Forecasting Financial Success of Hollywood Movies
A Comparative Analysis of Machine Learning Methods

Dursun Delen[1] and Ramesh Sharda[2]
[1]*Spears School of Business, Oklahoma State University, Tulsa, Oklahoma, U.S.A.*
[2]*Spears School of Business, Oklahoma State University, Stillwater, Oklahoma, U.S.A.*

Keywords: Prediction, Box-office Receipts, Hollywood, Machine Learning, Neural Networks, Sensitivity Analysis.

Abstract: Forecasting financial success of a particular movie has intrigued many scholars and industry leaders as a worthy but challenging problem. In this study, we explore the use of machine learning methods to forecast the financial performance of a movie at the box-office before its theatrical release. In our models, we convert the forecasting problem into a multinomial classification problem—rather than forecasting the point estimate of box-office receipts; we classify a movie based on its box-office receipts in one of nine categories, ranging from a "flop" to a "blockbuster." Herein, we present our comparative prediction results along with variable importance measures (using sensitivity analysis on trained prediction models).

1 INTRODUCTION

Forecasting box-office receipts of a particular motion picture has intrigued many scholars and industry leaders as a difficult and challenging problem. To some ana-lysts, Hollywood is the "land of hunch and the wild guess" (Litman and Ahn, 1998) due largely to the difficulty and uncertainty associated with predicting the product demand. Such unpredictability of the product demand makes the movie business one of the riskiest endeavors for investors in today's economy. In support of such observations, Jack Valenti, former president and CEO of the Motion Picture Association of America, once said "… No one can tell you how a movie is going to do in the marketplace… not until the film opens in darkened theatre and sparks fly up between the screen and the audience" (Valenti, 1978). Trade journals and magazines of the motion picture industry have been full of examples, statements, and experiences that support such a claim.

Despite the difficulty associated with the unpredictable nature of the problem domain, many researchers have attempted to develop models for forecasting the financial success of motion pictures, primarily using statistics-based forecasting approaches. Most analysts have tried to predict the total box-office receipt of motion pictures after a movie's initial theatrical release. However, most (Litman, 1983); (Sawhney and Eliashberg, 1996) did not get sufficiently accurate results to be used as decision aid. Litman and Ahn (1998) summarizes and compares some of the major studies on predicting financial success of motion pictures. Yet, these previous studies leave us with an unsatisfied need for a more accurate forecasting method, especially prior to a movie's theatrical release. Most studies indicate that box-office receipts tend to tail-off after the opening week. Research shows that 25 percent of total revenue of a motion picture comes from the first two weeks of receipts (Litman and Ahn, 1998). Thus, once the first week of box-office receipts are determined, the total box-office receipts of a particular movie can be forecasted with very high accuracy (Sawhney and Eliashberg, 1996). Therefore, the accurate estimate of the box-office receipts of motion pictures before its theatrical release is the most difficult and the most critical to the industry.

In this study, we explore the use of machine learning techniques, especially neural networks and decision trees, in forecasting the financial performance of a movie at the box-office before its theatrical release. In our models, we convert the forecasting problem into a classification problem. That is, rather than forecasting the point esti-mate of box-office receipts, we classify a movie based on its box-office receipts in one of nine categories, ranging from a "flop" to a "blockbuster."

The remainder of this paper is organized as

Delen D. and Sharda R..
Forecasting Financial Success of Hollywood Movies - A Comparative Analysis of Machine Learning Methods.
DOI: 10.5220/0004125006530656
In *Proceedings of the 9th International Conference on Informatics in Control, Automation and Robotics* (ANNIIP-2012), pages 653-656
ISBN: 978-989-8565-21-1

follows. The next section briefly reviews the literature on forecasting the box office success of theatrical movies. Section three provides the details of our methodology by specifically talking about the data, the model types, the experimental design used in this study. Next, the prediction results are presented and briefly explained. The last section of the paper discusses the overall contribution of this study along with its limitations and further research directions.

2 LITERATURE REVIEW

Literature on forecasting financial success of new motion pictures can be classified based on the type of forecasting model employed: (i) Econometric/Quantitative Models—those that explore factors that influence the box office receipts of newly released movies (Litman, 1983); (Litman and Kohl, 1989); (Sochay, 1994); (Litman and Ahn, 1998); (Elberse and Eliashberg, 2002), and (ii) Behavioral Models—those that primarily focuses on the individual's decision making process with respect to selecting a specific movie from a vast array of entertainment alternatives (Eliashberg and Sawhney, 1994); (Sawhney and Eliashberg, 1996); (Zufryden, 1996); (De Silva, 1998), (Eliashberg et al., 2000). These behavioral models usually employ a hierarchical framework where behavioral traits of consumers are combined (mostly in a sequential process) with the econometric factors in developing the forecasting models. Another classification is based on the timing of the forecast: (i) Before the Initial Release—that is forecasting the financial success of the movies before their initial theatrical release (Litman, 1983); (Litman and Kohl, 1989); (Sochay, 1994); (Zufryden, 1996); (De Silva, 1998); (Eliashberg et al., 2000), (ii) After the Initial Release—that is forecasting the financial success of the movies after their initial theatrical release where the first week of receipts are known (Sawhney and Eliashberg, 1996); (Ravid, 1999). Forecasting models that fall into the category of "after the initial release" tend to generate more accurate forecasting results due to the fact that those models have more explanatory variables including box-office receipts from the first week of viewership, movie critics, and word-of-mouth effects. Our study falls into the category of quantitative models for model type classification, and into the category of before the initial release in timing of the forecast classification. Following is a chronological review of the most relevant and the most cited literature published in the field of forecasting financial success of theatrical movies.

3 RESEARCH METHODOLOGY

In this section, we briefly explain (1) the nature of data SET used for the experimentations, (2) the machine learning methods selected and used, (3) the experimentation methodology utilized, and (4) the performance metrics used for prediction accuracy.

3.1 The Data

In our study, we used 386 movies released between 2009 and 2010. The sample data was drawn (partially purchased) from IMBD.com, ShowBiz Data Inc., among others. The dependent variable in our study is the box-office gross revenues, not including auxiliary revenues such as video rentals, international market revenues, toy and soundtrack sales, etc. Another important difference between our study and previous efforts is that we convert the forecasting problem into a classification problem. Rather than forecasting the exact amount of the dependent variable (box-office receipts), we classify a movie based on its box-office receipts in one of nine categories, ranging from a "flop" to a "blockbuster." This process of converting a continuous variable in a limited number of classes is commonly called in literature as "discretization" or "binning." In this study, we discretized the dependent variable into nine classes using the following breakpoints. These breakpoints are determined largely based on our consultations with several decision makers in the movie business.

Class No	1	2	3	4	5	6	7	8	9
Range (in Millions)	< 1 (Flop)	> 1 < 10	> 10 < 20	> 20 < 40	> 40 < 65	> 65 < 100	> 100 < 150	> 150 < 200	> 200 (Blockbuster)

We used a large number of independent variables. Our choice of independent variables is based partially on the previous studies conducted in the field. Each independent categorical variable is converted into an appropriate representation, which created a number of pseudo variables increasing the independent variable count.

3.2 The Machine Learning Methods Used

In this study, three most popular classification methods are used (and compared to each other): decision trees, artificial neural networks and logistic regression. These prediction methods are selected because of their superior capability of modeling classification type prediction problems and their popularity in recently published data mining literature. What follows is a brief description of these modeling techniques.

3.3 Experimental Design

In order to minimize the bias associated with the random sampling of the training and holdout data samples in comparing the predictive accuracy of two or more methods, researchers tend to use k-fold cross-validation. Figure 1 illustrated the k-fold cross validation for k having the value of 10 (which is a commonly practiced rule of thumb in comparative analyses of multiple prediction models).

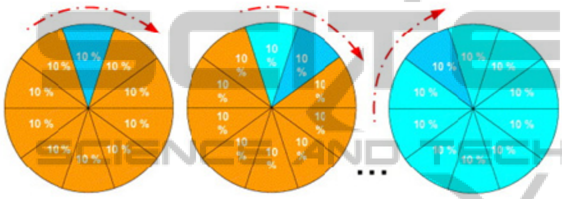

Figure 1: Depiction of *10*-fold cross-validation.

3.4 Performance Metrics

We used percent success rate to measure the predictive performance In our case, we have two different success rates: bingo (which measures the exact classification into the same class and the within one class) and 1-Away (which includes the neighboring classes as success). Algebraically, APHR can be formulated as shown in Eq. 1 and 2.

$$APHR_{Bingo} = \frac{1}{n}\sum_{i=1}^{g} p_i \quad (1)$$

$$APHR_{1-Away} = \frac{1}{n}\left((p_1 + p_2) + \sum_{i=2}^{g-1} p_{i-1} + p_i + p_{i+1} + (p_{g-1} + p_g)\right) \quad (2)$$

where, g is the total number of classes (= 9), n is the total number of samples (= 386), and p_i is the total number of samples classified as class i.

4 RESULTS

Table 1 shows the prediction results of all three machine learning methods as well as the results of the ensemble models. These results are obtained using a 10-fold cross validation methodology. The first performance measure is the percent correct classification rate, which we have called "bingo". We also report the 1-Away correct classification rate. As can be seen artificial neural networks performed the best among the individual prediction models, followed by decision trees and multinomial logistic regression. Ensemble models are developed using simple voting on already trained model types. In general, the ensemble models performed as good as the best individual prediction models. What is probably more important to decision makers is the significantly low standard deviation one could obtain from the ensembles compared to the individual models. Empirically proven by Seni and Elder (2010) that if done correctly ensembles produce more robust prediction outcomes.

Table 1: Tabulated prediction results for all model types.

	PREDICTION MODELS			
Performance Measure	ANN	DT	LR	Ensemble
Count (bingo)	203	186	161	198
Count (1-away)	334	317	293	327
Accuracy (% bingo)	52.59	48.19	41.71	51.30
Accuracy (% 1-away)	86.53	82.12	75.91	84.72

In the process of performing sensitivity analysis, the neural network learning is disabled so that the network weights are not affected. The basic idea is that the inputs to the network are perturbed slightly, and the corresponding change in the output is reported as a percentage change in the output (Principe et al., 2000). The sensitivity analysis results are summarized and presented as a column plot in Figure 2.

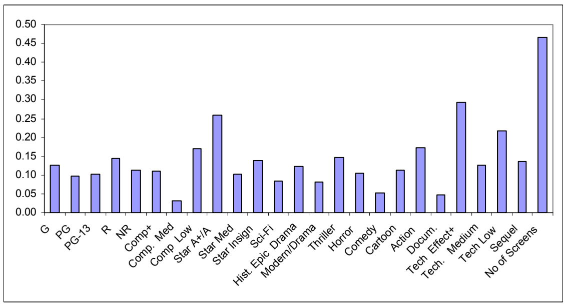

Figure 2: Sensitivity analysis results for all variables.

5 CONCLUSIONS AND DISCUSSION

Even though it is hard to objectively compare (because of the use of different data sets, different variables, different metrics), to the best of our knowledge these prediction results are better than any reported in the published literature for this problem domain. Beyond the attractive accuracy of our prediction results of box-office receipts, models could also be used to forecast the success rates of other media products. The particular parameters used within the model of a movie or other media products could be altered using the already trained prediction models in order to better understand the impact of different parameters on the end results. During this experimentation process, the decision maker of a given entertainment firm could find out, with a fairly high accuracy level, how much a specific actor, a specific release date, or the addition of more technical effects, mean to the financial success of a film.

The accuracy of the data mining models presented in this study can be improved by adding some of the other determinant variables such as production budget and advertising budget, which are known to be industry secrets and are not publicly released. Another method to improve the predictive accuracy of a system is through more sophisticated ensemble models (combining multiple classifiers into a single predictive model by considering their historical accuracy levels).

REFERENCES

Eliashberg, J. and Sawhney, M. S.: Modeling Goes to Hollywood: Predicting Individual Differences in Movie Enjoyment, *Management Science* 40(9), (1994) 1151-1173.

Eliashberg, J., Junker, J. J., Sawhney, M. S. and Wierenga B.: MOVIEMOD: An Implementable Decision Support System for Prerelease Market Evaluation of Motion Pictures. *Marketing Science*, 19(3), (2000) 226-243.

Litman B. R. and Ahn, H.: Predicting Financial Success of Motion Pictures. In *The Motion Picture Mega-Industry* by B. Litman. Allyn & Bacon Publishing, Boston, MA. (1998).

Litman, B. R.: Predicting Success of Theatrical Movies: An Empirical Study, *Journal of Popular Culture*, 16(9), (1983) 159-175.

Principe, J. C., Euliano, N. R. and Lefebre, W. C.: *Neural and Adaptive Systems: Fundamentals Through Simulations*. New York: John Wiley and Sons (2000)

Ravid, S. A.: Information, Blockbusters, and Stars: A Study of the Film Industry, *Journal of Business*, 72(4), (1999) 463-492.

Sawhney, M. S. and Eliashberg, J.: A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures, *Marketing Science*, 15(2), (1996) 113-131.

Seni, G. and Elder, J.: *Ensemble Methods in Data Mining: Improving Accuracy through Combining Predictions*. Morgan & Claypool Publisher (2010).

Sochay, S.: Predicting the performance of motion pictures, *The Journal of Media Economics* 7(4), 1-20.

Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. *Int. J. Digit. Libr.* 1 (1994) 108–121.

Valenti, J.: Motion Pictures and Their Impact on Society in the Year 2000. *Speech given at the Midwest Research Institute*, Kansas City, April 25, (1978) 7.

Zufryden, F. S., Linking Advertising to Box Office Performance of New Film Releases: A Marketing Planning Model," *Journal of Advertising Research*, July-Aug. (1996) 29-41.