Casting Actors: a $$$ Story

by Niall Murphy

October 7, 2016

The Question: Who do we cast for our movie?

The Goal: Cast actors who bring in $$$.

Data: BoxOfficeMojo

Models:

- ElasticNet and Lasso for coefficients
- Similar results for top actors, Lasso cuts the bottom ~150 actors (from ~400)
- Random Forest for feature significance

Results:

- Using Non-Animated Movies; Actors with 10+ movies
- Vast majority of model's explanatory power is contributed by 'expected' factors like budget
- Coefficients for Actors: grain(s) of salt are needed when interpreting them
- Overall R-squared of .35 with ElasticNet
- A brief look at the top actors/actresses confirms some assumptions, shows some weaknesses of model
- Examples: Production Budget Coefficient <1, intercept, negative coefficients

Top 13 Actors

- Harrison Ford (44m)
- Tom Hanks (44m)
- Samuel L Jackson (32m)
- Andy Serkis
- Elijah Wood
- Jack Nicholson
- Tyrese Gibson
- Robert Downey Jr.
- Chris Pratt
- Bill Murray
- Viggo Mortenson (15m)
- Orlando Bloom (15m)
- Bradley Cooper (15m)

Top 13 Actresses

- Carrie Fisher (36m)
- Elizabeth Banks (27m)
- Natalie Portman (24m)
- Kathy Bates
- Bonnie Hunt
- Julia Roberts
- Liv Tyler
- Anna Kendrick
- Sandra Bullock
- Scarlett Johansson
- Cate Blanchett
- Jennifer Lawrence (17m)
- Kiera Knightley (14m)

Future Enhancements:

- Differentiation between “role levels” and types - a leading role is probably different from a supporting actor in terms of their effect on the film
- Access to actor-price listings, demographics, movie-specific marketing data could allow for a “Sabermetrics” esque approach to some casting decisions

Reminder:

- Not a replacement for the casting process, but a potentially powerful supplemental tool