EM624 EX09. Author: Lei Sun

**Web page mining report**

**Report Description:**

In this project, I choose a latest movie,” FANTASTIC BEASTS: THE CRIMES OF GRINDELWALD”, which is also one of my favorite movie series “Harry Potter” and the second film of the series of “FANTASTIC BEASTS”. Since I am a big fan of “Harry Potter” and I want to know how the public like this movie, I use web page mining approach to search the reviews, which is written in text, and user-report score for rating, in order to identify how popular the movie is, and how good/bad the movie present for the original fiction novel.

**Data Preparation and mining process:**

**Section1: Web Mining:**

1. Start with a home page of link of the website, identify the total pages of reviews, read the content of the page, and find addresses “review” in the page;

2.loop each page through these link addresses, and then continue to loop until all the pages of the site have been crawled.

**Sectioin2: Review and rating data collecting:**

1. Create a list of lists [review, score], where “review” is the review text content and “score” is the rate given for each review;

2. Rank reviews by score;

3. Read list of review in to a word list;

4. Cleaning stop words, and eliminating words providing no insight.

**Research observations:**

According to rottentomatoes.com, the average rating is 3.6/5 which equals to 0.72. Then I curious about details of how audience like this film. So, I search for the reviews of text sorting from the highest score.

Surprisingly, all the highest score views are obviously from fans of Harry Potter, many mentions roles in the movie and recommend deserve to watch for people who understand it.

**Top 5 reviews sort from highest score:**

1.The immense wonder of the Potterverse lingers but the bulk of Grindelwald feels like two movies haphazardly sewn together and the camera is far too interested in the wrong side of this creation.

2. Perhaps the best thing from the onset would have been to leave the fanfic to the people who understand it: the fans.

3."The Crimes of Grindelwald" is going to leave you with more questions than answers, but for Rowling\s fans, this movie is a treat and proves she is still very much on top of her game.

4. "You don’t feel the horror or the charisma. I love the Harry Potter universe and I wish I could recommend this film with greater enthusiasm. Its competent but not enchanting enough.

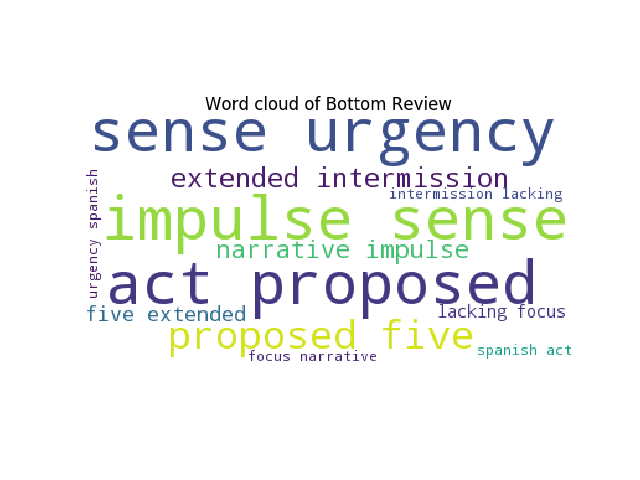
5. It ‘s a callous, perplexingly frustrating way to tell a story, and for large portions of the feature I sat scratching my head in appalled disbelief.

The word cloud shows the highest appreciation. (After eliminating characters’ names, which frequency is high as well)

In the contrast, people who don’t like the movie express the complains. It is obvious that they think the movie is too long and lacking focus, because they don’t know detail scenario in the original novel.

**Reviews the review of lowest score:** More than "Act 2" of the proposed five, the film feels like an extended intermission, lacking focus, narrative impulse and any sense of urgency.

From the word cloud, we can easily see that many mention “extended intermission” or ”lacking focus”.

****

They seem just feel ”sense urgency” or “impulse sense” and consider the film as a normal action movie, not like fans who passionate to follow the adventures of those characters in wizard magic world.

In conclusion, FANTASTIC BEASTS: THE CRIMES OF GRINDELWALD is an interesting film and rate 0.72 in general. However, the reviews and rating are extremely different. The comments divided into two opposite opinions: Fans’ latest favorite film, but not interested enough for new audiences.