ISSN:2790-1688 DOI: 10.56028/aetr.4.1.582.2023

Research on Machine Learning and Intelligent Decision Support System Based on Risk Prediction

Haoshu Qin

School of Computer and Software Engineering, Huaiyin Institute of Technology, Huaian, 223001，China

Email: 1181301134@smail.hyit.edu.cn

Abstract. Decision Support System (DSS) is a new technology developed in recent years. Intelligent Decision Support System (IDSS) is the product of the combination of management decision science, operational research, computer science and artificial intelligence. For decision-makers, it takes a long time to accumulate experience to reach a rich enough level of decision-making. On the other hand, the decision-making process is also influenced to a great extent by the subjective consciousness of the decision-maker and the interference of external factors, which makes it easy to make analysis with large deviation, and then affects the accuracy of decision-making. With its powerful ability to learn data automatically, extract complex patterns and make intelligent decisions, machine learning method has attracted more and more attention in biomedical and financial research fields, providing a new way to effectively solve decision support problems in these fields. Based on risk prediction, this paper analyzes the application of machine learning method in intelligent decision support system.

Keywords: Decision support system, Risk prediction, Machine learning.

1. Introduction

At present, data classification has become one of the most active and important research topics in the fields of data mining, pattern recognition and machine learning, and has a very wide range of applications in practice [1]. Many important decision-making problems in real life, such as disease diagnosis and financial risk prediction, can be attributed to classification problems. Decision support system (DSS) is a new technology of computer support management decision-making developed in recent years. Intelligent decision support system (IDSS) is a combination of management decision science, operational research, computer science and artificial intelligence [2]. By using expert system technology, the decision-maker's modeling economic experience is sorted into the knowledge expressed by computer in advance, and organized into the knowledge base. A group of programs called inference engine are used to simulate the decision-making expert's thinking reasoning, forming an intelligent component [3]. For decision-makers, it takes a long time to accumulate experience in order to achieve sufficient decision-making level. On the other hand, the decision-making process is also largely affected by the subjective consciousness of the decision-maker and the interference of external factors, which is easy to make a large deviation analysis, thus affecting the accuracy of decision-making [4]. Intelligent decision support system not only gives full play to the characteristics of expert system solving qualitative analysis problems in the form of knowledge reasoning, but also gives full play to the characteristics of decision support system solving quantitative analysis problems with model calculation as the core, which fully achieves the organic combination of qualitative analysis and quantitative analysis, not only expands the scope of DSS to solve problems, but also improves the ability to solve problems [5-6].

With the development and application of artificial intelligence technology, knowledge has become the most important part of information system. But in the use of decision support system, users hope it can be intelligent, so that it can accumulate experience and improve its performance in the continuous use process [7]. The original medical diagnosis method and financial risk prediction method obviously can not meet the requirements of the development of modern society. The rapid development of machine learning technology provides the basis for the formation of a new generation of intelligent decision support system. Therefore, it has become an important development direction

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.582.2023
to develop an efficient and high accurate disease diagnosis and financial risk intelligent decision model or system by giving full play to the ability of computer reasoning and learning [8]. The machine learning algorithm mainly uses the known data to capture the important characteristics of the unknown and potential probability distribution. Data can be regarded as samples to reveal the relationship between observed variables [9]. Machine learning method, with its powerful ability to automatically learn data, extract complex patterns from them and make intelligent decisions, has attracted more and more attention in biomedical and financial research and other fields, providing a new way to effectively solve decision support problems in these fields [10]. Based on risk prediction, this paper analyzes the application of machine learning method in intelligent decision support system.

2. Intelligent decision support system based on machine learning

Machine learning is one of the most intelligent frontier research fields in artificial intelligence. The prediction method based on machine learning tries to use the characteristics of human thinking to deal with some complex prediction and evaluation problems, and has been successfully used in many fields, such as venture capital prediction, stock market prediction, bidding and evaluation, etc. In recent years, researchers have proposed a DSS structure composed of knowledge system, language system and problem processing system. The problem processing technology of mature expert system (ES) is introduced into the architecture of DSS, which overcomes the general lack of knowledge in traditional DSS. Cash flow is the main indicator of repayment ability. According to the borrower's spot cash flow, and compared with the foreign debt paid in the same period, we can see whether it is enough to repay the due debt [11]. Non financial factors are used to evaluate the direction and degree of their impact on financial indicators such as cash flow. Credit support provides the second source of repayment for the borrower's credit support. The same model using fewer input features can make the diagnosis process more transparent and understandable, and provide a better explanation for doctors' diagnosis, which is particularly necessary in medical diagnosis. In addition, less feature input will make the model more compact, faster, and avoid the risk of over learning.

When the traditional IDSS components provide intelligent support, their learning behavior is mostly static and passive, that is, learning according to the predetermined heuristic strategy, lack of active learning mechanism, unable to formulate dynamic learning strategies according to the actual environmental needs, so the flexibility and adaptability of intelligent assistance are limited. The structure of machine learning system is shown in Figure 1.

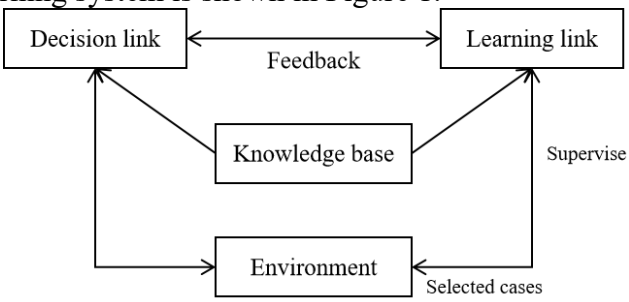

Figure 1 Machine learning system structure

In the decision-making stage, the knowledge in the knowledge base is used to make decisions or take actions to complete various tasks. The feedback information provided by this link to the learning link is very important, because the learning system needs some evaluation method to evaluate the learning effect and improve the system performance accordingly. By answering questions, decision-makers can be assisted to determine the boundary conditions and environment of the problem, thus effectively solving semi-structured or unstructured problems. Because the risk, time, difficulty, cost and inconvenience limit the collection of medical training samples, the medical data usually has the situation of high dimension and relatively few training samples. Studies have pointed out that the number of disease characteristics is increasing year by year, and the development speed is getting faster and faster. The existence of these redundant features and irrelevant features greatly affects the

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.582.2023
whole decision-making process. Therefore, it is necessary to use feature dimension reduction technology to eliminate redundant features and irrelevant features in data before constructing classifiers to make decisions on data samples. Figure 2 shows the structure of intelligent decision support system based on polyphase support vector machine.

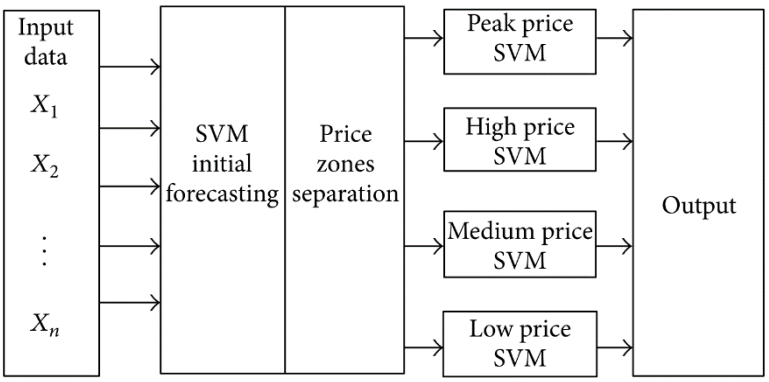

Figure 2 Structure of intelligent decision support system

Usually, the retrieved cases only match the current problems to a certain extent. At this time, the retrieved case is only used as a reference, or a certain correction algorithm is used to correct the retrieved case to match the current problem. It is precisely because of the existence of such adaptive formulas and rules for correcting cases in the system that the system has the characteristics of self-learning. Environment and knowledge base represent external information sources and learning results respectively, and interact with each other by processing knowledge through decision-making and learning links. Therefore, environment, knowledge base, learning link and decision-making link constitute the necessary components of the learning system. Based on the needs of consumers and the relevant experience of previous construction projects, these companies have made cost forecasts and reduced costs as much as possible according to the forecasts [12]. In order to improve the accuracy of case prediction by CBR, every new project is divided into many sub-problems to solve in the experiment. Another embodiment of intelligence in IDSS is the introduction of knowledge reasoning structure, which uses reasoning mechanism to simulate the thinking process of decision makers by expressing their thinking methods and decision-making experience in a certain way.

3. Machine learning and improving the performance of intelligent decision support system

Machine learning is one of the earliest topics in artificial intelligence research, and it is also one of the most intelligent and cutting-edge research topics in artificial intelligence. The great progress of machine learning means that artificial intelligence research has taken a solid step forward. Machine learning has been widely used in natural language understanding, non-monotone reasoning, machine vision, pattern recognition and other fields, especially in knowledge systems and decision science. IDSS still faces semi-structured problems. Its solution strategy is to decompose the semi-structured inter-problem layer by layer through machine or human reasoning with the help of the experience knowledge of management decision experts, until multiple structured sub-problems are obtained and corresponding sub-models are found, and then the solution model of the inter-problem is constructed. Establish a knowledge structure that can represent the past observation experience, and support the establishment and application of domain knowledge. For repetitive decision-making scenarios, establish an example template with standard solution steps, and directly use experience to solve similar decision-making problems. As entities, the basic sub-models stored in the model base should correspond to the basic sub-problems, which are the product of the decomposition of the inter-problem solving model [13]. One criterion of model decomposition is to eliminate model anomalies and standardize the model. Another criterion is to eliminate model redundancy and realize model sharing.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.582.2023

The problem recognition function in the language subsystem and problem handling subsystem of IDSS is a problem of natural language understanding, which means that the computer system extracts its semantics from the natural language request input by the user. Figure 3 is the conceptual model of the mechanism of the influence of enterprise policies on enterprise behavior.

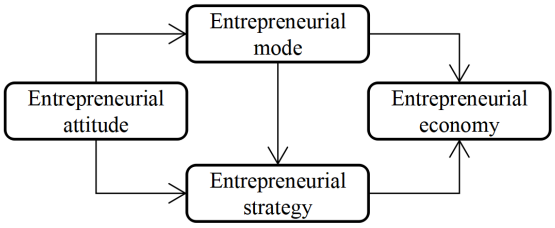

Figure 3 The conceptual model of the influence mechanism of corporate policies on corporate behavior

Parameter optimization and feature selection play an equally important role in SVM modeling, because they influence and interact with each other, which means that these two tasks need to be executed synchronously and in parallel to ensure the optimal generalization performance of SVM. Traditional machine learning methods are based on the assumption that the sample is infinite, but the actual problem is more limited sample learning, while the statistical machine learning method is specialized in studying the machine learning rule under the condition of small sample. It is obviously superior to the traditional machine learning method in some learning performance, so the statistical machine learning method can be applied to the construction of intelligent decision support system. The field of IDSS often involves a large amount of data processing. Overcoming the shortcomings of the inductive algorithms mentioned above is very important to improve the working environment of IDSS, because IDSS needs to strengthen its ability to express and deal with the common uncertainties in most decision support tasks. One of the criticisms of machine learning comes from the fact that it is difficult to properly evaluate the reliability of algorithm quality and the effectiveness of generating solutions, because machine learning algorithms are usually greedy algorithms for large-scale instance space search under the guidance of hill-climbing heuristic. Intelligent decision support system is mainly composed of knowledge base, database, model base and man-machine interaction interface [14].

Through the integration of these four components, decision support for semi-structured and unstructured problems can be realized. For decision makers, whether to mine useful knowledge or patterns from empirical data is an important index to evaluate the performance of an IDSS. Because the object of IDSS is an unstructured decision problem, it is usually difficult to find a standard solution to solve every decision problem. The data mining process in financial analysis is shown in Figure 4.

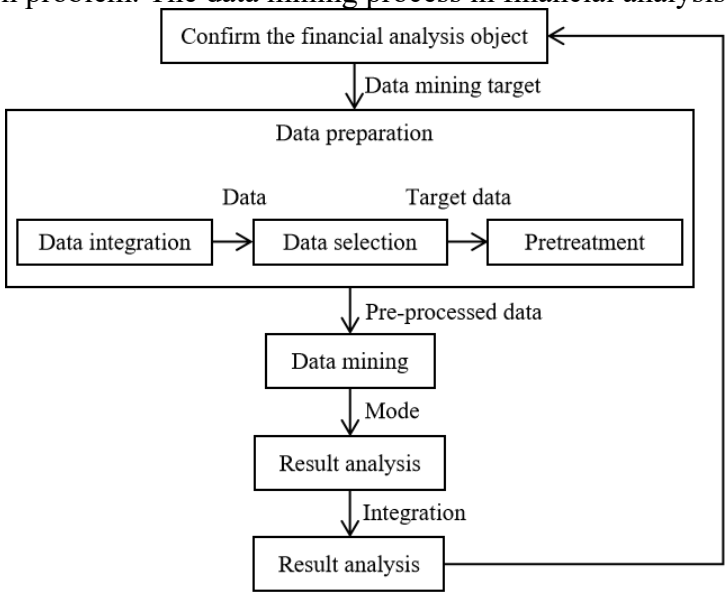

Figure 4 Data mining process in financial analysis and management

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.582.2023

There are S sample sets. Assuming that the number of distinct values of a class label attribute is m, the definition of m being of the same kind is:

$$I(s_1, s_2, \ldots s_m) = -\sum_{i=1}^{m} P_i \log_2 P_i \tag{1}$$

Among them, $P_i = S_i / S$ is the probability that any sample belongs to C_i. Let attribute A have v different values $\{a_1, a_2, \ldots, a_v\}$. Attribute A can be used to divide S into v subsets $\{S_1, S_2, \ldots, S_v\}$, where the samples in S_j have the same value $a_j (j = 1,2,\ldots,v)$ on attribute A. Let s_{ij} be the number of samples of class C_i in subset S_j. For a given subset S_j, the information expectation is:

$$E(S) = -\sum_{i=1}^{m} P_{ij} \log 2(P_{ij}) \tag{2}$$

Among them, $p_{ij} = s_{ij} / s_j$ is the probability of belonging to C_i in the sample S_j. Obtain the information gain on attribute A as $Gain(A) = I(s_1, s_2, \ldots, s_m) - E(A)$.

IDSS must be able to use different methods to solve different decision-making problems according to their characteristics. This heavy solving task can be simplified by accumulating solving experience and learning solving steps through learning methods. When reinforcement learning is used in CBI's IDSS, reinforcement learning can improve CBI's IDSS performance. In CBI's operation, reinforcement learning can be used to quickly search for similar examples that are most suitable for correction. Self-adaptability means that the system has the ability to improve its own performance with the change of environment. In order to improve this ability, IDSS needs to observe the quality of the solution frequently, and constantly revise the content of the knowledge base. After IDSS is developed, verifying the knowledge base and revising the rules is another task of the learning component [15]. This task can be triggered by identifying unsatisfactory solutions, so as to correct the rules with problems. Reinforcement learning regards learning as a tentative evaluation process. Agent chooses an action to act on the environment. After the environment accepts the action, its state changes. At the same time, it generates a reward and punishment signal and feeds it back to Agent to select the next action according to the reward and punishment signal, the environment signal and the current state of the environment. Database and data warehouse are the data base of the intelligent decision support system. Statistical machine learning mining module extracts the data sets to be mined from the database and data warehouse, stores the useful knowledge or patterns obtained from mining in the knowledge base, and returns them to the decision-maker in a suitable form by the man-machine interface.

4. Conclusions

Complex decision-making problems, such as disease diagnosis and financial risk prediction, require high knowledge and ability for decision makers. To a great extent, these decision-making problems mainly depend on the practical experience of the individual decision-makers. For decision-makers, it takes a long time to accumulate experience to reach a rich enough level of decision-making. In terms of scalability and ease of use of the model, compared with the traditional parameter model method, the heterogeneous big data-driven machine learning monitoring and prediction method does not need to conduct complex modeling for different application scenarios according to complex prior knowledge, and its dependence on parameters is not particularly high. Database and data warehouse are the data base of the intelligent decision support system. Statistical machine learning mining module extracts the data sets to be mined from the database and data warehouse, stores the useful knowledge or patterns obtained from mining in the knowledge base, and returns them to the decision-maker in a suitable form by the man-machine interface. At present, the trend of economic globalization is accelerating, and social development is increasingly dependent on the development of information technology. The modern lifestyle characterized by network technology is gradually

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.582.2023

being accepted by people, and the communication mode of young students, especially college students, is changing from the traditional communication mode to the online communication mode.

References

[1] Liu Shengwa, Sun Junming, Gao Xiang, et al. Drilling fluid big data analysis and intelligent decision support platform construction based on machine learning and situation awareness technology[J]. Internet of Things Technology, 2019, 009(005): 46-48+ 51.

[2] Lin Xin, Yuan Renguo, Han Xueyin, et al. Implementation and application of intelligent decision-making for geosteering while drilling[J]. Petroleum Drilling and Production Technology, 2020, 042(001):1-5.

[3] Yao Junxian. Application of Machine Learning in Information Technology Risk Identification[J]. Financial Electronics, 2018, 279(12):87-88.

[4] Yang Lingbo. Application of machine learning in intelligent auxiliary collision avoidance system[J]. Ship Design Communications, 2019, 157(02):83-86.

[5] Li Bing, Lin Wenzhao, Luo Zhengyin. Design and implementation of intelligent agricultural decision-making system based on machine learning[J]. Information and Computer: Theoretical Edition, 2018, 418(24):79-80.

[6] Li Haopeng. Research on intelligent robots based on machine learning methods[J]. Communication World, 2019, 26(04):247-248.

[7] Wu Xinghui, Zhou Yuping, Xing Haihua, et al. Application research of machine learning classification algorithm in diabetes diagnosis[J]. Computer Knowledge and Technology, 2018, 014(035):177-178+195.

[8] Zhang Guangnan, Rao Yuan. Intelligent decision-making algorithm for greenhouse breeding based on integrated learning [J]. Journal of Southwest University of Science and Technology, 2017, 032(004): 78-81.

[9] Zhang Yubin, Wei Zhengying, Zhang Shuai, et al. Research on Intelligent Irrigation Decision Support Technology Based on Knowledge Engineering [J]. Review of China Agricultural Science and Technology, 2017, 019(008): 47-54.

[10] Zhao Xuetong, Yang Yadong, Qu Hongzhu, et al. Application of machine learning methods in clinical decision support in the era of omics[J]. Heredity, 2018, 040(009):693-703.

[11] Song Yineng, Wang Jianan. Research on Intelligent Perception and Decision Control Based on Complex Physical and Chemical Processing Technology[J]. Automatic Instrumentation, 2019, 40(05):52-56.

[12] Wu Yafei, Fang Ya. Progress in the application of machine learning methods in the research of chronic diseases[J]. Chinese Health Statistics, 2020, 37(04):146-150.

[13] Cheng Guojian, Liu Lianhong. A review of machine learning interpretability[J]. Intelligent Computers and Applications, 2020, 10(05):16-18+23.

[14] Yan Xuemei, Fan Haiyan, Lu Zhifeng, et al. Shipping production analysis and decision system based on big data and artificial intelligence [J]. Port Science & Technology, 2019, 156(02): 17-22.

[15] Jiang Na, Yang Haiyan, Huang Jiya, et al. Application research of machine learning in campus security access control system[J]. Modern Information Technology, 2019, 3(01):168-169.