# Sentimental Analysis on Reviews of Christopher Nolan’s Works

Introduction

I have always considered myself a cinephile. Normally on the weekend, I will pick an intriguing film to enjoy with my friends as a weekly ritual. One of the ways that we determine if a film is watchable or not is by looking up its review from various sources. However, we often found ourselves spending too much time online browsing through reviews that we did not leave enough time for enjoying the actual film. As a result, it will be super helpful if there is a way to easily identify different sentiments within a series of film comments, making every film lovers’ life much easier by not wasting their precious time.

Plus, I am also a big fan of Christopher Nolan and his directing style. Nolan is an eclectic director who has worked on films ranging from sci-fi to historical. I am super interested in his work not only because he is one of the best directors in our time but also due to the controversy surrounding his works. No matter is his most received film like *Inception* (2010) or some of his less-known works such as *Insomnia* (2002) and *Following* (1998,) there are always different voices in response to his cinematic language, making us even harder to make a clear judgment about his film’s quality.

Hence, in this research paper, I would love to conduct a sentimental analysis based on reviews of Nolan’s films throughout his entire career. With the scientific approach, I hope to correctly identify different sentiments in reviews of Nolan’s work and to find out features in the film that are most welcomed/rejected by the audience. By finally evaluating the performance of different methodologies, I determine to offer a sentiment classifier that can accurately reflect audiences’ views and offer pertinent guidance to people who are struggling to find the right movie to watch.

Related Research

I found in total three existing works that are apropos of the purpose of my research:

1. [Analyzing Movie Reviews using Sentiment Analysis, Gayathri Devi Nagalapuram](https://medium.com/nerd-for-tech/analysing-movie-reviews-using-sentimental-analysis-77e28e463b1b)
2. [Analyzing Star Wars Movie Scripts, XAVIER](https://www.kaggle.com/xvivancos/analyzing-star-wars-movie-scripts)
3. [Sentiment Analysis for Movie Reviews, Ankit Goyal & Amey Parulekar](https://cseweb.ucsd.edu/classes/wi15/cse255-a/reports/fa15/003.pdf)

The first article is a major reference to my completed project. Since the research topic is identical to mine, the paper treats as a useful reference about directions that I should consider to complete the project, guiding me to structure the analysis in the right way step by step. On the flip side, instead of scrapping the review online, the author utilizes a pre-existing dataset from Kaggle to conduct her analysis. Compared to her dataset which contains film reviews from the past several decades, my scope is much smaller.

    The second project on Kaggle is also about conducting a text (sentimental) analysis pertaining to film-related text documents, in this case, a movie script. The author hopes to perform a statistical analysis text analysis on the Star Wars script from the Original Tribology Episodes to showcase the most frequent feature words. However, since the entire text analysis is based on R programming, there are not many things that I can reference directly from the author’s code. What is more, instead of analyzing the overall sentiment of one large chunk of text, I will be dealing with a large amount of short and fragment reviews. Plus, the text (Star Wars script) that the author uses for analysis is also more structural which requires disparate approaches in terms of cleaning and modeling compared to mine.

    Unlike the previous two projects which both are self-published online, the third related work is much more formal. It is a rigorous academic paper authored by two college graduates. One of the paper’s goals is to explore the performance of various statistical models (Logistic Regression, k-Nearest Neighbor classifier) applied to the film reviews. Such comparison is something that I want to include in the latter part of my project. It will be extremely useful to envision my own approach, by observing some of the menologies they adopt. In addition, since it is an academic research paper, its general structure, as well as its style of writing, will liken my final project write-up. Despite the similarity, the paper’s methodology is more rigorous where the author creates different versions of bags of words to test out the model’s accuracy. For the purpose of my research, on the other hand, I am going to use only two bags of words to extract meaningful features from the review text.

[Corpus](https://github.com/zejiachen9912/letterboxd-sentimental-analysis/tree/main/DataForAnalysis)

Movie-rating websites are often used by critics to post comments and rate movies which help viewers decide if the movie is worth watching. Instead of finding the data on conventional movie rating website, like IMDB or Rotten Tomatoes, where most of reviews are made by professional critics, I decide to come up with a more specific dataset which is more tailored to my research direction.

The website that I am going to get the film reviews is called [*Letterboxd*](https://letterboxd.com/) – a social platform focused on sharing opinions about, and love of, film. The platform allows users, many of whom are just film lovers, to rate and record their opinions about films which makes it ideal to conduct my desired sentimental analysis. Plus, I want to keep the scope of my research topic more specific. Hence, I am planning to scrap reviews of Nolan’s film on Letterboxed.

Films that I selected from Nolan’s career are as followed:

|  |  |  |
| --- | --- | --- |
| **Film Title** | **Year** | **Number of Reviews** |
| *Inception* | 2010 | 370 |
| *The Dark Knight* | 2008 | 374 |
| *Interstellar* | 2014 | 326 |
| *Dunkirk* | 2017 | 320 |
| *Tenet* | 2020 | 255 |
| *Memento* | 2000 | 389 |
| *The Prestige* | 2006 | 377 |
| *Insomnia* | 2002 | 364 |
| *Following* | 1998 | 418 |

Thanks to the clean formatted website URL, I was able to scrap all the reviews I need with `Request` and `BeautifulSoup`. To ensure the quality of reviews, I decided to scrap the review in an order of “Review Activity.” (amount of likes, comments received) To store the data locally, I first put the review in a data frame, and subsequently stored the data frame as a local csv file. In addition, I preprocess the data while I was scrapping it form *Letterboxd*. After examining some pages of reviews, I have come to an understanding that there are non-English reviews on Letterboxd. As a result, I took this into account while I was collecting the review from the website. I designed a function `isEnglish()` that takes an addition step to check if a review is consisting of English characters, ensuring the corpus is free from reviews that consist with non-English characters or emojis.

In terms of the actual dataset, it is consisted with in total two columns of data – movie rating and review text. Both the data are inextricable part of my project. The review text forms the basis of my text analysis whereas the rating is treated as a critical metrics to characterize reviews’ sentiment which I will conver in more detail when expiating my methodology.

Process & Methods

My text analysis consists of two parts:

1. Review Sentimental Analysis using VADER & exploratory feature analysis using Tf-Idf
2. Binary Classification with Logistic Regression

Via this two-part process, I first wish to have some sorts of preliminary result and insights using hand-on method such as VADER and Tf-Idf. Then, with a cleared direction in mind, I am going to refine my existing analytics pipeline using logistic regression and hope to gauge the performance between the two.

Above all, to evaluate the NLP’s methods efficacy, I must have some sorts of benchmark to validate their prediction. Hence, I manually classify the review sentiments into two general categories, positive and negative, based on the rating given reviewers. A piece of review is considered as a positive review if its rating is large or equal to four star and is seen as negative if the rating is smaller or equal to three star.

After having a clear guideline, the first step of the analysis is to use VADER to evaluate each review sentiments grouped by the film names. For a review that has a VADER compound score greater than 0.35, it is treated as a positive review while a VADER score below 0 will get a review classified as negative. Finally, I determined the VADER’s accuracy by calculating the share of positive/negative reviews that are identified by VADER correctly.

Subsequently, I subset the corpus into two sub-corpus that each contain the only positive/negative reviews based on rating, hoping to discover top features that are most frequently mentioned in the positive/negative review in each Nolan’s film using Tf-Idf. Except for following the regular procedures, I took some more step to refine the analysis further. First, I removed possible stemmed worlds when tokenizing the review text. Then I created a bag of words for both positive and negative review of Nolan’s film by calculating the total word counts for each word across all the reviews. Since bag of words ignores the semantic context of the review and concentrates primarily on frequency of each word, I also tried n-gram modeling where I tokenized the bi-gram features in each of the review document to offer more contextual information to the review text.

The next part of the study is to perform a binary classification of reviews as favorable or unfavorable. Here, I used a simple logistic regression to classify the text sentiments and trace back to determine prediction accuracy. Finally, I again used logistic regression to construct a model to make prediction to each word’s weight on determining the sentiment, offering us insights about the features that are most likely to tip our classification model.

Result & Discussion