

A Machine Learning Approach to the Indian Comics Industry

Siddhartha Bose¹, Dr. Pritpal Singh², Rakhi Nagpal³

Submitted: 22/08/2023

Revised: 09/10/2023

Accepted: 21/10/2023

Abstract: The comic book industry in India is growing quickly, and publishers and stores are always looking for ways to improve their marketing so they can get more customers. The use of machine learning programming in the industry has shown that it has a lot of potential to boost sales, but there hasn't been much research done on the specific things that make Indian comic book buyers want to buy. In this paper, a machine learning algorithm based on logistic regression is proposed to predict whether or not Indian comic book buyers are going to buy. A review of the literature was done to find out what research had already been done on how machine learning programming is used in the Indian comic book industry. Most of the research that had been done before was about using machine learning for recommendation systems and predictive analysis, the review found. But it's important to figure out what makes Indian comic book buyers want to buy. The proposed machine learning algorithm uses price, genre, author, and publisher to predict whether or not someone will buy something. Logical regression, a statistical method used for two-way classification problems, is used to create the algorithm. The algorithm is trained on a set of data about how many comic books were sold, and its accuracy is tested on a separate set of data. The results of the proposed algorithm show that it can accurately predict whether or not an Indian customer will buy a comic book. The algorithm can figure out which specific factors affect a person's decision to buy, which can help retailers and publishers make better marketing campaigns and better products. The proposed algorithm could make marketing strategies in the Indian comic book industry more efficient and effective. In conclusion, the proposed machine learning algorithm could make a big difference in the marketing strategies of the Indian comic book industry. By figuring out the specific things that affect a person's decision to buy, retailers and publishers can improve their marketing campaigns and make their products better. The proposed algorithm can also be used to keep track of inventory and analyze sales. In the future, research can be done to add more factors to the dataset and make machine learning algorithms for the Indian comic book industry that are more advanced.

Keywords: Indian Comics Industry, Machine Learning, Buying Intention, Marketing, Strategy

1 Introduction

People of all ages have used comics to have fun and learn new things. With the development of technology, the comic book business has changed in a big way. Over the years, Indian comics like Amar Chitra Katha, Tinkle, and Chacha Chaudhary have become more and more popular. Using machine learning programming can be a game-changer for the Indian comic book industry because it can have a big effect on whether or not customers want to buy. This research paper will look at how machine learning programming affects Indian customers' plans to buy comic books.

1.1 Background

In India, the comic book business has grown a lot over the years. The Indian comic book market was worth \$97 million in 2020, and it is expected to grow at a compound annual growth rate (CAGR) of 15% from 2021 to 2026. (Mordor Intelligence, 2021). The Indian comic book market is growing, but it still has problems like low availability, high

prices, and few ways to get comics to people. Also, the rise of digital platforms has made the market for physical comic books very competitive. Machine learning programming can help solve these problems and change the minds of Indian comic book buyers in a good way.

1.2 Machine learning programming

Machine learning programming is a branch of artificial intelligence that lets computers learn from data and get better at what they do without being told to do so. It uses algorithms that can look at data, find patterns, and make predictions. Programming that uses machine learning has been used in many fields, such as e-commerce, healthcare, and finance, to name a few. Using machine learning programming in the comic book industry can have a big effect on how likely Indian customers are to buy comics.

Effect of machine learning programming on Indian customers' plans to buy comic books

1.2.1 Personalized recommendations

Personalized suggestions can have a big impact on a customer's decision to buy. Machine learning software can look at a customer's past purchases, how they browse the site, and their demographic information to make personalized suggestions. For example, if a customer has bought a Tinkle

Mittal School of Business, Lovely Professional University
Phagwara, Jalandhar, Punjab, India – 144411

Correspondence Author: Siddhartha Bose1

siddhbos@gmail.com ; ORCID: 0000-0002-5353-7069

prtipal.16741@lpu.co.in ; ORCID: 0000-0001-5152-200X

rakhi.42000089@lpu.in ; ORCID: 0000-0001-7981-9980

comic book before, the machine learning algorithm can recommend other Tinkle comics that the customer might like. Customers are more likely to buy if you give them personalized recommendations.

1.2.2 Dynamic pricing

Personalized suggestions can have a big impact on a customer's decision to buy. Machine learning software can look at a customer's past purchases, how they browse the site, and their demographic information to make personalized suggestions. For example, if a customer has bought a Tinkle comic book before, the machine learning algorithm can recommend other Tinkle comics that the customer might like. Customers are more likely to buy if you give them personalized recommendations.

1.2.3 Inventory management

Keeping track of inventory is important for the success of any business. Programming that uses machine learning can look at past sales data and patterns of demand to figure out how to best manage inventory. For example, if there is a lot of demand for a certain comic book, the algorithm can suggest buying more of it to meet the demand. On the other hand, if a certain comic book isn't selling well, the algorithm can suggest cutting down on its stock to avoid losses. Inventory management that works well can make customers happier and boost sales.

2 Literature Review

The market for comic books in India has expanded gradually over the years. According to Mordor Intelligence (2021), the Indian comic book market was valued at \$97 million in 2020 and is projected to grow at a CAGR of 15% during the forecast period of 2021-2026. However, difficulties such as restricted supply, high prices, and few outlets for distribution persist in this sector. These issues can be remedied with the help of machine learning software, which in turn will have a beneficial effect on the purchasing decisions of Indian comic book readers.

For comic books, Vemula, Nalamothu, and Reddy (2021) suggested a machine learning-based recommendation system. In order to tailor suggestions to each individual user, the system combined techniques from joint filtering and content-based filtering. Collaborative filtering analyses users' past preferences to suggest similar items, while content-based filtering analyses the characteristics of items to recommend items with similar features. On a dataset consisting of comic books, the suggested system was found to perform better than preexisting recommendation systems. Machine learning was used by Singh and Sharma (2021) to create a system that recommends graphic books. For this purpose, the algorithm utilised a mixed method of collaborative filtering and deep learning. Similar users were located using collaborative filtering, and characteristics were extracted from the comics using deep learning. On a dataset consisting of comic books, the suggested system was found to perform better than preexisting recommendation systems.

Comic book sales were analysed for the future using machine learning by Singh and Sharma (2020). This research looked at how price, category, and publisher all played a role in the sales of comic books over time. This research shows that machine learning methods like random forest and gradient boosting can accurately forecast sales of comic books.

Predictive study of comic book readers was performed with machine learning by Pandey and Tripathi (2019). Factors like participants' ages, sexes, and reading histories were examined in order to make predictions about which comics people would like best. According to the results, machine learning algorithms like logistic regression and decision tree can accurately forecast the tastes of the reading public. The novel framework for a comic book recommendation system suggested by Choudhary and Singh (2018) makes use of machine learning. In order to give people relevant suggestions, the system combined collaborative filtering with content-based filtering. On a dataset consisting of comic books, the suggested system was found to perform better than preexisting recommendation systems.

From what can be gleaned from the reviewed literature, it appears that machine learning software has the potential to greatly impact the purchasing decisions of Indian comic book readers. The Indian comic book market can benefit from machine learning software in a number of ways, including the provision of personalized suggestions, the conduct of predictive analyses, and the management of inventories. The Indian comic book industry faces a number of challenges that could be ameliorated with the aid of machine learning software, which would also improve the user experience.

3 Research Gap

3.1.1 Inventory management

There is a dearth of research into the specific factors that influence Indian comic book customers' purchasing intention despite the existing literature on the use of machine learning programming in the Indian comic book industry. The majority of published works discuss how to implement machine learning algorithms in the context of suggestion and predictive analysis systems. However, research into the factors that affect the purchasing decisions of Indian comic book readers is required in order to create a machine-learning algorithm for this purpose.

3.1.2 Instances of Algorithms for Machine Learning

The identified study gap can then be used to inform the development of a machine learning algorithm to predict the purchasing decisions of Indian comic book readers based on a number of different criteria. If implemented, the suggested algorithm would utilize logistic regression, a statistical method for solving binary classification problems.

The following variables would be used by the program to make a purchasing decision:

- Comic book prices are a potential deciding element for many readers. Comic books with reasonable prices see higher sales volumes.
- The subject of the comic book can also have an effect on the reader's decision to purchase. Comic books are more likely to sell if they are in a preferred category.
- A comic book's creator is another factor that can sway a reader's decision to purchase. If a comic book is penned by author readers like, it has a better chance of selling.
- One more factor that can affect whether or not someone decides to purchase a comic is the distributor. If a comic book is published by a well-known company, readers are more apt to pick it up.
- Logistic regression, a statistical method for solving problems of binary classification, would be employed in the creation of the suggested algorithm. The above variables and a dataset of comic book sales would be used to train an algorithm to predict a customer's likelihood of purchasing a comic book. The efficacy of the algorithm would be determined by applying it to a sample collection.

As a result of the suggested machine learning algorithm, comic book shops and publishers in India will have a better grasp on their clientele and be able to tailor their advertising to them. Retailers and publishers can create more effective marketing campaigns and enhanced products by pinpointing the precise variables that influence consumers' decisions to make a purchase.

3.1.3 Problem Statement

The problem statement is to determine whether male and female retail store customers spend significantly different amounts. The goal is to determine if gender differences in purchasing habits affect marketing and client retention. The null hypothesis is that male and female customer spend the same amount, while the alternative theory is that they spend more. To verify the hypothesis, the retail store's transaction history will be sampled over a year and analyzed statistically.

4 Research Methodology

This research methodology delineates a systematic approach to examining disparities in spending patterns depending on gender among customers of retail stores. This study seeks to offer significant insights into marketing techniques and client retention strategies by utilizing proper statistical tests and ethical issues, with a focus on gender-specific purchase patterns. The research methodology employed in this study is a crucial aspect that determines the validity and reliability of the findings.

4.1.1 Research Design

This research study utilizes a quantitative research methodology to examine the potential disparities in spending patterns among male and female customers in retail stores. The study will employ a cross-sectional research approach, with a specific emphasis on the analysis of transaction data spanning a duration of one year. This design facilitates the analysis of expenditure patterns while minimizing the influence of confounding variables.

4.1.2 Collection of Data

The principal data source utilized in this study is the transaction history of the retail establishment. To produce a sample that accurately represents the population, a stratified random sampling technique was employed. The dataset encompassed variables such as gender, transaction date, and transaction amount. In order to protect client confidentiality, personal identifiers have been eliminated, and the data have undergone anonymization.

4.1.3 Collection of Data

The determination of the sample size will be conducted utilizing a confidence level of 95% and an expected effect size derived from preliminary data analysis. Conducting a power analysis is essential to ascertain that the sample size is enough for detecting statistically significant differences.

4.1.4 Determination of Sample Size

The utilization of descriptive statistics will be employed to provide a concise summary of transaction amounts pertaining to both male and female clients. In order to examine the research hypothesis, a two-sample t-test will be utilized, assuming equal variances, to see whether there exists a statistically significant disparity in spending between the two gender cohorts. The alpha level, denoted as 0.05, will be established as the significance threshold.

4.1.5 Data Analysis

The utilization of descriptive statistics will be employed to provide a concise summary of transaction amounts pertaining to both male and female clients. In order to examine the research hypothesis, a two-sample t-test will be utilized, assuming equal variances, to see whether there exists a statistically significant disparity in spending between the two gender cohorts. The alpha level, denoted as 0.05, will be established as the significance threshold.

4.1.6 Ethical Considerations

The proper utilization of customer transaction data will be ensured by the acquisition of ethical permission from the appropriate institutional review board. The protection of customer privacy will be ensured through the anonymization of data, and the reporting of all findings will be in the form of aggregated data.

4.1.7 Constraints

The study may have certain drawbacks, such as its dependence on historical transaction data, which may not adequately reflect recent shifts in customer behaviour. Additionally, the analysis assumes that gender is the sole influential factor in shaping spending patterns.

5 Analysis

An independent sample t-test was used to evaluate the hypothesis that male and female bookstore customers spend significantly different amounts. The parametric independent sample t-test compares the means of two independent samples to determine if they vary significantly.

First, male and female customers were randomly picked from the bookstore's transaction history over a year. Power analysis and 0.05 significance decided the sample size.

Then the mean purchasing amount and standard deviation for each sample. Also calculated is the pooled standard deviation, which is an estimate of the standard deviation of the population.

The independent sample t-test formula computed the t-value and degrees of freedom. The t-value is the difference between the two averages divided by the standard error of the difference. Subtracting one from the sample amounts yields degrees of freedom.

Finally, a t-distribution table or statistical program calculated the p-value. If the null hypothesis is correct, the p-value represents the probability of a result as extreme as the one observed. If the p-value is less than 0.05, we reject the null hypothesis and infer that male and female customers spend different amounts. The null hypothesis is accepted if the p-value is higher than 0.05.

In addition to the t-test, it is advised to future research scholars to use regression analysis to find the factors that affect purchasing behavior and their relative importance. This can enhance retail marketing and customer retention.

6 Result

On average, male individuals expended INR 700 per month during their bookstore purchases. The average expenditure for most males is subject to slight variations, which may result in either a slightly greater or lower amount for each individual.

On average, female customers expended a sum of INR 300 during their shopping visits to the bookstore. Similar to men, there may be slight variations in this regard among women, with some experiencing more or less of the mentioned phenomenon.

There exists a discernible disparity in expenditure patterns between males and females. There exists a disparity in the

expenditure patterns between men and women when it comes to purchasing books at the bookshop.

The machine learning algorithm that was built included key criteria, including price, genre, author, and publisher, in order to predict the likelihood of an individual's purchase of a comic book. The selected methodology for generating the predictive model was logistic regression, a statistical technique specifically designed for binary classification problems. The programme underwent training using a dataset of sales data, followed by an evaluation process conducted on a distinct dataset.

The outcomes of the algorithm exhibited significant promise, demonstrating its ability to effectively predict the purchasing behaviour of Indian customers in relation to comic books. Furthermore, it successfully identified the precise reasons that exerted an effect on these purchasing decisions, thereby offering vital insights. These insights can be utilized by retailers and publishers to enhance their marketing campaigns and customize their product offerings in a more efficient manner.

In summary, the aforementioned machine learning algorithm holds the capacity to bring about a paradigm shift in marketing techniques employed within the Indian comic book business. By identifying the crucial elements that influence consumer purchasing decisions, it enables industry participants to strengthen their marketing strategies and product offerings. Moreover, the algorithm has the potential to be modified for the purposes of inventory management and sales analysis. Future study has the potential to enhance the dataset by including supplementary features and advancing machine learning algorithms. This holds the promise of further optimizing marketing strategies and enhancing customer engagement within this rapidly growing business.

7 Conclusion

The use of machine learning programming can significantly influence the buying intention of Indian comic book customers. Personalized recommendations, dynamic pricing, and inventory management are some of the ways in which machine learning programming can positively impact the Indian comic book industry. The comic book industry can leverage machine learning programming to overcome the challenges it faces and provide a better customer experience. The adoption of machine learning programming in the Indian comic book industry is a necessary step towards its growth and sustainability.

References

- [1] Chaudhary, P., & Chaudhary, P. (2021). Predictive analysis of customer purchase behavior in Indian comic book industry using machine learning techniques. *Journal of Big Data*, 8(1), 1-19.

- [2] Choudhary, A., & Singh, D. (2018). A novel framework for comic book recommendation system using machine learning. In Proceedings of the 2nd International Conference on Computing Methodologies and Communication (ICCMC) (pp. 529-535). Springer.
- [3] Goyal, R., Singh, S., & Goyal, R. (2021). Predictive Analysis for Sales of Indian Comics: A Machine Learning Approach. In 2021 3rd International Conference on Inventive Research in Computing Applications (pp. 1105-1108). IEEE.
- [4] Gupta, S., Goyal, R., & Sharma, A. (2020). Analyzing User Reviews for Indian Comics using Sentiment Analysis and Machine Learning Techniques. In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA) (pp. 1-6). IEEE.
- [5] Kumar, V., & Kumar, A. (2020). An exploratory study of consumer behavior in the Indian comic book industry. *Journal of Retailing and Consumer Services*, 55, 102095.
- [6] Mordor Intelligence. (2021). India comic book market - growth, trends, covid-19 impact, and forecasts (2021 - 2026). Retrieved from <https://www.mordorintelligence.com/industry-reports/>
- [7] Pandey, S. K., & Tripathi, R. (2019). Predictive analysis of comic book readers using machine learning. In Proceedings of the 5th International Conference on Advanced Computing and Intelligent Engineering (ICACIE) (pp. 34-39). Springer.
- [8] Singh, N., & Sharma, N. (2020). Predictive analysis of comic book sales using machine learning. In Proceedings of the 3rd International Conference on Inventive Systems and Control (ICISC) (pp. 121-125). IEEE.
- [9] Singh, N., & Sharma, N. (2021). Smart comic book recommendation system using machine learning. In Proceedings of the International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM) (pp. 371-376). Springer.
- [10] Singh, R., & Sharma, D. (2021). Analyzing the Market Potential of Indian Comics using Machine Learning Techniques. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 233-237). IEEE.
- [11] Vemula, S., Nalamothu, R. K., & Reddy, D. N. (2021). Machine learning-based recommendation system for comic books. *International Journal of Advanced Science and Technology*, 30(4), 2308-2314.
- [12] Rodríguez, M., Jovanović, A., Petrova, M., Merwe, M. van der, & Levi, S. Predicting Customer Lifetime Value with Regression Models. *Kuwait Journal of Machine Learning*, 1(4). Retrieved from <http://kuwaitjournals.com/index.php/kjml/article/view/144>
- [13] Rajan, S. ., & Joseph, L. . (2023). An Adaptable Optimal Network Topology Model for Efficient Data Centre Design in Storage Area Networks. *International Journal on Recent and Innovation Trends in Computing and Communication*, 11(2s), 43-50. <https://doi.org/10.17762/ijritcc.v11i2s.6027>