An Investigation into How Movie Ratings Vary Following a Movie's Release

TEAM 11 | CITADEL DATA OPEN WEST COAST REGIONAL

Summary of Insights

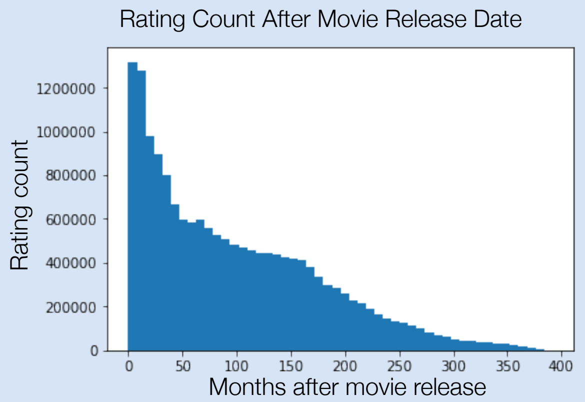
>The monthly average ratings of a movie tends to decrease

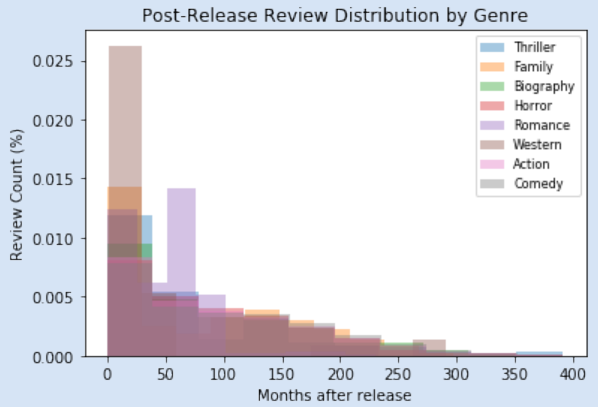
BACKGROUND

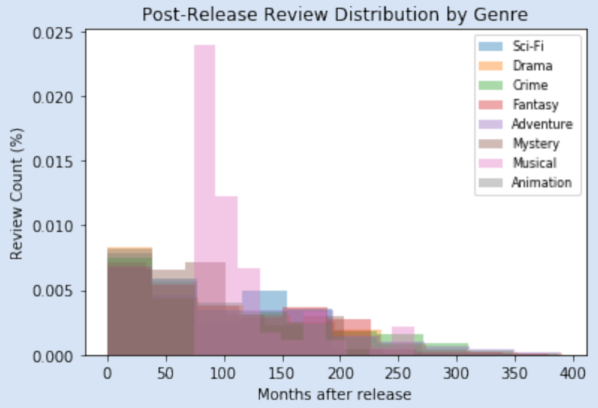
Consumers read product ratings and reviews all the time to make informed consumption decisions. As the film industry enters a Second Golden Age with the emergence of streaming giants, movie ratings are playing an increasingly prominent role in consumer choice. Nevertheless, these ratings are inherently inconsistent across reviewers. In particular, More importantly. How might we better understand the evolution of consumers' perception of a movie following its release?

Motivation

While one would presume most consumers tend to watch movies within a year or two after its release, we found that a bulk of consumers actually continue watching movies many years after (right). These distributions also seem to vary across genres (bottom).







DATA

Features for ~60,000 Movies
Industry metadata such as genre, budget, profit margin, runtime, year released

~28,000,000 User Rating Data from MovieLens
Unique user ratings for movies with timestamp

1,128 "Genome Tags" for ~13,000 Movies
Each tag-movie pair has a relevance score based on user feedback

MODEL 1: HOW DO MONTHLY AVERAGE RATINGS DEPEND ON TIME AFTER RELEASE AND CONSUMER PROFILE?

>MODEL

We ran a multilinear regression model for movies released during and after 1996 using the following equation.

$$y_t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon_i$$

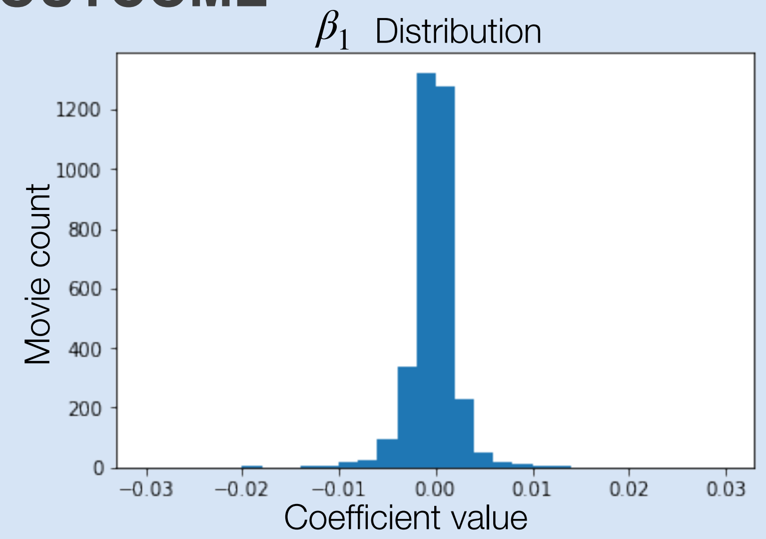
y_t monthly mean rating of a specific movie

x_1 number of months since the movie was first rated

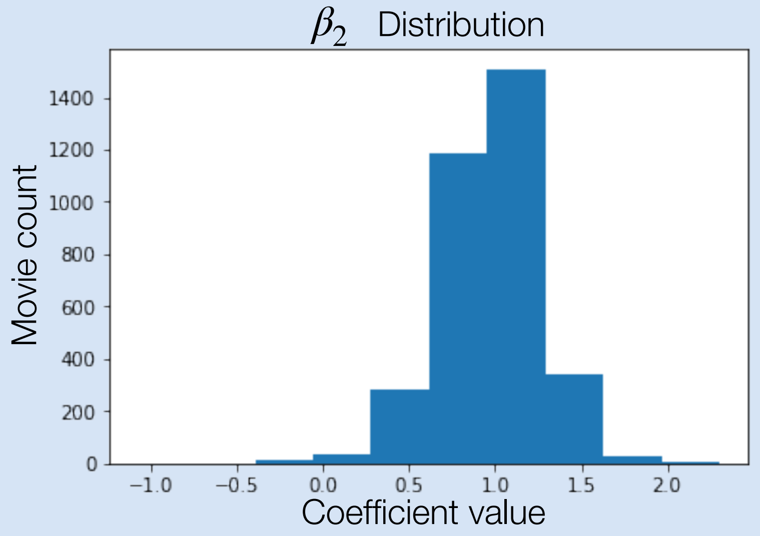
x_2 mean user rating (across all movies rated by each user) across users who rated the movie in the given month

In short, we are using time elapsed and overall user rating behavior (predictors) to predict a movie's monthly ratings.

>OUTCOME



For the impact of time elapsed since movie release, there is an almost even split on the direction of this impact, with a slightly negative skew, suggesting that monthly ratings decrease slightly over time. This distribution also suggests that impact may vary across movies.

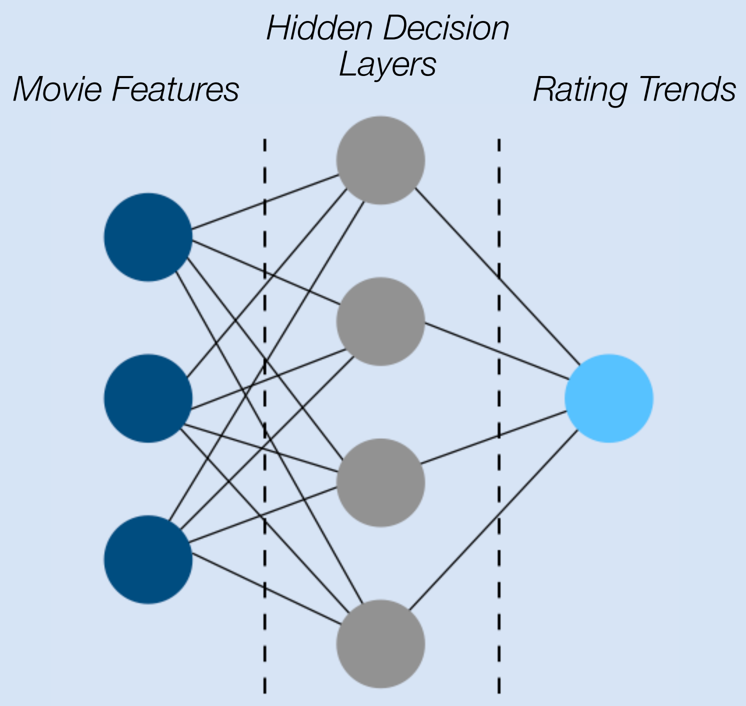


For the impact of user rating profile, the coefficient is generally positive, suggesting that the rating of most movies are in line with people's expectations. We can also see that there is an uneven distribution near 1.0, suggests

>DISCUSSION

Both of our predictors have high statistical significance across our movies dataset. In particular, the

MODEL 2: DO MOVIE FEATURES DETERMINE THEIR RATING TRENDS OVER TIME?



Simplified illustration of our artificial neural network, which we use to predict monthly rating trends using movie features

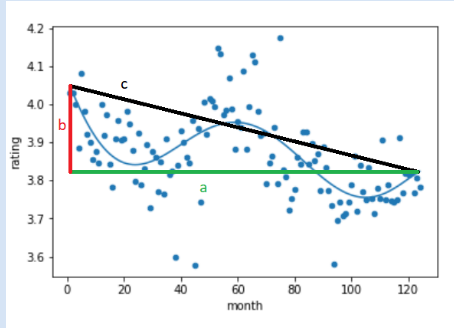
>MODEL

Based on the insights from the first model, we decided to further investigate the impact of time elapsed on monthly movie ratings, this time emphasizing the movie features as a predictor for the monthly rating trends over time.

We selected two particular characteristics in evaluating rating trends over time:

>Total variance in ratings over time insulated against random noise

>Net rating change (magnitude and direction) over a specific time period



>E.g. This is a movie with high variability (compare blue and black lines) and negative net rating change (red line).

Using the various features for our movies, we ran a neural network in an attempt to predict the rating trends based on movie features.

>OUTCOME

>DISCUSSION

Time-Series Trends

Trend

Movie