# Unearth the intrinsic values of movies

Arvind Krishna, Georgia Institute of Technology, Statistics in ISyE,
Haesong Choi, Georgia Institute of Technology, Statistics in ISyE,
Namjoon Suh, Georgia Institute of Technology, Statistics in ISyE,

## 1. ABSTRACT

The “goodness” of a movie is measured by the ratings given to the movie on a variety of reviewing websites. By using this information, we tried to figure out an awesome way to guess the goodness of a movie before it is released in public. We developed the following methodology to accomplish this goal. First, we collected the rating data of 30 movies from nation’s 5 biggest movie rating groups (Rotten tomatoes, Flixter, Metacritic, MRQE, IMDB). Second, we built an EM model to find the bias of each movie rating groups and the “true” value of the movie. Then, we use this true value to develop an elastic net model, which enables us to estimate the “true” value of an unreleased movie. Now, we add the true value of the movie to the bias of the movie-rating group to predict the rating of the movie with respect to the corresponding group. Thus, we provide the users with an estimate of the movie rating according to their preferred rating group even before the movie is released.

## 2. INTRODUCTION

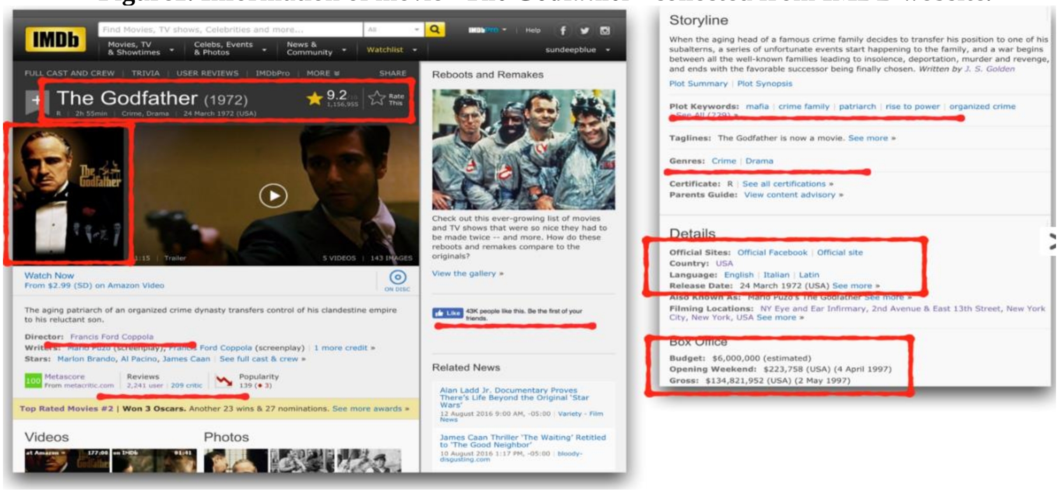
Movies are primarily a source of entertainment. In today’s busy world where people have limited time for entertainment, they should not be disappointed when they take out some time to watch a movie. Public as well as critics rate movies as soon as they released, which helps people to make a choice on the movie they want to watch. However, these ratings take time to stabilize. As soon as the movie releases, people who are film enthusiasts of the particular genre of the movie or the star cast usually rate it. Therefore, it is obvious that they tend to give a higher rating to the movie that skews the movie ratings. Among the critic-rating websites as well, the ratings of the movie keep changing until a large number of critics have rated it. This inspired us to address the problem of predicting the stabilized rating of a movie just before the movie is released.

In order to predict the stabilized rating of a movie on a particular reviewing website, the intuitive way is to build a regression model for each of the reviewing website. However, that will involve many distinct models, each of which will have no connection with each other. Therefore, in order to develop a single coherent prediction model, we split the rating of the movie on each website into the “true value” of the movie and the bias of the rating website. This way we get a single rating for the movie irrespective of the rating website. We use the true value of the movie to build a prediction model based on movie features such as budget, genre, popularity of actors etc. Since this model uses features that are available before the movie is released, we can predict the true value of the movie before its release date. After we obtain the true value from the regression model, we simply add the bias of the rating website to obtain the rating of the movie with respect to the rating website. Thus, users who prefer a particular rating website get to know the rating of the movie on that website even before the movie is released. Users who have no particular preference for a rating website can choose to see the true value of the movie in order to make a choice. Thus, we build a single model, but cater to all user preferences and rating websites.

## 3. DATA Description

We got a data-set from the nation’s biggest data-science website, “Kaggle”. Dataset originally contains more than 5000 movies released from the year 1916 to 2016 in 66 countries. The data scientist who provided this data-set, conducted a scraping by using a Python library called “Scrapy”. Data-set contains many important movie information, scraped from IMDB website. (e.g. movie title, director, name, cast list, genres, name of actors etc). Our purpose is to collect the data from as many movie reviewing websites as possible, so that we can get the bias of each websites and provide audiences with reliably estimated scores of unreleased movies on each websites. We randomly chose 30 out of 5000 movies, which were released since 2012, and collected ratings from 5 nation's biggest movie reviewing websites: Rotten tomatoes (Audiences, Critiques), Flixter, Metacritic, MRQE, IMDB. In order to build a regression model, we needed to introduce dummy variables to categorical variables like color, which indicates whether the movie is black or colored movie, genre, language, country, content_rating. Since there were no directors or actors, who shoot two or more different movies, out of our 30 selected movies, we did not take into account these factors while building regression model. (We provide the listings of 30 movies at the appendix part.)

**Figure1. Information of movie "The Godfather" collected from IMDB website.**




# 4. Methodology and Results
## *4.1 EM Algorithm*
### *4.1.1. Model assumption*

To determine the true value of the movie, and the bias of the reviewers, the EM algorithm was implemented. There were 30*(m)* movies, and 6*(r)* review scores found in each different review websites. $x^{(mr)}$ is denoted as the score that reviewer *r* gave to movie *m*. Each movie is assumed to contain "intrinsic" true value, which is denoted as $y^{(mr)}$ and "bias" of each reviewers is denoted as $Z^{(mr)}$. So, $x^{(mr)}$ is assumed to be combined with the "intrinsic" true value and reviewer bias with all sorts of different random factors influencing the reviewing process. Both $y^{(mr)}$ *and* $Z^{(mr)}$ are assumed to follow normal distribution and to be independent with each other. And we also assume that reviewer's scores are generated by a random process given as follows.

$$y^{(mr)} \sim \mathrm{N}(\mu_m, \sigma_m^2)$$
$$Z^{(mr)} \sim \mathrm{N}(\mu_r, \sigma_r^2)$$
$$x^{(mr)} | \, y^{(mr)}, Z^{(mr)} \sim \mathrm{N}(y^{(mr)} + Z^{(mr)}, \sigma^2)$$

The reviewer's scores ($x^{(mr)}$) are only observed ($\sigma^2$ is 2.5), and the $y^{(mr)}$ and $Z^{(mr)}$ are all latent random variables. $\mu_m, \sigma_m^2, \mu_r, \sigma_r^2$ are estimated with the EM algorithm by maximizing the conditional likelihood of the data, $x^{(mr)}$. According to our model assumption, $\mathrm{p}(x^{(mr)}, y^{(mr)}, Z^{(mr)})$ can be expressed as $p(x^{(mr)} | \, y^{(mr)}, Z^{(mr)}) * p(y^{(mr)}) * p(Z^{(mr)})$. In order to calculate the expectation of conditional log likelihood function, we first need to get joint density function of $y^{(mr)}, Z^{(mr)}$, conditioned on $x^{(mr)}$, denoted as $Q_{mr}$. The joint density function of $x^{(mr)}$, $y^{(mr)}, Z^{(mr)}$ is as follows ,

$$y^{(mr)}, z^{(mr)}, x^{(mr)} \sim \mathrm{N}\left( \begin{bmatrix} \mu_m \\ \mu_r \\ \mu_m + \mu_r \end{bmatrix}, \begin{bmatrix} \sigma_m^2 & 0 & \sigma_m^2 \\ 0 & \sigma_r^2 & \sigma_r^2 \\ \sigma_m^2 & \sigma_r^2 & \sigma^2 + \sigma_r^2 + \sigma_m^2 \end{bmatrix} \right)$$

We attach the detailed derivation steps of EM at the "appendix" of our report. Following is the result of Maximization step.

**[M-step]**

$$\mu_m = \frac{1}{R}\sum_{r=1}^{R}\mu_{mr,Y}$$
$$\mu_r = \frac{1}{M}\sum_{m=1}^{M}\mu_{mr,Z}$$
$$\sigma_p{}^2 = \sum_{r=1}^{R}(\Sigma_{mr,YY} + \mu^2{}_{mr,Y} - 2\mu_{mr,Y}\mu_m + \mu_m^2)$$

$$\sigma_r{}^2 = \sum_{m=1}^{M}(\Sigma_{mr,ZZ} + \mu^2{}_{mr,Z} - 2\mu_{mr,Z}\mu_r + \mu_r^2)$$

### *4.1.2. Analysis & Result*

4.1.2.1. Normality test on $x^{(mr)}$ data

Before implementing the EM algorithm, we need to check if the data we collected follow the model assumption or not. All we need to do is to check whether $x^{(mr)}$ follows normal distribution. For the 180 data we collected, we performed the wilk - shapiro test, and verify the data follow normal. Note that Shapiro-Wilk test's null hypothesis is "Data follow normal". (Here p-value is much larger than 0.05)

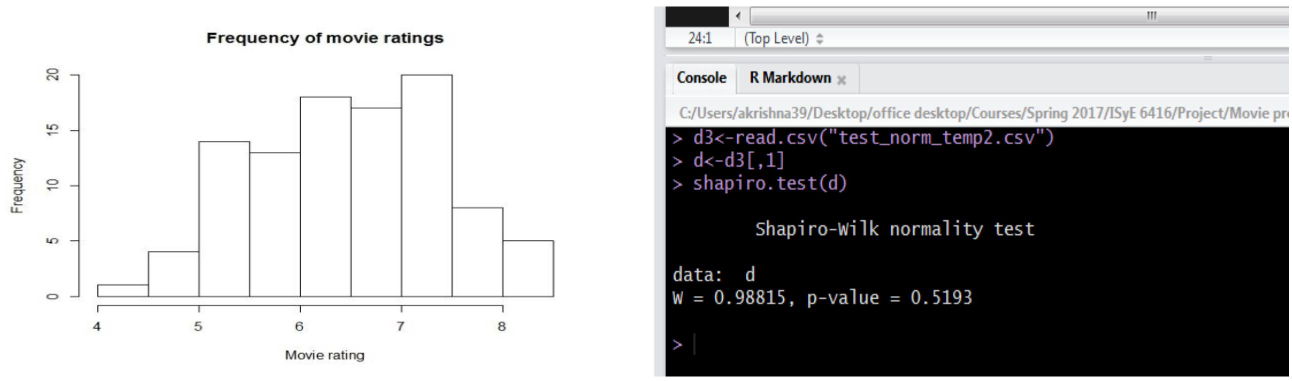



**Figure2. Histogram of 180 $x^{(mr)}$ data and result of wilk-shapiro test**

4.1.2.2. Result of EM algorithm & Analysis of result

After justifying the raw ratings follow normal distribution, we checked the Maximum Likelihood value of expected log-likelihood function converges over iterations. After 72th iteration, ML converges. (ML(i) – ML(i-1) < ε, $\varepsilon < 10^{-6}$, i $\geq$ 72)

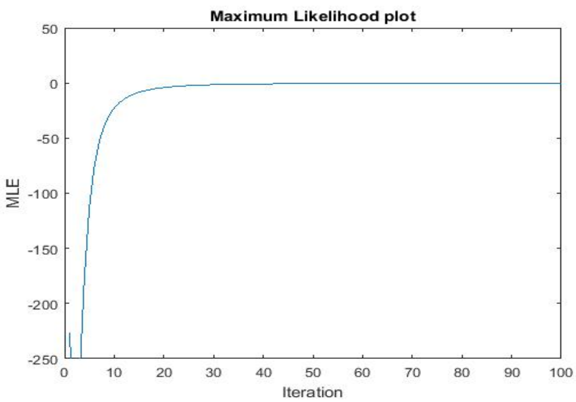
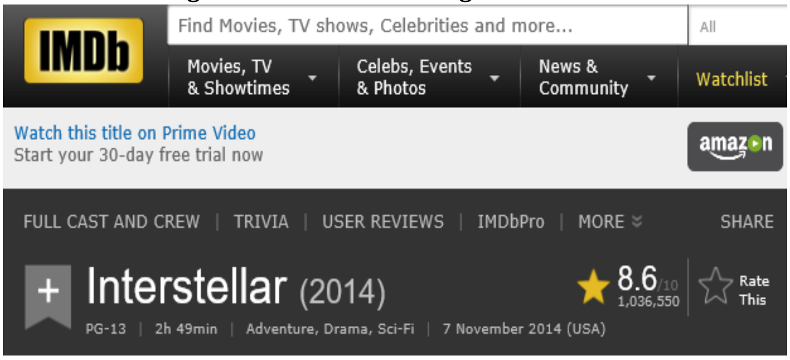



**Figure3. Graph showing the convergence of ML over iterations**

Table 1 shows the true values and the average scores of each 30 movies. 6 reviewing groups' scores are averaged. Confidence intervals of the true values are shown in Column 4 and 5. Here we found an interesting relationship between true values of movies and averaged scores. As you can identify at Figure 1, the gap between true value and averaged score for each movie is invariant as

0.2. And here, we tried to think why this happens: Though it seems like the movie rating on each reviewing group indicates a personal opinion of individual reviewer, it actually represents hundreds to thousands of people's opinion on that movie.

**Figure 4. Number of ratings of Interstellar**




As you can check at here, the total number of people who have rated the "Interstellar" so far is 1,036,550!! We can apply the Law of Large Number (LLN)

$$x^{(mr)} = y^{(mr)} + z^{(mr)} + \sigma_\epsilon$$

$$\frac{1}{R}\sum_{r=1}^{R} x^{(mr)} = \frac{1}{R}\sum_{r=1}^{R} y^{(mr)} + \frac{1}{R}\sum_{r=1}^{R} z^{(mr)} + \frac{1}{R}\sum_{r=1}^{R} \sigma_\epsilon$$

$$\mathrm{E}\left(x^{(mr)}\right) = \mathrm{E}\left(y^{(mr)}\right) + \mathrm{E}\left(z^{(mr)}\right)$$

As you can verify at table 3, the averaged bias of total reviewing group is equal to 0.2. ($\frac{1}{R}\sum_{r=1}^{R}\mu_r = \mathrm{E}\left(z^{(mr)}\right) = 0.2$) This result is consistent with that of the actual data we get through our model. (You can check that the difference between true and averaged value of each movie is always 0.2).

$$x^{(interstellar,\ IMDB)} = \frac{x^{(interstellar,\ 1)} + x^{(interstellar,\ 2)} + \cdots + x^{(interstellar,\ 1{,}036{,}550)}}{1{,}036{,}550}$$

We can explore some interesting facts regarding on bias analysis, as well. Table 3 and Figure 2 represent bias of each reviewing group. As you can see, while the ordinary audiences in Rotten tomato, Fixter, Meta-Critics, MRQE, and IMDB have positive bias, critiques from Rotten Tomato and Meta Critics have relatively more negative bias than the ordinary audiences. The audiences from two biggest movie reviewing groups in US, IMDB and rotten-tomato have greater than 0.5 positive biases, while four other groups have relatively smaller bias, which is almost near to zero.

**Table1** Comparison between the True values, and the Average Scores of the Reviewing Groups

| No. | Average score | True value | Low value | High value |
|---|---|---|---|---|
| 1 | 7.83 | 7.6316 | 7.5227 | 7.7405 |
| 2 | 8.83 | 8.6315 | 8.5235 | 8.7396 |
| 3 | 8.37 | 8.1647 | 8.0582 | 8.2713 |
| 4 | 8.82 | 8.6149 | 8.5021 | 8.7278 |
| 5 | 8.10 | 7.8984 | 7.7907 | 8.006 |
| 6 | 7.88 | 7.6817 | 7.5684 | 7.795 |
| 7 | 8.02 | 7.8149 | 7.7049 | 7.925 |
| 8 | 7.92 | 7.7146 | 7.6068 | 7.8223 |
| 9 | 6.50 | 6.2985 | 6.1795 | 6.4175 |
| 10 | 7.83 | 7.6315 | 7.5219 | 7.7412 |
| ... | … | … | … | … |

**Figure5. Comparison between the True Value and the average Reviewer Score of 5 movies**

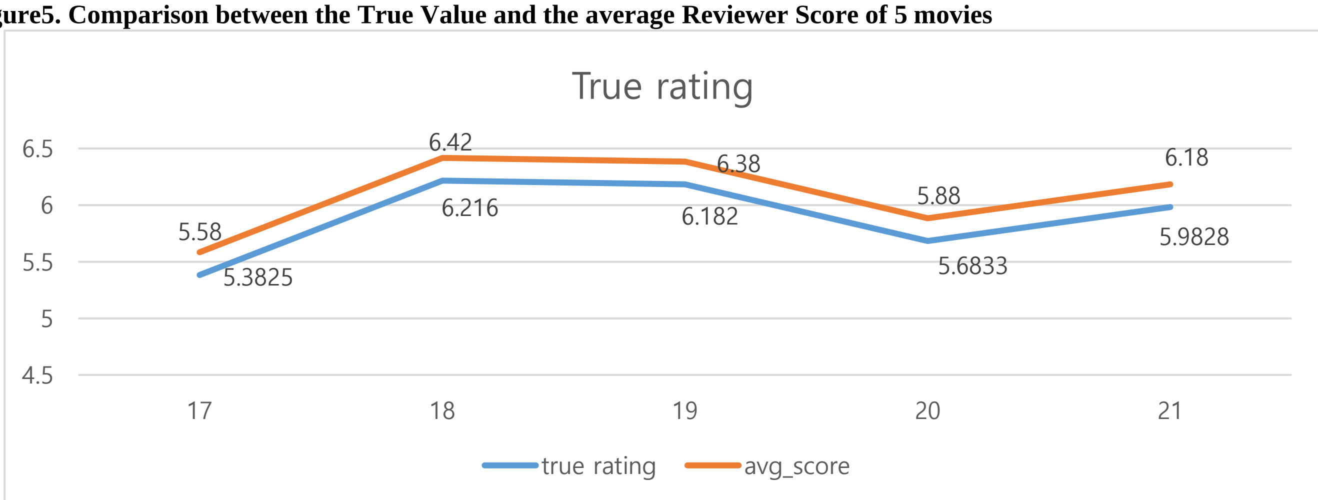
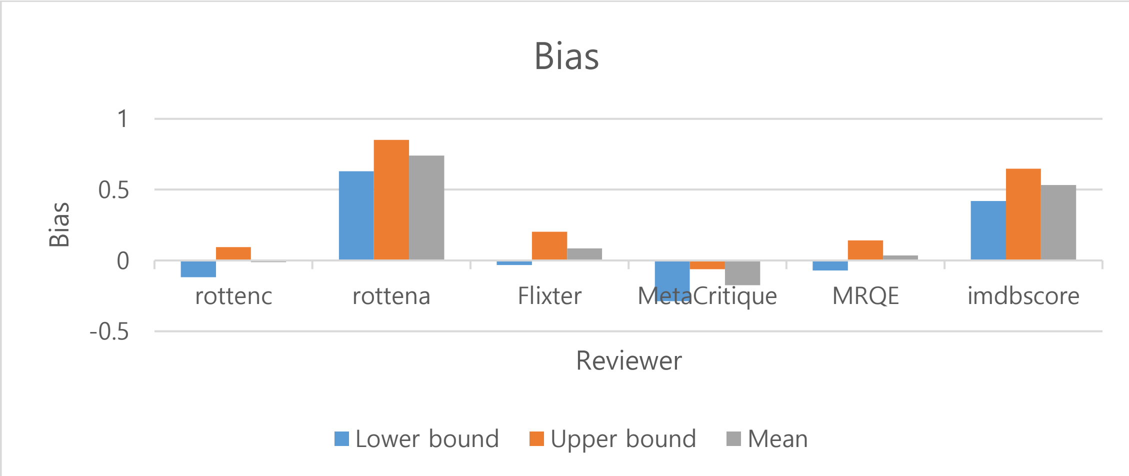


**Table2 Bias of the 6 Reviewer Groups**

| | Rotten Tomatoes_Critics | Rotten Tomatoes_Audience | Flixter | Meta Critics | MRQE | IMDb |
|---|---|---|---|---|---|---|
| Bias-mean | -0.0117 | 0.7403 | 0.0842 | -0.1755 | 0.0354 | 0.5338 |
| Bias_High | -0.1186 | 0.6299 | -0.0333 | -0.2887 | -0.0706 | 0.4207 |
| Bias_Low | 0.0952 | 0.8508 | 0.2017 | -0.0624 | 0.1414 | 0.6469 |

**Figure6. Confidence Interval of Bias of Review Groups in the 6 sites**




*4.2 Predicting the unreleased movies' true value by using Elastic net model.*

*4.2.1. Elastic-net penalty*

With the "intrinsic" values of movie and biases of respective reviewing groups that we found at the previous step, we tried to construct a predictive model using elastic-net penalized regression model. Before jumping into the main part of procedure, let us first explain why we use this special kind of penalty in our model. Though lasso penalty guarantees the model's sparseness, in other words, performs variable selection using the property of L1-norm, it quite often brings forth multi-collinearity issues on the model when covariates of model are linearly correlated with each other. In this case, the lasso estimates tend to exhibit somewhat erratic behavior as the regularization parameter $\lambda$ is varied, whereas the elastic net used to pull out the correlated groups together. We can build an elastic-net penalty by introducing $L_2$ norm of estimates to Lasso type of $L_1$ penalty. Here in our model, there are some variables, which are correlated with each other. Following is the form of elastic-net penalized objective function of regression model.

$$\hat{\beta} = arg\min_{\beta} \frac{1}{2N}\|y - X\beta\|_2^2 + \lambda/2\|\beta\|_1 + (1-\lambda)/2\|\beta\|_2^2$$

*4.3.2. Result and Analysis*

*4.3.2.1. Result*

With data of 30 movies that we already have, we first partitioned the data set into training dataset to build a model and testing dataset to validate the model. (We chose 25 data points randomly out of 30 data points as a training dataset). We performed 10-fold cross-validation to obtain an optimal value of lambda. Figure 3 shows the solution path of non-zero beta coefficient with respect to the values of lambda. Figure 4 represents the Cross-Validation (CV) error curves, showing that 0.02327 is the value of optimal $\lambda$ resulting in smallest mean squared error among all lambdas. Given the model with optimal $\lambda$, Table 5 shows selected variables which are likely to have great impacts on true values of unreleased movies. In Table 5, we reordered the selected variables in an ascending orders, with respect to the absolute values of coefficients, so as to easily identify the most important factors in our model. 16 variables are kept in this Elastic-Net regression.

**Figure 7 - Path of the Elastic-Net Regression Coefficients**

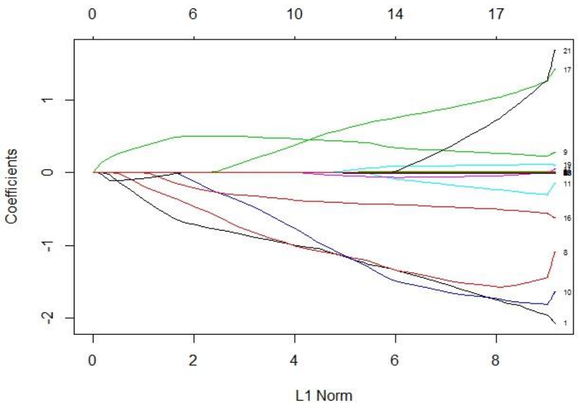



**Figure 8 - Cross-Validation Error Curve**

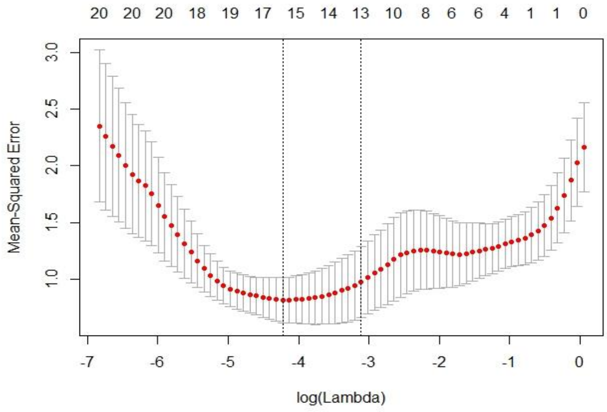



**Table 3 - Fithteen Coefficients Selected using Elastic Net Regression**

| No. | Data_Name | Coefficients |
|---|---|---|
| 10 | Action | -1.66 |
| 1 | Color | -1.55 |
| 8 | Sci_Fi | -1.49 |
| 17 | PG_13 | 0.89 |
| 16 | USA | -0.47 |
| 21 | Aspect_ratio | 0.35 |
| 9 | Drama | 0.30 |
| 11 | Thriller | -0.17 |
| 19 | Title_tear | 0.098 |
| 14 | Face_number_in_poster | -0.057 |
| 3 | Duration | 0.0157 |
| 2 | Num_critic_for_reviews | 0.00976 |
| 4 | actor_2_facebook_likes | -3.99E-05 |
| 20 | actor_1_facebook_likes | -5.21E-09 |
| 18 | budget | -4.27E-06 |

### *4.3.2.2. Analysis*

The top 5 highest absolute coefficient values are as following: Action(-1.66), Color(-1.55), Sci-Fi(-1.49), PG_13(0.89), USA(-0.47). Among the top 5 highest influential factors, except for PG_13, all the other four factors turn out to have negative impacts on the ratings. In contrast with our expectation, Action, Science_Fiction and USA have negative impacts. These days, there are hundreds of thousands of Hollywood movies being released annually, and a huge part of them is Action and/or Science_fiction related movies, and they are even colored!! As expected, factors like face_number_in_poster and actor_facebook_likes have negligible effects on the ratings, which means that the number of facebook likes of actors/actresses does not necessarily reflects the ratings of

| | |
|---|---|
| Drama\|Music | Drama |
| Action\|Thriller | Action |
| Adventure\|Animation\|Comedy\|Drama\|Family\|Fantasy | Drama |
| Action\|Adventure\|Sci-Fi | Action |
| Biography\|Crime\|Drama\|History\|Music | Drama |
| Drama | Drama |
| Comedy\|Fantasy\|Romance | Romance |
| Drama\|Romance | Romance |

**Figure 9. Original "Genre" variable (Left) & modified variable (Right)**

Here, through the result of modeling, we can also identify the effect of elastic-net, which is able to pull out correlated groups of variables. Usually movies being released today cannot be categorized in one specific genre. It can be a fantasy, horror, drama, and sci-Fi at the same time. However, as you can see at Figure 8, in order to build a regression model, we needed to modify this genre variables, so that movie can only have one genre at a time. The left column is the original data and right column is what our team subjectively select among many genres which each movie has. As you can identify at both table 3 and Figure 6, three genres (Sci-Fi, Action, thriller) have all negative and the biggest effects on scores, furthermore, all the Action and Sci-Fi movies are colored movies, which means that variables like Sci-Fi, Thriller, and Color are correlated with each other.

### *4.3. Prediction of the movie rating with respect to reviewing websites.*

Each user prefers some movie-rating website to judge a movie and then decide whether to watch it or not. However, the ratings take time to stabilize. In a user-rating website, the early raters of the movie are mostly the fans of the genre or the cast of the movie and tend to give higher ratings. There is a lot of variation in ratings among the critics of a critic-rating website as well, and so they too take time to stabilize. So, here we solve the problem of estimating the stabilized rating of the movie with respect to each of the major reviewing websites as soon as it is released or even before its release.

We estimate the rating of an unreleased movie as per the following equation.

$$E(\text{Rating})_{\text{Rating-website}} = (\text{True value})_{\text{Movie}} + (\text{Bias})_{\text{Rating-website}}$$

We find $(\text{True value})_{\text{Movie}}$ using the elastic-net model we developed in step 2. We got the bias of the rating-website using the EM algorithm we used in step 1.

We test this approach to estimate the ratings of 5 movies on each of the 6 major reviewing websites. Table 7 compares our estimated ratings with the actual ratings.

| Movie | Rotten tomatoes (A) | | Rotten tomatoes (C) | | Flixter | |
|---|---|---|---|---|---|---|
| | Actual rating | Our estimation | Actual rating | Our estimation | Actual rating | Our estimation |
| Whiplash | 8.6 | 8.2 | 9 | 8.9 | 9.4 | 8.3 |
| Mission: Impossible - Rogue Nation | 7.5 | 7.5 | 8.2 | 8.3 | 8.7 | 7.6 |
| The Finest Hours | 6.1 | 5.9 | 7.2 | 6.7 | 6.6 | 6.0 |
| Magic Mike | 5.9 | 5.6 | 6.8 | 6.4 | 5.6 | 5.7 |
| Oculus | 6.5 | 6.0 | 6.4 | 6.7 | 5.3 | 6.0 |

| Movie | Meta Critic | | MRQE | | IMDB | |
|---|---|---|---|---|---|---|
| | Actual rating | Our estimation | Actual rating | Our estimation | Actual rating | Our estimation |
| Whiplash | 8.8 | 8.0 | 8.5 | 8.2 | 8.5 | 8.7 |
| Mission: Impossible - Rogue Nation | 7.5 | 7.4 | 7.4 | 7.6 | 7.4 | 8.0 |
| The Finest Hours | 5.8 | 5.7 | 6.8 | 6.0 | 6.8 | 6.4 |
| Magic Mike | 6.0 | 5.5 | 5.7 | 5.7 | 5.7 | 6.1 |
| Oculus | 6.1 | 5.8 | 6.5 | 6.0 | 6.5 | 6.5 |

**Table 3: Actual ratings vs estimated ratings**

We observe that our estimated ratings are close to the actual ratings on each of the reviewing websites. The average root mean square error in our estimation is 0.47. Thus, our estimated rating is less that a unit away from the actual rating around 95% of the times!

## 5. CONCLUSION AND DISCUSSION

We conclude the following from this project:

a) We infer that the audience is positively biased and tends to give a higher rating than critics do. This is evident from the high positive bias if IMDB and Rotten Tomatoes (Audience score) and negative bias of Rotten Tomatoes (Critic score) and Metacritic

b) The “true rating” of the movie is proportional to the average rating of all the reviewing websites which is consistent with our modeling assumptions

c) We estimate the true value of a movie from the elastic net model. Since the values of the covariates in the elastic net model are available before the movie is released and we already know the bias of each reviewing website, we can predict the rating of any movie on each of the nation’s top 5 reviewing website even before the movie is released

< APPENDIX 1.>
Derivation of Expectation step and Maximization step.
The conditional joint density function of $y^{(mr)}, Z^{(mr)}$ can be obtained as follows :

$$Q_{mr}\left(y^{(mr)}, Z^{(mr)} \mid x^{(mr)}\right) \sim N\left( \begin{bmatrix} \mu_{mr,\ Y} \\ \mu_{mr,\ Z} \end{bmatrix}, \begin{bmatrix} \Sigma_{mr,YY} & \Sigma_{mr,YZ} \\ \Sigma_{mr,ZY} & \Sigma_{mr,ZZ} \end{bmatrix} \right)$$

$$\mu_{mr} = \begin{bmatrix} \mu_{mr,\ Y} \\ \mu_{mr,\ Z} \end{bmatrix} = \begin{bmatrix} \mu_m + \frac{\sigma_m^2}{\sigma_m^2+\sigma_r^2+\sigma^2}\left(x^{(mr)} - \mu_m - \mu_r\right) \\ \mu_r + \frac{\sigma_r^2}{\sigma_m^2+\sigma_r^2+\sigma^2}\left(x^{(mr)} - \mu_m - \mu_r\right) \end{bmatrix},$$

$$\Sigma_{mr} = \begin{bmatrix} \Sigma_{mr,YY} & \Sigma_{mr,YZ} \\ \Sigma_{mr,ZY} & \Sigma_{mr,ZZ} \end{bmatrix} = \frac{1}{\sigma_m^2+\sigma_r^2+\sigma^2} \begin{bmatrix} \sigma_m^2(\sigma_r^2+\sigma^2) & -\sigma_r^2\sigma_m^2 \\ -\sigma_r^2\sigma_m^2 & \sigma_r^2(\sigma_m^2+\sigma^2) \end{bmatrix}$$

## 1) E-step

$$\theta = \arg\max_{\theta} \sum_{m=1}^{M}\sum_{r=1}^{R} E_Q[\log P\left(x^{(mr)}, y^{(mr)}, Z^{(mr)};\ \theta\right) \mid x^{(mr)}, \theta^{(old)}]$$

$$= \arg\max_{\theta} \sum_{m=1}^{M}\sum_{r=1}^{R} E_Q[\log[\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x^{(mr)}-y^{(mr)}-Z^{(mr)})^2} \frac{1}{\sqrt{2\pi}\sigma_m} e^{-\frac{1}{2\sigma_m^2}(y^{(mr)}-\mu_m)^2} \frac{1}{\sqrt{2\pi}\sigma_r} e^{-\frac{1}{2\sigma_r^2}(z^{(mr)}-\mu_r)^2} \mid x^{(mr)}, \theta^{(old)}]$$

$$= \arg\max_{\theta} \sum_{m=1}^{M}\sum_{r=1}^{R} E_Q[\log\frac{1}{(2\pi)^{3/2}\sigma\sigma_m\sigma_r} - \frac{1}{2\sigma^2}\left(x^{(mr)} - y^{(mr)} - Z^{(mr)}\right)^2 - \frac{1}{2\sigma_m^2}\left(y^{(mr)} - \mu_m\right)^2 - \frac{1}{2\sigma_r^2}\left(z^{(mr)} - \mu_r\right)^2]$$

$$= \arg\max_{\theta} \sum_{m=1}^{M}\sum_{r=1}^{R} E_Q[\log\frac{1}{\sigma_m\sigma_r} - \frac{1}{2\sigma_m^2}\left(y^{(mr)} - \mu_m\right)^2 - \frac{1}{2\sigma_r^2}\left(z^{(mr)} - \mu_r\right)^2]$$

$$= \arg\max_{\theta} \sum_{m=1}^{M}\sum_{r=1}^{R} E_Q[\log\frac{1}{\sigma_m\sigma_r} - \frac{1}{2\sigma_m^2}\left((y^{(mr)})^2 - 2y^{(mr)}\mu_m + \mu_m^2\right) - \frac{1}{2\sigma_r^2}\left((z^{(mr)})^2 - 2z^{(mr)}\mu_r + \mu_r^2\right)]$$

$$= \arg\max_{\theta} \sum_{m=1}^{M}\sum_{r=1}^{R} \log\frac{1}{\sigma_m\sigma_r} - \frac{1}{2\sigma_m^2}\left(E_Q[(y^{(mr)})^2] - 2E_Q[y^{(mr)}]\mu_m + \mu_m^2\right) - \frac{1}{2\sigma_r^2}\left(E_Q[(z^{(mr)})^2] - 2E_Q[z^{(mr)}]\mu_r + \mu_r^2\right)$$

$$= \arg\max_{\theta} \sum_{m=1}^{M}\sum_{r=1}^{R} \log\frac{1}{\sigma_m\sigma_r} - \frac{1}{2\sigma_m^2}\left(\Sigma_{mr,YY} + \mu^2{}_{mr,Y} - 2\mu_{mr,\ Y}\mu_m + \mu_m^2\right) - \frac{1}{2\sigma_r^2}\left(\Sigma_{mr,ZZ} + \mu^2{}_{mr,z} - 2\mu_{mr,z}\mu_r + \mu_r^2\right)$$

where $E_Q\left[(y^{(mr)})^2\right] = \Sigma_{mr,YY} + \mu^2{}_{mr,Y}$ , $E_Q\left[y^{(mr)}\right] = \mu_{mr,\ Y}$, $E_Q\left[(Z^{(mr)})^2\right] = \Sigma_{mr,ZZ} + \mu^2{}_{mr,Z}$ , $E_Q\left[Z^{(mr)}\right] = \mu_{mr,Z}$

## 2) M-step

Setting derivatives w. r. t parameters $\mu_m$ , $\sigma_m^2$ , $\mu_r$ , $\sigma_r^2$ to 0

$$-\frac{1}{2\sigma_m^2}\sum_{r=1}^{R}(2\,\mu_m - 2\mu_{mr,Y}) = 0 \implies \mu_m = \frac{1}{R}\sum_{r=1}^{R}\mu_{mr,Y}$$

$$-\frac{1}{2\sigma_r^2}\sum_{m=1}^{M}(2\,\mu_r - 2\mu_{mr,Z}) = 0 \implies \mu_r = \frac{1}{M}\sum_{m=1}^{M}\mu_{mr,Z}$$

$$\sum_{r=1}^{R}[-\frac{1}{\sigma_p} - \frac{1}{\sigma_m^3}\left(\Sigma_{mr,YY} + \mu^2{}_{mr,Y} - 2\mu_{mr,\ Y}\mu_m + \mu_m^2\right)] = 0 \implies \sigma_m{}^2 = \sum_{r=1}^{R}\left(\Sigma_{mr,YY} + \mu^2{}_{mr,Y} - 2\mu_{mr,Y}\mu_m + \mu_m^2\right)$$

$$\sum_{m=1}^{M}[-\frac{1}{\sigma_r} - \frac{1}{\sigma_r^3}\left(\Sigma_{mr,ZZ} + \mu^2{}_{mr,Z} - 2\mu_{mr,\ Z}\mu_r + \mu_r^2\right)] = 0 \implies \sigma_r{}^2 = \sum_{m=1}^{M}\left(\Sigma_{mr,ZZ} + \mu^2{}_{mr,Z} - 2\mu_{mr,Z}\mu_r + \mu_r^2\right)$$

< Appendix 2.>

movie_title

1. Interstellar
2. Whiplash
3. The Dark Knight Rises
4. Inside Out
5. Captain America: Civil War
6. Straight Outta Compton
7. The Great Beauty
8. Midnight in Paris
9. Me Before You
10. Mission: Impossible - Rogue Nation
11. Ida
12. Frances Ha
13. St. Vincent
14. The Place Beyond the Pines
15. Blue Jasmine
16. The Neon Demon
17. Now You See Me 2
18. The Finest Hours
19. The Shallows
20. Money Monster
21. Oculus
22. Our Kind of Traitor
23. Teenage Mutant Ninja Turtles: Out of the Shadows
24. Deliver Us from Evil
25. Noah
26. Magic Mike XXL
27. Pixels
28. Point Break
29. Aloft
30. The Boy Next Door