Application of data mining and machine learning in management accounting information system

Xiaofang Zhang[1*]

[1] *Department of Economics, Handan Polytechnic College, Hadan056000, Hebei, China*
[*] *Corresponding author. E-mail: happyzhang2004@126.com*

Received: Feb. 02, 2021; Accepted: Mar. 04, 2021

In the traditional management accounting information processing method, the method used to solve the problem is often fixed due to excessive assumptions. In order to improve its operating efficiency, combined with artificial intelligence information technology, this paper uses data mining algorithms to conduct data acquisition and rule exploration. Moreover, this paper uses statistics, machine learning and other techniques to analyze the correlation between attribute values and transform data into knowledge needed for decision-making. In addition, this paper combines machine learning algorithms to build an intelligent management accounting information system and realizes the close connection between corporate finance and business, which helps to form a closed-loop management between financial analysis, risk management, performance management, and decision-making. Finally, this paper designs experiments to verify the performance of the model. The research results show that the system constructed in this paper satisfies the intelligent demand of accounting information.

Keywords: Data Mining, Machine Learning, Management Accounting, Information System

http://dx.doi.org/10.6180/jase.202110_24(5).0018

1. Introduction

In the era of artificial intelligence, data has exploded. For traditional management accounting, there is a certain degree of difficulty in processing large amounts of data, and the difficulty in obtaining effective data from large amounts of data is also increasing. In order to improve the operating efficiency of the management accounting information system, it is necessary to improve the intelligence of the management accounting information system through artificial intelligence technology, and combine data mining and machine learning to improve the system.

Nowadays, not only the level of social and economic development continues to improve, but with the rapid development of information technology, the management of enterprises is increasingly dependent on the development of information technology. In the management of enterprises, the position of management accounting in enterprise financial management is becoming more and more important. Through the combination of financial data and business data, through the research and analysis of information technology methods, management decisions that are conducive to the development of corporate strategy can be found. These scientific management decisions are of great significance to the improvement of the future economic profit of the enterprise, the expansion of the scale of operation, and the continuous strengthening of comprehensive strength. At the same time, the scale of enterprise development is getting bigger and stronger, and the business is getting more and more complicated. Then, it will rely more on the management needs of management accounting [1].

The application of management accounting information system in our country generally starts from a single field of practice, such as comprehensive budget management, performance management, cost management, etc.

The construction of management accounting information system is positively related to the level of enterprise management. The overall construction system of management accounting information system in my country is not perfect enough. Therefore, the management of enterprise management has not reached the ideal level. Therefore, Chinese enterprises should establish an integrated analysis system for corporate finance and various businesses through the construction of a complete management accounting information system. In this system, the enterprise can realize the information management of various departments and improve the efficiency of management work [2].

The management accounting information system relies on a superior development environment, and is technically supported by the development of modern information technology, and has strong government support and publicity in policy. To respond to the rapid economic development, enterprises need management accounting to provide refined business management. The construction of management accounting informatization requires vigorous development and extensive application. Therefore, it is meaningful to study its application, value creation and policy promotion [3].

2. Related work

The most authoritative and earlier definition of the concept of management control system by foreign scholars may be Anthony's classic framework. It is defined as: "The procedures used by managers to ensure the efficient use of resources for the organization to achieve the intended purpose is the management control system". Moreover, Anthony added strategic theories, psychology, and control systems to the management control system and used it as a management accounting control system. This system decomposes the corporate goals and turns them into the goals of multiple subjects, compares the actual performance and budget performance of each subject to correct the behavior of the business subjects, and enters these data information into the management control system [4].

The literature [5] believed that management accounting seems to be a strategic tool involving loosely organized company supervision, and that it is possible to deploy a defender strategy in a specific competitive and technological environment, which is different from the previous views. The literature [6] believed that the accounting information system is the most important part of the management information system, and obtaining accounting information from the accounting information system is the basis of an efficient management system. The development of accounting information system has laid the foundation for the development of management accounting informatization. The literature [7] used e-commerce accounting information system to maintain the competitive advantage of enterprises, which has become the trend of enterprise development. The literature [8] believed that the responsibility center is the key to the implementation of management accounting, and management accounting is gradually formed and developed to meet the economic needs of operation and management within the organization. Moreover, its main function is to improve the operating efficiency and effectiveness of various established internal accounting control systems, and to prepare and deliver various internal management systems. The literature [9] realized that management accounting practice has a great influence on innovation management. In the specific application of management accounting, the literature [10] realized that an important reason for establishing a management accounting system is to determine the manufacturing cost of a single product or a batch of products. The balanced scorecard is one of the new ways of contemporary management accounting development [11]. It provides a framework to transform business goals and objectives into a series of key performance metrics and goals. By using a balanced scorecard, each company can understand how its performance contributes to its indicators. The literature [12] proposed that revenue realization is a process rather than a point in time, and proposed methods for management accounting revenue forecasting in an e-commerce environment, and the feasibility of increasing revenue recognition time in financial accounting.

3. Neural networks

The neural network model is divided into training and recognition processes. The training process mainly includes collecting training data, designing the network structure, determining the number of layers of the neural network, the number of hidden nodes and the activation function of each layer, and the activation function and loss function of the output layer. That is, the parameters of each layer are learned through training samples, and the prediction results of the current network are constantly compared with the target value we really want, and the weight matrix of each layer is updated according to the difference between the two. Among them, the equation used to measure the difference between the predicted value and the target value is the loss function or the objective function. Generally, we use mean square error (MSE) to evaluate the difference between the predicted value and the target value. The smaller the MSE, the better the prediction effect. The training process is the process of continuously reducing the

MSE [11].

The error back propagation neural network model is a feedforward neural network model with three or more layers. According to the gradient algorithm, the error between the calculated output and the actual output is propagated to correct each connection weight. The BP network learns from training samples with correct results, and randomly selects samples from the training samples. Moreover, the data propagates from the input layer to the output layer through the intermediate layers. After that, the output result is compared with the target value, and each connection weight is corrected layer by layer from the intermediate layer to output layer to reduce the deviation. Finally, the algorithm returns to the input layer. With multiple selection and training of training samples, the model weights are continuously optimized, so that the deviation between the final output result and the target value converges. The calculation steps of BP neural network are as follows:

1. Initialization of the three-layer neural network: We set the number of nodes in the input layer as n, the input as $x_1, x_2, ..., x_n$, the number of hidden layer nodes as l, and the number of nodes in the output layer as m. Moreover, ω_{ij} is the weight from the input layer to the hidden layer, ω_{jk} is the weight from the hidden layer to the output layer, $a_j i$ is the bias from the input layer to the hidden layer, and b_k is the bias from the hidden layer to the output layer. In addition, the learning rate is η and the excitation function is $g(x)$. Among them, the excitation function $g(x)$ is a Sigmoid type function, and its form is [11]:

$$g(x) = \frac{1}{1+e^{-x}},\ x\epsilon(-\infty,+\infty),\ g(x)\epsilon(0,1) \quad (1)$$

2. The output H_j of the hidden layer is:

$$H_j = g\left(\sum_{i=1}^{n}\omega_{ij}x_i + a_j\right),\ j = 1,2,...,l,\ H_j\epsilon(0,1) \quad (2)$$

3. The output O_k of the output layer is

$$O_k = g\left(\sum_{j=1}^{l}H_j\omega_{jk} + b_k\right) \quad (3)$$

4. Error calculation: The error function is:

$$E = \frac{1}{2}\sum_{k=1}^{m}(Y_k - O_k)^2 \quad (4)$$

4. Model building

The business intelligence (BI) process is based on large amounts of data. Moreover, after the processing of ETL technology, the data is stored in DW in an orderly manner. After OLAP analysis, various relationships within the data can be discovered, which is very helpful for managers to acquire knowledge. Sometimes managers need to further understand the hidden information in the data, which can be achieved through DM technology.

DM can provide corresponding data source support through methods such as data crawlers and model libraries. The result of the decision can be stored in DW as a data source to provide a basis for the next decision. In this way, the management activities of the enterprise become intelligent with the support of BI. Fig. 1 shows this process [12].

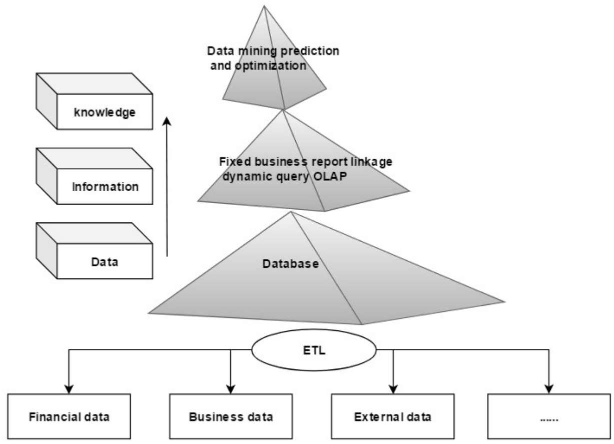

Fig. 1. BI process.

Data Warehouse (DW) is an enterprise's data storage center, which provides a global, subject-oriented analytical environment for the enterprise. DW models mainly include conceptual models, logical models, and physical models. Fig. 2 shows the mapping relationship of the three-tier model of the data warehouse.

1. Conceptual model: The conceptual model is expressed by topics, and the topics are expressed by dimensions and measures. Dimension is the specific angle from which people observe the world, and measurement is data information related to the dimension, such as sales volume. The conceptual model can be represented by an information package diagram, including the multi-dimensional characteristics of the subject, the level of each dimension, and analysis indicators.

2. Logical model: There are generally two logical models, namely star model and snowflake model. The star model consists of a fact table and a dimension table, and the fact table is connected to multiple dimension tables.

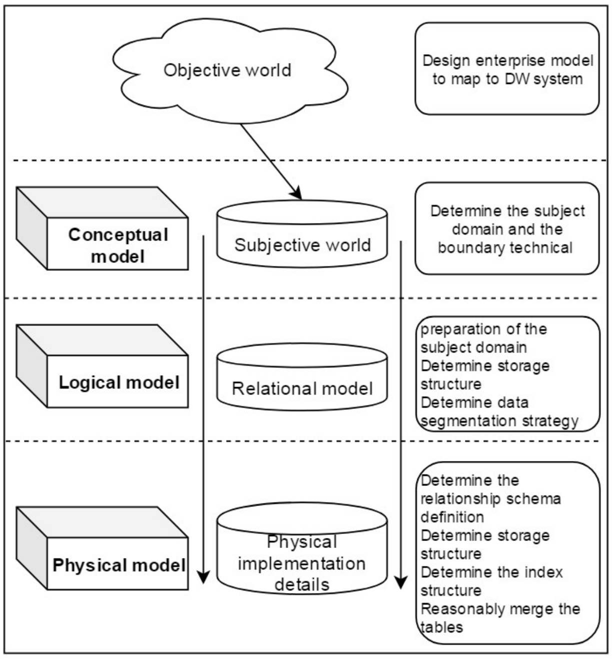

Fig. 2. The mapping relationship of the three-tier data model of the data warehouse and the design details of each layer.

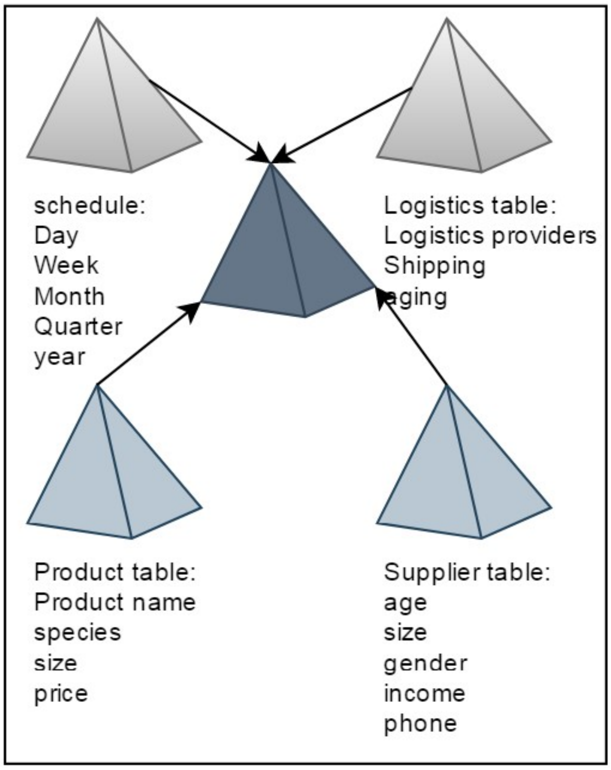

Fig. 3. Star model.

Fig. 3 is a star model of a company's procurement. The procurement table is a fact table, while the date table, product table, supplier table, and logistics table are all dimension tables.

The snowflake model is an extension of the star model. The data in some dimension tables can be subdivided into additional tables to reduce redundancy. The physical model is the realization of the logical model in the data warehouse, such as data storage structure, data indexing strategy, data storage strategy and storage optimization allocation.

DW does not simply pile data together, but passes the data through a certain process (Fig. 4). (1) Data extraction. The data in DW is subject-oriented, so it is necessary to identify the data related to decision support before importing the source data into the data warehouse.

The extraction is divided into the following five steps. A) The data of the identified data source and its meaning. B) Extraction. Which files or tables in the source data are accessed are determined. C) The extraction frequency is determined. D) Output. The destination of the data and the output format. E) Exception handling, that is, how to deal with data when the required data cannot be extracted. (2) Data cleaning. The noise data and the silent data are filtered. (3) Data conversion. The source data becomes the required target data, which mainly includes format conversion, summary calculation, etc. (4) Data loading.

The cleaned and converted data is loaded into the DW, which can be loaded through strategies such as direct append and full coverage.

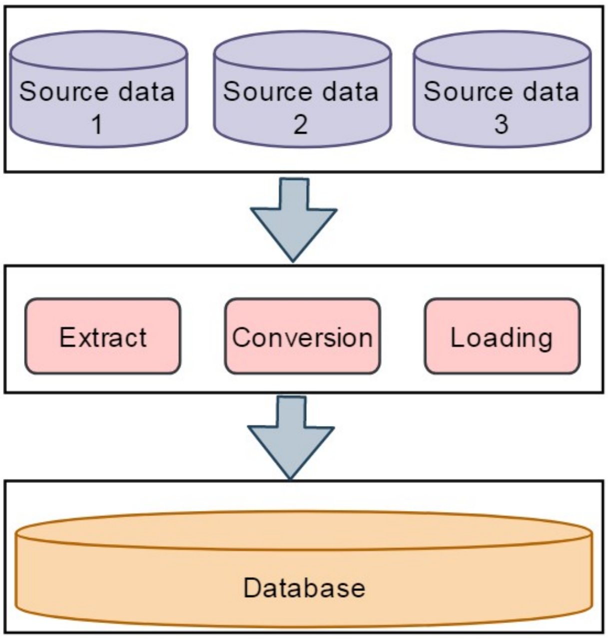

Fig. 4. ETL process.

The main operating structure in OLAP is based on a concept called a cube. The cube in OLAP is a multi-dimensional data structure (real or virtual) that supports rapid data analysis, and it can perform multi-dimensional and efficient operation and data analysis. The data structure in the cube is designed to overcome the limitations of relational databases. Fig. 5 shows a cube structure.

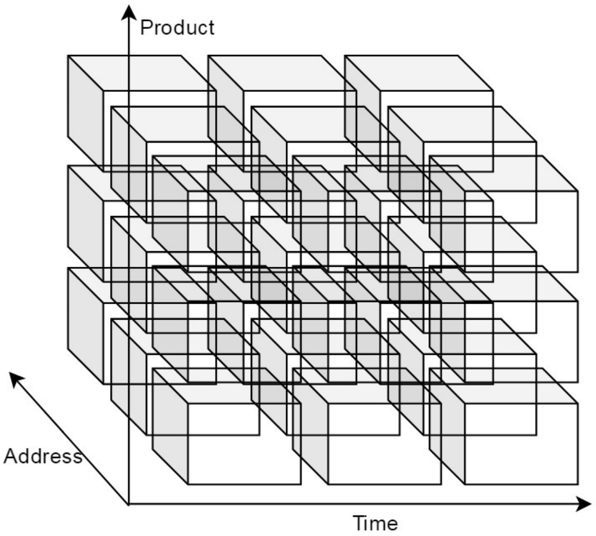

Fig. 5. Cube structure.

Data Mining (DM) refers to the use of automated or semi-automated tools to analyze data to discover hidden patterns in the data. Unlike the OLAP mentioned in the previous section, DM does not test the correctness of a certain hypothesis, but discovers unknown patterns in the data. Therefore, its essence is a process of inductive learning. The DM process is a complete process. Fig. 6 shows the data mining process.

In DW, OLAP can meet the data analysis needs of managers. However, some patterns exist in large amounts of data. At this time, OLAP cannot estimate it. However, forecasting is playing an increasingly important role in management accounting, and managers often need more forecasting work to support them in making decisions in an increasingly changing environment. At this time, DM is particularly important in the management accounting information system. Data mining is a combination of statistics, databases, machine learning and other technologies. It analyzes the correlation between attribute values to find hidden patterns in data sets, transforms data into information to help decision-making, and helps companies gain advantages in competition.

This article believes that an intelligent modern management accounting information system should include four

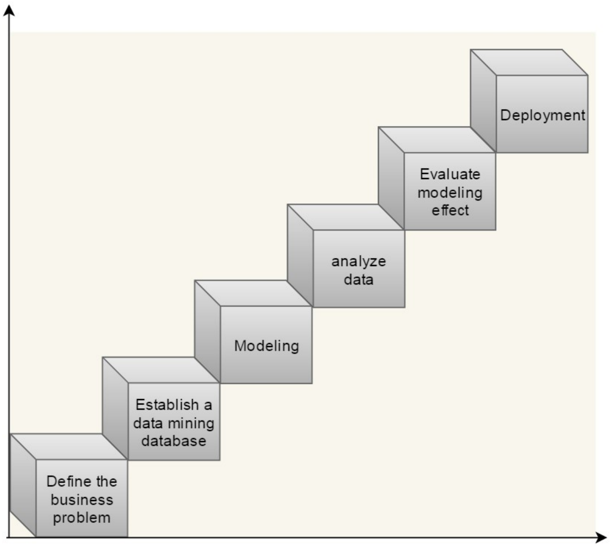

Fig. 6. Data mining process.

major subsystems (Fig. 7). The four subsystems are;

1. Accounting Analysis Management System
2. Risk Management Information System
3. Performance Management Information System
4. Accounting Decision Support System.

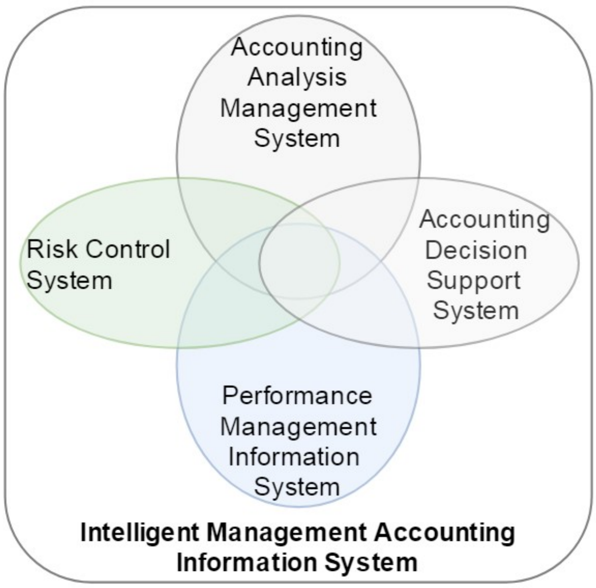

Fig. 7. Intelligent management accounting information system.

In response to the previous narrative, this paper attempts to build an intelligent management accounting information system as shown in Fig. 8. The system contains 5 major modules, which are the support platform including functions such as data interface and system services, the accounting analysis management system covering modules such as accounting report management, the closed-loop risk management information system for risk, the performance management information system for performance evaluation, and the accounting decision support system for decision makers. Moreover, each system is independent of each other, fulfills its own responsibilities, and is connected with each other to provide beneficial support for the management of the enterprise.

5. Model performance test

After constructing the above model, this paper verifies the performance of the model, and performs data processing analysis based on the actual needs of management accounting. First, this paper scores the effect of model data mining, and evaluates the performance of model data mining through expert scoring methods.

In this paper, 75 sets of data are used for experiments. The results are shown in Table 1 and Fig. 8.

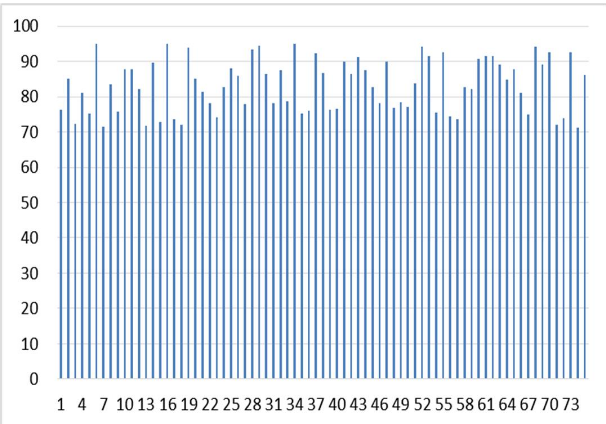

Fig. 8. Statistical diagram of scores of data mining of management accounting information system.

Through the above analysis, we can see that the management accounting information system constructed in this paper has a good data mining effect and can obtain effective information from massive data. Next, this paper evaluates the accuracy of data processing in the management accounting information system. The results are shown in Table 2 and Fig. 9.

It can be seen from the above chart that the accuracy rate of the management accounting information system in this paper is as high as 100% in data processing, which meets the rigorous requirements of accounting for data processing, so it can be applied in practice.

Table 1. Statistical table of scores of data mining of management accounting information system.

NO.	Scores	NO.	Scores	NO.	Scores
1	76	26	86	51	84
2	85	27	78	52	94
3	72	28	93	53	91
4	81	29	94	54	75
5	75	30	86	55	92
6	95	31	78	56	74
7	71	32	87	57	74
8	84	33	79	58	83
9	76	34	95	59	82
10	88	35	75	60	91
11	88	36	76	61	92
12	82	37	92	62	92
13	72	38	87	63	89
14	90	39	76	64	85
15	73	40	77	65	88
16	95	41	90	66	81
17	74	42	86	67	75
18	72	43	91	68	94
19	94	44	88	69	89
20	85	45	83	70	93
21	81	46	78	71	72
22	78	47	90	72	74
23	74	48	77	73	92
24	83	49	79	74	71
25	88	50	77	75	86

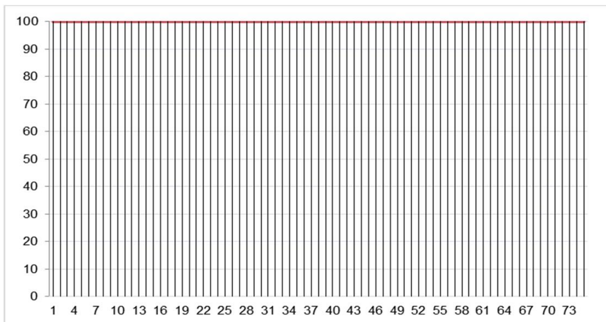

Fig. 9. Statistical diagram of the accuracy rate of data processing of management accounting information system.

6. Conclusion

This article uses data mining technology and machine learning technology to construct a management accounting in-

Table 2. Statistical table of the accuracy rate of data processing of management accounting information system.

NO.	Scores	NO.	Scores	NO.	Scores
1	100	26	100	51	100
2	100	27	100	52	100
3	100	28	100	53	100
4	100	29	100	54	100
5	100	30	100	55	100
6	100	31	100	56	100
7	100	32	100	57	100
8	100	33	100	58	100
9	100	34	100	59	100
10	100	35	100	60	100
11	100	36	100	61	100
12	100	37	100	62	100
13	100	38	100	63	100
14	100	39	100	64	100
15	100	40	100	65	100
16	100	41	100	66	100
17	100	42	100	67	100
18	100	43	100	68	100
19	100	44	100	69	100
20	100	45	100	70	100
21	100	46	100	71	100
22	100	47	100	72	100
23	100	48	100	73	100
24	100	49	100	74	100
25	100	50	100	75	100

formation system, which is the inevitable direction of the development of management accounting information systems in the era of artificial intelligence. Moreover, this paper believes that management accounting informatization is an inevitable trend of accounting development and that a comprehensive management accounting information system should include subsystems such as accounting analysis management system, risk management information system, performance management information system, and accounting decision support system. The sub-systems are independent and connected to each other. The accounting analysis management system realizes the basic functions of accounting data processing and analysis, the risk management information system can prevent the possible risks of the enterprise in all aspects, the performance management information system lays a foundation for the realization of corporate strategic goals through effective target decomposition and layer-by-layer implementation, and the accounting decision support system can effectively improve the efficiency and effect of decision-making. In general, the synergy of the four major subsystems can effectively enhance the business management capabilities of the enterprise.

References

[1] Jie Cai, Jiawei Luo, Shulin Wang, and Sheng Yang. Feature selection in machine learning: A new perspective. *Neurocomputing*, 300:70–79, 2018.

[2] J. N. Goetz, A Brenning, H Petschko, and P. Leopold. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. *Computers and Geosciences*, 81:1–11, 2015.

[3] Hamid Darabi, Bahram Choubin, Omid Rahmati, Ali Torabi Haghighi, Biswajeet Pradhan, and Bjørn Kløve. Urban flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques. *Journal of Hydrology*, 569:142–154, 2019.

[4] Alvin Rajkomar, Jeffrey Dean, and Isaac Kohane. Machine learning in medicine. *England Journal of Medicine*, 380(14):1347–58, apr 2019.

[5] Yang Xin, Lingshuang Kong, Zhi Liu, Yuling Chen, Yanmiao Li, Hongliang Zhu, Mingcheng Gao, Haixia Hou, and Chunhua Wang. Machine Learning and Deep Learning Methods for Cybersecurity. *IEEE Access*, 6:35365–35381, 2018.

[6] Logan Ward, Ankit Agrawal, Alok Choudhary, and Christopher Wolverton. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Computational Materials*, 2, 2016.

[7] Puyu Feng, Bin Wang, De Li Liu, Cathy Waters, and Qiang Yu. Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia. *Agricultural and Forest Meteorology*, 275:100–113, 2019.

[8] Konstantina Kourou, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis, and Dimitrios I. Fotiadis. Machine learning applications in cancer prognosis and prediction, 2015.

[9] Saleema Amershi, Maya Cakmak, W. Bradley Knox, and Todd Kulesza. Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4):105–120, 2014.

[10] V Rodriguez-Galiano, M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas. Machine learning predic-

tive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, 71:804–818, 2015.

[11] Connor W Coley, Regina Barzilay, Tommi S Jaakkola, William H Green, and Klavs F Jensen. Prediction of Organic Reaction Outcomes Using Machine Learning. *ACS Central Science*, 3(5):434–443, may 2017.

[12] Aritra Chowdhury, Elizabeth Kautz, Bülent Yener, and Daniel Lewis. Image driven machine learning methods for microstructure recognition. *Computational Materials Science*, 123:176–187, 2016.