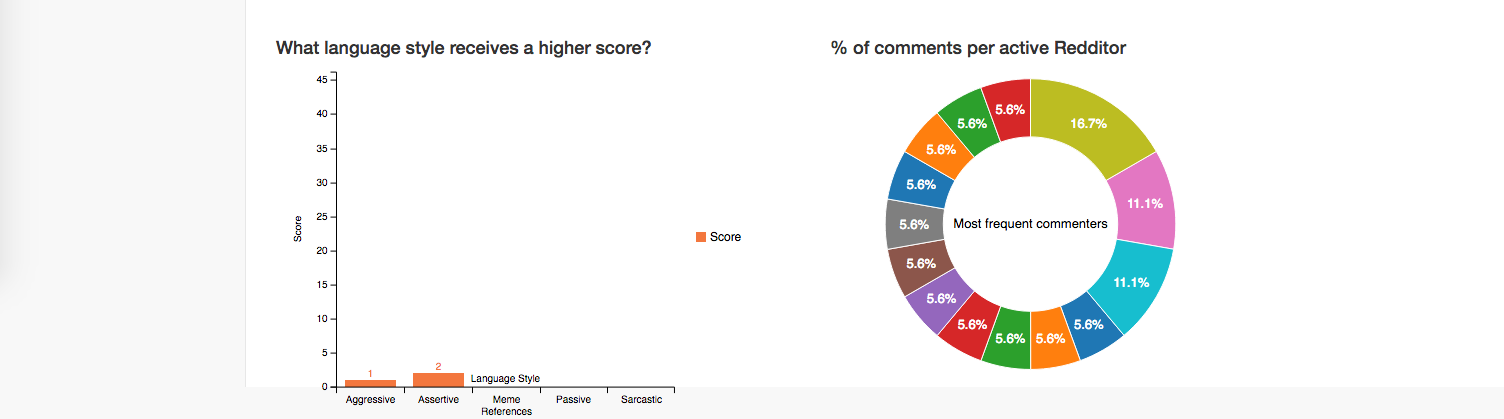
If the user opts to, perhaps, take a quick browse as to what the latest updates/public’s response for a certain topic they can do so by navigating back to the original homepage and entering their desired subreddit into the field provided.

### *Fig 2.3.1 Entering in subreddit query* *2.3.1 Subreddit Results*

Upon clicking the *Let’s Go!*  Button, the user is directed to a single web page which aims to inform the user as to what the overall sense of how other users are responding lately. For instance, a user can query the Subreddit for a certain TV show after the latest episode has aired in order to see a nice overview of the majority response. It displays the most frequently used words, the overall Sentiment, the number of new comments, the most popular language style and also it provides an overview of what proportion of comments belongs to a single user. This can help with highlighting whether or not a certain subreddit had been hijacked by a certain user.



*Fig 2.3.2 Subreddit Query Results ½*



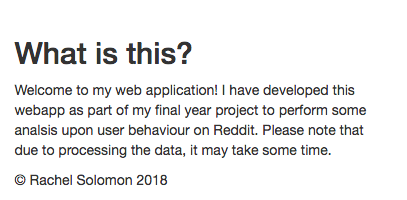
*Fig 2.3.3 Subreddit Query Results 2/2*  
  
After viewing this all, the user is able to navigate back to the homepage by clicking the “Back to Home” button on the left-hand side of the screen in the menu.



*Fig 2.3.4 Redirect back to homepage*

### *2.4 About*

Upon navigating back to the homepage, the user has the option for viewing the About section in order to get a greater overview of what the purpose of the application is.

  
 *Fig 2.4.1 About section*

### *2.5. Journey*

The user also has the option to learn more about the whole process of how the web application was developed. There is a link that will lead the user to the Gitlab where the source code and blog are for if they are interested.



*Fig 2.4.1 Journey section*

## 3. Ending a Session

To end a session, you can simply exit your browser by clicking the red ‘X’ on the top left of the window.