Opening Big in Box Office? Trailer Content Can Help

Adarsh Tadimari $^1$, Naveen Kumar$^2$, Tanaya Guha$^3$, Shrikanth S. Narayanan$^2$

$^1$Electrical Engineering, Indian Institute of Technology Madras
$^2$Signal Analysis and Interpretation Lab (SAIL), University of Southern California, Los Angeles
$^3$Electrical Engineering, Indian Institute of Technology Kanpur
The opening weekend is crucial for a movie’s financial success. Movie makers try to maximize revenue in early stages. Promotion via print, web, social media, TV. 4% of marketing budget spent on trailers.
Past Work

Researchers predict movie’s success using

- **Meta data** - actor, runtime, genre [Chang & Ki 2005]

- Sentiment analysis from language - critics reviews, weblogs [Joshi et al. 2010], [Mishne & Glance 2006]

- Social media - twitter chatter before movie’s release [Asur & Huberman 2010]

- Google search trends - before movie’s release [Panaligan & Chen 2013]
Past Work

Researchers predict movie’s success using

- **Meta data** - actor, runtime, genre [Chang & Ki 2005]

- **Sentiment analysis** from language - critics reviews, weblogs
  
  [Joshi et al. 2010], [Mishne & Glance 2006]
Past Work

Researchers predict movie’s success using

- **Meta data** - actor, runtime, genre [Chang & Ki 2005]

- **Sentiment analysis** from language - critics reviews, weblogs
  [Joshi et al. 2010], [Mishne & Glance 2006]

- **Social media** - twitter chatter before movie’s release [Asur & Huberman 2010]
Past Work

Researchers predict movie’s success using

- **Meta data** - actor, runtime, genre [Chang & Ki 2005]

- **Sentiment analysis** from language - critics reviews, weblogs
  
  [Joshi et al. 2010], [Mishne & Glance 2006]

- **Social media** - twitter chatter before movie’s release [Asur & Huberman 2010]

- **Google search trends** - before movie’s release [Panaligan & Chen 2013]
Importance of Content

- Past work did not use trailers
Importance of Content

- Past work did not use trailers
- Trailers are created to invoke viewers’ interest
Importance of Content

- Past work did not use trailers
- Trailers are created to invoke viewers’ interest
- Can trailer’s content predict movie’s financial success? If yes, to what extent?
Importance of Content

- Past work did not use trailers

- Trailers are created to invoke viewers’ interest

- **Can trailer’s content predict movie’s financial success?**
  - If yes, to what extent?

- Audiovisual features → predict a movie’s opening weekend gross
Our Database

Our Database is a collection of movie trailers from the years 2010 to 2014, which includes 474 movies. We use the IMDb dataset to gather information about the movies, including opening weekend gross, budget, actors, genre, MPAA rating, and number of screens. This data is adjusted for inflation and is available on GitHub at https://github.com/tadarsh/movie-trailers-dataset.
Our Database

- Budget, ticket price adjusted for inflation.
- https://github.com/tadarsh/movie-trailers-dataset
Method Overview

- Predict: Opening weekend’s gross ($$)

- Obtain:
  - Metadata features: budget, actor's experience, genre, MPAA rating, number of screens
    - Chang & Ki 2005
  - Audio and video features from trailer

Tadimari et al.
Method Overview

- **Predict:** Opening weekend’s gross ($$)

- **Obtain:**
  - **Metadata features:**
    - budget, actor’s experience, genre, MPAA rating, number of screens
  [chang & ki 2005]
Method Overview

- **Predict**: Opening weekend’s gross ($$)

- **Obtain**:
  - Metadata features: 
    - `budget`, `actor’s experience`, `genre`, `MPAA rating`, `number of screens`
    - [Chang & Ki 2005]
  - Audio and video features from trailer
Method Overview

- **Predict**: Opening weekend’s gross ($$)

- **Obtain**:
  - **Metadata features**:
    - *budget*, *actor’s experience*, *genre*, *MPAA rating*, *number of screens*
    - [Chang & Ki 2005]
  - Audio and video features from trailer

- Linear regression analysis
Prediction using Meta Data (replicating [Chang & Ki 2005])

Meta data explains a similar amount of variance as reported earlier.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of movies</td>
<td>431</td>
<td>474</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.617</td>
<td>0.631</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td><strong>0.611</strong></td>
<td><strong>0.613</strong></td>
</tr>
</tbody>
</table>
Observations

- **Sci-Fi Movie** $5.5 million $\uparrow$
- **Thriller Movie** -$3.2 million $\downarrow$
- **Budget** 0.17 per $ invested $\uparrow$
- **Christmas Release** -$7 million $\downarrow$
- **Movie Runtime** $100,000 per minute $\uparrow$
- **Screens** $5,000 per screen $\uparrow$
- **Sequel** $12$ million $\uparrow$
Audio Features

- Interspeech09 emotion challenge features [Schuller09]
- Designed to capture prosodic, spectral and voice quality features.

<table>
<thead>
<tr>
<th>Low level descriptors</th>
<th>Functionals</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZCR</td>
<td>mean</td>
</tr>
<tr>
<td>energy</td>
<td>standard deviation (std)</td>
</tr>
<tr>
<td>F0</td>
<td>kurtosis, skewness</td>
</tr>
<tr>
<td>HNR</td>
<td>extremes: value, rel.position, range</td>
</tr>
<tr>
<td>MFCC 1-12</td>
<td>offset, slope, MSE</td>
</tr>
</tbody>
</table>

- 384 dimensional features reduced to 50 using PCA.
Video Features

- Intensity histogram
Video Features

- Intensity histogram

- Color histogram (Hue)
Video Features

- Intensity histogram

- Color histogram (Hue)

- Number of shots
Designed Video Features

- Overall motion activity
Designed Video Features

- Overall motion activity

- Percentage of close up shots
## Comparison with Other Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budget</td>
<td>0.47</td>
</tr>
<tr>
<td>Screens</td>
<td>0.41</td>
</tr>
<tr>
<td>Sequel</td>
<td>0.12</td>
</tr>
<tr>
<td>Audiovisual features</td>
<td>0.11</td>
</tr>
<tr>
<td>Sci-Fi</td>
<td>0.09</td>
</tr>
<tr>
<td>Runtime</td>
<td>0.06</td>
</tr>
<tr>
<td>Christmas Release</td>
<td>0.05</td>
</tr>
<tr>
<td>MPAA Rating</td>
<td>0.05</td>
</tr>
<tr>
<td>Thriller</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Prediction using Content

<table>
<thead>
<tr>
<th>Feature</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata</td>
<td>0.61</td>
</tr>
<tr>
<td>Audiovisual features</td>
<td>0.11</td>
</tr>
</tbody>
</table>
Content Improves Prediction

<table>
<thead>
<tr>
<th>Features</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>audiovisual content</td>
<td>0.11</td>
</tr>
<tr>
<td>metadata</td>
<td>0.61</td>
</tr>
<tr>
<td>metadata + content</td>
<td><strong>0.65</strong></td>
</tr>
</tbody>
</table>

- ~6% improvement in explained variance.

Outliers: Iron man 3, The hunger games, Alice in wonderland (among the highest grossing movies)
Content Improves Prediction

<table>
<thead>
<tr>
<th>Features</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>audiovisual content</td>
<td>0.11</td>
</tr>
<tr>
<td>metadata</td>
<td>0.61</td>
</tr>
<tr>
<td>metadata + content</td>
<td><strong>0.65</strong></td>
</tr>
</tbody>
</table>

- ~6% improvement in explained variance.
- outliers: *Iron man 3*, *The hunger games*, *Alice in wonderland* (among the highest grossing movies)
Summary

- Studied the contribution of trailer content on a movie’s financial success
- Demonstrated that content carries additional information
- Created a database of more than 400 movie trailers and associated metadata.
- Future work: Design audio features, investigate ‘hype’
Thank you