The Construction of Consumer Buying Behavior Analysis System Based on Data Mining in Social Marketing

Di Zhang[1,2], Dongphil Chun[1] and Minghao Huang[3*]

[1]*Pukyong National University, Graduate school of Management of Technology, Busan, 48548, Korea*
[2]*Economic and Management School, Hulunbuir University, Inner Mongolia Hulunbuir 021008, China*
[3]*Seoul School of Integrated Sciences & Technologies, Seoul, 03767, Korea*

Abstract: This paper takes data mining technology as the core, completes the design and construction of machine learning models such as K-mediod algorithm, Decision Tree algorithm and Apriori algorithm by using Numpy module in Python language environment, and realizes the development of user classification, feature prediction and correlation analysis of consumer purchasing behavior data under social marketing. It combines Django development framework to realize the integration and encapsulation of various functional modules, and finally forms a web-based consumer purchasing behavior analysis system. The system adopts B/S architecture, and completes the deployment of all levels of the system and the planning and design of business logic according to MVC mode. It can facilitate users to discover the influence of social marketing, product attributes, service methods and other factors on consumers' purchasing behavior through concise and efficient operation, and take reasonable measures to improve social marketing strategies, adjust service methods, enrich service content, and maximize the benefits of enterprises.

1 INTRODUCTION

With the continuous change and improvement of network information technology, the social marketing is a "new marketing" model based on social relations, which uses the daily social activities of social media users and the dissemination of their own content to spread the brand information of enterprises. The social marketing will no longer focus on products, but turn to consumers, and its content and form are obviously different from the traditional marketing. More information outside the product, such as application scenarios, content expression, community groups, and psychological sense of belonging, will be relied on to promote the generation of consumers' purchasing behavior. So, the logical chain of consumer buying behavior has changed from "attention, choice and purchase" in the traditional marketing mode to "sharing, purchasing and paying attention". (Liu, 2021)

For enterprises, how to choose a suitable platform, formulate a perfect social marketing strategy, and realize the interaction, locking and transformation with the target audience has become the key

to "stand out". So, this paper thinks that taking data mining technology as the core, in Python language environment, using Django frame architecture to complete the construction of consumer buying behavior analysis system. Through the analysis and mining of consumers' purchasing behavior, the system enables the e-commerce platform to comprehensively obtain consumers' real needs, develop customer value, improve customer service quality, and provide guidance for the formulation, implementation and management of its social marketing strategy, and finally realize the growth and efficiency of enterprises.

2 OVERVIEW OF KEY TECHNOLOGIES

2.1 Data Mining Technology

As a kind of computer science and technology, data mining technology aiming at the complex process of extracting and mining hidden and valuable patterns

or laws from a large number of incomplete and noisy actual data. (Zhang, 2021) The realization of general data mining involves three steps: data preparation, data mining and result presentation. In the stage of data mining, we will focus on the selection of data samples, sample features and the establishment of algorithm models.

2.2 Python

The Python is a high-level scripting language that combines interpretive, compiler, interactive and object-oriented. 3]. The Django is a custom Web development framework based on Python. Its overall structure is similar to MVC pattern, but not exactly the same as MVC pattern. Under Django framework, we focus on Model, Template and Views, also known as MTV pattern. (Wang, 2021)

2.3 Development Process

In the data preparation stage, the data collection comes from the data of consumers' purchasing behavior in the enterprise's own e-commerce platform on the one hand, and the feedback from the social questionnaire on the other hand. After pre-processing, the two kinds of data will collect a total of 56 attribute variable values from three aspects: consumers, e-commerce and social influence, covering the whole process of e-commerce. In the stage of data mining, according to the data of consumers' purchasing behavior and questionnaire, three attribute variables of purchasing frequency, purchasing amount and latest purchasing time are selected, and the RFM model of K-mediod algorithm is adopted to divide consumers into three categories: important value consumers, general value consumers and worthless consumers. The key code of function realization is shown in Figure 1.

In the face of numerous attribute data, based on the calculation results of the dependency model, 10 key attributes are selected as the input attributes of feature prediction, and the purchase behavior rules of different types of consumers are further mined through the Decision Tree algorithm model. The implementation of Decision Tree algorithm model will be realized by C4.5 algorithm, and the key code is shown in Figure 2.

```
def distEclud(vecA, vecB):
    return np.sqrt(np.sum(np.power(vecA-vecB,2)))
# def cal_dist(dataMat, centroids, k):
#     n = np.shape(dataMat)[0]
#     dist = []
#     for i in range(n):
#         dist.append([])
#         for j in range(k):
#             dist[i].append(distEclud(dataMat[i, :], centroids[j]))
#     dist_array = np.array(dist)
#     return dist_array
def total_cost(dataMat, medoids):
    med_idx = medoids["cen_idx"]
    k = len(med_idx)
    cost = 0
    medObject = dataMat[med_idx,:]
    dis = cdist(dataMat, medObject, 'euclidean')
    cost = dis.min(axis=1).sum()
    medoids["t_cost"] = cost
```

Figure 1: Cluster analysis results of K-Mediod algorithm [Original].

```
def createTree(dataSet, labels):
    classList = [example[-1] for example in dataSet];
    if classList.count(classList[0]) == len(classList):
        return classList[0];
    if (len(dataSet[0]) == 1):
        return majorityCnt(classList);
    bestFeat = chooseBestFeatureToSplit(dataSet);
    bestFeatLabel = labels[bestFeat]
    myTree = {bestFeatLabel:{}}
    del(labels[bestFeat])
    featValues = [example[bestFeat] for example in dataSet]
    uniqueVals = set(featValues)
    for value in uniqueVals:
        subLabels = labels[:]
        myTree[bestFeatLabel][value] = createTree(splitDataSet(dataSet, bestFeat, value), subLabels) #
    return myTree
```

Figure 2: Key code of Decision Tree model implemented by C4.5 algorithm [Original].
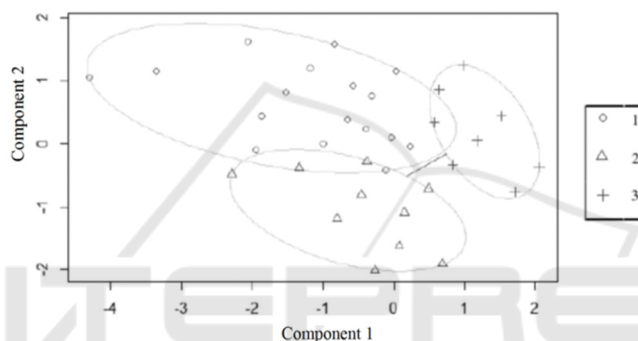

Figure 3: Consumer cluster analysis results of RFM model [Source: Network].

After further accurate acquisition of the consumer purchasing behavior rules, the future purchasing behavior will be predicted from the correlation rules analysis of the products, services and the marketing methods concerned about. The association rules are realized by Apriori algorithm, in which the key support and confidence indexes determine the strength of the association, and are used to guide the enterprise's social marketing decision and scheme formulation.

The Web application server is implemented by Django framework, the operating system is Windows10.0. The Web server is Nginx server, the project development language is Python 3.6.7, the development tool is PyCharm 2018.3.1 x64, and the database is MySQL5.7 to complete the construction and support of the system database system. The whole server is implemented by Django framework. Through the introduction of the above key technical theories, the overall environment of the system development, the configuration of related software and tools are determined, and the technical feasibility of

the overall project of the consumer buying behavior analysis system is also clarified.

3 FUNCTION REALIZATION

3.1 Data Import

When the user enters the system, under the data import function module, he selects the questionnaire data prepared in advance and the purchase behavior data of the enterprise's own e-commerce consumers to upload to the system database.

3.2 Customer Classification and Feature Selection

In this function module, the system will support three functions: consumer classification, attribute feature screening and purchase behavior rules mining of consumers. In the user interface, select the sample data, click on the consumer classification, and the system will automatically call the RFM

model of K-mediod algorithm to complete the consumer classification, and display the classification results, as shown in Figure 3.

After obtaining the classification results, users can see the corresponding chart contents under the details page. Meanwhile, we will also see all the attribute variables and 10 key attribute features contained in this part of sample data, including preferential way, price, search time, communication initiation time, collection number, comment number, e-commerce platform functionality, social recommendation, login mode, recommendation index, etc., will serve as an important basis for the prediction of consumers' purchase behavior rules, and will be input into the Decision Tree algorithm model for consumer purchase behavior prediction.

3.3 Preference Relevance Analysis

According to the purchase behavior rules of different types of consumers, select important value consumers and general value consumers for data analysis, and establish high-quality association rules from three aspects: products, services and concerned marketing methods. According to Apriori algorithm, the correlation data of different categories of consumers can be obtained, which provides scientific data basis for the subsequent formulation of social marketing strategies, the operation and sales of enterprise products and the improvement of service methods. (Zhang, 2021)

4 CONCLUDING REMARKS

In this paper, based on the difficulties faced by the social marketing strategy and decision-making process of e-commerce platform in the era of digital economy, an online interactive consumer buying behavior analysis system with data mining technology as the core and Web application technology as the framework is proposed. The system can realize user classification, feature prediction and correlation analysis, find social marketing, product attributes, service factors and other influence on consumer purchase behavior, and take reasonable measures to improve social marketing strategy, adjust the service mode, rich service content, to maximize the interests of the enterprise.

REFERENCES

Du Jiming. The analysis of Web development application of Python language [J]. Digital Communication World.2022.01

Liu Bingling. The Analysis and Research on the Influencing Factors of Social Media Marketing on Consumer Behavior [J]. China-Arab States Science and Technology Forum.2021.07

Wang Yang, Jiang Xintong. The Design Research of MVC Framework in Python and Django [J]. Computer and Information Technology.2021.02

Zhang Peng. The Data Mining Technology and Practice in the Age of Big Data [J]. China New Telecommunications.2021.11

Zhang Mengqi. The Analysis of Association Rules Based on Apriori Algorithm [D]. Dalian University of Technology .2021.04