# Applying BSC-Based Resource Allocation Strategy to IDSS Intelligent Prediction

Yun Zhang[1*], Yang Chen[2]

[1] School of Computer Science and Technology, Xi'an University of Science & Technology, Xi'an 710054, Shaanxi, P.R.China.
[2] Thought Works Software Technologies (Xi'an) Ltd, Room 06, 6th floor, Building E Huanpu Science and Technology Industrial Park No.211 8th Tiangu Road, Hi-tech Development District Xi'an, 710077, Shanxi P.R.China.

* Corresponding author. Tel.: 008613571980258; email: yunzhang710@163.com

**Abstract:** In this paper, a new data mining model is designed and constructed, the model is based on Balanced Score Card(BSC) and used for intelligent decision support system of data mining and prediction. On the one hand, based on data mining framework, the new data mining model come up with three different resource allocation mechanisms, emphasizing the understanding of the business process, focusing on data accuracy and covering the quality of data mining modeling. The BSC resource allocation mode and algorithm are designed with new insight perspective.On the other hand, the typical established case shows how the mining of resource allocation to improve the accuracy of intelligent prediction. Data mining based on BSC not only to provide an integration platform to support different mining components, but also be able to combine limited resources into reasonable mining process via using the resource evaluation analysis and resource allocation with purpose of improving the accuracy of the prediction and intelligent.

**Key words:** Intelligent decision support system, data mining, balanced scorecard, prediction, resource allocation.

## 1. Introduction

Intelligent decision support system (IDSS) is a combination of artificial intelligence(AI) and decision support system, which use expert system technique to help to solve the complex decision problems of aided decision system through the method of logical reasoning. The traditional intelligent decision is usually based on data mining technology, which can realize intelligent data mining, forecasting and decision making.

Data mining, also called knowledge discovery in database, is a complex process, which is extracted from a large number of data before mining, but has a lot of potential information[1]. Traditional data mining is the technology to analyze each data, to find the law from a large number of data . It contains 3 steps: data preparation, rule finding and rule representation [2].

The traditional issues and disadvnatages of Data Ming process are Non-standardized process of limited resource rational allocation. Limited resource are not reansonable allocated. Assumption in one project, because of no quantitative resource assessment, so most of the resources used in the Evaluation and validation（but it should not need too much）, and only a small number of resource used for data

pre-processing. This would cause significant reduction in the efficiency of data mining.

In various fields, the application, planning, processing of Data mining is used for different scnarios cause diverses. The reasons and root causes of this diverses are differences of technical analysis and professional knowledge involement. So the systematization, standardization and rationalization of resource allocation of the data mining process are particularly important.

The conception and theory of balanced score card in the resources allocation is first proposed, to combine BSC into Data Mining process. This new method effectively solves problems and standardize the process of data mining. The big contribution of this modification is reasonable allocation of limited resource in the process of Data Mining. This modification will be improving accuracy of intelligent prediction as well [3].

## 2. An Overview of Resource Allocation Strategies Based on BSC

A structured Balanced Scorecard [4], which should contain a series of links to the goals and measures, these measures and goals not only need consistent, but also strengthening each other. Fig. 1 shows the overall architecture of the traditional balanced scorecard.
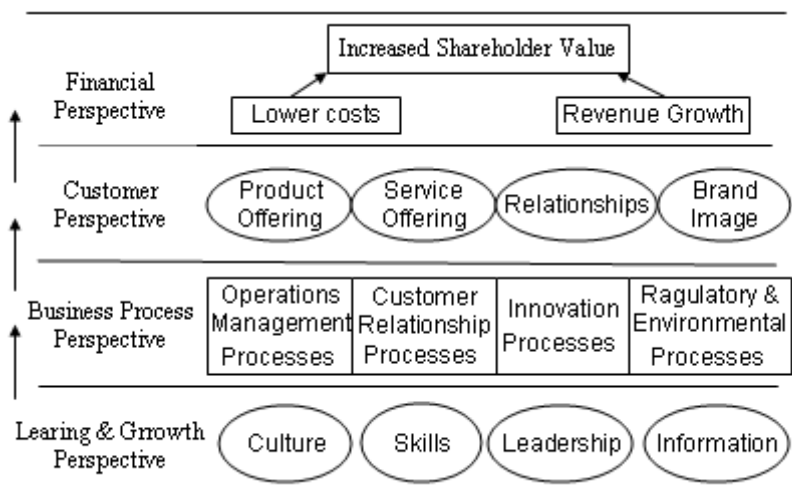


Fig. 1. Traditional balanced score card structure.

To establish a balanced scorecard need to observe the three principles [4]: 1 causality; 2 outcome measures; 3 value links.

Based on these three principles, Balance Scorecard interconnect with project strategically. The related links are instead of current process of decision, will make positive influence on achievement in the future. The purpose of these metrics is to express the new work-flow specification to the organization, and establish the logic process of the strategic priorities and strategic outcomes, to carry out the optimization of project process and the allocation of organizational resources . This will be more efficient to make strategic decisions.

This strategy applies the concept of Balanced Scorecard to the process of data mining. Among them, the data mining process is a "causal relationship", and the resource allocation of process is the "value" of the balanced score card . Reasonable allocation of limited resources to every stage of the data mining process is the key to improve the quality of data mining, but also the innovation of this paper.

## 3. Methodology and Framework of BSC of Data Mining Combination

BSC covers phases from strategy view, tactical design, and business operational output to execution[5].

Accordingly, based on the task–centric, a comprehensive methodology about how CDM-BSC provides whole life cycle support BSC transformation is developed and shown in Fig. 2.
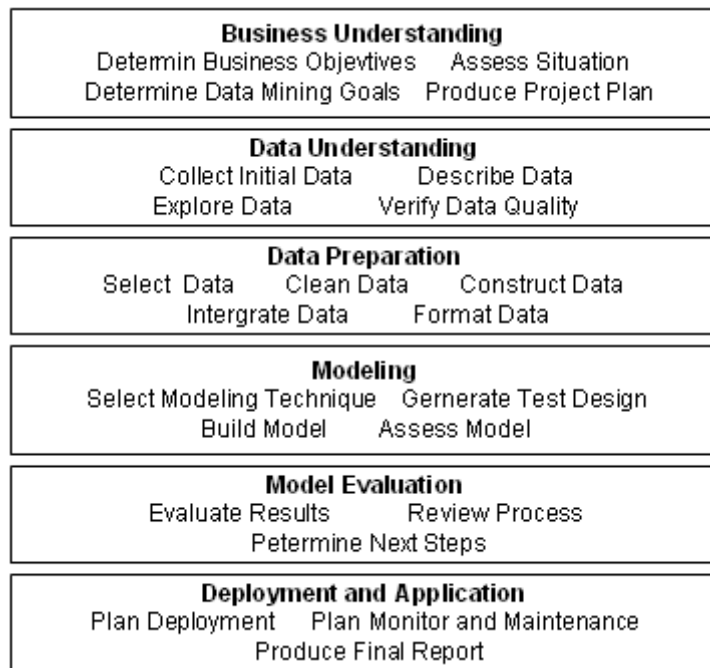


Fig. 2. BSC processing phases.

Fig. 2 illustrates how BSC facilities the tasks-centric BSC transformations phase by phase and end to end.

## 4. Allocation Weight for BSC

In order to realize different Tactical Business Values and leverage objectives of business values to enterprise limited resource, making levels of division of priority is perceived and intuitive. Fig. 3 figure out the weight of influence for each phase.
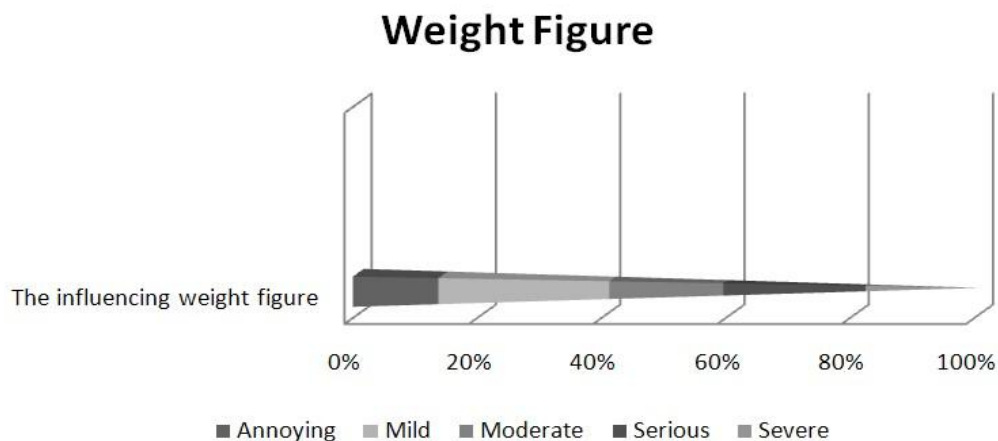


Fig. 3. BSC AWS weight diagram.

AWS 1: Understanding the project objectives and requirements from a business perspective, not only customers but also business analysis should pay much more time on digging into business understanding process. (Emphasis on Business understanding)

AWS 2: The accuracy and preparation of data is important for whole process without considering which kinds of data mining components would be used. The data operations phases cover all activities needed to

construct the final dataset. (Emphasis on Data Perspective)

AWS 3: At this level mining component is essential for the whole data mining process. Which kinds of mining components will be chosen will impact on predictive result and accuracy of data mining process (e.g. CHAID Algorithms: In practice, CHAID is often used in the context of direct marketing to select groups of consumers and predict how their responses to some variables affect other variables, although other early applications were in the field of medical and psychiatric research). (Emphasis on Modeling Stuff)

## 5. BSC Resource Allocation Task Model and Algorithm

There is a group of tasks T and a group of resource allocation processor S. Assuming that each task Ti is respectively independent and resource allocation processor can execute task simultaneously. Classify all the resource allocation processor and they are classified as three types, which are Saws1, Saws3, Saws3 (Si, $S_i = \{S_1, S_2, S_3\}$), according to three weighting scheme. The main problem need to be solved here is: judge which type of processor is suit for one task, which means to judge the task is executed by which processor have the highest income [3].

### 5.1. Model Design and Definition

Firstly, notate the set of resource allocation processors as $S_i$, the set of idle processors as $S_i'$ and the set of independent tasks as $T$, the set of unfinished tasks as $T'$, $T' \in T$, the set of resources as R (Assuming the size and form of resources have no difference). Definitions of some functions used for model definition are given as following [6]:

$R(t)$ notates the set of resources needed for executing task $t$, where $R(t) \subset R, t \in T$;

$R(s_i)$ notates the set of resources occupied by processor $S$, where $R(s_i) \subset R, s_i \in S_i$;

$R(r, s_i)$ notates whether processor $s_i$ occupied resource $r$, if resource $r \in R(s_i)$, $R(r, s_i) = 1$; otherwise $R(r, s_i) = 0$.

$C_{t,s}$ notates the total number of resources needed to be borrowed from other processors when processor $s_i$ executed task $t$.

For example, if a processor $s_i$ has already had resource $a, b$ and task $t$ need resource $a, b, c, d$, then $C_{t,s}$ is 2 because required resources are $c, d$.

Assuming that all the processors require the same price to acquire all the resources and only consider the acquire price of resource, not including the time price of computation.

Formula is expressed as follows:

$$C_{t,s_i} = \sum_{r \in R(t)} (1 - R(r, s_i)) \tag{1}$$

$C^{\min}(t, S_i')$ notates the minimum required price for executing task $t$ in the set of existing idle processors. Then $C^{\min}(t, S_i')$ is expressed as:

$$C^{\min}(t, S_i') = \min_{s_i \in S_i'}(C_{t,s_i}) \tag{2}$$

$C(T', S_i')$ notates an estimated value of minimum required price for the set of unfinished tasks $S_i'$ based on the set of existing idle processors $T'$.

## 5.2.  BSC task Allocation Algorithm

If there are set of idle processors $S_i'$ and set of unfinished tasks $T'$, task allocation algorithm based on BSC resource allocation is as follows:

First Step:  $C(T', S_i') = 0$, if $S_i' = \phi$, then $S_i' = S_i$; if $T' = \phi$, then stop；

Second Step:

$$C(T', S_i') = \min_{t \in T'} C^{\min}(t, S_i') + C(T', S_i') , \quad T' = T' - t' , \quad S_i' = S_i' - s_i' ,$$

where  $C_{t', s_i'} = \min_{t \in T'} C^{\min}(t, S_i')$ ;

Third Step: compute formula(3)

$$G_{T', S_i'}(t, s_i) = C(T', S_i') - C((T' - t), (S_i' - s_i)) - C_{t, s_i} \tag{3}$$

Fourth Step: if  $S_i' = \phi$, then $S_i' = S_i$; if  $T' = \phi$, then stop. Otherwise return to step 2.

Finally compared  $G_{T', S_i'}(t, s_i)$  under three weighting allocation scheme, the largest can used as the optimal allocation method to aid decisions.

| | Age | Sex | BP | Cholesterol | Na | K |
|---|---|---|---|---|---|---|
| 1 | 23 | F | HIGH | HIGH | 0.793 | 0.031 |
| 2 | 47 | M | LOW | HIGH | 0.739 | 0.056 |
| 3 | 47 | M | LOW | HIGH | 0.697 | 0.069 |
| 4 | 28 | F | NORMAL | HIGH | 0.564 | 0.072 |
| 5 | 61 | F | LOW | HIGH | 0.559 | 0.031 |
| 6 | 22 | F | NORMAL | HIGH | 0.677 | 0.079 |
| 7 | 49 | F | NORMAL | HIGH | 0.790 | 0.049 |
| 8 | 41 | M | LOW | HIGH | 0.767 | 0.069 |
| 9 | 60 | M | NORMAL | HIGH | 0.777 | 0.051 |
| 10 | 43 | M | LOW | NORMAL | 0.526 | 0.027 |
| 11 | 47 | F | LOW | HIGH | 0.896 | 0.076 |
| 12 | 34 | F | HIGH | NORMAL | 0.668 | 0.035 |
| 13 | 43 | M | LOW | HIGH | 0.627 | 0.041 |
| 14 | 74 | F | LOW | HIGH | 0.793 | 0.038 |
| 15 | 50 | F | NORMAL | HIGH | 0.828 | 0.065 |
| 16 | 16 | F | HIGH | NORMAL | 0.834 | 0.054 |
| 17 | 69 | M | LOW | NORMAL | 0.849 | 0.074 |
| 18 | 43 | M | HIGH | HIGH | 0.656 | 0.047 |
| 19 | 23 | M | LOW | HIGH | 0.559 | 0.077 |
| 20 | 32 | F | HIGH | NORMAL | 0.643 | 0.025 |
| 21 | 57 | M | LOW | NORMAL | 0.537 | 0.028 |
| 22 | 63 | M | NORMAL | HIGH | 0.616 | 0.024 |
| 23 | 47 | M | LOW | NORMAL | 0.809 | 0.026 |
| 24 | 48 | F | LOW | HIGH | 0.874 | 0.058 |
| 25 | 33 | F | LOW | HIGH | 0.858 | 0.026 |
| 26 | 28 | F | HIGH | NORMAL | 0.557 | 0.030 |
| 27 | 31 | M | HIGH | HIGH | 0.741 | 0.024 |
| 28 | 49 | F | NORMAL | NORMAL | 0.695 | 0.074 |
| 29 | 39 | F | LOW | NORMAL | 0.649 | 0.029 |
| 30 | 45 | M | LOW | HIGH | 0.754 | 0.042 |
| 31 | 18 | F | NORMAL | NORMAL | 0.554 | 0.063 |
| 32 | 74 | M | HIGH | HIGH | 0.715 | 0.075 |
| 33 | 49 | M | LOW | NORMAL | 0.626 | 0.057 |
| 34 | 65 | F | HIGH | NORMAL | 0.829 | 0.026 |
| 35 | 53 | M | NORMAL | HIGH | 0.645 | 0.046 |
| 36 | 46 | M | NORMAL | NORMAL | 0.526 | 0.072 |
| 37 | 32 | M | HIGH | NORMAL | 0.530 | 0.056 |
| 38 | 39 | M | LOW | NORMAL | 0.605 | 0.043 |
| 39 | 39 | F | NORMAL | NORMAL | 0.518 | 0.053 |
| 40 | 15 | M | NORMAL | HIGH | 0.642 | 0.071 |

Fig. 4. Part of the human body function index data of the patients (system screenshot).

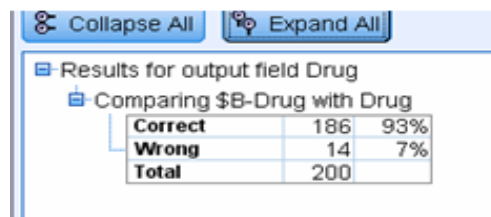## 6.  Validation of Effectiveness

In this paper, we use a real medical case to verify the accuracy of the BSC resource allocation strategy. In medicine, we can choose the drug through the analysis of all aspects of the patient's body function indicators. In other words, in drug selection, you can predict the need to use drugs, through the analysis of the patient's age, gender, blood pressure, fat content, the body's sodium content and potassium content and other indicators of human body function. In this paper, we choose 200 real patients for the intelligent prediction of the required drugs, due to the limited length of the article, figure 4 shows some of the

patient's functional indicators.

In this medical case, we need to dig through massive empirical data in order to make decisions in patients with the final use of drugs. In order to verify the BSC resource allocation strategy, we will use the traditional IDSS prediction and BSC resource allocation strategy based on the prediction to predict and make the final decision of the patient's real data.

## 6.1. Using the Traditional IDSS Prediction Method

By using the traditional IDSS to predict, the algorithm of system modeling is selected by industry experts. In this case, industry experts chose SLIQ algorithm (based on experience) for modeling, after the IDSS data pre processing, data mining, inference and evaluation. Accurate prediction success rate was 93% (prediction results as shown in Fig. 5).



Fig. 5. Accuracy of the traditional IDSS prediction results.

## 6.2. Using the Prediction Method of BSC Resource Allocation

In this example, the promotion of data mining is through the BSC data mining overall process, including business understanding, data understanding, data preparation, as well as modeling, evaluation and deployment. The resource allocation of BSC data mining strategy is based on AWS1, AWS2 and AWS3 three kinds of different resource allocation weight scheme and model task algorithm. The distribution is shown in the following table (as shown in Table 1).

Table 1. Resource Allocation Ratio Under Three Resource Weights

|  | AWS1 | AWS 2 | AWS 3 |
|---|---|---|---|
| Buisness Understanding | 45 | 20 | 20 |
| Data Understanding | 35 | 50 | 18 |
| Data Preparation | 20 | 50 | 12 |
| Modeling | 15 | 25 | 45 |
| Modeling Evalutation | 12 | 15 | 45 |
| Deployment and Applicatior | 18 | 10 | 10 |

According to this ratio, we analyze and describe the results by use of radar structure diagram [5], as shown in Fig. 6. You can clearly see that the allocation of resources in the distribution of the three modes of weight. Among them, BU represents Buisness Understanding; DU represents Data Understanding; DP represents Data Preparation; ME represents Modeling Evalutation; D&A represents Deployment and Applicatior.

Using three kinds of AWS resource allocation weights to carry out the effective division of resources, so as to predict, it can be concluded that:

1) To predict the distribution by using AWS1, the accurate success rate is 88%;
2) Using AWS2 to predict, the accuracy of the success rate is 94%;
3) Aws3 prediction, we put most of the limited resources allocated in the modeling and evaluation, In the system, the selection of the model algorithm, not by industry experts to choose, but to traverse all the algorithms. For the modeling with each algorithm , as well as the evaluation of all models, and finally select an optimal algorithm to carry out the final prediction. Using AWS3 to predict, the accurate success rate of up to 98%.
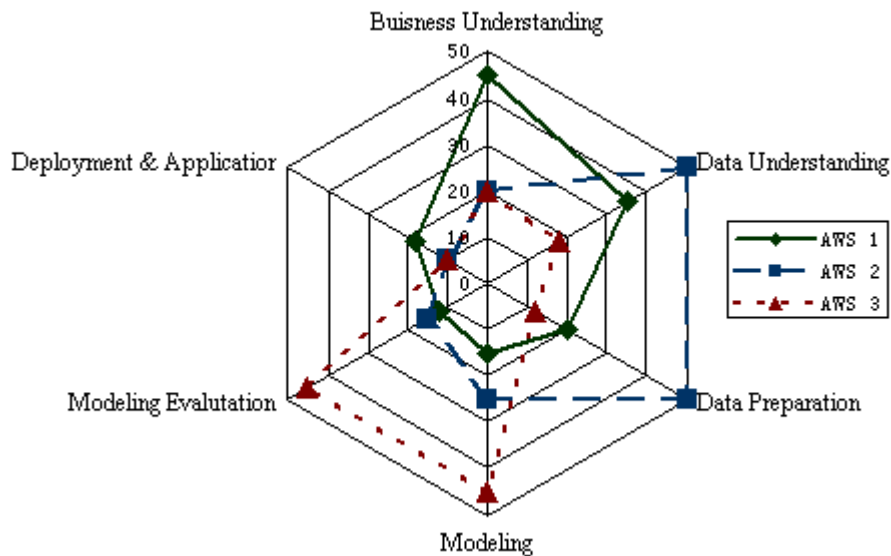


Fig. 6. Radar map of BSC resource allocation weight.

From the aspects of accuracy we can see, AWS3 far higher than the first two. This fully shows that the BSC resource allocation strategy, through the resource value analysis and reasonable allocation, to predict the final accuracy has a very important significance. That is, in this case, we must focus on the allocation of limited resources in the process of modeling and model assessment, the results of the prediction accuracy is the highest.

The experiment shows that the data mining method based on BSC resource allocation is really effective to improve the accuracy of data mining prediction.

## 7.  Conclusion

The process of data mining based on BSC can make the data mining process more intuitive, it will be a good idea for the balanced scorecard. In complex data environment, accurate and high quality data prediction and analysis is more important than ever. Therefore, the theory of resource allocation is introduced in the process of traditional IDSS data mining, which can be a high quality forecasting service.

At the beginning of the study, the aim is to improve the structure and accuracy of data mining. However, the feedback and massive experience data analysis as well as accumulation from the environmental information are more important for the intelligent decision-making system. Therefore, we need to do more in-depth research in this area, and this method is connected with the actual complex data environment, to fully improve the practicality of the strategy.

Acknowledgment

## References

[1] Domingos P., & Hulten, G. (2000). Mining high-speed data streams. *Proceedings of ACM SIGKDD* (pp. 71–80).

[2] Hulten, G., Spencer, L., & Domingos, P. (2001). Mining time changing data streams. *Proceedings of ACM KDD* (pp. 97–106).

[3] Zhang, Y., Li, W. H., & Yu. H. (2014). Applying balanced score card resource allocation strategy to intelligence decision support system. *Journal of Northwestern Polytechnical University*, *32(2)*, 328-333.

[4] Voelper, S., Leibold, M., Eckhoff R., & Davenport, T. (2006). The tyranny of the balanced scorecard in the innovation economy. *Journal of Intellectual Capital*, *7(1)*, 43–46.

[5] Norreklit, H. (2000). The balance on the balanced scorecard — A critical analysis of some of its assumption. *Management Accounting Research*, 65-88.

**Zhang Yun** was born in Xi'an, Shaanxi, China, in 1983. She graduated from School of Computer Science and Technology, Northwestern Polytechnical University(Xi'an Shaanxi, China), She obtained a PhD in computer science and technology in 2014, and the research field is intelligent decision support system .

She worked as a teacher at Xi'an University of Science and Technology after graduation.

She is with School of Computer Science and Technology, Xi'an University of Science & Technology, Xi'an 710054, Shaanxi P.R.China, and the research field is intelligent decision support system.