A Two-Stage Machine Learning Approach to Forecast the Lifetime of Movies in a Multiplex

Abhijith Ragav[1], Sai Vishwanath Venkatesh[1], Ramanathan Murugappan[2], and Vineeth Vijayaraghavan[3]

[1] SRM Institute of Science and Technology, Chennai, India
{abhijithragav, saivishwanathv}@ieee.org
[2] Madras Institute Of Technology, Chennai, India
ramanathanmurugappan@ieee.org
[3] Solarillion Foundation, Chennai, India
vineethv@ieee.org

Abstract. Collecting over $2.1 billion annually, the cinema exhibition industry contributes 55% of the total revenue towards the Indian film industry. Selection of films is one of the most economically crucial decisions in cinema exhibition. Film selection is incredibly complicated to execute in India owing to its diverse demographic across regions and the resulting behavioral complexity. Working with data from one of India's leading multiplexes, the authors offer a two-stage solution using machine learning to predict if a movie would proceed to be screened in the following week and the number of weeks it would continue to be screened if it does. The estimation of a movie's lifetime helps exhibitors to make intelligent negotiations with distributors regarding screening and scheduling. The authors introduce a new metric *MLE* to evaluate the error in predicting the remaining lifetime of a film. The approach proposed in this paper surpasses the existing system of lifetime prediction and consequent selection of movies, which is currently performed based on intuition and heuristics.

Keywords: Machine learning, Feature Engineering, Movie lifetime forecasting, Film Industry

1 Introduction

The media and entertainment sector in India represents approximately 1% of the country's GDP with a $15.6 billion value estimate. Amongst this, the Indian film industry grosses $3.8 billion and is constantly growing with a compound annual growth rate of over 10% in the past years. This paper focuses on the most economically valuable subset of the Indian film Industry - Cinema exhibition. According to a Deloitte report on the economic contribution of the film industry in the year 2017 [1], more than 50% of the gross share of the film industry is attributed to cinema exhibition. Cinema exhibition in India is carried out either through multiple screen theatres (multiplexes) or single screen theatres.

Organizing cinema exhibitions through multiplexes is an arduous task, especially in India due to the added complexity of the market. Cinema Exhibition in India deals with more than 2500 films produced nationally every year across 20 languages in different regions, apart from an average of 700 films produced internationally that are also screened. In contrast to the domestic box office of US and Canada which majorly consists of Hollywood, the Indian domestic box office is segmented based on region and language [1] with the majority of the box office revenue collected by Bollywood (Hindi Language), Kollywood (Tamil Language) and Tollywood (Telugu Language). Box office revenue collected by Bollywood accounts for 34% of the total revenue while Kollywood and Tollywood account for 15% and 13% respectively. In India, cinema exhibition is carried out through more than 2000 multiplexes across the country and 6000 scattered single screen theatres with factors that differ extensively from region to region. Typically, each multiplex will have to cater to movies from at least 3 distinct languages in addition to movies from languages specific to the multiplex's region.

According to Jehoshua Eliashber et al [2], whose research work is based off a multiplex in the Netherlands, the programming problem faced by a multiplex can consist of 2 stages, (i) the selection of movies to be screened and (ii) the scheduling of these movies over screens, days and times of the day. Conventionally (i) is handled by an expert at the multiplex on a weekly basis. It is one of the most economically valuable stages in a multiplex since the dividends of the ticket revenue of selected films are shared by both the exhibitors and the distributors. The factors that affect the expert's decision include the distributor's pitch for the film, analysis, intuition, internal policies, previous occupancy statistics, and existing exhibitor-distributor agreements. The expert consolidates this information to construct the projected lifetime of a film. Loss in revenue can be attributed to underestimating the lifetime of a movie and/or scheduling movies that won't provide maximum returns for a week. This also attributes to loss in potential profit from movies that could perform better but fail to be selected for scheduling in the week. This emphasizes the need for accurate lifetime estimation in multiplexes.

This paper focuses on one of the leading multiplexes operating at a metropolitan city in India. The multiplex under consideration schedules movies across the screens on a weekly basis. The authors of this paper define the *lifetime* of a movie to be the total number of weeks the movie is screened within a multiplex since its initial screening. The authors offer a two-stage solution using machine learning to predict if a movie would proceed to be screened the following week and the number of weeks it would continue to be screened if it does.

This paper is organized as follows: Section 2 discusses the relevant research done across the Film Industry. The dataset under consideration is briefly described in Section 3 and further explored and analyzed in Section 4, based on which intuitive features are engineered. Section 5 highlights the training methodology and approaches employed in this paper while Section 6 compares the efficiencies of the approaches discussed.

2 Related Works

The Film Industry has been an area of active research for several years. Ajay Shiva Santhosh Reddy et al [3] deal with box office performance of movies based on Hype Analysis using Twitter. Hype is calculated using information pertaining to the number of tweets relating to the movie, total number tweets from distinct users and the follower count for each user. Box-office collection is predicted by multiplying the hype factor with the number of shows screened during the first weekend. Ramesh Sharda and Dursun Delen [4] in one of their works have dealt with the prediction of how successful a movie turns out to be. The target variable has been divided into 9 classes, ranging from 'flop' to 'blockbuster' based on the movie's box-office receipts. The classification problem is tackled using a Neural Network architecture with features including presence of a star actor, the genre of the movie, number of screens allocated to the movie etc. Sameer Ranjan Jaiswal and Divyansh Sharma [5] have come up with a similar model specifically targeting Bollywood movies. They utilize a feature named 'music score', a characteristic factor for Bollywood movies that greatly improves the performance of the model. Andrew Ainslie et al [6] propose an interesting concept in which they analyze box-office sales in the context of a market share model. They claim that the number of screens allotted during the opening week are overestimated in traditional models. The work also specifies that the actors have a direct effect with the customers' movie choice while the director has an indirect effect.

Majority of the research work revolving around movies and theatres focus on predicting whether a movie is successful based on box office revenue. While most of the authors consider external factors such as tweets and market share, they do not consider local behavioral factors such as the behavior of the crowd, the operational pattern of the multiplex and seasonal characteristics. Each state in India has a diverse demographic with varying regional languages. Consequently, movie preferences change vastly from state to state. Hence, behavioral factors are crucial in analyzing the success of a movie in the locality of the multiplex.

The authors of this paper predict the lifetime of movies as a measure of success using local behavioral factors as illustrated in Section 4. Moreover, accurate estimation of the lifetime of movies within a multiplex helps the exhibitors maximize profits by making smarter negotiations with distributors. Additionally, selecting movies that would continue to screen next week helps the multiplex in scheduling.

3 Dataset

Operating 17 multiplexes across 10 cities, each with different cultural, language and ethnic backgrounds, the multiplex under consideration is one of India's leading cinema exhibitors. The dataset consists of over 15 million records and has been previously used for food sales forecasting [7]. Its dataset consists of information of transactions from 2015-2017 pertaining to movie ticket purchases from one of the 17 diverse multiplexes. A *show* is defined by the authors as a film being played on screen at a unique time. The fields considered from the dataset are illustrated in Table 1.

Table 1. Transaction Data

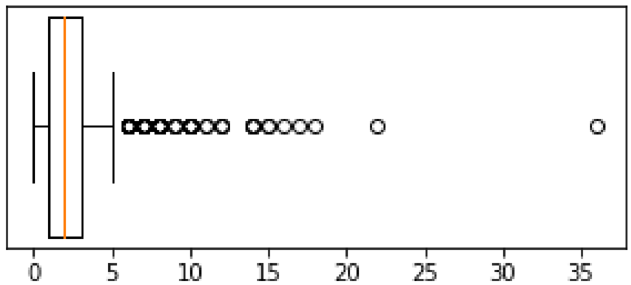
Fields	**Description**
Film String Code	Unique Identification Key for Film
Screen name	Unique name assigned to screen
Session date-time	Date and time of film Screening
Session seats reserved	Number of seats reserved by the multiplex for every show
Show number	The corresponding time slot of the show considered in the day
Transaction value	The payment amount remitted by the customer for the transaction
Transaction datetime	Date and time of transaction
Seats per transaction	The number of seats sold in the transaction
Transaction ID	Unique key of identification for each transaction

The multiplex offers special screenings and arrangements for films to screen exclusively for a maximum of 4 days. For this reason, the authors only consider movies with at least 5 days of screening. The transactional data also labels tickets that are cancelled by the multiplex due to unforeseen circumstances. Hence to maintain the integrity of the dataset, the cancelled tickets are removed for further analysis.

Since the multiplex releases movie schedules for the following week each Wednesday, the authors define a *business week* to start from Wednesday to deliver predictions prior to scheduling. The transactional records for over 2500 movies are aggregated based on our definition of a business week. Each record contains fields specific to a movie in a business week such as the number of seats filled in a week, weeks since movie release. Some of these fields will be discussed in detail in Section 4.

The box plot shown in Fig 1 represents the distribution of lifetime for all the films considered. The average lifetime is observed to be 2.5 weeks, with 75% of movies having a lifetime ≤ 3 weeks.

Fig. 1. Frequency distribution of lifetime in weeks

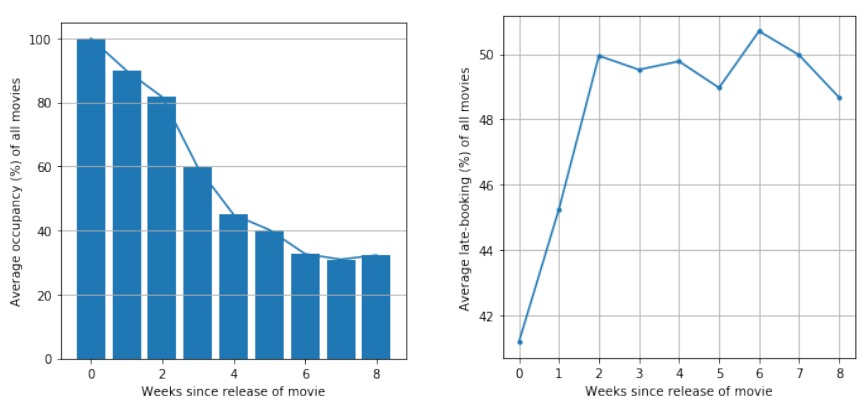


4 Feature Engineering, Extraction and Analysis

As discussed in Section 2, it is very important to consider local behavioral aspects while delivering predictions. Therefore, the authors have engineered features based on:

- Crowd behavior
- Screening behavior
- Seasonal and Regional behavior
- Movie specific aspects

4.1 Crowd Behavioral Features



(a) Variation of PSF

(b) Variation of Late-Bookings

Fig. 2. Variation of PSF and average LB across weeks since release of movies

Percentage of Seats Filled (PSF). PSF represents the ratio of the Number Of Occupants (NOC) to the total seats scheduled by the multiplex for the movie in a particular week (CAP). This is a strong indicator of the popularity of the movie among the cinema audience. As observed from Fig 2(a), the mean value of PSF across all movies tends to reduce as they tend to age. As the PSF declines for a movie, the multiplex reduces the CAP for the movie to minimize idle seats. When PSF tends to saturation, the movie is observed to be dropped from screening.

$$PSF_{\text{avg}}(i) = \frac{1}{n}\sum_{j=0}^{n}\frac{NOC_j(i) \times 100}{CAP_{\text{j}}(i)} \tag{1}$$

where:
i represents a unique movie,
n represents the max number of weeks the movie has been screened.

The mean of PSF (PSF_{avg}) for a movie is a measure of consistency for a movie's occupancy. While PSF is a short term indicator of the film's popularity, PSF_{avg} serves as a long term indicator for measuring the performance of a movie in theatres.

Late-Bookings (LB). LB represents the ratio of the seats booked at the closing of a show's transaction window (defined by the authors as after 3PM of the previous day of the show screening) to the total seats allocated for the screening. It is important to know that the multiplex facilitates offline in-person bookings besides the conventional online bookings. The rate with which the seats are filled typically represents the zeal for a film. Hence people tend to book tickets well in advance of the show to ensure they get a seat. Late bookings also account for last minute transactions that occur moments before a show is screened. Increase in the average number of late bookings per week (in Fig 2(b)) indicates a depleting interest for movies. This inference is drawn based on the comparable reduction in average occupancy per week as seen in Fig 2(a).

Relative Occupancy (RO). A higher number of occupants for a particular movie may require the multiplex to increase the capacity for it, consequently reducing the capacity for another movie in the week. $\text{RO}_{\text{j}}(\text{i})$ specifies the share of seats held by the i^{th} movie among other movies screening in the j^{th}week at the multiplex and is calculated by equation(2). If the $\text{RO}_{\text{j}}(\text{i})$ is 5%, it implies that the seats booked for the i^{th} movie holds 5% of all seats booked in the multiplex for the j^{th}week. The relative occupancy for the multiplex ranges from 0.007% to 81% with an average RO of 4% observed. Movies with consistently higher RO values have a higher probability of screening in the multiplex the next week. A smoothed histogram in Fig 3 illustrates the distribution of movies. It can be observed that for movies failing to screen the next week, the average RO is 0.8% and ranges from 0.007% to 18%. Meanwhile, the movies that screen the next week have a relatively greater average (5.2%) and a larger range (0.14% - 81%).

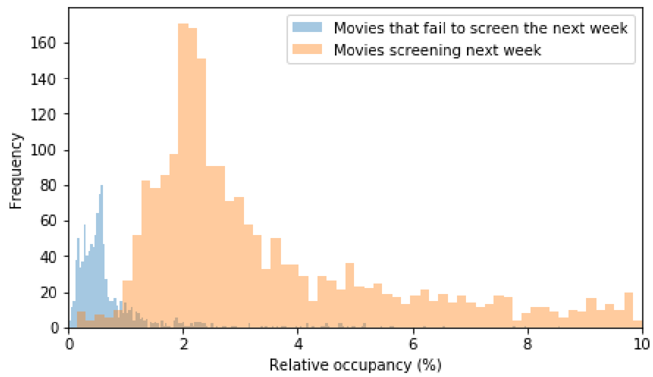


Fig. 3. Relative Occupancy histogram

$$RO_{\mathrm{j}}(i) = \frac{NOC_j(i) \times 100}{\sum_{k=i_0}^{i_{\mathrm{n}}} NOC_j(k)} \tag{2}$$

where:
i represents a unique movie,
j represents a particular week,
k represents a movie in the week,
n represents the total number of movies in the week.

Frequency of Seats Booked per Transaction (SBT). The number of seats booked in a single transaction-$\mathrm{SBT_m}$ (where m represents the number of bookings) sheds light into the response of the incoming crowd towards a film, and in turn characterizes the film. The multiplex facilitates the booking of multiple seats in a single transaction. Fig 4 represents the correlation between SBT and the probability of movies continuing to screen the following week. It can be observed that the occurrence of higher values of SBT ($\mathrm{SBT_7}$ and above) indicate a very high probability for movies to continue to be screened.

History Features. Since the model operates on a weekly basis to make predictions, it is important that we supply short term memory features to help the model understand the variations in the behavior across the past week. Therefore, 7 history points corresponding to the days in the prior week are provided for occupancy features such as RO, LB and PSF.

4.2 Screening Behavioral Features

Days Screened in a Week (DSW). It denotes the number of days a movie has been screened in the past week. The multiplex can arrange for a movie to screen

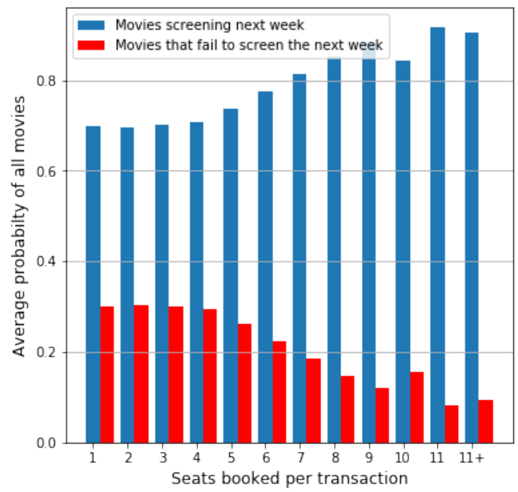


Fig. 4. Seats booked per transaction distribution

any number of days through the week based on factors including accommodating new releases in the week, holidays, as well as crowd behavioral factors discussed in Section 4.1. DSW captures these dynamics and upon observing its behaviour from Fig 5, it can be inferred that movies screened for just 1 or 2 days in the past week have a high probability of getting dropped from screens by the next week in contrast to films having higher days of screening (5-7 days), thus illustrating its capability to provide short-term forecasting insight.

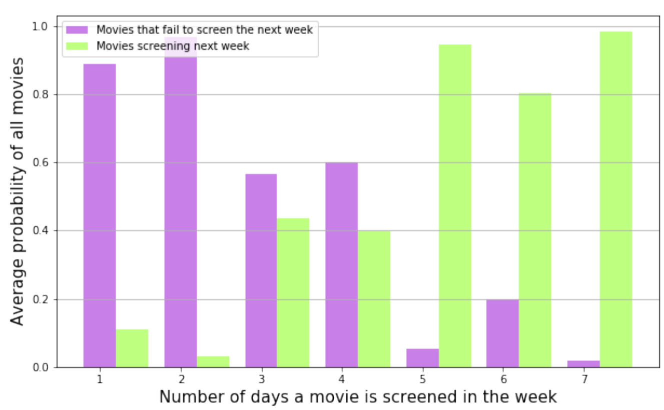


Fig. 5. Behavior of DSW

4.3 Seasonal and Regional Behavioral Features

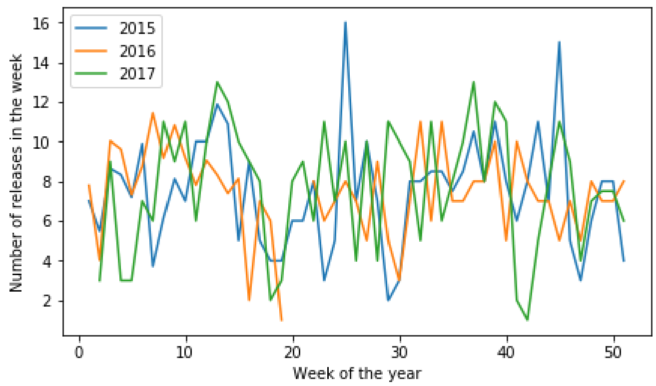


Fig. 6. Seasonality in weeks

Time of the Year. Seasonality is a vital factor when it comes to modelling regional crowd behavior across time. It refers to characteristics displayed by data that happen to recur in a defined periodic cycle. It can be observed in short term cycles (weekly) such as weekend and weekday behavior or in long term periodic cycles (yearly) such as festival holidays in a calendar year. The data showcases periodicity across a few weeks through the year as seen in Fig 6. Week-based seasonality is crucial since the multiplex requires scheduling information offered by our predictions in a weekly basis.

Language. The population is highly diverse owing to various demographic factors attributing to local behaviour. This feature describes language preference across the region. The languages considered were broadly split into Tamil (Regional language), English, Hindi (National language) and Others. The language split displays a dominance for the regional language followed by English as shown in Fig 7. This information assists in labeling the character of the incoming crowd.

4.4 Movie Specific Features

Genre and Runtime. Genres represent the themes showcased by a movie. Drama, Comedy and Romance were observed to be the most popular genres in the dataset and hence were considered for analysis. In contrast to movies with a single genre, movies with an intersection between one or more genres were observed to have a greater mean movie lifetime as shown in Fig 10. Additionally, Runtime indicates the scheduled running length of a movie.

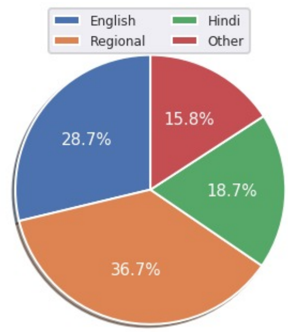


Fig. 7. Language split across the movies

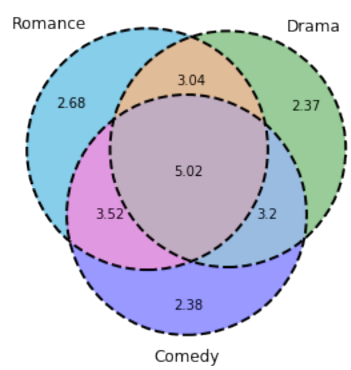


Fig. 8. Average lifetime of genres (in weeks)

New Releases in a Week (NRW). A new release can pivot the sales and demand of movies currently running. NRW refers to the number of new movie releases normalized over the total number of movies screening in the week.

5 Methodology

The authors make predictions at the start of a business week for two different use cases. The first is binary classification, to flag whether a particular movie screening in the week will continue to screen the following week (classification use case). For the movies predicted to screen the following week, a regressor engine predicts how long the movie will continue to be screened in weeks (regression use case).

5.1 Training and Testing

Transactional booking data from the years 2015, 2016 and 2017 were considered for analysis of movie lifetime prediction. The transactional booking behavior was observed to change within seasons through the years considered. The domain experts attributed this to the volatile nature of the data considered and further specified that this volatility consistently prevailed over the last decade. Therefore, the best way to model such data is to consider a uniform split across the 3 years for training and testing. Hence, the authors allocated 70% data from each year for training the model and the rest was considered as testing data. This way, the model is able to capture the behavior across all the years considered.

5.2 Feature Scaling and Standardization

Feature Scaling is the process of normalizing a feature to a defined scale. It is very important for predictive model optimization, especially deep networks as they offer faster means of convergence [8]. Occupancy features such as PSF, LB and RO need to be normalized based on the capacity allocated since a raw value of occupancy would not provide much insight to the predictive model without context. This is due to the fact that the multiplex under consideration has screens with capacities varying from 110 seats to 310 seats. Hence, the features have been normalized to values that lie between [0,1].

The features were standardized using Standard Scalar, which transforms the data based on the mean and standard deviation of the feature such that the resulting distribution has a mean of 0 and variance of 1.

5.3 Single Model Approach

In this approach, the authors aim to solve both the classification and the regression use cases using a single machine learning model. As the authors use a regressor for the classification task, the predictions are transformed based on a threshold due to the continuous nature of the regressor output (which is in terms of weeks).

Since the threshold used can be subject to bias, the authors consider all possible thresholds in a week to determine the best. A *threshold* labels a movie that screens the next week as a movie that screens a minimum of 'n' days, where $n \leq 7$. A threshold of '3' implies that a movie is considered to screen next week only if it is predicted to screen for a minimum of 3 days (0.428 weeks). Deep Neural Network (DNN), Extreme Gradient Boosting (XGB) and ET (Extra Trees) regressors were considered and their performance is summarized and presented in Section 6.2.

5.4 Two Model Approach

Although a single model for both use cases reduces the complexity of the solution, it fails to capture the specificity of the requirements posed by each (Section 6.3).

While the classification use case mainly requires short term features such as DSW (Section 5), regressors require a combination of both long term and short term information to make a continuous estimate. This approach deals with the use of separate classifier and regressor models for the discussed use cases.

Classification. A standalone model for classification is considered because of the volatile nature of the data. Since the prediction is pipe-lined, the error caused by the classifier is propagated onto the regressor engine. Therefore, it is very important to minimize the error caused by the classifier. An Extra Trees classifier is considered by the authors since it is the best performing model as seen in Section 6.2.

Regression. Regression is carried out to forecast the remaining lifetime for movies that are predicted to screen next week by the classifier. DNN, XGB and ET regressors were considered for this use case.

The DNN architecture used consisted of 3 dense hidden layers with varying number of neurons, each having a supporting dropout layer. In addition to preventing the proposed DNN to overfit, the dropout layers add controlled noise to the model to improve its capability of generalization [9]. The performance of these models are evaluated and compared in Section 6.2.

Table 2. Thresholds Applied to Regressors

		DNN			XGB			ET		
		Movie will not screen	Movie will screen	Mean	Movie will not screen	Movie will screen	Mean	Movie will not screen	Movie will screen	Mean
1 day	Precision	0.97	0.91	0.93	0.99	0.84	0.89	1	0.9	0.93
	Recall	0.78	0.99	0.93	0.55	1	0.87	0.74	1	0.92
	F1 score	0.86	0.95	0.92	0.71	0.91	0.87	0.85	0.95	0.92
2 days	Precision	0.96	0.94	0.95	0.98	0.88	0.91	0.99	0.92	0.94
	Recall	0.85	0.98	0.95	0.67	0.99	0.9	0.79	1	0.94
	F1 score	0.9	0.96	0.94	0.8	0.93	0.89	0.88	0.96	0.93
3 days	Precision	0.95	0.95	0.95	0.96	0.91	0.93	0.98	0.93	0.95
	Recall	0.88	0.98	0.95	0.77	0.99	0.92	0.82	0.99	0.94
	F1 score	0.91	0.96	0.95	0.86	0.95	0.92	0.9	0.96	0.94
4 days	Precision	0.93	0.95	0.95	0.94	0.93	0.93	0.98	0.94	0.95
	Recall	0.88	0.97	0.95	0.82	0.98	0.93	0.86	0.99	0.95
	F1 score	0.91	0.96	0.95	0.88	0.95	0.93	0.91	0.97	0.95
5 days	Precision	0.91	0.95	0.94	0.89	0.94	0.93	0.97	0.95	0.96
	Recall	0.89	0.96	0.94	0.86	0.96	0.93	0.88	0.99	0.96
	F1 score	0.9	0.96	0.94	0.88	0.95	0.93	0.92	0.97	0.95
6 days	Precision	0.88	0.96	0.94	0.82	0.95	0.91	0.96	0.96	0.96
	Recall	0.9	0.95	0.93	0.89	0.92	0.91	0.91	0.98	0.96
	F1 score	0.89	0.95	0.93	0.85	0.93	0.91	0.93	0.97	0.96
7 days	Precision	0.77	0.97	0.91	0.77	0.96	0.9	0.88	0.97	0.95
	Recall	0.9	0.91	0.9	0.9	0.89	0.89	0.93	0.95	0.94
	F1 score	0.9	0.88	0.9	0.83	0.92	0.89	0.91	0.96	0.94

6 Results and Observation

6.1 Metrics

Classification. Precision, Recall and F1 score are considered as the primary metrics for classication.

Regression. Since the regressors forecast the number of weeks a movie will continue to screen, the error in lifetime is measured in terms of weeks. The Movie Lifetime Error (MLE) is calculated as shown in equation 3.

$$MLE = |lifetime_{\text{actual}} - lifetime_{\text{predicted}}| \tag{3}$$

where,
$\text{lifetime}_{\text{actual}}$ represents the actual lifetime remaining for the movie,
$\text{lifetime}_{\text{predicted}}$ represents the predicted lifetime remaining for the movie.

6.2 Single Model Approach Classification Results

Various thresholds (defined in Section 5.3) were applied to convert the output of a single regressor to a classifier output. Table 2 represents the performance of Deep Neural Network (DNN), Extreme Gradient Boosting (XGB) and ET models across the 7 thresholds considered. The best results for each of the considered models are shaded in green as observed from Table 2. ET outperformed the other models with an average F1 score of 0.96 with an applied threshold of '6 days'.

6.3 Two Model Approach Classification Results

A standalone ET classifier provides the best results as seen from Table 3. To find the algorithm that poses minimum loss for the classification task, the percentage of wrong predictions are considered in Table 4. The ET classifier provides the least error with 1.21% of wrong predictions when predicting if a movie is screened the next week and 8% of wrong predictions when determining if a movie is not screened the next week. It is evident that the ET classifier provides exceptional performance when compared to other models tried out in Section 5.3. Not only does the ET classifier provide a marginally better F1 score, but also predicts with a lower percentage of wrong predictions. The ET classifier has the capability to be less sensitive to noise and volatile behaviour by offering variable smoothing using the η_{min} parameter [10].

6.4 Regression Results

The performance of three regressor models namely DNN, XGB and ET are compared in Table 5. The DNN clearly performs better than the other two models considered, predicting with less than 1 MLE 65% of the time and less than 2 MLE 85% of the time.

Table 3. Classification Best Results

	DNN regressor			XGB regressor			ET regressor			ET classifier		
	Movie will not screen	Movie will screen	Mean	Movie will not screen	Movie will screen	Mean	Movie will not screen	Movie will screen	Mean	Movie will not screen	Movie will screen	Mean
Precision	0.93	0.95	0.95	0.89	0.94	0.93	0.96	0.96	0.96	0.97	0.92	0.94
Recall	0.88	0.97	0.95	0.86	0.96	0.83	0.91	0.98	0.96	0.97	0.99	0.98
F1 score	0.91	0.96	0.95	0.88	0.95	0.93	0.93	0.97	0.96	0.97	0.97	0.97

Table 4. Percentage of wrong predicts

Models used	Movies screening next week (%)	Movies not screening next week (%)
XGB regressor	5.52	16.35
DNN regressor	4.17	9.94
ET regressor	2.02	10.26
ET classifier	1.21	8.01

7 Conclusion

Of the two approaches considered, the Two Model Approach (Section 5.4) provides maximum accuracy and is the optimal solution for the two business use-cases discussed. The standalone ET classifier performs better than the regressors transformed to do the classification task. The model provides an accuracy of 97% for the classification use-case, which helps the multiplex accurately schedule movies on a weekly basis. The regression use case is performed using the DNN due to its superior performance as shown in Section 6.2. The DNN is more robust to outliers and is able to capture the non-linear trends present in the data.

Currently, the multiplex scheduling experts consider the admissible range of error to be within 2 weeks. They estimate the lifetime and films screening the next week based on empirical methods and heuristics. Our approach estimates the remaining lifetime correctly with less than 2 MLE 85% of the time. These results set a benchmark for the experts in the domain regarding lifetime estimation. Since our method is the first of its kind it will be tested in real-world circumstances in the next revision of strategies by the considered multiplex.

Table 5. Regression Results

	DNN	XGB	ET
[0,1) MLE(%)	64.85	59.00	55.16
[1,2) MLE (%)	19.10	22.48	28.44
[2,3) MLE (%)	7.82	10.71	10.58
[3,4) MLE (%)	3.45	4.10	3.17
>4 MLE (%)	4.77	3.70	2.65

References

1. Deloitte (2018)."*Economic Contribution of the Film ad Television Industry in India, 2017*". Retrived from https://www.mpa-i.org/wp-content/uploads/2018/05/India-ECR-2017_Final-Report.pdf
2. Jehoshua Eliashberg, Quintus Hegie, Jason Ho, Dennis Huisman, Steven J. Miller, Sanjeev Swami, Charles B. Weinberg, Berend Wierenga, "*Demand-driven scheduling of movies in a multiplex*", International Journal of Research in Marketing, Volume 26, Issue 2, 2009, Pages 75-88, ISSN 0167-8116. DOI: 10.1016/j.ijresmar.2008.09.004
3. A. Sivasantoshreddy, P. Kasat, and A. Jain, "*Box-Office Opening Prediction of Movies based on Hype Analysis through Data Mining*," International Journal of Computer Applications, vol. 56, no. 1, pp. 1–5, 2012. DOI: 10.5120/8852-2794
4. R. Sharda and D. Delen, "*Predicting box-office success of motion pictures with neural networks*", Expert Systems with Applications, pp. 243-254, 2006. DOI: 10.1016/j.eswa.2005.07.018
5. Jaiswal, S. R., and Sharma, D. (2017). Predicting Success of Bollywood Movies Using Machine Learning Techniques. Proceedings of the 10th Annual ACM India Compute Conference. DOI: 10.1145/3140107.3140126
6. Andrew Ainslie, Xavier Drèze, Fred Zufryden, (2005) "*Modeling Movie Life Cycles and Market Share*". Marketing Science 24(3):508-517. DOI: 10.1287/mksc.1040.0106
7. Adithya Ganesan V., Divi S., Moudhgalya N.B., Sriharsha U., Vijayaraghavan V. (2020). "*Forecasting Food Sales in a Multiplex Using Dynamic Artificial Neural Networks*". In: Arai K., Kapoor S. (eds) Advances in Computer Vision. CVC 2019. Advances in Intelligent Systems and Computing, vol 944. Springer, Cham. DOI: 10.1007/978-3-030-17798-0_8
8. L. Huang, X. Liu, Y. Liu, B. Lang and D. Tao, "*Centered Weight Normalization in Accelerating Training of Deep Neural Networks*," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 2822-2830. DOI: 10.1109/ICCV.2017.305
9. Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. *"Dropout: a simple way to prevent neural networks from overfitting"*. J. Mach. Learn. Res. 15, 1 (January 2014), 1929-1958. URL: https://dl.acm.org/citation.cfm?id=2670313
10. Geurts, P., Ernst, D. Wehenkel, L. Mach Learn (2006) 63: 3. URL: https://doi.org/10.1007/s10994-006-6226-1