# Business Application of Machine Learning Technology in Data Mining

**Jin Zhao**[*]

*Lyceum of the Philippines University, Philippines*

*zhaojin20072010@126.com*

[*]*corresponding author*

*Abstract:* In today's business field, data mining technology has become an indispensable part, and it is also a hot research direction. This paper mainly discusses and analyzes machine learning algorithm. First, we will introduce the classification, clustering and classification models of biological samples. Secondly, we will briefly discuss two different methods based on artificial neural network and hybrid Bayesian tree algorithm in theory. Finally, we will draw a conclusion through experimental comparison that the performance of machine learning in business application models in data mining is relatively reasonable, and the model processing time is relatively short, which meets the needs of users. At the same time, this paper also puts forward relevant suggestions to enhance the application value of this technology in business, so as to provide some help for enterprises to mine effective information.

## 1. Introduction

In the business value of data mining, machine learning technology is a very important and ubiquitous application, which can store and have been recorded in some historical databases but not used or not found and used. Bhaskar G proposes deep learning as a skilled technique to examine hazards and react to attacks and security events. Thus, this business positions both security and energy productivity in the IoT, utilizing two new strategies aided by deep learning. This work ends up burning less energy, taking less time to transmit information, and giving greater security when considering existing frameworks [1]. Dogra A K proposes that data collected through embedded devices in objects are analyzed to take important decisions that help society to use these applications in different types of services. By applying machine learning (ML) techniques, these applications are becoming intelligent and can produce better results by solving traffic management

problems and providing safety for residents. The development of new applications that use data from IoT devices to provide meaningful results and make things intelligent through ML algorithms is presented[2]. Lu H studied the entrepreneurial model of remote intelligent classroom, based on machine learning technology, combined with intelligent image recognition technology to recognize the status and expressions of students in the distance education classroom. In addition, this paper conducts a more detailed study of face detection and expression recognition technology and tries to apply them to classroom teaching evaluation, which shows some feasibility in the experiment[3].

Machine learning research abroad began in the 1930s, mainly based on data mining, rather than real data mining[4]. At present, the United States and the United Kingdom are the leading countries in the world, and some research on image analysis has also been carried out in China. For example, for the image classifier model in terms of algorithm and technology, from the perspective of image retrieval, the method of mapping the content of interest to clusters by clustering (association scanning), and then judging whether it is the best mark or target information by similarity is an effective, fast and efficient method [5-6]. Some scholars put forward and applied it to data mining. He based on clustering theory and established a statistical sample classifier to deal with high-dimensional variable information using similar cosine similarity measure. Others believe that although some progress has been made in this area, it has not yet met people's expectations. Therefore, this paper studies data mining in commercial applications based on machine learning technology.

With the development of database technology, the amount of management data is increasing. How to obtain useful knowledge from massive information has become a topic of concern in the research of business value and market competitiveness. This paper mainly introduces the application of machine learning in the computer field to analyze items of interest to mining users, as well as the process of their prediction and classification. Through specific steps such as model training algorithm and clustering method, it proposes solutions to the problems existing in traditional data mining algorithms and effectively applies the improved data mining algorithms to the actual scene, obtaining practical results and bringing profits and considerable value to enterprises.

## 2. Discussion on the Business Application of Machine Learning Technology in Data Mining

### 2.1. Common Commercial Activities

In daily work, we can see some commercial activities based on data mining, such as advertising, sales promotion and others [7-8]. Figure 1 shows the application process of data mining technology in business activities.
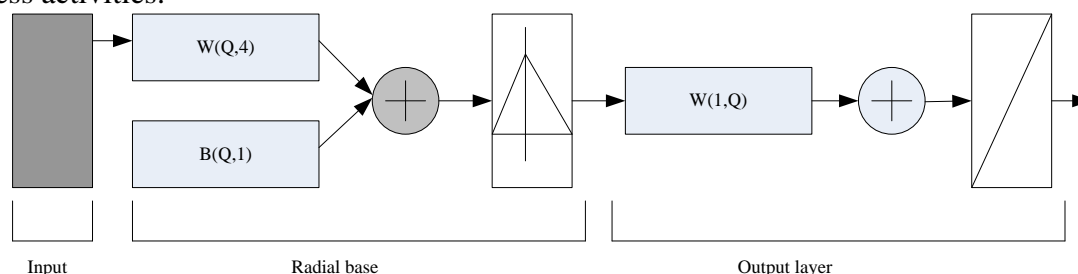


*Figure 1. The use process of the data mining technology in commercial applications*

(1) Advertising. It is one of the processes of building a database based on the business model to obtain income information from users and forming a customer group after analysis. It may also be used in marketing strategies to directly obtain the structure of the target group and make purchase

decisions for potential customers. Data mining technology is used in the work content, which is widely used in the field of data mining and deeply rooted in people's minds. That is, through the collection, classification and processing of a large amount of information, useful users are interested in or have been familiar with and trusted by them, or they may choose some buyers who do not know the product very well or do not like the business to find others to provide service support as potential customers, so as to generate new customer groups after market segmentation and establish good business reputation. When we browse the web, we can transmit information to the outside world through words and pictures [9-10]. However, due to a large number of unrecognized structural points, non-quantitative relations and other problems in the data warehouse, businesses cannot effectively classify and recommend these goods to users. However, it is easy to obtain satisfactory results by using machine learning technology to integrate them into the content stored in the database and the images at the corresponding locations for comparative analysis.

(2) Marketing activities. Predict the demand of potential customers through the consumption habits and behavior information of users. These potential consumers will generally, to a certain extent, decisively choose to use products or enjoy relevant experience and treatment when purchasing goods or services. The whole business process of obtaining customer information and making corresponding sales plans by using data for market research, prediction and purchase is called marketing activities [11-12]. In marketing activities, data mining technology is a typical application. Its main purpose is to obtain useful knowledge by analyzing and researching a large number of potential users' information. It can link some valuable or important business opportunities with commodities, and help enterprises effectively integrate the information they need to know. After analyzing the relationship between user needs, purchase motivation and other factors, we can formulate the next development direction and plan that is suitable for customer psychology and behavior habits as well as product strategies. We can also adjust the commodity structure or develop new product R&D activities according to the forecast results to meet market requirements and achieve sales revenue.

(3) Search engine. It uses a large amount of customer information in the database as a knowledge source to obtain market opportunities and profit points. Social networks (WeChat friends circle) are mainly used for interactive communication between consumers and potential consumers to contact with existing competitors or other websites. You can also use the rules of information to obtain the required content, and provide relevant services and supporting help when users perform corresponding retrieval, so as to achieve rapid and effective transmission, processing and sharing of information. Data mining technology finds potentially useful or useless parts by analyzing the characteristics of a large number of raw data in massive databases [13-14].

## 2.2. Data Mining Technology

Data mining is to obtain useful information through new statistical rules and methods, so as to predict things or phenomena. Therefore, we can carry out some prediction and reasoning tasks according to this feature, use clustering and organizational structure model to classify their relationship types, and then obtain the relevant feature information between different categories, and form a new category data set. We can use decision trees to achieve classification and de specify the regularization form, so that the training set can more effectively complete tasks. At this stage, data mining technology mainly has two classifications, rule analysis based and support vector machine algorithm [15-16]. According to different research objects or purposes, it can be divided into model types such as modeling and regression for specific problems, and model types defined for typical human decision-making processes with certain similarity. According to research tools, it can be roughly divided into machine learning methods, statistical inference methods, and artificial

neural networks. In the actual research, it is found that people's views on things will change over time. According to the business intelligence website, different types of users use their own needs and other product related attributes to make a simple description of the commodity category, and consider how data mining technology is implemented, and whether it can help enterprises make better management decisions, so as to get more effective prediction results. Conditional probability of occurrence of Y in transaction X:

$$\text{Confidence}(X \Rightarrow Y) = \frac{\text{Support}(X \cup Y)}{\text{Support}(X)} \tag{1}$$

Rule X=>Y's support in database D is the ratio of the number of transactions containing X and Y in the transaction set to the number of all transactions, which is recorded as:

$$\text{Support}(X \Rightarrow Y) = \text{support}(X \cup Y) \tag{2}$$

It includes several different types of algorithms, such as feature based and non-parameter grouping methods and decision tree based modeling, to pre train the data set, and then use online selection of viscosity coefficient, similarity index and information entropy to realize the probability estimation of machine learning.

## 2.3. Machine Learning Algorithm

Machine learning algorithm is a way to obtain data and knowledge through computers. It is mainly composed of three parts, sample classification, clustering and training set. Among them, sample analysis includes two aspects. First, a large amount of original information contained in the database should be extracted, and then the results of these samples should be converted into feature space values to represent the relationship between input and output. The structure of artificial neural network is one of the typical BP algorithm models. The algorithm learns and imitates the internal structure of the human brain neural system. In the process of classification, we should first extract the features of the samples. Then the potential information is mapped to the computer according to the inherent relationship between samples.
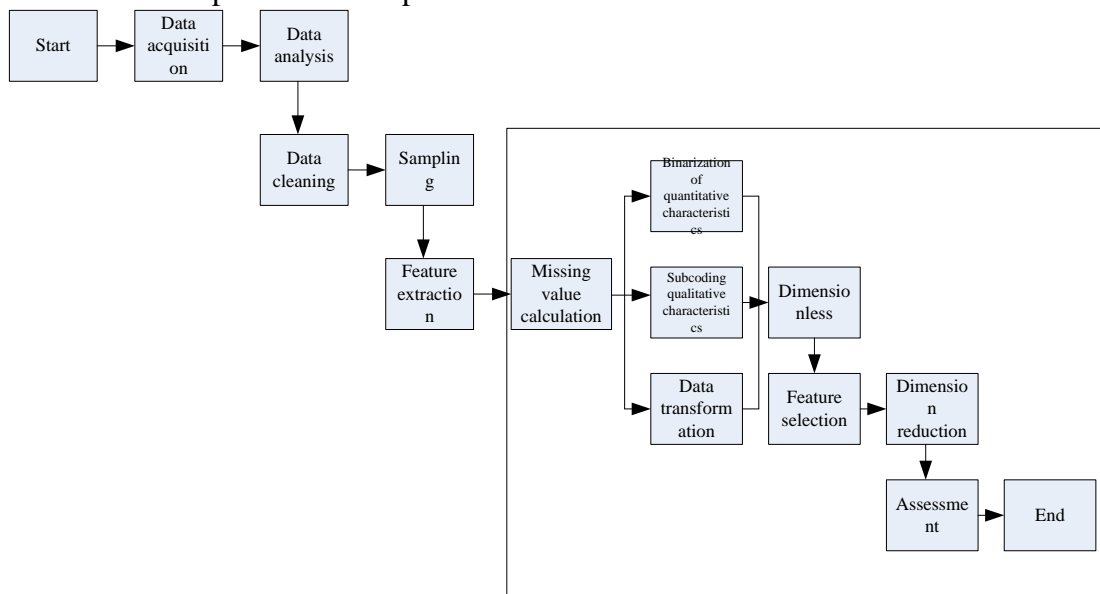


*Figure 2. The process of machine learning technology*

Through this process, we can get the output results of the training set and test set, that is, the model output (as shown in Figure 2). According to the above algorithm steps, we know that the

node is the most important in data mining. It determines whether a random event will occur or how much its probability will change by experimenting with it, calculates the weight required in these results, and then quantifies the process to get the final conclusion and compare it with the actual situation. Finally, make decisions based on the results obtained, and give implementation plans and methods [17-18]. In order to classify a given sample $x \in R$, we first need to obtain k nearest neighbor samples of each sample in the transformed feature space F, and then find an optimal separating hyperplane on the k nearest neighbor samples. In fact, that is to say, the local support vector machine algorithm builds the training model on the k nearest neighbor samples in the feature space F of each training sample x, unlike the traditional support vector machine, which builds the training model on all samples. The decision rules of local support vector machine are as follows:

$$\text{KNNSVM}(x) = \text{sign}(\sum_{i=1}^{k} \alpha YK(x_{r_x(i)}, x) + b)$$

(3)

The function of r reordering training samples is obtained from the dual optimization problem of support vector machines using nearest neighbor points as training sets.

## 3. Experimental Process of Business Application of Machine Learning Technology in Data Mining

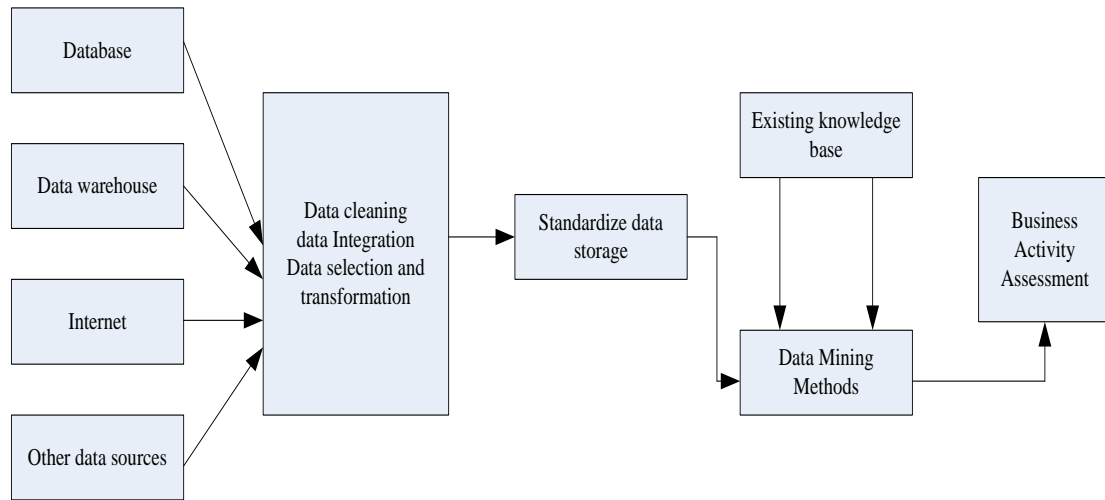### 3.1. Business Application Model Based on Machine Learning Technology



*Figure 3. Business application model based on machine learning*

The business application model based on machine learning (as shown in Figure 3) is a new system built based on data mining technology, statistical analysis, prediction analysis results as the research core, and optimization of existing algorithms. Through face-to-face communication and discussion with experts, mutual communication and interaction can be formed to improve team cohesion and enthusiasm. At the same time, data mining technology can also be used for purposeful learning to build models and achieve significant results in prediction. The algorithm analyzes and studies decision-making problems through data mining, so that better economic benefits, social effects and other related aspects depend to a large extent on the very accurate matching between the means used for data information processing and model parameters, and the business application model based on machine learning technology provides a more complete and accurate prediction system.

### 3.2. Performance Test of Business Application Model Based on Machine Learning Technology

In the process of testing the business application model, we need to analyze the performance and effect of the developed model, so as to obtain an accurate, reliable, correct use of data mining results and a high degree of prediction ability and timely feedback to users. First, it is necessary to determine whether the algorithm can obtain enough experimental samples, and judge whether its running speed, stability and other parameters can reach the expected value or exceed the standard through testing. Second, it is necessary to verify and explain the accuracy of the required data during the running process to ensure the validity and rationality of the model.

### 4. Experimental Analysis on Business Application of Machine Learning Technology in Data Mining

### 4.1. Performance Test Analysis of Business Application Model Based on Machine Learning Technology

Table 1 shows the performance test data of the commercial application model.

*Table 1. Commercial application model performance test*

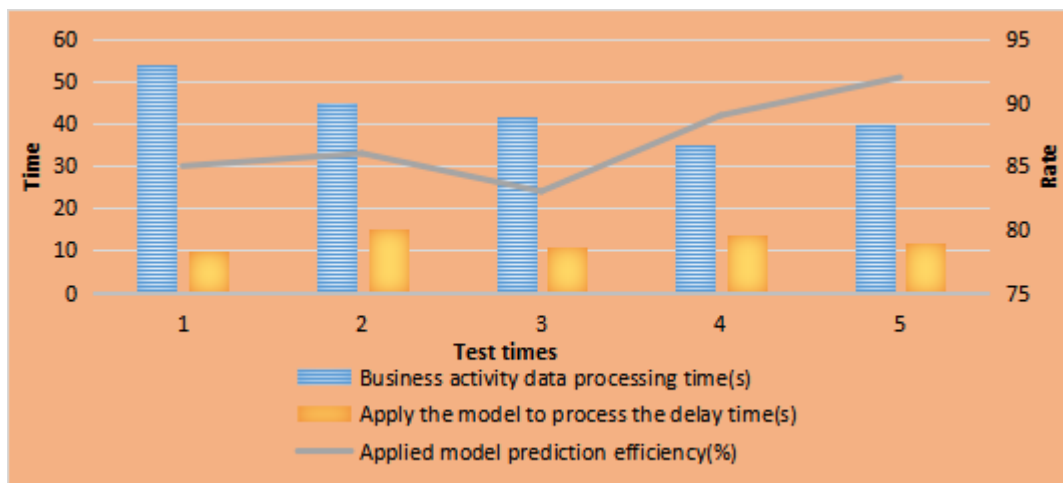| Test times | Business activity data processing time(s) | Apply the model to process the delay time(s) | Applied model prediction efficiency(%) |
|---|---|---|---|
| 1 | 54 | 10 | 85 |
| 2 | 45 | 15 | 86 |
| 3 | 42 | 11 | 83 |
| 4 | 35 | 14 | 89 |
| 5 | 40 | 12 | 92 |



*Figure 4. Performance test of commercial application models based on machine learning technology*

In the process of testing, the performance of the data mining model mainly depends on the experimental results, while the commercial application software has high requirements for algorithms and parameters, so it is feasible to use machine learning technology to test. At present, this paper uses the simulation evaluation method to calculate the accuracy, experimental efficiency and analysis accuracy of sample items, and uses a large number of data sets and tools to test

whether there are significant differences in the performance of data models under different conditions. It can be seen from Figure 4 that the performance of machine learning in the business application model in data mining is relatively reasonable, and the model processing time is relatively short, which meets the user's needs.

## 5. Conclusion

In the computer field, data mining is an important research direction, and one of the main applications is machine learning technology. This paper will analyze and discuss it from several aspects. Firstly, this paper introduces some research achievements and applications in this field in our country at present. Secondly, it outlines what problems exist in the development process of current machine learning technology and summarizes the reasons. Finally, based on this, this paper proposes a model to effectively reduce the prediction error rate, which provides a certain degree of application in the relationship similarity between research objects through data mining algorithms.

## Funding

This article is not supported by any foundation.

## Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

## Conflict of Interest

The author states that this article has no conflict of interest.

## References

[1]   Bhaskar G, MD Reddy, Thatikonda M. A Review on Secure Data Transmission for Banking Application using Machine Learning. International Journal of Engineering and Advanced Technology, 2021, 10(5):182-186.

[2] Dogra A K, Kaur J. Moving towards smart transportation with machine learning and Internet of Things (IoT): a review. Journal of Smart Environments and Green Computing, 2022, 2(1):3-18.

[3] Lu H. Application of wireless network and machine learning algorithm in entrepreneurship education of remote intelligent classroom. Journal of Intelligent and Fuzzy Systems, 2021, 40(2):2133-2144.

[4]   Krischke U. Rhona Alcorn, Joanna Kopaczyk, Bettelou Los and Benjamin Molineaux (eds.). 2019. Historical Dialectology in the Digital Age. Edinburgh: Edinburgh University Press, xvii + 274 pp. 42 figures, 33 tables, 80.00.. Anglia, 2020, 138(1):166-170.

[5] S Sigurjonsdottir, Nowenstein I. Language acquisition in the digital age: L2 English input effects on children's L1 Icelandic:. Second Language Research, 2021, 37(4):697-723.

[6]   Sihite M, Manullang S O, Nugroho B S. Relevance of mastery of information systems skills and success of business management in the digital age: a systematic review. International Journal of Social Sciences and Humanities, 2021, 5(2):68-78.

[7] M Pap k, L Pap ková Application of selected data mining techniques in unintentional accounting error detection. Equilibrium, 2021, 16(1):185-201.

*[8] Dadouh A, Aomari A. Moroccan TV Broadcasters and Viewership Changes in the Digital Age: An Exploratory Study. European Journal of Business Management and Research, 2021, 6(1):232-237.*

*[9] Wang J, Cao S J, Yu C W. Development trend and challenges of sustainable urban design in the digital age:. Indoor and Built Environment, 2021, 30(1):3-6.*

*[10] Mandal P C. Public Policy Issues and Technoethics in Marketing Research in the Digital Age. International Journal of Technoethics, 2021, 12(1):75-86.*

*[11] Rahmat A, Syakhrani A W, Satria E. Promising online learning and teaching in digital age: systematic review analysis. International Research Journal of Engineering IT & Scientific Research, 2021, 7(4):126-135.*

*[12] Razmetaeva Y S, Sydorenko O O. Abortion, Human Rights And Medical Advances In Digital Age. Wiadomości lekarskie (Warsaw, Poland: 1960), 2021, 74(1):132-136.*

*[13] Vartouni A M, Teshnehlab M, Kashi S S. Leveraging Deep Neural Networks for Anomaly-Based Web Application Firewall. IET Information Security, 2019, 13(4):352-361.*

*[14] Kadiri K, Oluwatoyin I N, Akanbi I I. Design of deep learning system for agricultural purpose. International Journal of Information and Communication Technology, 2021, 2(1):30-35.*

*[15] Cybulski J L, Scheepers R. Data science in organizations: Conceptualizing its breakthroughs and blind spots. Journal of Information Technology, 2021, 36(2):154-175.*

*[16] Kohsaka R, Fujihira Y, Uchiyama Y. Biomimetics for business? Industry perceptions and patent application. Journal of Science and Technology Policy Management, 2019, 10(3):597-616.*

*[17] Rana M E, Wang W. A Machine Learning based Software Project Schedule Management Solution. Test Engineering and Management, 2020, 83(1):307-321.*

*[18] Mimoun J. Technology Focus: Well Testing (February 2021). Journal of Petroleum Technology, 2021, 73(2):51-51.*