International Journal of E-Business Research
Volume 19 • Issue 1

Assessing the Predictive Performance of Machine Learning in Direct Marketing Response

Youngkeun Choi, Sangmyung University, South Korea*
Jae W. Choi, University of Texas at Dallas, USA

ABSTRACT

This paper intends to better understand the pre-exercise of modeling for direct marketing response prediction and assess the predictive performance of machine learning. For this, the authors are using a machine learning technique in a dataset of direct marketing, which is available at IBM Watson Analytics in the IBM community. In the results, first, among all variables, customer lifetime value, coverage, employment status, income, marital status, monthly premium auto, months since last claim, months since policy inception, renew offer type, and the total claim amount is shown to influence direct marketing response. However, others have no significance. Second, for the full model, the accuracy rate is 0.864, which implies that the error rate is 0.136. Among the patients who predicted not having a direct marketing response, the accuracy that would not have a direct marketing response was 87.23%, and the accuracy that had a direct marketing response was 66.34% among the patients predicted to have a direct marketing response.

KEYWORDS

Decision tree, Direct marketing response prediction, Machine learning

1. INTRODUCTION

In the realm of product advertising and promotion, there exist two distinct approaches, namely, mass marketing and direct marketing (Roddy, 2002). Mass marketing utilizes various mass media channels to disseminate product-related information to both current and potential customers. These channels typically include television, radio, magazines, and newspapers, and the marketing messages are targeted toward large customer segments. Mass marketing does not discriminate against individual customers within the group, and the information delivered remains consistent for all. On the other hand, direct marketing focuses on targeting specific individuals or households with tailored marketing messages.

Direct marketing has become an increasingly important strategy for businesses seeking to reach their target customers (Poon et al., 2017). However, the success of direct marketing campaigns depends on the

DOI: 10.4018/IJEBR.321458
*Corresponding Author

ability to accurately predict customer behavior and preferences. Inaccurate predictions can lead to wasted resources, ineffective campaigns, and a poor return on investment. Therefore, predicting direct marketing outcomes has become a critical area of research in marketing. Accurate predictions can help businesses optimize their campaigns, targeting the right customers at the right time with the right message. By using advanced data analytics and machine learning techniques, businesses can identify patterns in customer behavior and tailor their marketing strategies accordingly. Furthermore, predicting direct marketing outcomes can also help businesses stay ahead of the competition. As the marketing landscape becomes more competitive, businesses need to find ways to differentiate themselves and provide more personalized experiences to their customers. Predictive analytics can help businesses gain a deeper understanding of their customers, allowing them to deliver more targeted and effective marketing campaigns.

Direct marketing relies on the creation of precise predictive models within databases and is an area that has the potential to benefit significantly from such models (Berger & Magliozi 1992). As an increasing number of companies adopt direct marketing as a distribution strategy, there has been a surge in spending on these channels, with direct marketing staff prioritizing consumer response modeling to enhance revenue, minimize costs, and improve overall profitability. In addition to the traditional statistical approach to forecasting consumer purchases, researchers have recently applied data mining techniques to large, noisy databases, which offer several unique advantages.

Especially, there are several reasons why a new data analysis method is needed for predicting direct marketing (Sharma et al., 2022). First, traditional methods such as logistic regression have limitations in handling large and complex datasets. With the increasing amount of data being generated in today's digital age, there is a need for more powerful and scalable algorithms. Second, consumer behavior and preferences are constantly evolving, making it challenging to build accurate predictive models that can adapt to these changes. Third, the rise of new technologies such as artificial intelligence and machine learning has opened up new possibilities for predictive modeling, and there is a need to explore these methods to improve the accuracy and efficiency of direct marketing predictions. Overall, a new data analysis method for direct marketing prediction is needed to keep up with the changing landscape of consumer behavior and technology, and to improve the effectiveness of marketing campaigns.

In recent years, there has been a growing trend to use machine learning techniques to perform tasks associated with artificial intelligence (AI) (Simon, 2019). These tasks include recognition, analysis, planning, robot control, and forecasting, among others. One example of a well-known US company that has applied AI-powered data analysis methods for predicting direct marketing outcomes is IBM. IBM's Watson Marketing Insights platform uses machine learning algorithms to analyze customer data and predict which marketing campaigns are most likely to generate the desired response from customers (Hassan, 2022). In a real-world application of this technology, a large insurance company used Watson Marketing Insights to analyze their customer data and identify the most effective channels for reaching different customer segments. The analysis revealed that email marketing was the most effective channel for one customer segment, while direct mail was more effective for another segment. Armed with this knowledge, the insurance company was able to tailor their marketing campaigns to each customer segment and achieve better results.

Especially, Machine learning is a method of automatically learning models and making predictions using data in the field of artificial intelligence. Machine learning is used in many fields and has various definitions. Samuel (1959) argues that machine learning is defined as "the field that enables computers to learn from experience to perform specific tasks." This definition highlights one of the key features of machine learning: learning based on experience. Michalski et al.(1984) suggest that machine learning is "building a computer program that learns about a measure of performance P from experience E on some task T". This definition emphasizes the importance of improving performance, which is the goal of machine learning, and computer programs, which are the output of learning. Ng & Jordan (1921) define machine learning as "a branch of artificial intelligence that allows computers to learn automatically from experience." This definition highlights a key feature of machine learning: automatic learning.

Machine learning involves the study and development of algorithms that can make predictions based on data. It is a powerful tool that allows the creation of programs with tunable parameters that adapt to earlier data to improve their performance. Machine learning techniques have been rapidly advancing and have been shown to emulate the human mind by representing multi-level records and effectively addressing the selectivity-invariance dilemma (Cun et al., 2015). In the context of direct marketing, machine learning can be used to predict whether a particular customer is available, and can automate the process of feature validation. However, a disadvantage of this model is that it assigns different weights to each factor. In real life, direct marketing responses may sometimes depend on a single strong factor only, which cannot be captured by this system.

To overcome the shortcomings of these models, the purpose of this study is to find and analyze the prediction of direct marketing response so that companies can use the study to formulate possible solutions for this strategy. Therefore, this study has the following important meanings. First, this methodology can also be considered a roadmap for the readers. They can follow the steps and apply a procedure to identify the causes of many other problems. Second, this paper aims to provide a quick, immediate, and easy way to choose deserving customers. It can offer particular advantages to companies. The direct marketing response prediction system can automatically calculate the weight of each feature taking part in direct marketing response processing. On new test data, the same features are processed concerning their associated weight. Finally, a time limit can be set for the applicant to check whether the direct marketing response can be sanctioned or not. The direct marketing response prediction system allows jumping to a specific application so that it can be checked on a priority basis.

To this end, first, we will comprehensively review the literature on existing studies and methodologies on direct marketing. Second, the dataset used in this study, machine learning technique, media mining model, and performance evaluation method are presented. Third, we present the analysis results of the decision tree, which is a machine-learning technique used in this study. Finally, in the conclusion part, a discussion of the results of this study, academic contributions and practical implications of the study, limitations of the study, and future research directions are presented.

2. RELATED STUDY

2.1 DIRECT MARKETING

Direct marketing refers to the practice of communicating with potential customers directly and individually, using a variety of channels such as mail, email, social media, or telemarketing (Lim, et al., 2023). It has been extensively studied in the field of marketing, and below are some key findings and trends. First, from the perspective of personalization, customizing marketing messages based on individual preferences, behaviors, and demographic information can significantly improve response rates and conversion rates (Srivastava, & Bag, 2023). Second, by a data-driven approach, direct marketing heavily relies on customer data to identify potential customers, tailor messages, and measure the effectiveness of campaigns. Advances in data analytics and machine learning have made it easier to collect, analyze, and apply customer insights (Suoniemi, et al., 2022). Third, by multi-channel approach, combining multiple channels in a coordinated manner (e.g., email, social media, and direct mail) can enhance the reach, frequency, and impact of direct marketing campaigns (Bellaaj, 2021). Fourth, considering timing and frequency, the timing and frequency of direct marketing messages can have a significant impact on customer engagement and response rates. Marketers need to carefully balance frequency with the risk of over-communicating and irritating customers (Swain et al., 2023). Fifth, in the perspective of trust and privacy, direct marketing involves collecting and using customer data, which can raise privacy concerns and erode trust if not handled transparently and ethically. Compliance with data protection regulations such as GDPR and CCPA is critical (Calzada, 2022). Sixth, by measuring and optimizing, direct marketing campaigns need to be constantly monitored and optimized based on performance metrics such as response rates, conversion rates, and customer lifetime value. Testing, segmentation, and attribution modeling are some common techniques used to improve campaign effectiveness (Lim, 2022).

2.2 DIRECT MARKETING RESPONSE ANALYSIS MODEL

The principal objective of direct marketing consumer response modeling is to identify customers who are most likely to respond to promotional campaigns. Researchers have developed several direct marketing response models based on consumer data. One of the classic models, the Recency, Frequency, Monetary Value (RFM) model, assesses the probability of customer response by considering the recency of the last purchase, frequency of purchases made over the past years, and the monetary value of a customer's purchase history (Berger & Magliozzi, 1992). More advanced models can be created by analyzing customer demographic and psychographic variables, credit histories, and purchasing patterns, thereby enhancing understanding of consumer responses and improving purchase prediction accuracy.

Until recently, statistical methods such as logistic regression and discriminant analysis have predominated direct marketing response modeling (Berger & Magliozzi, 1992). These statistical methods, while powerful, impose strict assumptions on data types and their distribution and can only manage a limited number of variables. Regression-based methods rely on a fixed-form equation and assume a single best solution, limiting researchers' ability to compare alternative solutions manually. Furthermore, when applied to actual data, critical assumptions underlying research methods are often violated (Bhattacharyya, 1999). Recently, researchers have developed more advanced models, including beta-logistic models (Rao & Steckel, 1995), tree-generating techniques such as Classification and Regression Trees (CART) and Chi-squared Automatic Interaction Detection (CHAID) (Haughton & Oulabi, 1997), and hierarchical Bayes models (Allenby et al., 1999). Several studies have addressed selection and endogeneity biases in existing models to enhance predictive accuracy (Gönül et al., 2000).

Recently, there has been a rapid accumulation of customer and transactional data, leading to the creation of vast databases. The voluminous amount of consumer data provides researchers with unique opportunities to use data-mining methods to gain insight into consumer behavior. For example, researchers have utilized consumer lifetime and transaction variables, in addition to the RFM variables, to enhance the performance of models (Bhattacharyya, 1999; Venkatesan & Kumar 2004). However, with the increasing amount and variety of customer data, it may be impossible to manually optimize response models (Bitran & Mondschein 1996). This presents new challenges in how to leverage the different types and increasing volumes of customer data to support management decision-making.

Innovative methods, such as machine learning, offer a way to perform data mining with large databases to provide decision support to managers. Machine learning is a computer-based approach that can extract patterns or knowledge from data and perform optimization tasks with minimal human intervention. These methods are rooted in artificial intelligence and dynamic programming and have been widely adopted as useful data-mining tools in various fields to discover "interesting" and non-obvious patterns or knowledge hidden in a database that can improve the bottom line. Examples of these methods include association rules, decision trees, neural networks, and genetic algorithms. Business researchers have utilized some of these techniques to address classification problems such as predicting bankruptcy and loan default and modeling consumer choice (Hu et al., 1999). Additionally, these methods can be highly valuable in situations where researchers have observable data, but the model structure is unknown, enabling them to learn new knowledge.

Especially, using machine learning for direct marketing response analysis has several benefits (Ding, & Goldfarb, 2023). Firstly, machine learning algorithms can process and analyze large volumes of data much faster and more accurately than humans, which is essential when dealing with marketing campaigns that often involve millions of potential customers. Secondly, machine learning can help identify patterns and trends in customer behavior, allowing marketers to better understand what types of messaging and offers are most likely to resonate with their target audience. Thirdly, machine learning can help optimize marketing campaigns in real-time by constantly monitoring and adjusting messaging, targeting, and other variables based on customer responses. Overall, machine learning enables marketers to make more data-driven decisions and create more effective and efficient direct marketing campaigns that are more likely to achieve their desired outcomes.

3. METHODOLOGY

3.1 Dataset

The dataset used for experiments in this paper was related to direct marketing and is available at IBM Watson Analytics in IBM Community. The details of 23 attributes are shown in Table 1. IBM Watson Analytics in IBM Community has several advantages because it has datasets in various fields. First, IBM Watson Analytics in IBM Community has datasets from multiple disciplines. This helps users easily find and utilize data in the desired field. Second, IBM Watson Analytics in IBM Community has large datasets. This helps to derive more accurate analysis results by providing various data required for analysis. Third, IBM Watson Analytics in IBM Community has high-quality datasets. This not only increases the reliability of the analysis results but also prevents incorrect conclusions by ensuring

Table 1.
The variables in each category

Variables		Description
Independent Variables	State	Categorical (California 34%, Oregon 28%, Other(3) 37%)
	Customer Lifetime Value	Numeric
	Coverage	Categorical (Basic 61%, Extended 30%, Other(1) 9%)
	Education	Categorical (Bachelor 30%, Cllege 29%, Other(3) 41%)
	Effective to date	Numeric
	EmploymentStaus	Categorical (Employed 62%, Unemployed 25%, Other(3) 12%)
	Gender	Binary (Female 51%, Male 49%)
	Income	Numeric
	Location Code	Categorical (Suburban 53%, Rural 19%, Other(1) 17%)
	Marital Status	Categorical (Married 58%, Single 27%, Other(1) 15%)
	Monthly Premium Auto	Numeric
	Months Since Last Claim	Numeric
	Months Since Policy inception	Numeric
	Number of Open Complaints	Numeric
	Number of Policies	Numeric
	Policy Type	Categorical (Personal Auto 74%, Corporate Auto 22%, Other(1) 4%)
	Policy	Categorical (Personal L3 38%, Personal L2 23%, Other(2) 27%)
	Renew Offer Type	Categorical (Offer1 41%, Offer2 32%, Other(2) 27%)
	Sales Channel	Categorical (Agent 38%, Branch 28%, Other(2) 34%)
	Total Claim Amount	Numeric
	Vehicle Class	Categorical (Four-Door Car 51%, Two-Door Car 21%, Other(4) 29%)
	Vehicle Size	Categorical (Medsize 70%, Small 19%, Other(1) 10%)
Dependant Variable	Response	Binary (Yes or No)

the accuracy of the data. Finally, IBM Watson Analytics in IBM Community provides pre-cleaned datasets. This helps users perform analysis faster by reducing the time it takes to clean the data.

3.2 Decision Tree

Among various analysis techniques, a decision tree (DT) is a powerful and popular machine-learning algorithm to this date for predicting and classifying big data (Gonzalez-Cava et al., 2018). The decision tree has the following advantages over other classifiers. First, a decision tree has a tree structure, and each node divides data based on specific attributes. This structure makes it easy to understand how the model works. You can also figure out which attributes play an important role in each node. Second, decision trees work well with a variety of datasets, from simple to complex. In addition, the prediction time is very fast and shows high speed even on large datasets. Third, decision trees can handle various types of data. It also requires fewer data preprocessing than other classifiers. This makes model development and maintenance easy. Finally, decision trees can prevent overfitting by setting appropriate hyperparameters. In addition, you can further improve the performance of your model when used with ensemble techniques.

The utility of decision trees extends to both classification and regression problems. The question may arise as to why we prefer to use decision tree classifiers over other classifiers. There are two key reasons for this preference. Firstly, decision trees attempt to imitate the thinking process of the human brain. Thus, it is quite easy to understand the data and draw valid conclusions or interpretations. Secondly, unlike Support Vector Machines (SVM), Neural Networks (NN), and other black-box algorithms, decision trees enable us to see the logic behind the data and its interpretation. This feature renders decision trees highly desirable among modern-day programmers due to its simplicity and clarity. Having established the benefits of decision trees, we shall delve further into the decision tree classifier. The decision tree is essentially a tree where numerous nodes represent features (attributes). Each link (branch) represents a decision or rule, and each leaf of the tree represents an outcome or categorical/continuous value. The objective is to create a tree that spans the entire dataset and obtains an outcome at every leaf (Canfield et al., 2005). Having gained a better understanding of what a decision tree entails, we shall proceed to discuss how to build a decision tree classifier. Two different algorithms can be utilized for this task: Classification and Regression Trees (CART) and Iterative Dichotomiser 3 (ID3).

For ID3 first, we take the x value in the column and a y value, which stays at the last position of the column and only has a "YES" or "NO" value. For the chart above, we have (outlook, temp, humidity, and windy) as our x values and play, which only has two options either 'YES' or "NO" is at the last position of the column, or is our y value. Now we need to do the mapping of x and y. As we can see, it is a binary classification problem, so let us build the tree using the ID3 algorithm. Now to create a tree, we need a root node first, and we need to pick one first to be the root node. A general rule of thumb is to choose the feature which has the most influence on the value y first as the root node. Then we move on and choose the next most influential feature to be the next node. Here we are going to use the concept of entropy, which is the measure of the amount of uncertainty in the data set. We need to calculate the entropy for all categorical values for the binary classification problem (Goetz, 2010). So to sum it all up, we can say that we need to compute the entropy for the data set first. Then for every attribute/feature, we need first of all to calculate entropy for all the categorical values, then take the average value information entropy for the current attribute and finally calculate how much we have gained for the current attribute. After that, we need to pick the highest gain attribute and repeat until we get our desired tree. Now that is the process of ID3.

As we have discussed above decision tree classifier has been made on another algorithm known as CART short for classification and regression trees. In this algorithm, the Gini index is used as a cost function used to evaluate splits in the dataset. Here our target variable is indeed binary, so it will take two values (yes and no). And as we all know, there can be 4 combinations. Now we need to figure out the Gini score which will give us a good idea of how we can split the data. If we can get a

Gini score of 0 we can consider it to be a perfect separation, whereas the worst-case scenario would be a split of 50/50. Now the question arises of how we can calculate the Gini index value.

Now, if the target variable is a categorical variable with multiple levels, the Gini index will still be similar. So the steps for this method are the first compute of the Gini index for the dataset. Then for every feature, we need to calculate the Gini index for all categorical values and take the average information entropy for the current attribute and, in the end, calculate the Gini gain. After we are done with that, we can pick the best Gini gain attribute, and we need to repeat it until we get our desired tree. And that is how the decision tree algorithm works.

DT classification methods involve building tree models that consist of a series of predictors. Each of these predictors (attributes) within a training set is split repetitively until pure subsets are obtained. This process of repetitive splitting is influenced by a particular entity's (i.e., customer) characteristics (James et al., 2013). The basic anatomy of a DT comprises both a leaf node and a decision node. The leaf node represents a predictor variable and signifies the point where binary splits transpire. Leaf nodes are also known as internal nodes (Mendez et al., 2008). The decision node, also known as the terminal node (Mendez et al., 2008), represents the output variable (binary outcome variable) and graphically is depicted as the end of the branch. It is the terminal node that serves as the basis for churn prediction for it reports the category with the majority of cases. Extant literature has revealed that four major DT machine-learning algorithms are commonly utilized. 1) Classification and Regression Trees (CART) 2) C4.5 3) chi-squared automatic interaction detection (CHAID) and 4) C5.0. DTs serve as the foundation of other tree methods like random forests and ensemble forests, which essentially involve aggregating multiple decision trees (James et al., 2013).

The process of binary splitting an attribute relies on selecting the right attributes to split. Correct attribute selection is dependent on calculating either entropy measures (C4.5) or choosing the Gini criterion (CART) based on the type of DT algorithm (Verbeke et al., 2012). DT analysis is quite popular due to its simplicity, graphical layout, and ease of interpretation (Höppner et al., 2017). DTs provide an appropriate schematic to model both quantitative and qualitative decision-making questions without needing to create dummy variables or transformations (Höppner et al., 2017). Moreover, DTs are also able to monitor non-linearities and are easy to compute (Höppner et al., 2017). However, DTs also have their disadvantages. DT results may not always be as predictively accurate as other methods. Furthermore, minor changes in the dataset can result in non-robust predictions (James et al., 2013). However, this classification technique has been used frequently to model churn (De Caigny et al., 2018).

3.3 Data Mining Models

To survive in an increasingly competitive marketplace, many companies are turning to data mining techniques for decision prediction analysis. To manage customers effectively, we need to build a more effective and accurate decision prediction model. Statistical and data mining techniques have been utilized to construct decision prediction models. The data mining techniques can be used to discover interesting patterns or relationships in the data, and predict or classify the behavior by fitting a model based on available data. In the case where the learning dataset and the test dataset are separated for machine learning, the test dataset must satisfy the following requirements. First, the training dataset and the test dataset must be created in the same format. Second, the test dataset should not be included in the training dataset. Third, the training dataset and the test dataset must be consistent in data. However, it is challenging to create a test data set that meets these requirements. In data mining, various verification frameworks using one dataset have been developed to solve this problem. This study uses the Split Validation operator provided by RapidMiner to support this. The operator splits the input dataset into a training dataset and a test dataset to support performance evaluation. This study selects relative segmentation among the segmentation method parameters of this operator and uses 70% of input data as learning data.

3.4 Performance Evaluation

Performance assessment uses training data to determine how well the generated model works. Performance measures can be divided into technical performance measures and heuristic measures. The technical performance measures to be used in this study show performance results by generating models from training data, processing test data into models, and comparing the class labels of original verification cases with predicted class labels. Measuring technical performance can be divided into supervised and unsupervised learning. The supervised learning used in this study is classified and regressed. The data used for this learning and test all have original class values. The performance is obtained by comparing and analyzing the original class values with the prediction results.

The classification problem is the most common data analysis problem. Various metrics have been developed to measure the performance of classification models. For classification problems of category type, accuracy, precision, recall, and f-measure are used a lot. RapidMiner includes Performance (Classification), which measures performance indicators for common classification problems, and Performance (Binominal Classification), which provides performance indicators specific to binomial classification problems. Table 2 shows how these indicators are calculated.

4. RESULTS

Figure 1 shows the classification tree for the full model after pruning the tree using cross-validation to avoid overfitting (Kuhn, & Johnson, 2013). The key variables in the full model analysis consist of 10 ones, as shown below, based on the criterion established with each of these variables. Among all variables, customer lifetime value, coverage, employment status, income, marital status, monthly premium auto, months since the last claim, months since policy inception, renew offer type, and the total claim amount is shown to influence direct marketing response. However, others have no significance.

Tables 3 illustrate each of the confusion matrix measures. For the full model, the accuracy rate is 0.864, which implies that the error rate is 0.136. Among the patients who predicted not to have a direct marketing response, the accuracy that would not have a direct marketing response was 87.23%, and the accuracy that had a direct marketing response was 66.34% among the patients who predicted to have a direct marketing response.

5. CONCLUSION

5.1 Discussion

The primary aim of this research paper is to evaluate the accuracy of existing models and develop a new model for predicting direct marketing response. The research encompasses two primary objectives: 1) to gain a better understanding of the variables' role in predictive modeling for direct marketing response, and 2) to assess the predictive performance of decision trees. Based on the

Table 2.
Key performance indicators of binomial classification

		Actual class (as determined by Gold Standard)	
		True	**False**
Predicted class	Positive	True Positive	False Positive(Type I error)
	Negative	False Negative(Type II error)	True Negative

Precision = TP/(TP+FP), Recall = TP/(TP+FN), True negative rate = TN/(TN+FP), Accuracy = (TP+TN)/(TP+TN+FP+FN), F-measure = 2·((precision·recall)/(precision + recall))

Figure 1.
Classification tree for the full model

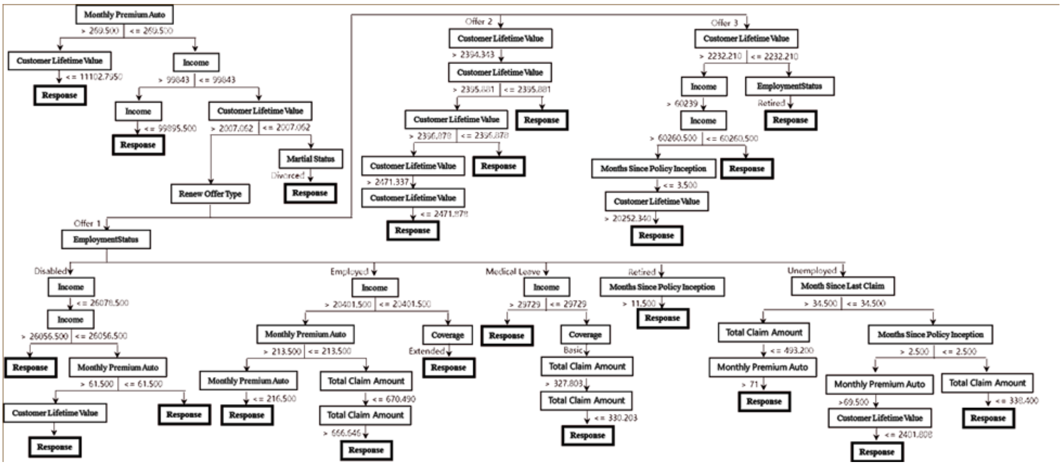

Table 3.
Performance evaluation

	True N	True Y	Class precision
Pred. N	2,302	337	87.23%
Pred. Y	34	67	66.34%
Class recall	98.54%	16.58%	

findings, several implications are derived. The results suggest that assessing the variables' role is a complex process, and their impact varies depending on the classification techniques used. Decision tree methods prioritize explanatory power as the most crucial factor in analysis. Therefore, reaching a unanimous conclusion about the most critical explanatory variables for direct marketing response prediction is not feasible. Nonetheless, this research provides additional insights into customer profiles. Companies should use classification techniques to predict direct marketing responses. The study reveals that customer lifetime value, coverage, employment status, income, marital status, monthly premium auto, months since the last claim, months since policy inception, renewal offer type, and total claim amount impact direct marketing response, while other variables are insignificant. Secondly, the full model's accuracy rate is 0.864, indicating an error rate of 0.136. Among patients predicted not to have a direct marketing response, the accuracy rate of not having a direct marketing response is 87.23%, while the accuracy rate of having a direct marketing response is 66.34% among patients predicted to have a direct marketing response.

5.2 Research Contributions and Practical Contributions

The present study makes research and practical contributions. Firstly, it expands the existing literature by empirically investigating the combined impact of variables on direct marketing response modeling. Although many studies have been conducted on direct marketing response prediction, no universal tool has been developed to predict it due to its complexity and interrelation with multiple factors. As a result, researchers tend to focus on a limited number of elements and neglect the effects of other factors, such as changing customer demographics and privacy issues. This study contributes to the literature on direct marketing response prediction by offering a comprehensive model summarizing the determinants of direct marketing response prediction, including customer factors. Secondly, the methodology employed in this paper serves as a guide for readers to follow the same steps taken in this case study and apply the one-day procedure to diagnose other problems. This paper strives to develop the best-performing model for predicting direct marketing response based on a restricted set of features, including customer factors, by utilizing machine learning techniques such as decision tree and feature importance analysis to achieve greater accuracy. Using this methodology, the study reveals a pattern of direct marketing response prediction.

In practical terms, the application discussed in this paper assists companies in managing customer personal records, allowing for faster decision-making if the user's report is already on file. The paper presents a prototype model that organizations can utilize to make informed decisions about approving or rejecting direct marketing requests from customers. Additionally, this study is focused solely on the managing authority of the company, ensuring that the entire prediction process remains private and immune from stakeholder interference. Results for a specific direct marketing response ID can be transmitted to various departments within the company, enabling them to take appropriate action on the application and facilitating other formalities across all departments.

5.3 Limitations and Future Research Directions

In the proposed system, we need to have a database to store the records of the customers, and when the count of customers increases means, more data will be generated, and the storage will become a problem. Therefore, in a future release, there will be a cloud facility to store all the records in the cloud. Therefore, the data is protected safely and can be retrieved from anywhere if we have the right to access the data. The smart device will be synced with our application in the future release. Therefore, the customer's real-time financial condition will be monitored, and in the case of financial needs, the companies will get alerted.

In the future, first, the machine learning model will make use of a larger training dataset, possibly more than a million different data points maintained in an electronic financial record system. Although it would be a huge leap in terms of computational power and software sophistication, a system that will work on artificial intelligence might allow the financial practitioner to make the best-suited decision for the concerned customers as soon as possible. Second, the purpose of this study is to empirically demonstrate the process of solving issues in the field of direct marketing with machine learning. And this study is the first such attempt. In the field of direct marketing, the predictive power required to solve problems based on limited data is high. However, the prediction performance in this study has limitations of the current study and should be improved in the future.

COMPETING INTERESTS

All authors of this article declare there are no competing interest.

REFERENCES

Allenby, G. M., Leone, R. P., & Jen, L. (1999). A dynamic model of purchase timing with application to direct marketing. *Journal of the American Statistical Association*, *94*(446), 365–374. doi:10.1080/01621459.1999.10474127

Bellaaj, M. (2021). Why and how do individual entrepreneurs use digital channels in an emerging market? Determinants of use and channel coordination. *International Journal of Emerging Markets*. doi:.10.1108/IJOEM-08-2020-0882

Berger, P., & Magliozzi, T. (1992). The effect of sample size and proportion of buyers in the sample on the performance of list segmentation equations generated by regression analysis. *Journal of Direct Marketing*, *6*(1), 13–22. doi:10.1002/dir.4000060105

Bhattacharyya, S. (1999). Direct marketing performance modeling using genetic algorithms. *INFORMS Journal on Computing*, *11*(3), 248–257. doi:10.1287/ijoc.11.3.248

Bitran, G., & Mondschein, S. (1996). Mailing decisions in the catalog sales industry. *Management Science*, *42*(9), 1362–1381. doi:10.1287/mnsc.42.9.1364

Calzada, I. (2022). *Trends, Aftermaths, Emancipations. Emerging Digital Citizenship Regimes*. Emerald Publishing Limited. doi:10.1108/9781803823317

Canfield, J., Hansen, M. V., & Rackner, V. (2005). Heart disease. *Health Communication*.

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, *269*(2), 760–772. doi:10.1016/j.ejor.2018.02.009

Ding, M., & Goldfarb, A. (2023). The Economics of Artificial Intelligence: A Marketing Perspective. In Sudhir, K. and Toubia, O. (Ed.) Artificial Intelligence in Marketing (Review of Marketing Research, Vol. 20), pp. 13-76. Emerald Publishing Limited, Bingley.

Goetz, T. (2010). *The decision tree: How to make better choices and take control of your health.* Rodale Books.

Gönül, F. F., Kim, B., & Shi, M. (2000). Mailing smarter to catalog customers. *Journal of Interactive Marketing*, *14*(2), 2–16. doi:10.1002/(SICI)1520-6653(200021)14:2<2::AID-DIR1>3.0.CO;2-N

Gonzalez-Cava, J. M., Reboso, J. A., Casteleiro-Roca, J. L., Calvo-Rolle, J. L., & Pérez, J. A. M. (2018). A Novel Fuzzy Algorithm to Introduce New Variables in the Drug Supply Decision-Making Process in Medicine. *Complexity*, *2018*, 1–15. doi:10.1155/2018/9012720

Hassan, A. (2022). The Role of Artificial Intelligence in Human Resources. In A. Hamdan, A. Harraf, P. Arora, B. Alareeni, & R. Khamis Hamdan (Eds.), *Future of Organizations and Work After the 4th Industrial Revolution. Studies in Computational Intelligence* (Vol. 1037). Springer. doi:10.1007/978-3-030-99000-8_4

Haughton, D., & Oulabi, S. (1997). Direct marketing modeling with CART and CHAID. J. *Direct Marketing*, *11*(4), 42–52. doi:10.1002/(SICI)1522-7138(199723)11:4<42::AID-DIR7>3.0.CO;2-W

Höppner, S., Stripling, E., Baesens, B., & Verdonck, T. (2017). Profit Driven Decision Trees for Churn Prediction. arXiv:1712.08101. https://arxiv.org/abs/1712.08101

Hu, M. Y., Shanker, M., & Hung, M. S. (1999). Estimation of posterior probabilities of consumer situational choices with neural network classifiers. *International Journal of Research in Marketing*, *16*(4), 307–317. doi:10.1016/S0167-8116(99)00018-X

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning (Vol. 112)*. Springer.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). Springer. doi:10.1007/978-1-4614-6849-3

Le Cun, Y., Bengio, Y., & Hinton, G. E. (2015). Deep learning. *Nature*, *521*(7553), 436–444. doi:10.1038/nature14539 PMID:26017442

Lim, W. M., Kumar, S., Pandey, N., Rasul, T., & Gaur, V. (2023). From direct marketing to interactive marketing: A retrospective review of the Journal of Research in Interactive Marketing. *Journal of Research in Interactive Marketing*, *17*(2), 232–256. doi:10.1108/JRIM-11-2021-0276

Mendez, G., Buskirk, T. D., Lohr, S., & Haag, S. (2008). Factors associated with persistence in science and engineering majors: An exploratory study using classification trees and random forests. *Journal of Engineering Education*, *97*(1), 57–70. doi:10.1002/j.2168-9830.2008.tb00954.x

Michalski, R. S., Carbonell, G., & Mitchell, T. M. (1984). Machine Learning. An Artificial Intelligence Approach. Springer-Verlag.

Nachev, A. (2014). Application of Data Mining Techniques for Direct Marketing. *Ithea*, 86–95.

Ng, A., & Jordan, M. (2001). On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in Neural Information Processing Systems*, 14.

Poon, P., Albaum, G., & Yin, C.-Y. (2017). Exploring risks, advantages and interpersonal trust in buyer-salesperson relationships in direct selling in a non-western country. *International Journal of Retail & Distribution Management*, *45*(3), 328–342. doi:10.1108/IJRDM-08-2016-0124

Quarteroni, A. (2022). A Bit of Maths (Behind Artificial Intelligence and Machine Learning). In *Algorithms for a New World*. Springer. doi:10.1007/978-3-030-96166-4_5

Rao, V. R., & Steckel, J. H. (1995). Selecting, evaluating, and updating prospects in direct mail marketing. *Journal of Direct Marketing*, *9*(2), 20–31. doi:10.1002/dir.4000090205

Roddy, M. (2002). *Direct Marketing: A Step-by-step Guide to Effective Planning and Targeting*. Longdog, Kogan Page.

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. doi:10.1147/rd.33.0210

Sharma, A., Fadahunsi, A., Abbas, H., & Pathak, V. K. (2022). A multi-analytic approach to predict social media marketing influence on consumer purchase intention. *Journal of Indian Business Research*, *14*(2), 125–149. doi:10.1108/JIBR-08-2021-0313

Simon, J. (2019). Artificial intelligence: Scope, players, markets and geography. Digital Policy. *Regulation & Governance*, *21*(3), 208–237. doi:10.1108/DPRG-08-2018-0039

Srivastava, G. & Bag, S. (2023). Modern-day marketing concepts based on face recognition and neuro-marketing: a review and future research directions. *Benchmarking: An International Journal*. doi:.10.1108/BIJ-09-2022-0588

Suoniemi, S., Zablah, A., Terho, H., Olkkonen, R., Straub, D., & Makkonen, H. (2022). CRM system implementation and firm performance: The role of consultant facilitation and user involvement. *Journal of Business and Industrial Marketing*, *37*(13), 19–32. doi:10.1108/JBIM-08-2021-0380

Swain, S., Jebarajakirthy, C., Maseeh, H.I., Saha, R., Gupta, N. & Grover, R. (2023). Permission marketing: a systematic review of 22 Years of research. *Marketing Intelligence & Planning*. doi:10.1108/MIP-05-2022-0187

Venkatesan, R., & Kumar, V. (2004). A customer lifetime value framework for customer selection and resource allocation strategy. *Journal of Marketing*, *68*(4), 106–125. doi:10.1509/jmkg.68.4.106.42728

Verbeke, W., Dejaeger, K., Martens, D., Hur, J., & Baesens, B. (2012). New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. *European Journal of Operational Research*, *218*(1), 211–229. doi:10.1016/j.ejor.2011.09.031

Young Keun Choi Division of Business Administration, School of Business and Economics, Sangmyung University, Seoul, The Republic of Korea. His research areas are business analytics, data science, entrepreneurship, technology management, and etc.

Jae Won Choi Department of Computer Science, Erik Jonsson School of Engineering and Computer Science, University of Texas at Dallas, TX, USA. His research interests are data science, artificial intelligence, blockchain, game, and etc.