

Name: Viraj Patel

Date: 04/28/2022

EX09: Rotten tomatoes movies review analyses

OBJECTIVE

The goal of this program is to extract reviews from the rotten tomatoes website (if the movie/review exists) for the movie the user enter, the user can search for as many movies as they like and enter "DONE" to exit the program. The program will than print the top 5 and the bottom 5 reviews for the entered movie, generate a word cloud for the top 75% reviews extracted and will also mention the overall prevalent sentiment of the movie. For purpose of our analyses, I have chosen the movie "spiderman", which is the first spiderman movie.

PROCEDURE

Firstly, after importing the required libraries like requests, string, matplotlib, pandas, re and word cloud, I defined two functions "get_reviews", which will extract reviews, score and sentiment for the entered movie and "review_analyses" which will give us the top 5 and bottom 5 reviews, the word cloud and the overall prevalent sentiment.

The "get_reviews" function takes "url" as an argument, which is the URL of the reviews page for the entered movie, in our case, "spiderman". This function than iterates over all the pages of reviews and stores the reviews and their score and sentiment in our global variables.

Our "review_analyses" function than uses these global variables to create a local data frame with 3 columns, "reviews", "score", and "sentiment", to further analyse that the data for our desired output. It also performs the cleaning to the data frame, as in, the scores column had scores of different formats, so I had to normalize those to one format, which was done using the "map" function. Also, the basic cleaning of the reviews column was done, to remove stops words and punctuations for giving us a better word cloud.

On running the program, it asks the user to enter the name of the movie they want to get details for. Once the user enters the movie name, the program outputs the top 5 reviews for that movie, the bottom 5 reviews for that movie, the overall prevalent sentiment of the movie and a word cloud for the reviews. The user can than look the same things for another movie or they can type "DONE" to exit the program.

Let's, run our program and look for reviews for the movie "spiderman".

```
The movie name requirements are:  
Use all lowercase alphanumeric characters
```

```
Enter the movie name you want to get reviews for or enter "DONE"  
to end:spiderman
```

Our program then gives us the output as follows:

1. The top 5 reviews for the movie spiderman are:

```
The Top 5 reviews are:  
54   final sequence leaving theater feeling dizzy giddy looking  
nearest building wishing shoot  
62   raimi tobey maguire bring hero cinematic life fast  
actionpacked blockbuster  
64   terrifically entertaining blockbuster that once gets right  
67   movie swings literally physical emotional energy hours  
89   review threatens turn session gushing film type experience  
watching kept wanting applaud
```

2. The bottom 5 reviews for the movie spiderman are:

```
The bottom 5 reviews are:  
131  just boring bythenumbers interchangeable mess next  
55   good movie mediocre means passing hours classic example  
money special effects diluting artistic vision talent  
148  captures letter spirit spiderman comic book cartoon fame  
171  superhero movie leave audiences flat feeling  
214  reaction word yawn
```

3. The overall prevalent sentiment of the movie spiderman is:

```
The overall sentiment of the movie is Positive.
```

4. Word cloud for the top 75% of reviews is:



CONCLUSION

As we can see from the word cloud that spiderman was a blockbuster movie, with good effects and good entertainment. We can also conclude that lot of people have compared it to the spiderman comic book, the actor selection also played an important role in the movie. As the overall sentiment of the movie is positive, we can conclude that the movie is worth watching.