Marketing Strategies Based on Machine Learning Approaches

Yuanzheng Yin(✉)

Optoelectronic Information Science and Engineering, University of Shanghai for Science and Technology, Shanghai, China
1935031628@st.usst.edu.cn

Abstract. Marketing is the activity, collection of institutions, and processes for producing, conveying, providing, and exchanging value-added solutions for consumers, consumers, collaborators, and society overall. The development of emerging fields (e.g., the Internet of Things (IoT) and online social media) has dramatically changed the marketing field's data collection modes and quantity of the data. Traditional measurement tools cannot tap the market value of big data, while machine learning approaches can fully lead the market based on the bigdata analysis. This paper briefly summarizes the usage of machine learning in marketing and demonstrates the procedure for implementation. Some commonly used machine learning algorithms related to marketing are firstly introduced, including random forest, logistic regression, and artificial neural networks. Subsequently, the state-of-art applications of machine learning are discussed including market segmentation, customer behavior prediction, and recommendation systems. Finally, limitations and future guidelines for the current applications are demonstrated. Overall, these results shed light on further development of marketing strategies design in terms of machine learning approaches.

Keywords: Marketing · Machine Learning · Algorithm · Application

1 Introduction

Generally, marketing is commonly considered as advertising, whereas advertising is only a part of marketing, even not the most crucial element. Marketing is the process of attracting more customers to buy product or service, which refers to all activities a company does to promote and sell products or services to consumers. The traditional marketing mix refers to four broad marketing decision levels: product, price, promotion, and place [12]. Contemporarily, if marketers want to have a deeper understanding of customers' needs, develop customized products and effective pricing, make distribution and promotion, it is necessary to carry out bigdata analysis. Marketers had to collect the data and determine whether it was useful in this case or not. Then, marketers create hypotheses, test them, evaluate them, and analyze them. The capable calculation size of human beings and the relatively short working hours are also the disadvantages compared to the way of using machine learning to solve target marketing issues. Sometimes the results are even incorrect because information varies every second.

D. Qiu et al. (Eds.): ICBBEM 2022, AHIS 5, pp. 400–409, 2023.
https://doi.org/10.2991/978-94-6463-030-5_42

Machine learning (ML) is the science of computer algorithms that can learn and develop on their own with experience and data. Based on machine learning, marketers can enlarge the analysis capacity of data (with considerable values), which could be too complicate to be dealt with by a human. In general, machine is able to work 24 h every day to sift and analyze massive amounts of data and then pinpoints minute patterns and contrasting bits of data, which allows one to automate marketing processes and avoid routine work. It can also improve the quality of data analysis. With the benefit of adapting to changes and new data, the marketing model is upgrading nowadays.

In order to provide a thorough review, the descriptions of some commonly used machine learning algorithms in marketing are introduced, including random forest, logistic regression, and artificial neural networks. Subsequently, machine learning applications in marketing are demonstrated, mainly about market segmentation, customer behavior prediction, and recommendation systems. Afterwards, the limitations of current scenarios are discussed and future guidelines are offered since machine learning is still developing. Eventually, a brief summary of the paper is given.

2 Machine Learning Approach

In most marketing cases, the independent variables "x" are mainly the customer's information, and the dependent variable "y" refers to the results marketers want to find out. Among the machine learning algorithms in the marketing field, random forest, logistic regression, and artificial neural networks are three of the most widely and commonly implemented models.

2.1 Random Forest

Random Forest uses a mixture of bootstrap aggregation [10] and random segmentation techniques to create numerous decision trees and fuse them to get a more accurate and stable model. The structure of the random forest is shown in Fig. 1. The random forest generates several different decision trees. Trees respond to consecutive questions by continuing down a certain branch of the tree based on the response. The model responds to "if this, then that" circumstances, resulting in a certain outcome. When a sample needs to be forecasted, the predictive performance of each tree in the forest for the sample is recorded, and the result is chosen by voting from these prediction results.

The randomness can be obtained in two ways: randomly picked characteristics and randomly selected samples, resulting in each tree in the forest having both similarities and differences. The random forest training algorithm is aided by tree learning, which employs the generic bagging approach. The bagging approach repeats selecting from the training sample with substitution and then trains a tree model on these data, providing a training set X and a goal Y. Results for an unknown x can be obtained after training by averaging the estimates of all single trees on x, as shown in Eq. (1), or by selecting the class with the most votes in a classifier:

$$\hat{f} = \frac{1}{B}\sum_{b=1}^{B} f_b\left(x'\right) \tag{1}$$

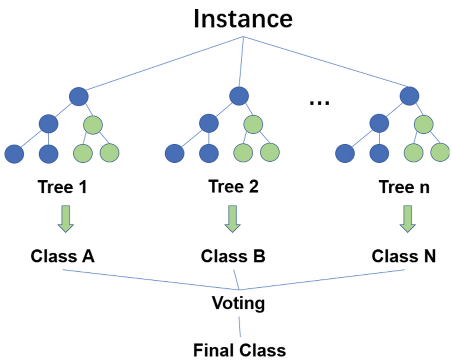

Fig. 1. Structure of Random Forest Classification [1]. (Photo credit: Original)

The random forest's accuracy of the prediction depends on two things. First of all, the accuracy of every two forest trees is determined by their similarity. To be specific, the higher their similarity, the lower the reliability. Additionally, the identification ability of each tree also affects the accuracy: the more strong a single tree's classification ability, the more powerful the forest's classification ability.

While random forests generally outperform single decision trees in terms of accuracy, they do so at the cost of decision trees' inherent interpretability. Along with linear models, rule-based models, and attention-based models, decision trees are part of a minority group of machine learning models that are comprehensible. The most significant aspects of decision trees are their interpretability. It enables developers to verify whether the model has acquired useful information from the data, and it enables end-users to have faith in the model's judgments [7]. Using random forest may not improve the precision of the base learner if the predictive qualities are linearly associated with the attribute value [14]. Furthermore, the random forest might not always be capable of improving the performance of the basic learner in cases with several continuous variables [13].

2.2 Logistic Regression

The logistic model is designed to ensure any risk estimation obtained is limited between zero to one. As a result, one can never receive a risk estimate higher than 1 or lower than 0. The output of the model will similar to Fig. 2. In Fig. 2, the horizontal axis is the measured value, and the vertical axis is the probability of whether the event will happen. Because this isn't always the situation with other models, the logistic model is commonly used to estimate probabilities.

The logistic model takes into account the general epidemiologic study framework as follows. A group of data was learned and turned into independent factors X. The D represent whether the participant is sick or not, one stand for that they are with the disease, and zero means they are healthy. The equation can be described as:

$$P(X_1, X_2, \ldots, X_k) = \frac{1}{1+e^{-(\alpha+\Sigma\beta_k X_k)}} \tag{2}$$

The α and β_k in the Eq. (2) are unknown parameters that one must estimate based on data gathered for a set of participants X_k and on D [8].

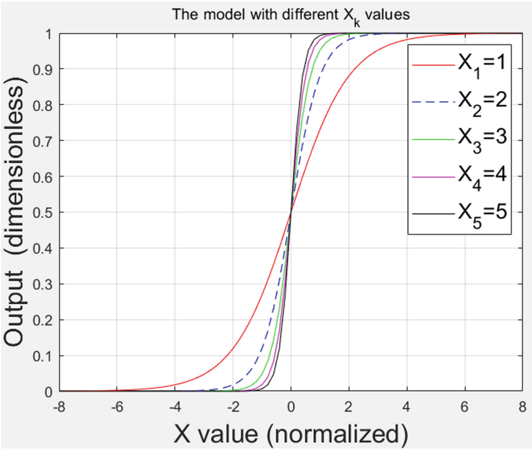

Fig. 2. Shape of Logistic Regression Model

Having two mutually exclusive phases, logistic regression allows for the analysis of binary findings. On the other hand, logistic regression allows for the inclusion of continuous or categorical factors and the adjustment of numerous predictors. It makes logistic regression particularly effective when analyzing observational data and adjusting for any bias stemming from variations in the groups being compared.

As a regression function for the predictors, the logistic regression model uses the natural logarithm of the odds. This can be obtained as:

$$\ln\left[\frac{\alpha}{1-\alpha}\right] = \beta_0 + \beta_1 X \quad (0 < \alpha < 1) \tag{3}$$

where α represents the possibility of the event Y occurs, $\alpha/(1-\alpha)$ is the formula of odds ratio which is a metric for the relationship between an exposure and a result. The odds ratio compares the chances of an event occurring in the absence of a certain exposure against the odds of that outcome occurring in the presence of that exposure; β_0 stands for intercept term while β_1 is the value of regression coefficient.

There is no technique for assessing in logistic regression, unlike linear regression. Finding the most accurate estimates entails fine-tuning rough approximations until they are steady. While this is straightforward to do on a computer, it makes logistic regression less intelligible and more of a "black box" approach for many researchers [11].

2.3 Artificial Neural Network

The artificial neural network abstracts the neuron network of the human brain from the perspective of information processing, establishes a certain simple model, and forms different networks according to different connection methods. There are two forms of neural networks: single-layer perceptron and multi-layer perceptron (MLP). The input, hidden, and output layers are the three types of layers in a multi-layer perceptron, and information is conveyed through weighted connections. Certain neurons in the input layer get information from the data set, combine it by their weights, and transfer it to

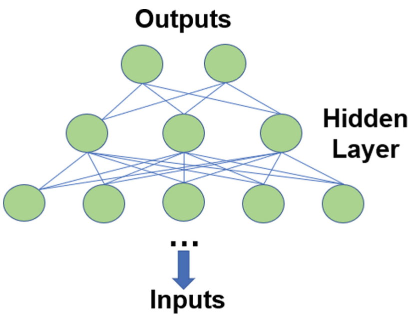

Fig. 3. Architecture of a neural network [16]. (Photo credit: Original)

the hidden layer afterwards, which is multiplied by a new weight matrix. Layers that are hidden also have neurons which accept data from the input layer and map it between zero and one using the sigmoid function. The outputs of the hidden layers are subsequently sent to the output layer through weighted connections. Figure 3 illustrates a typical architecture of a neural network, the lines drawn between the layers are the connection of the neurons.

The neuron output depends on following expression:

$$h_a = \sigma\left(\sum_{b=1}^{N} V_{ab}x_b + T_a\right) \tag{4}$$

Where h_a is the output of neuron a in the hidden layer, N represents how many layers there are of input neurons, V_{ab} represents the weights, x_b represents the value of the input neurons, and T_a represents the threshold terms of the hidden neurons which means the neuron will not fire unless the sum up of the inputs is greater than the threshold. The aim of σ is to limit the output of the neuron, i.e., the artificial neural network won't be collapsed by widely different neurons, in addition to introducing nonlinearity into the neural network.

3 Application in Marketing

In this paper, the usage of machine learning in the marketing area was divided into three components: market segmentation, recommendation system, and predicting consumers' behavior, which based on the two stages in the consumers' shopping decision process: before and during purchasing.

3.1 Market Segmentation

Market segmentation is a common marketing strategy. Its objective is to discover and characterize market segments on which companies will focus their business activities. Marketing segmentation has the advantage of being easy to partition total demand into relatively homogenous categories that can be identified by a few key indications. These characteristics are critical for defining and forecasting consumer reactions to marketing behaviors in a certain market group [15].

Market segmentation for consumers is primarily based on consumers' statistical information and classification of product-related psychological and behavioral differences. For example, Award collected data on consumers' opinions about green consumption. Factor analysis and K-means clustering scale results for attitudes towards green consumption clustering was carried out [2]. According to the analysis, group characteristics about green consumption attitudes were discovered through machine learning approaches.

Market segmentation for products is mainly based on consumers' consumption of a product or brand. There are various products and brands in the market, and machine learning methods can cluster product submarkets effectively. France and Ghose introduced a method of identifying, analyzing and visualizing submarkets in a specific product category through techniques such as hierarchical clustering, and illustrated and tested the results of submarkets made through machine learning methods [5].

To sum up, machine learning overcomes the inherent inefficiency and poor accuracy of traditional market segmentation methods and other methods. It solves the multidimensional clustering problem of consumers and product markets, which is helpful for enterprises to analyze market conditions more thoroughly and understand consumer product needs well.

3.2 Recommendation System

A recommendation system's goal is to provide suggestions to a group of customers about items that might be of interest to them. With the rapid development of e-commerce these years, the usage of recommendation systems has become more often, which is usually implemented through collaborative filtering methods [4]. In addition, as non-numeric data (e.g., texts, photos, and videos) become more prevalent in marketing contexts, machine learning clustering and classification algorithms play a more significant role in recommender systems than traditional methods. Text mining and nonlinear classification problems can be solved using topic models, K-Means clustering, Bayesian algorithm, and other techniques, which can help enhance the accuracy of the recommendation system.

One of the most extensively utilized algorithms approaches in customized consumer recommendations is collaborative filtering. Users are recommended by collaborative filtering based on their similarity to other users or items. Collaborative filtering can

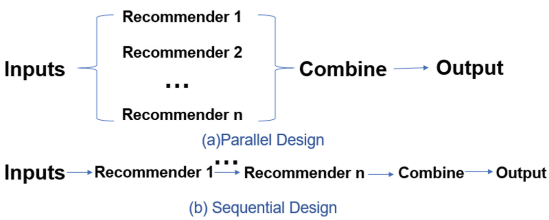

Fig. 4. Architecture of two types of hybrid recommendation systems [3]. (Photo credit: Original)

be separated into two types: user-based and item-based. Machine learning is used to categorize consumers in the recommender system.

Different recommendation systems each have their own set of advantages and disadvantages. When employed in isolation, many of these strategies can appear to be restrictive, especially when multiple data sources are available for the problem. Hybrid recommender systems are built to provide reliable judgments from various data sources. Parallel and sequential are the two most common designs for hybrid recommendation systems. The parallel design feeds many recommendation systems with data, and each recommendation is integrated to produce a single result. The sequential architecture gives a single recommendation engine the input parameters, and the output is handed on to the next recommender in the series. A graphic illustration of both designs is depicted in Fig. 4.

3.3 Customer Behavior Prediction

As the dynamic changes of consumers on the mobile terminal and the Internet are widely tracked nowadays, consumer data has undergone significant changes in terms of type and volume. Machine learning methods can effectively process a large amount of data that changes with geographic location and time, help marketers to understand consumer demand fluctuations in an all-around way, and improve the accuracy of consumer behavior prediction. Scholars in marketing mainly use classification and regression algorithms to predict consumer preferences and churn.

Client retention is the most basic and vital concern for commercial businesses, notably shopping malls. Companies' business activities trying to recruit new clients usually acquire a high cost. Therefore, the way to keep the loyalty of the customers appears to be a critical problem. Customer churning means due to the implementation of various marketing methods of enterprises, the phenomenon that customers and enterprises suspend cooperation [9]. Machine learning has made it easier to detect the causes and timing of customer churn, make strategies to bring the customers back, and avoid further loss of customers.

Machine learning can sort through a significant volume of complicated consumption data and extract useful insights, helping marketers to discover consumers' deeper content preferences. Marketers used a convolutional neural network to filter redundant information in Amazon review content, reducing the amount of consumer information that needs to be manually identified [6]. The accuracy rate was 74.2%, which is good. All in all, machine learning improves user demand identification efficiency and accuracy.

4 Limitation and Future Guideline

Although machine learning is widely used in research in the field of marketing, its' related studies are not yet fully developed, and there are still new challenges and some issues that need to be resolved.

4.1 Personal Privacy Protection

Human beings are benefited greatly due to the rapid development of the Internet in the era of big data. Data has a high financial value for Internet service providers, but the increasing usage in the business analysis will make privacy safety more challenging to be protected. Moreover, the more easily accessible to get customers' data allows enterprises' overuse it for illicit purposes such as reselling, fraud, etc., not only for people. Thereby, the path to deal with security and privacy issues comes into the top priority in modern society [18].

Enterprises must have more vital management to protect consumers' privacy while using big data and machine learning tools to conduct related research. Legal and technical measures will also become beneficial guarantees and rely on the efforts of companies themselves. On the one hand, legal institutions should develop and improve appropriate legislation and data privacy protection and employ legal power to prevent businesses from improperly obtaining data. On the other hand, there are technologies (e.g., federated learning with differential privacy [19] that assure data privacy while data mining and machine learning are being done. Future machine learning systems might be able to strike a balance between privacy protection and research.

4.2 Extend Type of Data

Existing research typically uses quantitative or text data, but as technology advances, new data forms (e.g., photographs, videos, and voice) have emerged in the short video, social media, and live broadcast platforms. While machine learning can handle and analyze new sorts of data, these new types of data will undoubtedly bring fresh viewpoints and directions to marketing research. Meanwhile, the emergence of these new technologies has fundamentally altered traditional marketing thinking or methods. As a result, researchers using machine learning in the future will be able to understand consumer needs better and develop more effective marketing strategies by utilizing these data. Besides, the challenges come with these new types of data. More effective ways still need to be developed in machine learning to identify these characteristics and convert them into readable information for marketers. Moreover, the connection between these new data and the commercial strategies needs to be clarified.

4.3 Establish Evaluation Standards

In the subject of machine learning, there are numerous measures for evaluating results. The accuracy of model is a sufficient evaluation criterion for the model's quality from the perspective of computer science, while customer happiness is a more significant benchmark from the perspective of marketing research. For example, while accuracy is an essential sign of a suggesting system's efficiency, with consumers' seeking feeling of freshness and other consumption psychology, the most accurate recommending system may not lead to increased searching behaviors. Consequently, the model's evaluation should not be restricted to indicators (e.g., accuracy in the field of computer science [17]) but should also use professional knowledge in the field of marketing to create new evaluations.

5 Conclusions

In summary, the state-of-art models of machine learning for marketing are illustrate and discussed. Specifically, the descriptions and processes of the models as well as applications in marketing are demonstrated. Besides, the drawbacks of the approaches are also discussed based on current research of machine learning in marketing. Machine learning technology breaks new ground in marketing and provides a great potential for researching and development for numerous marketers. Although there are still some challenges with this technique in safety and its effect evaluation, it is believed that the moment of machine learning taking over the classical approaches is just around the corner due to its higher rate of efficiency and high accuracy. Overall, these results offer a guideline for future development of the marketing based on machine learning.

References

1. Afroz Chakure. "Random Forest Classification and its implementation in Python." [Online]. Available: https://medium.com/swlh/random-forest-classification-and-its-implementation-d5d840dbead0 [Accessed:6-Jul-2019].
2. Awad, Tamer A. "Environmental segmentation alternatives: buyers' profiles and implications." Journal of Islamic Marketing (2011).
3. Aggarwal, Charu C. Recommender systems. Vol. 1. Cham: Springer International Publishing, 2016.
4. Breese, John S., David Heckerman, and Carl Kadie. "Empirical analysis of predictive algorithms for collaborative filtering." arXiv preprint arXiv:1301.7363 (2013).
5. France, Stephen L., and Sanjoy Ghose. "An analysis and visualization methodology for identifying and testing market structure." Marketing Science 35.1 (2016): 182–197.
6. Fischer, Thomas, and Christopher Krauss. "Deep learning with long short-term memory networks for financial market predictions." European Journal of Operational Research 270.2 (2018): 654–669.
7. Hastie, Trevor. "Tibshirani, r. and Friedman, J. (2009): The elements of statistical learning. Data mining, inference, and prediction." (2008).
8. Kleinbaum, David G., et al. "Logistic regression." New York: Springer-Verlag, 2002.
9. Khodabandehlou, Samira, and Mahmoud Zivari Rahman. "Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behavior." Journal of Systems and Information Technology (2017).
10. Lee, Tae-Hwy, Aman Ullah, and Ran Wang. "Bootstrap aggregating and random forest." Macroeconomic Forecasting in the Era of Big Data. Springer, Cham, 2020. 389–429.
11. LaValley, Michael P. "Logistic regression." Circulation 117.18 (2008): 2395–2399.
12. McCarthy, E. Jerome. "Basic marketing: a managerial approach." RD Irwin, 1978.
13. Piryonesi, S. Madeh. "The application of data analytics to asset management: Deterioration and climate change adaptation in Ontario roads." Diss. University of Toronto (Canada), 2019.
14. Smith, Paul F., Siva Ganesh, and Ping Liu. "A comparison of random forest regression and multiple linear regression for prediction in neuroscience." Journal of neuroscience methods 220.1 (2013): 85–91.
15. Tynan, A. Caroline, Jennifer Drayton. "Market segmentation." Journal of marketing management 2.3 (1987): 301–335.
16. Wang, Sun-Chong. "Artificial neural network." Interdisciplinary computing in java programming. Springer, Boston, MA, 2003. 81–100.

17. Wu Wenshuai. “Evaluation and application of data mining method based on multi-objective decision-making.” Chengdu: University of Electronic Science and Technology of China (2015).
18. Zhang, Dongpo. “Big data security and privacy protection.” 8th International Conference on Management and Computer Science (ICMCS 2018). Vol. 77. Atlantis Press, 2018.
19. Zhang Xiaojian, Meng Xiaofeng. “Differential Privacy Guarantee for Data Release and Analysis Nursing.” Chinese Journal of Computer (2014): 197–219.