

Tracing the Role of Big Data in the Digital Success of Bollywood Films: A Dual Case Study

Rashi Srivastava*

Abstract

The world is continuously generating enormous data and new technological tools are making analysis of such data easier than ever. Big Data is one such tool, gradually changing the Entertainment Industry, which is now driven by audiences' word of mouth owing to the availability of several online platforms. Therefore, fan engagement has become the key to successful marketing and Big Data is creating numerous opportunities for it. In this case study, we analyze two examples demonstrating how the technology of data crunching and social listening is making its way into Bollywood. The first half of this study focuses on how IBM predicted a 73% success rate for the Hindi movie, Ram Leela (Sanjay LeelaBhansali, 2013) using the Social Sentiment Index (SSI) methodology based on right selection of features. Secondly, we study how Shah Rukh Khan's blockbuster film Chennai Express (Rohit Shetty, 2013), made extensive use of Big Data & Analytics solutions to drive its digital marketing campaigns using the eMee Social Radar which gave it an edge over the traditionally marketed movies. These movies had a runaway success at the box office, proving Big Data to be a potential game changer that can revitalize the entire entertainment industry.

Copyright © 2022 International Journals of Multidisciplinary Research Academy. All rights reserved.

Keywords:

Big Data;
Digital Marketing;
Sentiment Analysis;
Actionable Insight;
Entertainment Marketing.

Author correspondence:

Rashi Srivastava,
Bachelors Program, School of Information Technology Engineering
Vellore Institute of Technology, Vellore, Tamil Nadu, India.
Email: rashi.k.srivastava@gmail.com

1. Introduction

Nowadays, the audience to any form of entertainment is more than just a passive receptacle. It possesses the ability and power to react, in its own way, to creative works as they unfold. Audience reactions have not remained just a measure of the success or failure of a film, they have instead become the fuel for running the analytics machinery to produce meaningful, actionable insights into what works and what doesn't. The emergence of movies in general, interrupted the feedback mechanism of theatre audience of the previous eras. But it would be wrong to think that audiences have no role in the Bollywood machinery today. Rather, the role has shifted. We now throw our tomatoes online—in blogs and twitter feeds and Facebook timelines. An overview of the Internet will tell us that today's audience is just as vocal as ever. Consumers are constantly giving signals to content makers on what they like and what they don't like, what they will/ will not pay for, and whom they like to see. All of this generates significant levels of data, which companies are starting to tap, analyze, derive collective insights and use to improve the quality of engagement with the customers. This leads to informed decision making, time and location based content optimization, effective competition monitoring and above all, reaching specific target audience.

Persistent System's eMee Social Radar is a social media listening tool. The internal structures and mechanics involved with this tool are analogous to the spiders and bots used by search engines that crawl through web

*Bachelors Program, School of Information Technology Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India.

pages. Social Radar crawls through the accessible contents of web pages in real time, continuously or on certain time intervals depending on the user's goals and objectives. It searches these web pages to look for specific "mentions" of specific keywords or phrases and then summarizes and reports them all back to the user in a form where relevant insights can be obtained. In addition to statistics, the Social Radar platform uses carefully calibrated metrics to gauge engagement levels of the audience. It is then pushed to big data clusters, where different data processors chew out insights like Sentiment & Mood Analysis, Campaign Optimization, Competition Analysis and Influencer Targeting.

The IBM Social Sentiment Index^[1] uses advanced analytics and natural language processing technologies to analyze large volumes of social media data in order to assess public opinions. A striking feature of the software used is that it can distinguish between sarcasm and sincerity, and applies machine learning to identify which social media commentary is important and which is just background noise. Through this sentiment analysis, IBM creates real-time public opinion snapshots, identifying trends and gauging how consumers feel about a variety of topics from retail and sports, to movies and urban traffic. The Social Sentiment Index can identify and measure positive, negative and neutral sentiments shared by people in public forums such as Twitter, Facebook, blogs, message boards and other social media platforms, and provide useful, quick and relevant insights into online consumer conversations.

Big data platforms like Hadoop and Hive provide the computing infrastructure needed for detailed analysis of the humungous amount of social data generated. This is complemented by advanced data analysis by social data scientists to provide 'Actionable Insights' to brand managers. Hereafter, brand managers need to be able to take users up the 4A levels of engagement - from aware to active, and from addicted to amplified.

The film industry is also witnessing reforms in the traditional ways of data collection. For example, in in-cinema advertising, business models on contracts are shifting from lump sum deals to pay per view deals, which is a better way to measure the efficiency of the medium. According to a recent KPMG report^[2] in 2015, digital advertising continued its strong run with a 38.2% growth over 2014 as a growing internet user base and usage was supplemented by increased spend allocation by marketers. In such a scenario, accurate data analysis will definitely play a major role in the digital marketing and promotion of films in the coming years too.

2. Literature Review

Shailja Gupta^[3], Chief Digital Strategist for Chennai Express, says, "One of the biggest factors for Chennai Express' digital success was the role of social media integration with the different digital products and the daily collection and analysis of huge data to understand the pulse of the digital audience. I totally believe that acquiring the data on a daily basis is completely worthless unless you analyze it correctly, interpret it intuitively, strategize it creatively and implement it immediately." Ashok Karania^[4], Vice President - Europe Business at InfoStretch Corporation, points out that Big Data can definitely be a boon to the entertainment and sports industry. It can improve the profitability of the movies - always a high risk business. The green-lighting of the story to the cast selection to the timing of release can be determined. It can help to pick the right players for the sporting leagues - allowing talent to win. Sanjeev Mishra and Mansha Sharma^[5] opined that the emergence of new technology and the era of 'big data' has enabled such unconventional sectors as film industry to reap the benefits of analytics. Moreover, the latest trends witnessed by this industry have brought in a plethora of platforms to connect with the consumers, a scenario that calls for a more systematic approach to increase both the efficiency and efficacy of promotions through each platform in order to achieve maximum possible ROI from minimum possible investments. According to Michael Gold, Ryan McClarren, and Conor Gaughan^[6], data science will be adopted as integral to critical manifestations in the media and entertainment sector. For film, music, digital media, and television, big data and predictive analytics will offer some value-added deployments.

3. Audience Analytics- Targeting viewers the right way with social listening

There have been a lot of overwhelming changes on how people communicate and interact with each other online in recent years – thanks to Social Media. It practically revolutionized the Internet into something that is more personal and participatory. That's why social networking is now the top online activity, overtaking email, 60% doing so through their tablets and smartphones, according to the latest Intelligence report from Business Insider^[7]. For content creators, it's absolutely vital to understand what a target market wants. Social media listening can help the entertainment industry get a better idea about which properties are likely to catch on. As shown in Figure 1, social listening involves not only monitoring the data but measurement and reporting are also essential for marketing success. Businesses rely on various forms of analytics to refine the targeting profiles of their audiences and to identify the most efficient way to reach them at scale. Earlier, location and population based optimization was sufficient to produce effective marketing strategies. In today's world, it is important to go deeper than just demographics when it comes to customer evaluation.

Demographics help us understand only the most basic, general facts about our audience as a group and essentially ignore variances like values, beliefs and attitudes. Psychographics^[8] takes variables like values, opinions and attitudes into account, in an attempt to create a current, detailed profile of an individual, which is the representative of a larger group. Understanding these variables and plugging them into the content creation process does not only help to connect with the audience on a deeper level, it also informs a number of practical factors such as tone, style and language. Marketers are therefore required to collect and measure disparate data routinely as well as demonstrate program success on sales and revenue. “Technology plays a big role in figuring out what fans liked, and what is going through the sieve. Big Data analytics, a variety of in-house and off-the shelf tools and our unique approach based on the 4A model of engagement provided Actionable Insights to the social media marketing team. However, these insights needed to be culturally and contextually relevant, and to achieve this, the team had to keep an ear on the ground and understand every nuance before providing recommendations! While the overall marketing strategy is important, the key lies in observing social media trends and creating campaigns on the fly in response to these trends. This is more art than science, and technology, at best, is an enabler” says SiddheshBhobe, CEO, Persistent Systems.

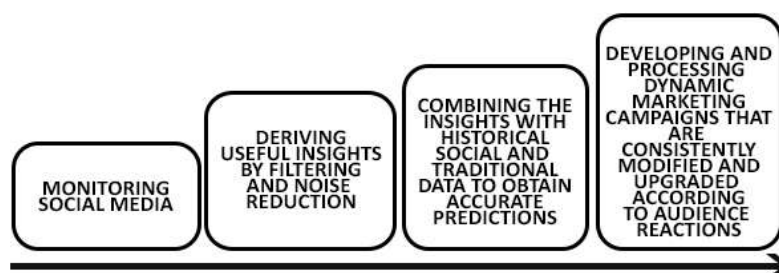


Figure 1. Social Listening and Audience Analysis Process

4. Actionable Insights - The link between marketing strategies and audience reaction

Actionable Insight is a term in data analytics and big data for information that can be acted upon or information that gives enough insight into the future such that the actions needed to be taken become clear for decision makers. With many companies struggling to make sense of their data and create value with their big data investments, the promise of actionable insights sounds wonderful. Forrester^[9] reports 74% of firms say they want to be “data-driven,” but only 29% are actually successful at connecting analytics to action. Actionable insights appear to be the missing link for entertainment companies that want to drive content marketing outcomes from their data. Actionable insights sit at the apex of the data pyramid. An insight that drives action is typically more valuable than one that simply answers a question—especially an insight that makes you rethink something and pushes you in a new direction. They are the highly treasured output of all the work that goes into collecting, preparing and analyzing your data. Maximizing the actionable insights you receive from your analytics investments is important to your data-driven success. There are six attributes to examine how actionable your data insights are: Alignment, Context, Relevance, Specificity, Novelty and Clarity. With social media ever evolving, such expert insights are necessary for a coordinated social media strategy. These Actionable Insights are often used to produce real time and contextually relevant recommendations to drive the social media campaigns for film promotions.

5. Case Study 1 – How IBM’s Social Sentiment Index predicted the success of Ram Leela

5.1 An Overview

IBM has established the IBM Social Sentiment Index (SSI)^[10] to aggregate and gauge public opinion from a range of social media. The software used can distinguish between sarcasm and sincerity, and applies machine learning to identify which social media commentary is important and which is just background noise. Through this sentiment analysis, IBM creates real-time public opinion snapshots, identifying trends and gauging how consumers feel about a variety of topics from retail and sports, to movies and urban traffic. The company says it had applied its analytics framework to the *Deepika Padukone*^[11], *Ranveer Singh*^[12] starrer much before its release, and had predicted with 73% confidence that it would have a successful first weekend. The film indeed had a great first weekend^[13] and went on to become a blockbuster. That prediction framework was developed after IBM analyzed over seven lakh posts across a variety of platforms including Facebook, Twitter, YouTube and blogs for 25 Bollywood films with considerable social buzz. These included films like *Ek Tha Tiger* (Kabir Khan, 2012), *Barfi* (Anurag Basu, 2012), *Bhaag Milkha Bhaag* (Rakeysh Omprakash Mehra, 2013), *Kai Po Che* (Abhishek Kapoor, 2013), *Kahaani* (Sujoy Ghosh, 2012) and *Agneepath* (Karan Malhotra, 2012). The analysis was used to arrive at a predictability of success by genres.

Figure 2 shows the *Ensemble Analysis Model* designed by IBM which listened to 700,000 social media conversations and analyzed 30 films covering every genre using film reviews and IMDb listings, over a span of two years to collect the right kind of data. This data was then processed using the SSI methodology to develop a predictive model for Ram Leela.

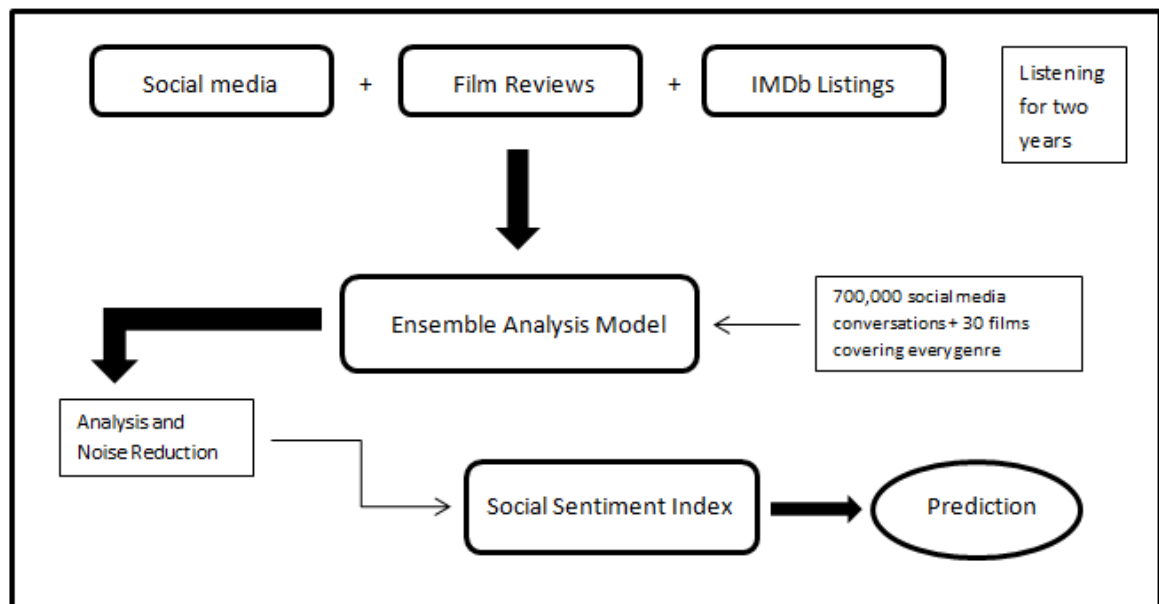


Figure 2. Flow chart of IBM's Social Sentiment Index Model

5.2 Understanding the Formula

The predicted confidence of success for the following genres was obtained along with subdivisions like remakes & dubs, choreography & set etc. as shown in Table 1.

Table 1. Predicted confidence of success for previous two year movies' data**

Category	Remakes & Dubs	Choreography & Set	Superstars	Star Director	Comedy Track	Music
Political Movies	37%	1.0%	0.4%	35%	0.0%	0.2%
Action Movies	14%	2.0%	15%	15%	10%	5.0%
Comedy Movies	17%	10%	10%	5.0%	30%	3.0%
Star Director	70%	60%	5.0%	-	45%	50%

Studying Audience Reactions prior to the release of the film, following responses as shown in Table 2 were recorded:

Table 2. Analysis of Audience reaction**

	Positive	Negative	Neutral	Ambivalent
Plot	40%	15%	40%	5.0%
Choreography & Set	50%	5.0%	40%	5.0%
Promotion	45%	10%	45%	-
Music	40%	20%	30%	10%
Glamour	50%	-	50%	-
Positive Stimulus	50%	10%	30%	10%

** Values are obtained by approximation from available infographic data.

5.3 How the Success of the film was predicted prior to its release:

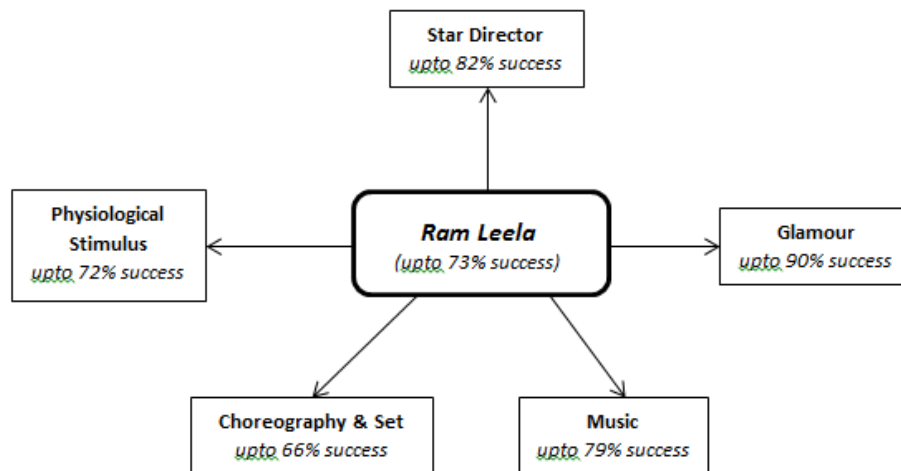


Figure 3. Success Prediction of Ram Leela

Spotting trends geographically and chronologically is common, advanced analytics can now spot trends across a third dimension: social sentiment and cultural domain. IBM foretold with 73% confidence based on the predictive model that the movie will have a strong opening. The SSI revealed some interesting trends. It showed that it could be said with 73% confidence that a film with a political flavour stands a strong chance of scoring well at the box office, with factors like remakes and a star director further improving its success rate, as shown in Figure 3. In this genre, a superstar and choreography & set design were found to make only minor contributions to success.

6. Case Study 2 – The Role of eMee Social Radar in the Digital Marketing of Chennai Express

6.1 An Overview

Social media differs from traditional marketing techniques in its immediacy and directness; it enables a direct relationship with the target audience, a one-to-one feedback mechanism and sharing platform. Film marketing is no longer about releasing a trailer or a few posters and waiting for the audience reaction on opening day. The market survey approach of using focus groups is also passé. After all, social media is the 21st century's perfect crowd-sourced focus group! It enables real-time, two-way communication, allowing marketing campaigns to be changed based on trends and feedback. eMee^[14] and its Social Radar platform was used to build, analyze and interpret the audience behavior online. It followed Chennai Express' online presence closely and implemented the system of Actionable Insights - a recommendation system that uses analytics, big data tools and calibrated metrics to gauge what people like and their engagement levels. A coordinated social media strategy was achieved with the help of contextually relevant recommendations provided by Social Radar. eMee worked closely with the Chennai Express team throughout the marketing lifecycle, including reaching out to the core fan base, influencer targeting, design of effective campaigns, marketing recalibration, the effective use of twitter hashtags and competitor analysis.

6.2 Understanding the formula

According to the documentation provided by Persistent Systems, the eMee Social Radar collected data at scheduled intervals from a variety of Social Media platforms like Twitter, Facebook, YouTube and a wide range of blogs. The unstructured JSON files received from these varied sources were then pushed into data ingestors for critical first level filtering, metadata enhancement and noise clean-up. Figure 4 depicts the functioning of this platform in detail. eMee's location and clout derivation tools also ran on top of this data to get a better understanding of each user's profile. Individual tweets or posts were mapped in relation to Chennai Express and important tags and keywords used in the posts were identified. The data was then transported to the Big Data Cluster on MongoDB for sentiment and mood analysis in an automated fashion. At the same time, critical insights and recommendations were added to the final report that was provided to the Chennai Express team at the end of each day for review and follow up action. The architecture of the Social Radar is as shown below-

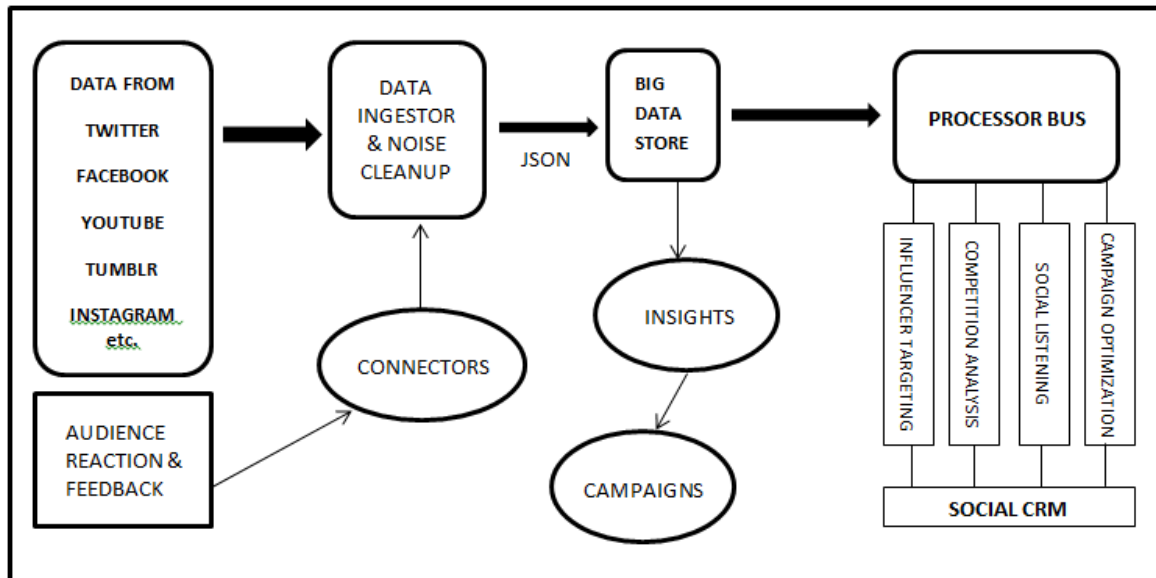


Figure 4. eMee's Social Radar Architecture

6.3 Sentiment and Mood Analysis

There is no content creator who would not want to tap into the thoughts and feelings of his target audience, for doing so can help him personalize his offerings to give customers exactly what they want. However it is very difficult to discover how people feel about something simply by learning what they're sharing through social media and other digital channels. Thus, using solutions that tap into the power of sentiment analysis becomes important. Sentiment and Mood Analysis helps organizations to identify pieces of content as positive, negative or neutral, providing valuable insights into audience feelings, preferences and characteristics that can help predict future behavior. eMee captured moods and helped judge pulse of the audience during the promotional campaign of Chennai Express. It collected useful data from public tweets and Facebook posts and processed it by performing phrase analysis, generating sentiment score with ontology and gave results based on pattern matching with progressive dictionary containing positive, negative and neutral words. Both twitter and Facebook data is capable of providing useful mood analysis data because of the hashtags in twitter and 'feeling' tags in Facebook along with other prefixed indicative words that are fed into the system. It is therefore, very easy to gauge public reaction to a particular event, in this case, it could have been a trailer release, song promo or an event related to the movie.

6.4 How Social Radar performed Social Sentiment Analysis for Brand Shah Rukh Khan:

In the month of March 2013, four months prior to the release of Chennai Express, a week long social sentiment analysis^[15] was performed by eMee for Brand Shah Rukh Khan (SRK)^[18] in order to provide useful insights about the upcoming film marketing campaign. Its data sources included Wikipedia, IMDB, Topix, Dailymotion, YouTube, Reddit, Blog, Twitter, Facebook and Discussion Boards. From 62 thousand mentions and a 19 million reach, social radar derived a 66% positive sentiment towards the star for the considered week. The analysis involved the then-trending topics like Magic of Frooti endorsed by SRK, Performance at IPL opening ceremony, Shah Rukh Khan's speech on women first and Jaago re campaign, *Om Shanti Om* (Farah Khan, 2007) release in Japan that week and SRK appearing on the Forbes magazine. Also, external events such as the political debate over Shah Rukh Khan's Outlook essay was considered as it created variations in the positive as well as negative sentiments of the people. The study also calculated eMee Social Index for Brands endorsed by Shah Rukh Khan as shown in Table 3.

Table 3. eMee Social Index for Brands endorsed by Shah Rukh Khan

Brand	Positive				Negative				eMee Social Index
	Mentions	Reach	Spread	Users	Mentions	Reach	Spread	Users	
Tag heuer	426	403218	23	250	11	5728	0	9	54.31
Hyundai	742	291156	27	382	77	32312	2	44	55.77
Frooti	9702	5538851	132	3286	393	266188	7	195	762.53
Pepsodent	530	440051	18	192	9	4164	1	7	55.97
Lux	487	320550	17	328	20	10569	2	15	49.73

Both positive and negative mentions, reach, spread and users contributed to the overall Social Index of the brand under observation. Owing to the recent “Magic of Frooti” campaign, the eMee social index came out to be highest for the brand Frooti whereas the other brands were more or else in the same range.

This Social sentiment analysis helped Persistent Systems to develop modules for effectively enhancing the marketing process of Chennai Express based on the audience reactions to the star of the movie. Ultimately, the overwhelming success of Chennai Express ^[16] has proved that social media is the channel of the future and that it presents unique opportunities to marketers and brands, at an unbeatable return on investment.

7. Results and Analysis

Unlike the last decade, the tastes of people in terms of entertainment are fast-changing in this age of digital revolution. Fresh concepts and ideas become monotonous and stale in no time. As a result, the movie makers need to make sure they keep themselves updated and paced according to the audience demands by using right kind of analytics. Traditional marketing approaches fail to dig deeper into the reactions received by the audience and are therefore becoming ineffective. Carefully assessing the loopholes in these approaches can help in the development of accurate, well-balanced digital marketing tools. Big data analytics has surely improved the overall revenues of the entertainment industry and knowing how will motivate more and more production houses to incorporate this technological boon to make their content a sure shot hit. Machine Learning^[17] techniques have been applied in the past to predict a movie’s success by using generic data. Along with the powerful and ever-developing analytical technologies, it can produce more accurate and helpful results which can benefit the entertainment industry from suffering enormous losses due to the failure of a film which sometimes goes unpredicted by Big Data.

8. Conclusion

Analytics allow studios to go beyond simple focus groups or established financial modeling to determine how audiences might respond to a given film. It's all about identifying patterns in past data, melding them with current data points that are readily available, and then taking action to improve business performance. Such analytics can help film producers understand better what will work and what will not, at the stage that a film is being scripted and conceptualized. Given that teasers of a film, including its music, come out several months before release and generate a lot of buzz, analyses of such buzz can also help film-makers fine tune their products before release, and decide on the most appropriate release dates and release locations. Making more successful movies requires getting a high-resolution picture of your audience. And until recently, this snapshot was very soft-focus. The blunt instrument of panels, polls and reviews is fast giving way to computers that can render us in fine detail by picking up the trails of digital breadcrumbs we leave online and building them into predictive models of what we like and don't like. As a result, the sensibility spectrum of a modern audience is more accurately represented now. Clearly the promotional strategies and analytics involved in the success of movies like *Ram Leela* and *Chennai Express* identify the impact of big data and data science to a large extent. The last decade has not only seen an increase in the diversity of devices on which entertainment is consumed but also in the profile of the audience that is now getting used to custom content. Profiling of customers for better targeting, personalization of content on select mediums etc. comprise endless opportunities to benefit from analytics. Therefore, the collaboration of entertainment Industry leaders with leading big data startups and visionaries is bound to create new success stories in the future too.

References

- [1] IBM, "Birth of A Trend: Steampunk," IBM Analytics – Conversations: Social Sentiment Index—United States Retrieved from <https://www.ibm.com/analytics/us/en/conversations/social-sentiment-steampunk.html>
- [2] KPMG, "The Future: now streaming," FICCI Media and Entertainment Industry Report 2016. Retrieved from <https://home.kpmg.com/content/dam/kpmg/pdf/2016/04/The-Future-now-streaming.pdf>
- [3] Gupta, S. (n.d.). "The Role of Digital Marketing & Social Media in success of Chennai Express." Retrieved from [www.emee.co.in: http://emee.co.in/case-studies/chennai-express-case-study.pdf](http://emee.co.in: http://emee.co.in/case-studies/chennai-express-case-study.pdf)
- [4] Karania, A. "How Big Data Is Changing the Entertainment Industry." Retrieved from [linkedin.com/pulse/how-big-data-is-changing-the-entertainment-industry](https://www.linkedin.com/pulse/how-big-data-is-changing-the-entertainment-industry)
- [5] Mishra, S. "Bringing Analytics into Indian Film Industry with Back Tracing Algorithm." Retrieved from <https://www.analyticsvidhya.com/blog/2016/08/bringing-analytics-into-indian-film-industry-with-back-tracing-algorithm>
- [6] Gold, Michael, Ryan McClarren, and Conor Gaughan. "The lessons Oscar taught us: data science and media & entertainment." *Big Data 1.2* (2013): 105-109.
- [7] Gregorio, J. "Social Listening- Its Importance and How it Works." Retrieved from <http://digitalmarketingphilippines.com/social-listening-its-importance-and-how-it-works>

- [8] Lovato, J. "Serving Nachos for Dessert; Why Audience Analysis matters."
Retrieved from <http://www.mediavisioninteractive.com/blog/copywriting/serving-nachos-for-dessert-audience-analysis>
- [9] Dykes, B. "Actionable Insights: The Missing Link between Data and Business Value."
Retrieved from <http://www.forbes.com/sites/brentdykes/2016/04/26/actionable-insights-the-missing-link-between-data-and-business-value>
- [10] IBM Big Data & Analytics Hub, "The great Bollywood Formula decoded."
Retrieved from <http://www.ibmbigdatahub.com/infographic/great-bollywood-formula-decoded>
- [11] Padukone, D. Indian Film Actress, "Wikipedia Page of DeepikaPadukone,"
https://en.wikipedia.org/wiki/Deepika_Padukone
- [12] Singh, R. Indian Film Actor, "Wikipedia Page of Ranveer Singh"
https://en.wikipedia.org/wiki/Ranveer_Singh
- [13] <http://www.ibtimes.co.in/articles/522749/20131117/ram-leela-box-office-collection-beats-krrish.htm>
- [14] Bhobe, S. (n.d.). "Digital and social business of Bollywood"
Retrieved from <http://emee.co.in/case-studies/chennai-express.html>
- [15] eMee, "Social Sentiment Analysis for Brand SRK"
Retrieved from <http://www.emee.co.in/demos/socialradar>
- [16] Times, E. (2013). "Big Data, Analytics help Chennai Express tap social media."
Retrieved from www.economictimes.in: http://articles.economictimes.indiatimes.com/2013-12-26/news/45593001_1_chennai-express-big-data-analytics
- [17] Ericson J., Grodman J., "A Predictor for Movie Success."
Retrieved from cs229.stanford.edu/proj2013/EricsonGrodman-APredictorForMovieSuccess
- [18] Khan, S. Indian Film Actor, "Wikipedia Page of Shah Rukh Khan"
https://en.wikipedia.org/wiki/Shah_Rukh_Khan