

Research on the Application of Decision Tree in Mobile Marketing

Huimin Chen

Digital Business School, Beijing Information Technology College, Beijing, 100018, China

Keywords: Decision tree; Mobile marketing; Application

Abstract: In recent years, the achievements of the mobile marketing market are much lower than expected. Mobile telecom operators must pay attention to efficiency and cost, and at the same time, improve the quality of service, innovate the direction of communication products, and give users brand-new excellent services and pleasant experience environment. This paper tries to study the application of mobile marketing market through decision tree. Data mining decision tree algorithm constructs a decision tree by training historical data, and the process of constructing the tree does not need parameter setting and knowledge in related fields. The algorithm can deal with high-dimensional data, and the results obtained are presented in the form of intuitive knowledge tree, which is easy to understand. In this paper, the data are summarized, sampled, and the training set is balanced to explore the data. The result is not obviously different from that of directly exploring the whole data set, but it can greatly reduce the time consumed.

1. Introduction

The competition between the mobile telecommunication industries has extended from the traditional voice and text fields to video calling, private lines, broadband, mobile phone reading, mobile phone payment, mobile phone games and other fields. At the same time, the emergence of Internet communication has a fierce impact on the traditional mobile telecommunications industry, and the mobile telecommunications market will usher in a more open, competitive and fierce new business collaboration environment. Mobile marketing is developed on the basis of the development of data business, which mainly uses mobile media to carry out various marketing activities, including collecting information, promoting sales, selling products and maintaining customer relations. At present, the main applications in China are SMS marketing, MMS marketing, WAP marketing, mobile game marketing, CRBT marketing, SMS website interactive marketing and so on [1]. In recent years, the achievements of the mobile marketing market are much lower than expected. Mobile telecom operators must pay attention to efficiency and cost, and at the same time improve the quality of service, innovate the direction of communication products, and give users brand-new excellent services and pleasant experience environment [2]. Marketing is the behavior that can bring value to customers, customers, partners and the whole society in the process of designing, exchanging products, delivering and communicating, and the application and research process and system of data mining in marketing. It is also a marketing behavior aimed at the whole market. The concept of marketing is divided into three categories: service consumption behavior is one kind of marketing; Marketing as a social phenomenon should be accepted; Marketing is very complicated from another angle, because it can connect the market and enterprises closely through some channels. Operators want to gain a foothold in the fierce market in the future, the key lies in improving their own industrial integration ability, increasing investment in scientific research and innovation and enhancing users' dependence on products and services [3]. This paper tries to study the application of mobile marketing market through decision tree. Decision tree algorithm of data mining constructs a decision tree by training historical data, and the process of constructing the tree does not need parameter setting and knowledge in related fields. The algorithm can handle high-dimensional data, and the results obtained are presented in the form of intuitive knowledge tree, which is easy to understand [4]. For the potential members of the mobile network, the system analyzes the call pattern characteristics of the potential group customer members obtained by classification, and obtains the users who have close contacts with these members. By finding out

which group user has the most contacts with this user in the stable communication circle, it can determine which group the user belongs to, and make targeted contact with it, and achieve the purpose of counterattacking by recommending suitable products and services [5]. Deepening the understanding of users' life scenes, expanding the scope of users' needs, and storing relevant big data liquidity, such as real-time mobile customer information under the base station, can provide decision-making basis for outdoor advertising location and advertising content matching. Based on the analysis results of users' data, behaviors and preferences, the whole customer labeling system intuitively describes users' basic characteristics and preferences, provides information support at the user level for key applications such as stock management and traffic management, and lays a foundation for new fields of data revenue generation in the future.

2. Optimization of Decision Tree Algorithm Based on Mobile Marketing Applications

2.1. The Defects of Decision Tree Algorithm in Practical Problems

The reason why decision tree algorithms have been widely applied in research is that they can learn from known historical data. The so-called decision-making is the process of selecting the optimal solutions for several feasible solutions that achieve the same goal. From this, it can be seen that decision-making is made among multiple feasible options, so decision-making is closely related to uncertainty, that is, decision-making has uncertainty [6]. Through this learning of data, a tree structure model with high information content can be obtained that can reveal the internal information and rules of the data. The target data can be classified based on this tree structured model. The decision tree algorithm can classify datasets very accurately and quickly; Can generate a simple, intuitive, clear, and understandable tree structure model. Due to the inherent shortcomings of decision tree algorithms, they may result in low accuracy or inability to solve business problems in practical applications. The decision tree algorithm selects split attributes based on the size of the entropy value of the classification information of that attribute [7]. However, in practical applications, attributes with more attribute values are often not sticky attributes that are of interest to practical problems. For example, in the mining of potential members of group customers studied in this article, operators are more interested in business bundle attributes such as "whether they are committed to the lowest consumption user", "whether they are group unified payment users", and "whether they are used for SMS packages". These business attributes only have two attribute values: "Yes" and "No", while the other user's own attribute "User uses service brand" has multiple attribute values.

2.2. Optimization of decision tree algorithm in terms of computational efficiency

China Mobile has accumulated a lot of mature experience in tag content planning, construction technology, analysis model and tag application. At the same time, it has continuously improved the tag library, enhanced the ability of tags to gain in-depth insight and timely grasp the dynamic needs of users, and continuously enhanced the work support for big data and micro-marketing. First, continue to enrich the user tags related to 5G services and provide a data foundation for the development of 5G+ services. The second is to adapt to the development of Vo LTE and broadband services, expand the source of tag data, enhance the analysis of users' Vo LTE communication behavior and broadband internet access behavior, and enrich related user tags. The precision of data and marketing has two sides. To make the direction of mobile marketing more accurate, it is necessary to put people first, accumulate from the most basic data, and value the unique advantages of media in this regard. Because on the internet, the details of any data are clearly visible, and the obtained data is massive and authentic, making every detailed need of netizens possible. The purpose of the algorithm is to minimize the error rate of classification. However, imbalanced data can lead to algorithms leaning towards the side with more data, meaning that classification algorithms have higher accuracy in classifying categories with larger amounts of data, while their recognition ability is poor for smaller categories. Moreover, the greater the data difference, the more obvious this bias becomes. Using classification algorithms on extremely imbalanced datasets,

even if all classes with a small number are divided into larger classes, the accuracy is still very high, which is meaningless for studying classification models with a small number[8]. By using decision tree algorithms to train historical datasets, a decision tree model is obtained, and then mobile users are classified using the decision tree model. The target attributes of the decision tree model training data only have two attribute values: ordinary users and group users. Due to the fact that the target attribute only has two values, the information entropy formula of the classification attribute A can be decomposed during calculation to achieve the goal of simplifying the calculation amount of the formula. Let the label set E of the target attribute values in dataset X have two values: (e_1, e_2) , $E = e_1$ with a dataset size of m , $E = e_2$ with a dataset size of n , and the sum of m and n with a total dataset size of $m+n=[X]$. The attribute A has v values (a_1, a_2, \dots, a_v) , and when the attribute $A = a_i$ has m_i $E = e_1$ data and n_i $E = e_2$ data, based on these conditions, the information entropy of the A attribute is:

$$Info_A(X) = \sum_{i=1}^v \frac{m_i + n_i}{m + n} \quad (1)$$

$$Info(X_i) = -\frac{m_i}{m_i + n_i} \log_2 \quad (2)$$

By substituting formula (2) into (1) and simplifying it, we can obtain:

$$Info_A(X) = \sum_{i=1}^v \frac{1}{(m+n) \ln 2} \quad (3)$$

After a series of mathematical calculations, it can be clearly seen that the calculation amount of the simplified formula will be greatly reduced compared with the original entropy calculation formula, and at the same time, the algorithm does not need to call the system function Math.log frequently to calculate the entropy value.

When using decision tree algorithm in the field of mobile marketing, the main channels to find business opportunities are to correctly distinguish customer categories, choose reasonable customers to communicate with and find relevant document information to match. It can provide powerful and stable data computing ability, gain real-time insight into consumer behavior and respond in time; The popularity of mobile terminals makes data analysis feasible anywhere. Decision tree algorithm analysis and marketing planning are driven by each other, which makes the objectives of marketing activities more clear, traceable, measurable and optimized, thus creating a closed-loop management model of marketing analysis with data as the core. It is far from enough to rely on the previous experience of the marketing department. Using data mining technology can find out the data information that improves the efficiency and competitiveness more accurately and quickly, which is the advantage of decision tree algorithm [9]. The optimized decision tree algorithm in this paper is mainly based on the improvement of the specific business of the potential false mining of mobile group customers. The following samples are selected from the mobile test data for the efficiency test of the algorithm. The test was carried out under the environment of Intel Core I5 4-core processor, 4G memory and Win764-bit. The test results are shown in Table 1.

Table 1 Comparison of Algorithm Accuracy

Number of test pieces	Original decision tree accuracy	Optimized decision tree accuracy
100	84.12	94.25
500	82.45	91.15
1000	78.54	88.47
5000	75.49	85.72
10000	74.15	78.57

From the above test results, it can be seen that the larger the dataset, the more significant the efficiency improvement of the optimized decision tree algorithm. Although the accuracy of both the original decision tree algorithm and the optimized decision tree algorithm decreases as the dataset becomes larger, the accuracy of the optimized decision tree algorithm has always been higher than that of the original decision tree algorithm. The results show that the optimized decision tree algorithm can effectively improve the computational efficiency of the algorithm, improve the bias problem of the decision tree algorithm, and improve the accuracy of the algorithm. Through the data analysis of old users, we can optimize the network configuration. For example, in the background of data analysis, we can clearly see the time and place when users talk and surf the Internet, and even the movement track. The visualization and integration of these data can bring practical guiding significance to resource allocation, including base station optimization. Second, customer perception is better. Based on a large number of factual data, we can analyze users' preferences and habits in terms of age, occupation, education, income and other dimensions, and set "labels" for users to know themselves better than users.

3. Application of Optimization Algorithms in Group Customer Potential Member Mining System

3.1. Module composition of group customer mining system

Marketing is one of the most important activities of enterprises, and any marketing activity cannot do without information, such as sales, customers, services, quality, personnel information, etc[10]. With the development of data collection, storage, management, analysis, mining, application and other technical systems, the answers to these questions can already be revealed. Through data mining and analysis, the treasures hidden in the data Wang Yang can be salvaged; The data fusion of each channel improves the accuracy of precision marketing; Visualization technology polishes complex data into intuitive graphics, making it a tool and means that is easy to understand and available to everyone. In order to effectively support marketing decisions, it is necessary to collect, process, and analyze this information. The potential false member mining system for group customers is divided into five modules: overall overview, potential customers, false customers, sticky customers, and contact records. The potential customers, false customers, and sticky customer modules all include two sub modules: member analysis and member mining. The functional modules of this system are shown in Figure 1.

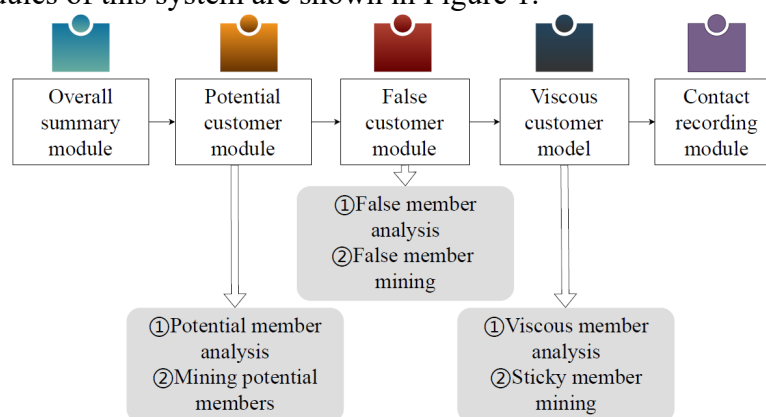


Figure 1 Distribution diagram of potential false mining system modules for group customers

The potential customer module, false customer module, and sticky member module are all based on improved decision tree algorithms to establish mining models, classify users based on the models, and display them using visual technologies such as report graphics. These three modules are further divided into two sub modules: member analysis and member mining. Member analysis is a list of various users obtained based on mining model classification, and each user is analyzed based on their own attributes, business attributes, and other related modules.

3.2. Mining model of potential members of group customers

The commercial operation support system, user segmentation system, key customer marketing system, group customer marketing system and customer service center system owned by mobile companies have accumulated a large number of user historical data, which provide reliable data sources for establishing data mining models. Sales service personnel compile the collected market information into short messages according to the specified format, and send them to the short message cat. The short message background service software on the server reads and identifies the short messages in the short message cat, extracts the short message data with correct format and writes it into the database of the system, and replies the prompt information to the marketing service personnel. The data involved are: user basic information, user business information, user billing information, user communication circle information and group customer information. Data is the object of data mining, so it is necessary to know the structure, attributes, types and other information of data before modeling. Therefore, technically, operators such as China Mobile and China Unicom can set up a security cordon to filter and monitor the content of each mobile phone short message, and strictly review the traffic that greatly exceeds the normal situation. Behind the data is the analysis and application work. China Mobile's data analysis has changed from equipment optimization to network optimization, with marketing data analysis and innovative application to reduce costs and increase revenue. In the future, the company will realize the innovation of enterprise products and services through big data analysis and realize the comprehensive upgrade of the company's industrial chain. Mobile phone manufacturers should also develop some products with the functions of blocking spam messages and short message firewall to provide technical support from receiving terminals. The data of mobile telecom operators is massive and huge, so it often takes a long time to explore the data from the beginning, which is very time-consuming and unnecessary. Here, we can explore the data after summarizing, sampling and balancing the training set. The result is not obviously different from that of directly exploring the whole data set, but it can greatly reduce the time consumed.

4. Conclusions

As more and more people use mobile devices, safety hazards are gradually being taken seriously. Although operators have made significant progress in the construction and improvement of security mechanisms, the reality is that consumers still cannot avoid being harassed and deceived by spam messages. As a large customer with high contribution to the operator, group customers can stably bind a large number of users through their own influence. After a thorough interpretation of the potential false mining system for mobile group customers, this article takes the basic knowledge of data mining as the starting point. By introducing the optimized decision tree algorithm into the potential false mining system for group customers, accurate classification storage processes can be obtained through this algorithm, and the target user dataset can be classified based on the storage process. When information is sent from one device to another, due to the incompatibility of mobile devices, there may be situations where the receiving device cannot receive it properly or incompletely, which greatly affects the effectiveness of mobile marketing. This article focuses on the research and analysis of decision tree algorithm and its improvement, while applying the algorithm to the potential false mining system for group customers. However, the selection of data and the analysis of data preprocessing are relatively single, and future research should focus on data preprocessing. By using more data mining techniques to process data more professionally and effectively, the data can fully reflect business issues.

References

[1] Pedro José Pereira, Cortez P, Mendes R. Multi-objective Grammatical Evolution of Decision Trees for Mobile Marketing user conversion prediction[J]. Expert Systems with Applications, 2020, 168(20):10-17.

- [2] Kapinga A F, Suero Montero C, Mbise E R. Mobile marketing application for entrepreneurship development: Codesign with women entrepreneurs in Iringa, Tanzania[J]. *Electronic Journal of Information Systems in Developing Countries*, 2019, 19(7):12-23.
- [3] Sarkar S, Khare A, Sadachar A. Influence of consumer decision-making styles on use of mobile shopping applications[J]. *Benchmarking*, 2020, 11(1):27-32.
- [4] Gu Z, He C. Application of Fuzzy Decision Tree Algorithm Based on Mobile Computing in Sports Fitness Member Management[J]. *Hindawi*, 2021, 18(2):11-19.
- [5] Awodiji T O. The Importance of Decision Tree Analysis on Strategic Management Practice: Evidence from Retail Industries in US[J]. *Journal of Marketing and Consumer Research*, 2022, 26(8):10-14.
- [6] Saranya V, Drr P. A Study and Analysis of Decision Tree Based Classification algorithms using R[J]. *Expert Systems with Application*, 2018, 26(11):25-34.
- [7] Pereira P J, Cortez P, Mendes R. Multi-objective Grammatical Evolution of Decision Trees for Mobile Marketing user conversion prediction[J]. *Expert Systems with Application*, 2021, 15(Apr.):168-175.
- [8] Zheng Y. Decision Tree Algorithm for Precision Marketing via Network Channel[J]. *International journal of computer systems science & engineering*, 2020, 11(4):35-41.
- [9] Lau E K W. Precision Marketing: A Mobile Application Case from Hong Kong[J]. *Journal of international business and economics*, 2020, 22(3):20-28.
- [10] Manjupriya R, Poornima A. Customer Churn Prediction in the Mobile Telecommunication Industry Using Decision Tree Classification Algorithm[J]. *Journal of Computational and Theoretical Nanoscience*, 2018, 15(9):2789-2793.