IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 4, August 2026, pp. 3528~3536
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i4.pp3528-3536

A machine learning framework for predicting and optimizing return on investment across marketing channels

Chandra Chathura, Keerthan Saya, Sathishkumar Mani
Department of Computer Science and Engineering, GITAM School of Technology, GITAM University, Bengaluru, India

Article Info

Article history:

Received May 2, 2025
Revised Jun 17, 2026
Accepted Jul 9, 2026

Keywords:

Business intelligence
Data-driven decision making
Machine learning
Marketing channel optimization
Return on investment

ABSTRACT

In the current fast-paced competitive marketing environment, firms require data-centric methods to maximize their investments in several avenues. This research work applies machine learning techniques to estimate the return on investment (ROI) for marketing costs, which helps organizations in budgeting more effectively. Four models including random forest, extreme gradient boosting (XGBoost), gradient boosting, and linear regression were utilized for their accuracy in making predictions. Results showed that the highest accuracy was achieved by linear regression at 99.39%, random forest at 99.07%, gradient boosting at 99.01%, and XGBoost at 98.81%. It was further noted that digital marketing avenues such as social media and online stores gave the highest ROI, indicating that companies should prioritize digital marketing more than traditional marketing. On a practical level, this approach helps marketing team for choosing high performing channels since it estimates expected returns from each marketing channels and make smarter budget allocation. Yet, the study is done by using Kaggle dataset. In order to improve its generality, future research may use larger real-world datasets and extensive visualization techniques.

Corresponding Author:

Sathishkumar Mani
Department of Computer Science and Engineering, GITAM School of Technology, GITAM University
Bengaluru, India
Email: sathishkumarmani17@gmail.com

1. INTRODUCTION

Marketing has undergone huge changes, moving from traditional marketing like television (TV), billboards, and newspapers to modern marketing like social media, online stores, and mobile applications. Technology development is one of the main reasons behind this and also change in customer behavior. Companies desire to reach consumers efficiently and grow their business while adapting to technological changes also adds to this. Marketing is one of the main aspects which brings consumers close to companies products and drives sales which results in profits. The biggest challenge for companies is to identify the right marketing channel to achieve high return on investment (ROI), and it is important to understand what traditional marketing and modern marketing are bringing to the table for companies. In order to cover mass audiences in the market companies usually go for traditional marketing like billboards, TV, and newspapers. These channels result in covering a large volume of audience, helps in increasing brand recognition but these channels fail in directly impacting targeted consumers. Further shaping marketing strategies in real-time to be precise personalized advertising, since companies are more focused on tailored campaigns according to targeted audiences. Instead of using data-driven insights most of the businesses rely on past experiences and industry trends while allocating marketing budgets. This may result in wasting resources and poor investment choices which again result in decreasing profitability. By determining the most efficient channels to invest in and adjusting the spending helps business to overcome these challenges. To make this possible, marketing campaigns require continuous evaluation to maintain long-term growth, further it should have the capacity to forecast accurately and find ROI from various marketing channels by using machine learning. Machine learning and predictive analytics present an opportunity for revolutionizing marketing decision-making by predicting potential returns using past data.

Yahia and ElBolok [1] offered a stochastic nonlinear programming approach for optimizing budget allocation for digital marketing campaigns; however, the paper did not provide a comparative study for different learning models in terms of marketing channels but focused more on mathematical optimization rather than predicting ROI using machine learning models. Nayyar *et al.* [2] explored data-driven machine learning techniques to optimize advertising campaigns in modern media and focused only on campaign optimization strategies and did not provide an analysis of machine learning models for ROI predictions with various marketing channels. Ma and Sun [3] discussed the importance of machine learning to improve human decision-making in marketing. This work emphasizes matching computational models with consumer psychology and approaches. It doesn't include forecasting of financial performances like ROI prediction or indicating quantitative marketing investment optimization or campaign budgeting. Miklosik and Evans [4] provide an overview of how big data and machine learning techniques are changing market sector. Brei [5] provides an overview of machine learning strategies in marketing. This paper does not focus on a particular application model nor quantitative assessment or algorithm comparison of ROI. Herhausen *et al.* [6] discuss the changing landscape of machine learning for marketing applications. This work requires more empirical studies with comparisons of machine learning models.

Chen [7] discusses machine learning for e-commerce strategy, but it doesn't concentrate on ROI forecasts or multi-channel marketing effectiveness, which restricts its usage in strategic budget planning. Nordin and Ravald [8] examine decision-making in changing marketing environments. Ngai and Wu [9] demonstrated a framework for understanding machine learning that can augment marketing efficiency. This model outlines many machine learning techniques for targeting and personalization. Miklosik *et al.* [10] explored the usage of machine learning tools. The ROI predictive model is not included in this paper, and tools are not tested against quantitative performance measures. Vashishth *et al.* [11] share trends in artificial intelligence adoption. Andayani *et al.* [12] examine application of big data but didn't discuss channel specific ROI forecasting. Saba *et al.* [13] surveys machine learning applications in marketing transformation. This paper lacks in elaborating model methods, evaluation metrics and practical implementation for ROI estimation. Kumar [14] described a marketing support hybrid model, which concentrates on data mining rather than financial return prediction and lacks in marketing channel breakdowns or investment analysis. Zaghloul *et al.* [15] research involves a comparison of traditional and deep methods in predicting consumer satisfaction in an e-commerce context, although these researchers specifically examine results regarding customer experience and not ROI prediction or budget distribution in a variety of online marketing channels. Sadrnia [16] focuses on predictive modelling for campaign strategy. Gooljar *et al.* [17] systematically reviewed the use of predictive models based on sentiments for purchases conducted online, in relation to "Marketing 5.0" specifically, although their research is centered on the analysis of consumer sentiments rather than directly regarding the prediction in terms of ROI return or budget distribution in various available online marketing channels. Sangsawang [18] discussed work on click-through rates (CTR) prediction rather than multi-channel investment analysis.

2. METHOD

The methodology plays a significant role in promoting user experience and decision making. This predictive model starts with the channels that the user can invest in, and users select one or more channels. Then the user inputs the amount to invest in the marketing channels. The model then predicts the ROI and the system produces a pie chart which represents the contribution of each marketing channel to overall ROI. This enables businesses to make informed decisions on the channels that they can invest in and avoid the channel where the ROI will be low, hence saving on costs. Figure 1 shows the systematic workflow of the predictive machine learning model developed for marketing ROI optimization. The process begins with the dataset loading. It contains historical marketing performance data, such as expenses on campaigns, metrics of customer engagement, and revenue generated. This information forms the foundation for training and testing machine learning models. Then the dataset is preprocessed, which includes cleaning and transformation of the raw dataset. Missing values are handled, numerical attributes are normalized, categorical variables are encoded, and key performance indicators are derived. Then feature extraction is done, based on profit performance and campaign analysis. The dataset is divided into testing data and training data after preprocessing. The training sample is used to build machine learning models, and the test sample is kept for later to measure the accuracy and generalization of the models.

To predict ROI from marketing expenses, the study relies on algorithms like random forest regressor [19], [20], gradient boosting [21], extreme gradient boosting (XGBoost) [22], [23], and linear regression [24], [25]. These models pick up on patterns and relationships between marketing costs and profits, helping the system make knowledgeable decisions based on the numbers. Upon completion of the training phase, the next step is the generation of the ROI prediction. In this phase, return estimates for each marketing channel are created based on the user's input regarding the amount to invest in each of them. This phase aims to provide businesses with insights into prospective returns before committing resources. Further, the accuracy evaluation phase is incorporated to ensure the reliability of the system.

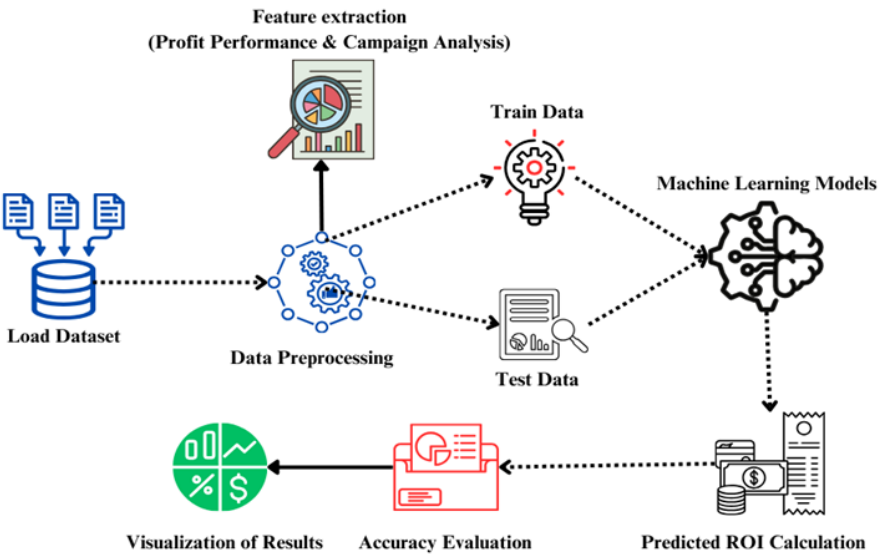

Figure 1. Machine learning framework for marketing ROI prediction

2.1. Experimental setup

The modern tools and languages like Python, Pandas, and Matplotlib are adopted for the proposed work for data manipulation and data visualization. Further scikit-learn for machine learning model which includes R-squared (R^2) score, mean squared error (MSE) for model evaluation. These libraries provide the foundational tools needed for data processing, visualization, and machine learning model development. Data preprocessing is carried out by ensuring there are no missing values. Column means are used to replace missing numerical values to avoid losing the data and in further it ensures consistency. The profit margin % is computed, which is one of the most important financial metrics for assessing profitability for different marketing campaigns. Selection of the feature set plays an important role in this work, so that variables which are relevant and significant in predicting ROI are identified. An open-source distribution called anaconda is used in this work to manage libraries required for machine learning models and python is written and executed using the open-source environment called Jupyter Notebook. This work was simulated using 16 GB RAM, Intel Core i5 processor, and 120 GB of hard disk space in Windows 10 operating environment.

2.2. Data source

The dataset for the analysis of marketing channels and predicting ROI is taken from Kaggle. The dataset contains some significant parameters like money invested, product type, and targeted people for specific marketing channels. Only the parameter variables and output variables are fed into the system to avoid logistical and privacy concerns. The dataset has 60,000 rows and 12 columns for 6 different marketing channels and the marketing channels are social media, TV, mobile applications, billboard, newspapers, and online store.

2.3. Algorithms and performance evaluation metrics

Machine learning and predictive analytics present an opportunity for revolutionizing marketing decision-making by predicting potential returns using past data. This work aims at creating a predictive machine learning model that predicts marketing ROI and helps companies to maximize their marketing efforts. This model uses machine learning methods, for instance, random forest regressor, gradient boosting, XGBoost, and linear regression to predict ROI accurately and consistently. This research work is intended to create a bridge between traditional marketing and digital marketing by providing a deep analysis of each of them. Digital marketing allows for advanced targeting and tracking, but traditional marketing is still good for reaching broad demographics. Companies that apply both approaches receive the best results because they create the most balanced marketing mix and get the highest ROI. Companies that use machine learning for marketing analytics have a competitive advantage because they can make their marketing investment much more efficient and receive the highest possible return.

The accuracy is evaluated by comparing the forecasted ROI to the actual test data. The performance indicators used for evaluation of machine learning algorithms accuracy are MSE as in (1), R^2 as in (2), and mean absolute error (MAE) as in (3). The mathematical models to compute the performance indicators are as follows.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{1}$$

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y})^2}{\Sigma(y_i - \bar{y})^2} \tag{2}$$

$$MAE = \frac{1}{n} \times \Sigma|y_i - \hat{y}_i| \tag{3}$$

Where n is number of data points, y_i is actual value, $\hat{y}_i$ is predicted value, $(y_i - \hat{y}_i)^2$ is squared error for the i-th data point, $\bar{y}$ is mean of actual values, and Σ is summation notation.

2.4. Data preprocessing

Data preprocessing is essential for data analysis and machine learning, involving steps like data collection, cleaning, integration, transformation, and reduction, ultimately leading to improved model performance through careful dataset preparation. Data cleaning involves identifying and rectifying errors and inconsistencies in data such as filling in missing values, correcting inaccuracies, and addressing outliers to enhance the dataset's quality and readiness for analysis or modelling. Data integration merges data from diverse sources into a coherent dataset, while data transformation prepares that data for analysis by converting it into an appropriate format and ensuring feature comparability. Data reduction involves minimizing data volume while maintaining key information, enhancing computational efficiency and model simplicity and feature selection, which help prevent overfitting and improve processing speed without greatly compromising performance. Data splitting involves dividing a dataset into distinct subsets for model development typically a training set for fitting the model, a validation set for parameter tuning, and a test set for final performance evaluation. This research work adopts 80% training and 20% testing ratio to ensure the model generalizes effectively to unseen data.

3. RESULTS AND DISCUSSION

The visualization aids businesses in identifying the most efficient channels and allocating resources accordingly. The significance of selected machine learning algorithms is justified in the following section. Random forest regressor is a model which represents many decision trees. Each tree on their own learns from a subset of dataset and gives its own prediction. The final answer is the average of all of the predictions. It allows to decrease overfitting, to make the model more robust [20]. In the context of predicting marketing ROI, this model demonstrates how complicated the relationship is between the invested money and revenue. XGBoost regressor is an advanced boosting algorithm. It improves decision trees via an iterative learning procedure. XGBoost constructs trees in serial manner and every next tree corrects the errors of previous trees.

The final prediction is the aggregated value that the algorithm provides after the necessary optimization [22], [23]. This approach is especially useful in marketing predictions. Since it is capable of handling large data and reducing error rates. Linear regression provides a straightforward and interpretable baseline for predicting ROI, reasonably effective at modeling subtle relationships [24], [25]. Figure 2 is the scatter plot of predicted ROI values versus actual ROI values when employing the linear regression model. The ideal fit y=x represented by a dashed red line signifies perfect prediction where predicted and actual values are equivalent. Blue points signify individual predictions, and the dot's position on the red line indicates the accuracy of the model's fit. Gradient boosting regressor builds several decision trees one after another the dataset (D) is divided into subsets (D1, D2, ..., Dn). Each subset is used for training the decision tree and each tree learns from the mistakes of the previous tree. The final result is obtained by averaging the predictions of all trees, which achieves higher accuracy [21]. This approach is highly recommended for marketing analytics because it helps to predict the ROI with high accuracy and also decreases the error rate.

To compare the effectiveness of the investment, product performance analysis (Figure 3) and marketing campaign analysis (Figure 4) are carried out upon preprocessing of the dataset. Identification of the most and least profitable products can be done using a bar chart that compares the total profit and total cost for each product. Figure 4 representing the percentage of total revenues versus total expenses, further it provides an overall perspective on firm profitability. Advertising strategy can be optimized by campaign analysis of marketing campaigns representing the overall cost incurred for each campaign and the number of channels utilized for marketing. The ROI contribution of each marketing channel and the breakdown of marketing costs on a channel-wise basis are represented using a pie chart; it helps the model to identify the most profitable marketing channel for investments. To ensure better decisions in budget allocation, this model forecasts the ROI from various marketing channels. Random forest regressor algorithm averages the outcomes and minimizes variance using decision trees.

The XGBoost regressor is an advanced gradient boosting model, and it can manage the missing data efficiently and continually enhance the precision. The linear regression is simple and understandable model to predict ROI, and the gradient boosting regressor enhances the predictions by correcting the errors of the preceding models. The quality of each model is measured by the R^2 coefficient, which indicates the proportion of variance in ROI. The MSE, indicates the error in prediction. The system compares all models and chooses the best prediction. This is achieved through a pie chart, which shows the percentage of the contribution of each marketing channel to the total ROI. The actual ROI and predicted ROI are calculated using (4) and (5).

$$Actual\ ROI\ (\%) = \frac{Return - Investment}{Investment} \times 100 \quad (4)$$

$$Predicted\ ROI\ (\%) = f(\text{marketing channel spend, cost of product, profit margin (\%)}) \quad (5)$$

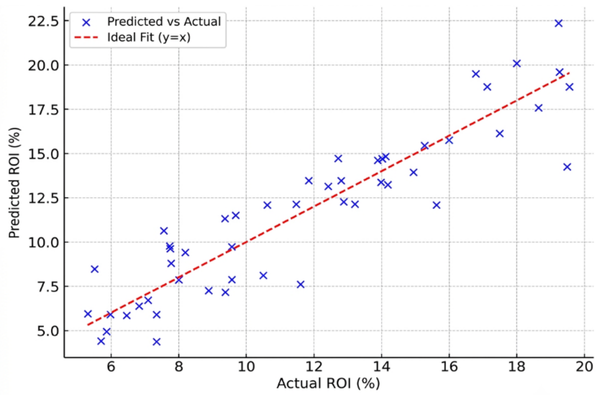

Figure 2. Predicted vs actual ROI using linear regression

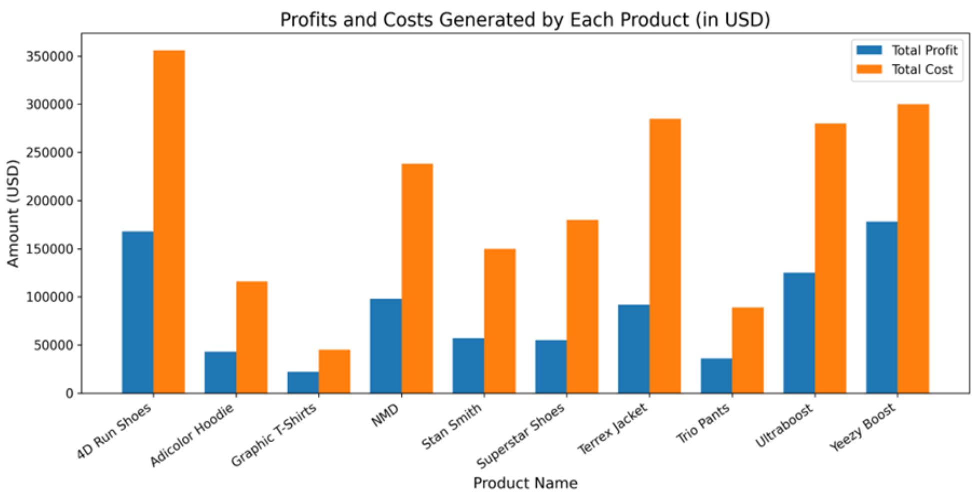

Figure 3. Product-wise profit and cost comparison

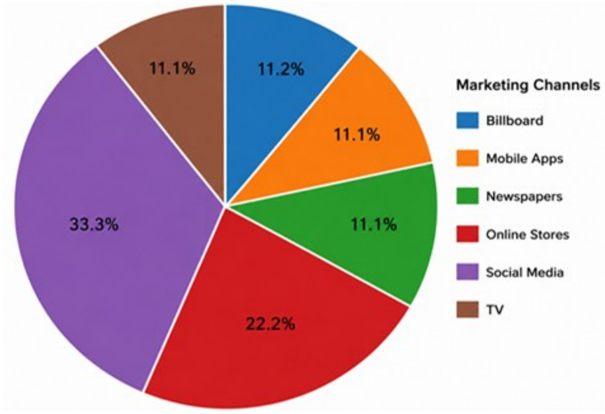

Figure 4. ROI distribution across marketing channels

The model accuracy is checked by forecasting the ROI and comparing it with the entered user value for actual ROI. For each machine learning model, absolute percentage error (APE) is computed by the system to identify the prediction accuracy. Table 1 has the actual ROI value, the predicted ROI value, and the percentage of accuracy. This final step aims at helping companies select the most accurate machine learning algorithm to increase their marketing return. Generally, this deployment merges data visualization and machine learning-based ROI prediction, enabling companies to make investment decisions based on data, reduce inefficiencies, and increase profitability. Linear regression gives high accuracy in predicting ROI, further it is evident that this algorithm can be used to model and demonstrate the interaction between marketing expenditure and ROI effectively. Performance of models is evaluated based on MSE and R^2 which explains how effectively each model reduces prediction error and explains variation of ROI. The performance of the random forest regressor lies between XGBoost, which had the lowest accuracy, and linear regression, which showed the highest. It is able to predict ROI outcomes with greater fidelity than XGBoost and gradient boosting, although it is slightly less precise compared to linear regression.

The random forest regressor exhibits good overall performance. In general, it performs well with both fitting and generalization models to the data. XGBoost has also shown strong generalization capabilities. It is very strong at finding general patterns in the dataset. However, it wasn't among the best ones in the competition. It had larger error rates compared to other models. This means that it might not be the best fit for the dataset. It is exceptional at handling complex features. The gradient boosting regressor has shown better performance than XGBoost. It has a lot lower prediction error demonstrated, the effectiveness in learning from data patterns, and reducing inaccuracies. Its ability to capture complex relationships within the data made it one of the most accurate and consistent models in the study.

Linear regression model delivered a surprise, as it was rated best in terms of prediction accuracy. The model has shown the smallest error, as it has almost perfectly embraced the data. This implies that the relationship between input features and ROI is highly linear, allowing the linear regression model to provide an almost accurate prediction. As per the performance, linear regression model appears to be the best model to predict the ROI in such circumstances. Although the variation in accuracies among models are very minimal, it is evident that all models were effective in ROI prediction for marketing expenditure, further supporting the efficacy of machine learning in marketing budget optimization. Table 1 presents a comparison of actual and predicted ROI of different marketing channels for the four machine learning models also provides more insights. The overall accuracy as in Figure 5 of each model is the average of the accuracies of all individual marketing channels.

Table 1. Actual and predicted ROI values across marketing channels

Marketing channel	Random forest regressor		XGBoost		Gradient boosting regressor		Linear regression	
	Actual ROI (%)	Predicted ROI (%)	Actual ROI (%)	Predicted ROI (%)	Actual ROI (%)	Predicted ROI (%)	Actual ROI (%)	Predicted ROI (%)
Billboard	11.22	11.1	11.22	11.1	11.22	11.2	11.22	11.1
Newspaper	11.22	11.1	11.22	11.1	11.22	11.1	11.22	11.2
Social media	33.29	33.5	33.29	33.4	33.29	33.3	33.29	33.2
TV	11.21	11.1	11.21	11	11.21	11	11.21	11.1
Online stores	21.87	22.1	21.87	22.3	21.87	22.3	21.87	22.1
Mobile applications	11.19	11.1	11.19	11.1	11.19	11.1	11.19	11.2
Overall accuracy	99.07		98.81		99.01		99.39	

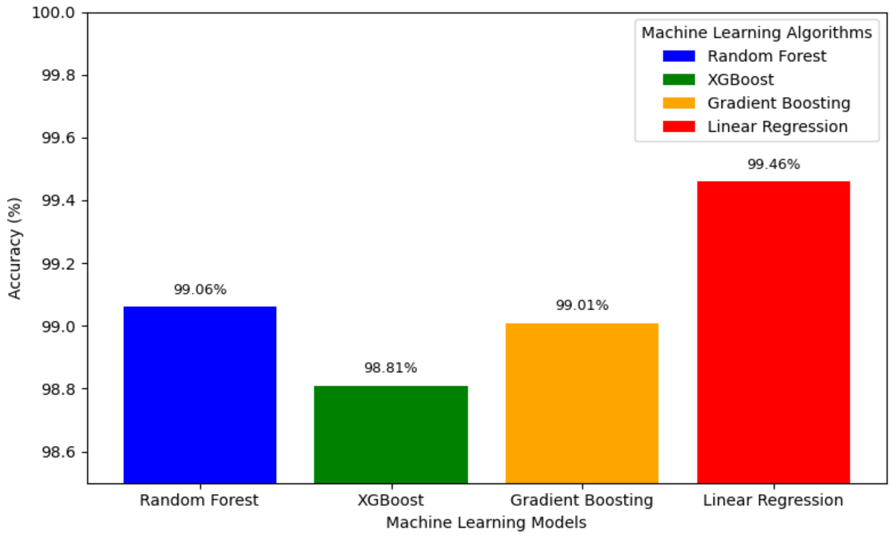

Figure 5. Accuracy comparison of machine learning models

The result shows that linear regression has the highest overall accuracy at 99.39%, then random forest at 99.07%, followed by gradient boosting at 99.01%, and lastly XGBoost at 98.81%. This is proven by the MSE and R^2 values that also prove that linear regression was the most effective in ROI prediction. The analysis is evident that machine learning is also effective in optimization of marketing budget. Machine learning algorithms performed well in analyzing marketing channels individually and provided accurate forecasts, further the variation in forecast is very minimal among algorithms. Channels of social media and online stores always had the highest ROI, which corresponds to the trend of higher profitability of digital marketing strategies over the traditional ones, such as billboards and newspapers.

The analysis gives an insight that investing in digital platforms is more profitable than investing in traditional advertising. The predicted ROI of the channels of TV, billboards, and newspapers coincided with the actual ROI most of all in all the models, which shows that the traditional channels of advertising have relatively stable ROI, and less profitable than digital channels, but the ROI of these channels is more predictable and less dynamic. Among the ensemble models, random forest and gradient boosting performed slightly better than XGBoost. Overall results are also validating that machine learning algorithms do have the ability to generate high accurate ROI estimates, enabling organizations to make solid investment decisions backed by data and high confidence levels. The detailed analysis indicates that digital marketing channels dominated the ROI predictions and suggests that businesses should prioritize investments in social media and online. The study shows that machine learning-based ROI predictions are useful for decision making in marketing, allowing businesses to distribute their budget properly, minimize financial risks and, therefore, maximize the ROI. Though results have been shown within the present experiment, it is possible to dive deeper into the results by conducting experiments on bigger datasets. Additional research work could be done to improve interpretability methods to further improve the application of the approach.

4. CONCLUSION

This research work was conducted to incorporate machine learning methods for the prediction of the ROI of marketing campaigns. The research provides businesses evidence-based recommendations on how to modify their budget distribution. Further this research work has shown that machine learning techniques can efficiently predict marketing performance with a high level of accuracy. The result shows that the best model for this type of analysis is linear regression, which has achieved the lowest MSE and the highest R^2 score, further it indicates that the dataset follows a highly linear relationship between input features and ROI. The highest ROI was regularly derived from digital marketing channels, for example, social media and online stores. This research work confirms digital marketing investments are ahead of traditional methods. Predictive analytics has proven to be really helpful for making marketing decisions. The best accuracy is achieved by linear regression at 99.39%, followed closely by random forest at 99.07%, gradient boosting at 99.01%, and XGBoost at 98.81%. The implemented model does not capture sudden behavioral changes since this research is based on historical data. Subsequent research can be done by using extensive datasets which include data collected over time, the behavior of consumers and the tactics of the price dynamics to enhance the accuracy and personalized predictions.

FUNDING INFORMATION

No funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Chandra Chathura	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
Keerthan Saya			✓	✓	✓	✓	✓	✓	✓	✓	✓			
Sathishkumar Mani		✓		✓	✓	✓				✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

No conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [SM], upon reasonable request.

REFERENCES

[1] Z. Yahia and M. ElBolok, "A stochastic nonlinear programming model for budget mix optimization of digital marketing campaigns under uncertainty," *Future Business Journal*, vol. 11, no. 1, Oct. 2025, doi: 10.1186/s43093-025-00664-x.

[2] G. Nayyar, S. Suman, and T. R. K. Lakshmi, "Optimizing ad campaigns with machine learning: data-driven approaches in modern media," *International Journal of Multiphysics*, vol. 18, no. 3, pp. 1746–1754, 2024, doi: 10.52783/ijm.v18.1489.

[3] L. Ma and B. Sun, "Machine learning and AI in marketing – connecting computing power to human insights," *International Journal of Research in Marketing*, vol. 37, no. 3, pp. 481–504, Sep. 2020, doi: 10.1016/j.ijresmar.2020.04.005.

[4] A. Miklosik and N. Evans, "Impact of big data and machine learning on digital transformation in marketing: a literature review," *IEEE Access*, vol. 8, pp. 101284–101292, 2020, doi: 10.1109/ACCESS.2020.2998754.

[5] V. A. Brei, "Machine learning in marketing: overview, learning strategies, applications, and future developments," *Foundations and Trends® in Marketing*, vol. 14, no. 3, pp. 173–236, Aug. 2020, doi: 10.1561/1700000065.

[6] D. Herhausen, S. F. Bernritter, E. W. T. Ngai, A. Kumar, and D. Delen, "Machine learning in marketing: recent progress and future research directions," *Journal of Business Research*, vol. 170, Jan. 2024, doi: 10.1016/j.jbusres.2023.114254.

[7] N. Chen, "Research on e-commerce database marketing based on machine learning algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, pp. 1–13, Jun. 2022, doi: 10.1155/2022/7973446.

[8] F. Nordin and A. Ravald, "The making of marketing decisions in modern marketing environments," *Journal of Business Research*, vol. 162, Jul. 2023, doi: 10.1016/j.jbusres.2023.113872.

[9] E. W. T. Ngai and Y. Wu, "Machine learning in marketing: a literature review, conceptual framework, and research agenda," *Journal of Business Research*, vol. 145, pp. 35–48, Jun. 2022, doi: 10.1016/j.jbusres.2022.02.049.

[10] A. Miklosik, M. Kuchta, N. Evans, and S. Zak, "Towards the adoption of machine learning-based analytical tools in digital marketing," *IEEE Access*, vol. 7, pp. 85705–85718, 2019, doi: 10.1109/ACCESS.2019.2924425.

[11] T. K. Vashishth, V. Sharma, K. K. Sharma, B. Kumar, S. Chaudhary, and R. Panwar, "Embracing AI and machine learning for the future of digital marketing," in *AI, Blockchain, and Metaverse in Hospitality and Tourism Industry 4.0*, New York: Chapman and Hall/CRC, 2024, pp. 90–117, doi: 10.1201/9781032706474-7.

[12] D. Andayani, M. Madani, H. Agustian, N. Septiani, and L. W. Ming, "Optimizing digital marketing strategies through big data and machine learning: insights and applications," *Journal of Computer Science and Technology Application*, vol. 1, no. 2, pp. 104–110, Aug. 2024, doi: 10.33050/corisinta.v1i2.29.

[13] N. S. Saba, R. Gandhi, S. R. Rajendran, and N. D. Abraham, "Revolutionizing digital marketing using machine learning," in *Contemporary Approaches of Digital Marketing and the Role of Machine Intelligence*, New York, United States: IGI Global Scientific Publishing, 2023, pp. 1–22, doi: 10.4018/978-1-6684-7735-9.ch001.

[14] T. S. Kumar, "Data mining-based marketing decision support system using hybrid machine learning algorithm," *Journal of Artificial Intelligence and Capsule Networks*, vol. 2, no. 3, pp. 185–193, Aug. 2020, doi: 10.36548/jaicn.2020.3.006.

[15] M. Zaghloul, S. Barakat, and A. Rezk, "Predicting E-commerce customer satisfaction: traditional machine learning vs. deep learning approaches," *Journal of Retailing and Consumer Services*, vol. 79, Jul. 2024, doi: 10.1016/j.jretconser.2024.103865.
[16] L. Sadrnia, "The future of marketing: how predictive modeling optimizes campaign strategies," *iBusiness*, vol. 15, no. 4, pp. 249–262, 2023, doi: 10.4236/ib.2023.154018.
[17] V. Gooljar, T. Issa, S. H. -Ramanan, and B. A. -Salih, "Sentiment-based predictive models for online purchases in the era of marketing 5.0: a systematic review," *Journal of Big Data*, vol. 11, no. 1, Aug. 2024, doi: 10.1186/s40537-024-00947-0.
[18] T. Sangsawang, "Predicting ad click-through rates in digital marketing using support vector machines," *Journal of Digital Market and Digital Currency*, vol. 1, no. 3, pp. 225–246, Dec. 2024, doi: 10.47738/jdmdc.v1i3.20.
[19] B. H. Hayadi and I. M. M. El Emary, "Predicting campaign ROI using decision trees and random forests in digital marketing," *Journal of Digital Market and Digital Currency*, vol. 1, no. 1, pp. 1–20, Jun. 2024, doi: 10.47738/jdmdc.v1i1.5.
[20] K. Jin, Z. Z. Zhong, E. Y. Zhao, "Sustainable digital marketing under big data: an AI random forest model approach," *IEEE Transactions on Engineering Management*, vol. 71, pp. 3566–3579, 2024, doi: 10.1109/TEM.2023.3348991.
[21] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, Oct. 2001, doi: 10.1214/aos/1013203451.
[22] C. Şahin, "Predicting base station return on investment in the telecommunications industry: machine-learning approaches," *Intelligent Systems in Accounting, Finance and Management*, vol. 30, no. 1, pp. 29–40, Jan. 2023, doi: 10.1002/isaf.1530.
[23] S. Lakshminarayanan, R. Seranmadevi, N. B. Muddangala, S. Susendiran, S. S. Sundar, and B. Namasivayam, "Advanced data science techniques for dynamic optimization of digital marketing campaigns: leveraging reinforcement learning, NLP, and gradient boosting," in *3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024)*, 2025, pp. 1–11, doi: 10.2139/ssrn.5086666.
[24] K. Chen, L. Dong, and L. Wang, "Linear regression model for business strategy: a case study of SMARTFOOD company," in *2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022)*, 2022, pp. 1670–1677, doi: 10.2991/aebmr.k.220307.273.
[25] Y. Shi, "Application of improved linear regression algorithm in business behavior analysis," *Procedia Computer Science*, vol. 228, pp. 1101–1109, 2023, doi: 10.1016/j.procs.2023.11.144.

BIOGRAPHIES OF AUTHORS

Chandra Chathura is a scholar of GITAM University Bengaluru, Karnataka, India and currently pursuing the Bachelor of Technology degree in Computer Science and Business Systems. Her academic and project work reflect a strong interest in data science and machine learning, with a particular focus on their applications in business. Her research interests lie at the intersection of data science, business intelligence, marketing analytics, and applied machine learning. She is is passionate about combining technical expertise with business insights. She can be contacted at email: workwithchathurachandra@gmail.com.

Keerthan Saya is a scholar of GITAM University Bengaluru, Karnataka, India and currently pursuing the B.Tech. degree in Computer Science and Business Systems. He has participated in projects exploring the intersection of business analytics and financial modeling, aiming to derive actionable insights for strategic planning. His interests lie in financial analytics, business intelligence, FinTech, and the application of data-driven methods in modern business environments. He can be contacted at email: keerthan775@gmail.com.

Sathishkumar Mani has obtained his B.E. degree in Computer Science and Engineering from Bharathiar University, Coimbatore, India and M.Tech. degree in Information Technology from Punjabi University, Patiala, India. He earned his Ph.D. in Computer Science and Engineering from Saveetha University, Chennai, India. He is currently an associate professor in the Department of Computer Science and Engineering at GITAM University in Bengaluru, India. He has over 25 years of experience in multiple domains like teaching, research and software development. His research area is network security, machine learning, and internet of things. He can be contacted at email: sathishkumarmani17@gmail.com.