



What makes a hit / flop Action movie?

Presented by: Shuheng.Ma, Yikang Wang, Nicole (Yue) Jiang, Rongzhi Xu, Yiqun Tian
Cohort: Section B

Agenda



**Problem
Statement**



Methodology



**Result &
Models**



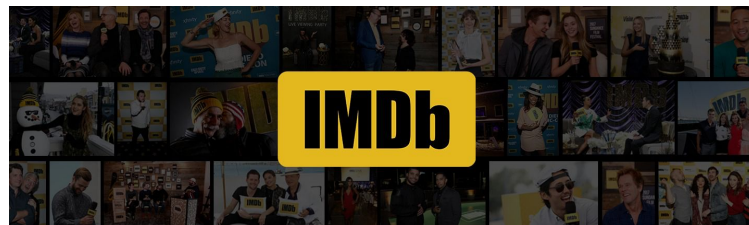
Analysis



Conclusion

- **Data Source**
- **Pre-processing**
- **Topic Modeling**

Problem Statement



- **Background:**

- For movie companies efficiently select the movie scripts worth their investment
- Help movie companies to find what are the popular topics in current market
- Movie companies to evaluate if a new movie will perform well in the market
- To do better in forecast movie' sales

- **Goal:** Identify the features that make Action Movies succeed or fail.

Methodology

- **Data Sources:** IMDB movie comments

- Website link:

[https://www.imdb.com/search/title/?title_type=feature&num_votes=25000.&genres=action&sort_user_rating.desc](https://www.imdb.com/search/title/?title_type=feature&num_votes=25000.&genres=action&sort=user_rating.desc)

- **Scrape features**

- Top 20 best action movies and top 20 worst action movies, with 200 comments from each movie

Feature Film, Rating Count at least 25,000, Action (Sorted by IMDb Rating Descending)

1-50 of 1,450 titles. | [Next »](#)

View Mode: [Compact](#) | [Detailed](#)

Sort by: [Popularity](#) | [A-Z](#) | [User Rating ▼](#) | [Number of Votes](#) | [US Box Office](#) | [Runtime](#) | [Year](#) | [Release Date](#) | [Date of Your Rating](#) | [Your Rating](#)



1. **The Dark Knight** (2008)

PG-13 | 152 min | Action, Crime, Drama

★ 9.0

☆ Rate this

84 Metascore

When the menace known as the Joker wreaks havoc and chaos on the people of Gotham, Batman must accept one of the greatest psychological and physical tests of his ability to fight injustice.

Director: [Christopher Nolan](#) | Stars: [Christian Bale](#), [Heath Ledger](#), [Aaron Eckhart](#), [Michael Caine](#)

Votes: 2,258,158 | Gross: \$534.86M



2. **The Lord of the Rings: The Return of the King** (2003)

PG-13 | 201 min | Action, Adventure, Drama

★ 8.9

☆ Rate this

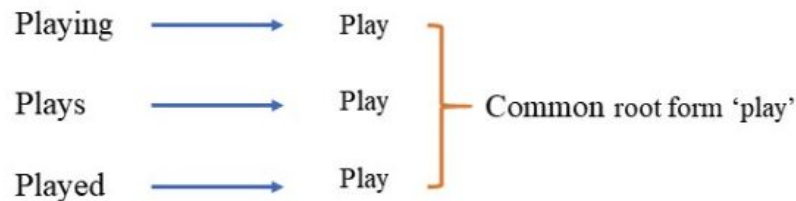
94 Metascore

Gandalf and Aragorn lead the World of Men against Sauron's army to draw his gaze from Frodo and Sam as they approach Mount Doom with the One Ring.

Director: [Peter Jackson](#) | Stars: [Elijah Wood](#), [Viggo Mortensen](#), [Ian McKellen](#), [Orlando Bloom](#)

Votes: 1,612,774 | Gross: \$377.85M

Methodology



- **Pre-processing:**

- Comments Lowercase
- Lemmatization
- Stopwords with additional movie common phrases (E.g. movie, film)
- Regular expression - > Tokenization

```
> stopwords("english")
[1] "i"      "me"      "my"      "myself"  "we"
[6] "our"    "ours"    "ourselves" "you"     "your"
[11] "yours"  "yourself" "yourselves" "he"      "him"
[16] "his"    "himself" "she"      "her"     "hers"
[21] "herself" "it"      "its"      "itself"  "they"
[26] "them"   "their"   "theirs"   "themselves" "what"
[31] "which"  "who"     "whom"     "this"     "that"
[36] "these"  "those"   "am"       "is"       "are"
[41] "was"    "were"    "be"       "been"     "being"
[46] "have"   "has"     "had"      "having"   "do"
```

Methodology

- **Topic Modeling:**

- For the best/worst 20 action movies, obtain 5 topics for each
- Use Excel to obtain the specific five topics

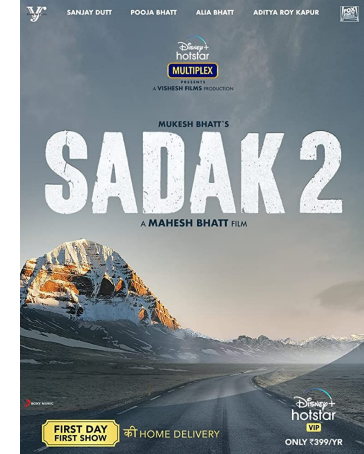


Topic_ID	Best 20 Action Movie	Keywords
1	Sci-Fi	Star wars, Jedi
2	Family	Samurai, family, father
3	Fantasy	Lord of rings, frodo
4	Superhero	Terminator, batman, joker
5	Drama	Story, characters, scenes

Methodology

- **Topic Modeling:**

- For the best/worst 20 action movies, obtain 5 topics for each
- Use Excel to obtain the specific five topics



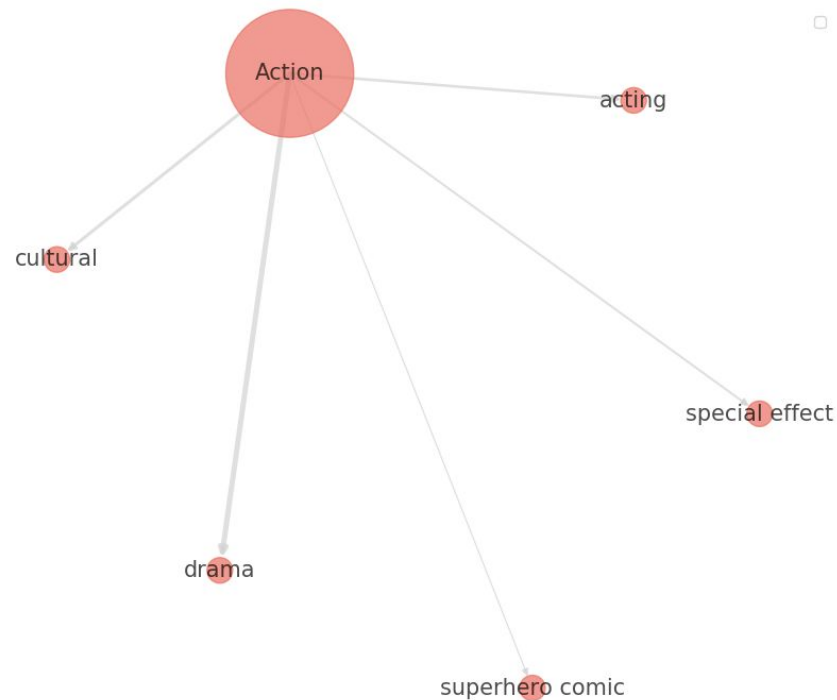
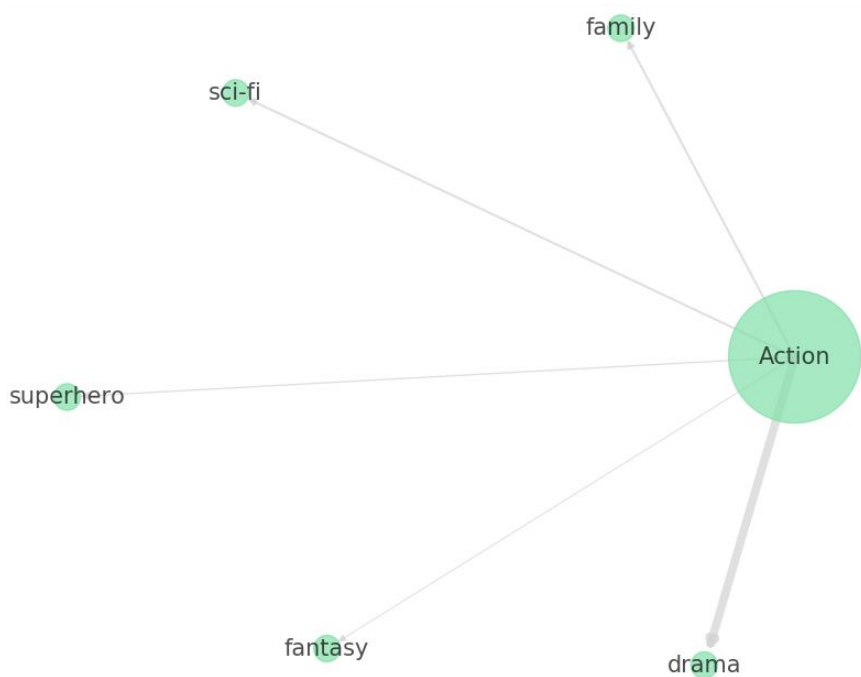
Topic_ID	Worst 20 Action Movie	Keywords
1	Special Effect	Special Effects, Explosive
2	Acting	Bad/poor Acting, Performance
3	Superhero Comic	Batman, Catwoman, Robin
4	Cultural	History, Tradition
5	Drama	Story, Plot

Results and Analysis

Word frequency for best & worst movies



Results and Analysis



Conclusion

- For the top 20 action movies, they shared the common features:
 - **Family, fantasy, sci-fi, drama, and superhero** as the flags of successful movies.
- For the worst 20 action movies, they are:
 - **Special effect, acting, superhero comic, cultural, and drama**
- Film companies could use these as indicators to locate potential successful scripts and to avoid highly difficult ones
- Film companies can also use these indicators as factors to forecast movie sales.

Q&A

