

Proposed Plan of Trailer Research

Mike Ma¹*

¹SIRC. Ontario Tech University, ON, CA

In this proposed plan I described the motivation of choosing data and research direction in the first part. In the second part I discussed the possible problems before literature reviews. In the third part I discussed my proposed plan with time-frame, including a short summary of data collected.

Motivation.

The success or failure of a movie is often determined in its first weekend of play. In order to make this opening successful, movie producers must employ a number of promotional strategies including movie trailers, to publicize the movie for a significant period of time prior to its release. A movie trailer, as described in Wikipedia, is an advertisement or a commercial for a feature film that will be exhibited in the future. Of some ten billion videos watched online annually, film trailers rank the third, after news and user-created video. It is clear that businesses have a strong interest in tapping into this huge data source to extract information that might improve their decision making process. For example, predictive models derived from movie and its trailers may facilitate filmmakers making more profitable decisions.

Possible Problems.

1. Given meta-data about a movie and trailer, predict the number of views the trailer will receive after a fixed period of time.
2. Given meta-data about a movie and trailer, create a method to measure the preference (like, dislike, tweet, share, and preference of comments) of audience toward trailer, and predict the preference after a fixed period of time.
3. Analyze the relationship between preference of trailer and rating of movie, and predict the rating of a movie before it published.
4. Analyze the relationship between preference of trailer and movie grosses, and predict the grosses and movie stock price of the movie.
6. Test whether the sentiment of user-generated comments correlates with the movie's box office information better than simple counts of like/dislike from Youtube. Make predictions on movie popularity and stock price in terms of sentiments of user-generated comments. Dataset would be processed using a set of positive and negative words and then classified as positive, negative or neutral. Use sentiments data as new predictors.

Proposed Plan.

1. Collect Related Data (July 23 - Sep 1)
 - Description, duration, viewCount, likeCount, dislikeCount, favoriteCount, commentCount, Stars, Direc-

tors, ReleaseTime, Synopsis from about 10K trailers from 10 channels in youtube, including different language trailers.

- About 100K Comments of all 8 categories trailers from the biggest channel in youtube.
- About 3K all 8 categories trailers videos from the biggest channel in youtube, 720K, 40GB.
- About 1000 trailers Sharing info from the trailer channel
- Title, Link, Certificate, Content, rating, runtime, Director, Stars, Writers, Genre, PublishYear, Metascore, Gross revenue, Headline, Plot Keywords, certificate, Official Sites, Country, Language, Filming Location, Gross USA, Cumulative World Gross, Production Co, Sound Mix, Color, Aspect Ratio, Discription, TrailerRuntime, UK Release Time, Budget, Storyline of about 1K Movies from 2011-2020 movie lists in IMDb

2. Data Cleaning and save data to Cloud (Sep 1-Sep 8 if possible)
3. Read Papers and finish literature review. (Aug 19-Sep 1)
4. Learn deep learning, NLP and build models (Aug 25- Sep 10)
5. Begin Experiment. (Sep 5-Sep 15)
6. Write a formal paper. (Sep 15-Oct X)