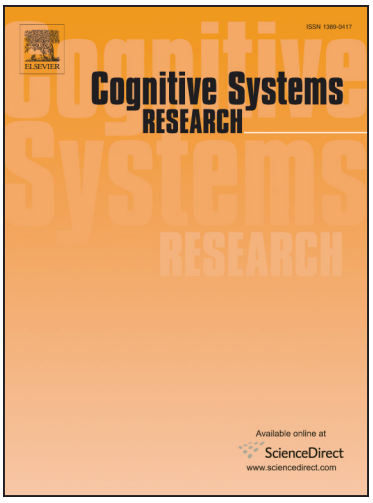

Government Accounting Optimization Based on Computational Linguistics

Jiyou Li

PII: S1389-0417(18)30465-0
DOI: https://doi.org/10.1016/j.cogsys.2018.10.017
Reference: COGSYS 758

To appear in: *Cognitive Systems Research*

Received Date: 9 August 2018
Revised Date: 3 October 2018
Accepted Date: 6 October 2018

Please cite this article as: Li, J., Government Accounting Optimization Based on Computational Linguistics, *Cognitive Systems Research* (2018), doi: https://doi.org/10.1016/j.cogsys.2018.10.017

Government Accounting Optimization Based on Computational Linguistics

Jiyou Li[1*]

[1.] Glorious Sun School of Business and Management, Donghua University, Shanghai, Post Code 200051, China

Abstract: The level of moral development and moral intensity in cognitive psychology will not only affect the ethical behavior of accountants, but also have a direct impact on the quality and level of accounting work. Therefore, in this paper, the ethical behavior of accountants was analyzed from the perspective of cognitive psychology. Computer-aided data mining techniques were introduced, and government accounting risk assessment management of financial accountants was studied. In this paper, the principle of cognitive psychology to measure the ethical level of accountants was first described. The predicament of moral judgments was analyzed and an optimization plan to improve the ethical intention of accountants was proposed. Support Vector Machine classification technology in data mining was studied to explore how to conduct effective and reliable evaluation, so as to provide a scientific basis for decision-making in improving accounting management. After the simulation experiment, it is proved that continuously improving the ethical standards of accountants and strengthening the forecast of accounting risks can continue to optimize the accounting office management.

Keywords: computational linguistics; accounting optimization; research

1. Introduction

Cognitive psychology is a subject that understands people's behavior and scientifically analyzes people's mental and intellectual processes, memory structures and so on (Skavronskaya L, et al. 2017) [1]. The level of moral development and moral intensity in cognitive psychology will not only affect the ethical behavior of accountants, but also have a direct impact on the quality and level of accounting work. Ethical behavior of accountants refers to the moral and unethical behavior of accountants through professional judgment in accounting work (Johnson M K, et al. 2016) [2]. This study originated in the psychological study of the 90s of last century. The main theoretical basis is the macro study of the cognizance judgment and decision-making in psychology, which is a medium and micro study of the business managers, accountants, analysts and other professional people in the capital market (Pappas T N, et al. 2015) [3]. Many scholars studied the influence of moral development and moral intensity of cognitive psychology on the ethical behavior of accountants. For example, scholars used the DIT to study accountants' influence factors on occupational environment when they made moral decisions, such as supervisor's pressure and supervision regulations (Quinlan P, et al. 2016) [4]. Scenario experiments are often used in the study of individual moral issues. However, when moral dilemmas occur, accountants may consciously control others' perceptions of themselves and inevitably show a state of self-moral modification. People are only willing to show and display external appearances widely recognized by the society. If people have to answer their views on certain matters, they are often not the true thoughts at heart, but rather the images after management. Such management of impression inevitably leads to moral deviation. Therefore, the impact of real accounting ethical behavior on occupational quality will be studied in this paper (Wright D B, et al. 2016) [5].

The prevention of financial risk is an important part of the optimization of government accounting. The evaluation of financial credit risk is mainly to evaluate the possibility of adverse events from the probability and frequency of adverse events and to assess the degree of impact and consequences of risks (Turker M, et al. 2015) [6]. In the past, the financial risk assessment mainly adopted the traditional evaluation methods such as "5C" elemental analysis and linear discriminant analysis. These methods tend to be used only in the evaluation of certain restrictions, there will be limitations (Kong Y, et al. 2016) [7]. After entering the new century, with the development of computer aided technology, artificial intelligence technology has brought more scientific methods for financial risk assessment. Intelligent algorithms such as genetic algorithms and support vector machines are used in the construction of financial credit risk assessment models (Abd A M, et al. 2016) [8]. These intelligent algorithm models can effectively solve complex nonlinear conditions and have significant advantages

in dealing with some specific complex models. In view of the fact that the amount of governmental financial data sets is not very large, but the data types and attributes involved are quite complex, in this paper, data mining technology is introduced, the prediction model of optimizing financial credit risk is established to help the government optimize the financial management space and prevent financial risk. The principle of SVM algorithm is based on statistical learning theory, which adopts the concept of VC dimension and uses the principle of structural risk minimization to solve the problem optimally (Li G, et al. 2015) [9]. The learning ability of support vector machine algorithm is outstanding, which maps a space that can't be linearly divided into a highly dimensional linear separable space through a prior specified nonlinear mapping and uses the principle of structural risk minimization to construct the optimal classification hyperplane in this high dimensional space, that is, to find the optimal solution (Yong Y, et al. 2016) [10].

2. Analysis of moral behavior of accountants based on cognitive psychology

From the perspective of cognitive psychology, the research on the process of moral decision in the field of financial accounting mainly starts from the perspective of moral reasoning. Moral reasoning is the process of decision making by personal morality, and it is the process of judging and determining the right and wrong of a matter by personal logic (Proctor R W, et al. 2016) [11]. Generally speaking, individuals will proceed through four stages: moral identification, moral judgment, moral intention and moral behavior, but these four stages will change frequently (Smith D J, et al. 2015) [12]. Different logical sequences represent the analytical framework for the generation of individual moral behavior. Moral recognition is the sensitivity and perception of a person to moral behavior in a particular environment. Moral judgment is a person's non - Interpretation of the specific behavior of a particular situation. Moral intent is the intention of a person to follow moral values and decide to conduct a moral and correct act (Houwer J D, et al. 2017) [13]. Moral behavior is the ability of the individual to have the perseverance, the execution and the intention to act after the moral intention. Moral reasoning is a conceptual process. The above four elements of morality constitute the main framework of moral reasoning (Ifenthaler D, et al. 2015) [14]. In evaluating the ability of individual moral judgment, the "defining issues test" (DIT) has been adopted to measure the level of individual moral development. The DIT test studies the moral dilemma with 3 to 6 stories and the title after the story. Through the DIT test, each participant's "postponement level" score can be obtained, denoted by P, which is also referred to the principled moral score. The higher the P value, the higher the degree of moral development of the tested person. Through pre-experiment, accountants fill out the instructions, and manually integrate after completing the questionnaire test. According to the theory of moral development, the DIT method is used to calculate the scores from the second to sixth stages. The P score here is the score after converting the sum of the scores of the individuals in the three stages of 5A, 5B and 6. P score represents the depth of thinking of personal moral philosophy, and the average score P of the adult is below 40 points, and the effective score is between 0 and 95. M score is used to test the consistency of the test paper. If the M score exceeds 8 points, the test is not serious, or the answer is not understood, so the paper will be abandoned.

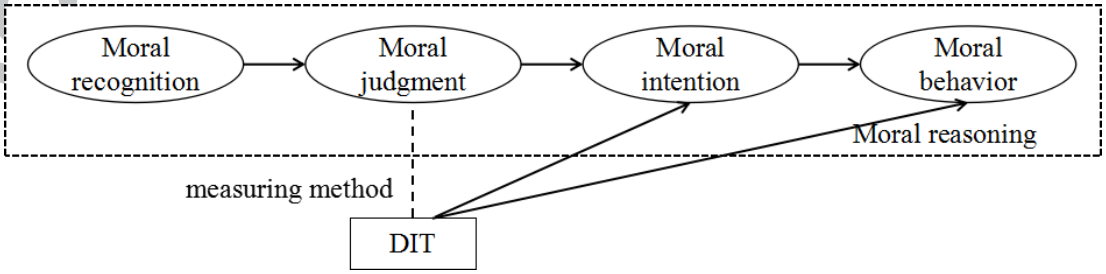

Fig.1 The relationship diagram of DIT, four component model and moral reasoning

Moral intensity refers to that the characteristics of the unnoticed events also affect the individual's moral or immoral behavior. Moral intensity refers to the characteristics that affect the individual affairs, mainly including six aspects: social consensus, urgency of events, probability of outcome, severity of results, proximity and effect concentration. At present, the research on moral decision-making mainly

includes the four-component model, the personal scene interaction model, the moral decision-making model and the five-stage synthesis model based on the theory of reasoning behavior. The model of personal moral decision-making based on moral intensity theory is also called the issue-contingent model. That is to say, the process of moral reasoning is determined by the intensity of morality, which has an impact on the whole process of moral recognition of moral behavior. The structure of the issue-contingent model is shown in Fig. 2.

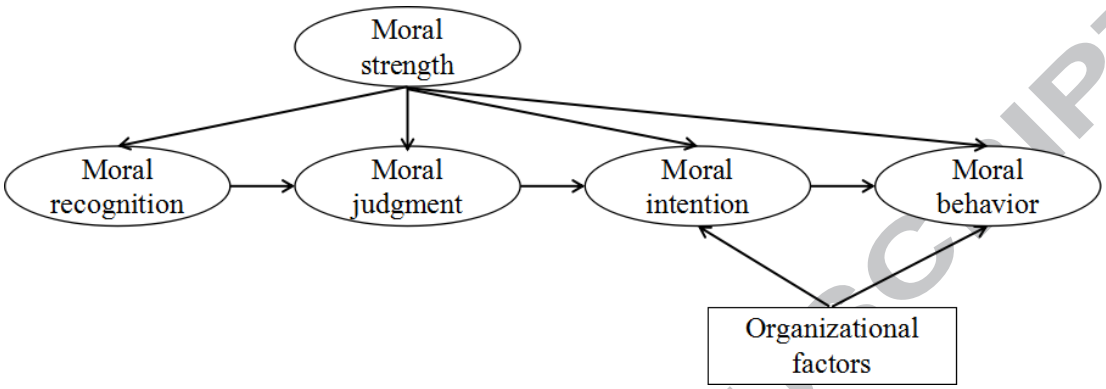

Fig.2 Structure diagram of problem contingency model

Professional ethics is a certain moral standard that professionals should abide by, that is, the basic restrictions on ideological understanding, working style and working style from occupation. The professional ethics of accountants is the bottom line of moral principles and behavior norms that should be carried out in the daily work of a person engaged in the profession of accounting. High-quality accounting information is directly linked to the ethics of accountants. In practical work, when accountants want to abide by professional ethics, they will conflict with the interests of the management, which is the moral dilemma. When accountants make psychological considerations in the face of moral considerations, they often use cost benefit method to do algorithm analysis, so as to maximize personal utility. The cost of accounting behavior mainly includes transaction costs and psychological costs. The cost function is shown in equation (1). C is the total cost of accounting activities, BC is transaction cost, w is material cost, r is labor cost, c is financial cost, PC is psychological cost, p is detected probability, q is the number of accounting behaviors, d is the degree of possible punishment. Formula (2) is the income function of accounting behavior, which is made up of material factors and mental factors. R represents the total revenue of accounting activities, and SR is the income of mental factors, which is a kind of spiritual reward. MR is the material factor income. From the formula (1) and (2), the expression equation of the cost - benefit model of the accountants can be obtained as shown in formula (3).

$$C=\sum_{i=1}^{n}\left(BC_i\right)+\sum_{i=1}^{n}\left(PC_i\right)=\sum_{i=1}^{n}f_i\left(r,w,c\right)+\sum_{i=1}^{n}f_i\left(p,q,d\right) \qquad (1)$$

$$R=\sum_{i=1}^{n}f_i\left(SR\right)+\sum_{i=1}^{n}f_i\left(MR\right) \qquad (2)$$

$$E=\frac{\sum_{i=1}^{n}\left(R_i-C_i\right)}{\sum_{i=1}^{n}C_i} \qquad (3)$$

After calculating the cost and benefit of accounting, there is a scientific basis for the evaluation of the moral level of the accountants. Afterwards, according to the current environmental factors in China's accounting profession, a model of accountants' ethical behavior is established, as shown in Fig. 3. Among them, it is assumed that the main factors that affect the moral level of accountants are the level of moral development and strength, the pressure of superiors and the supervision. The secondary factor is demographic variables. Through this model, the influencing factors can be analyzed in depth.

Among them, H1 to H11 represent different influencing factors respectively.

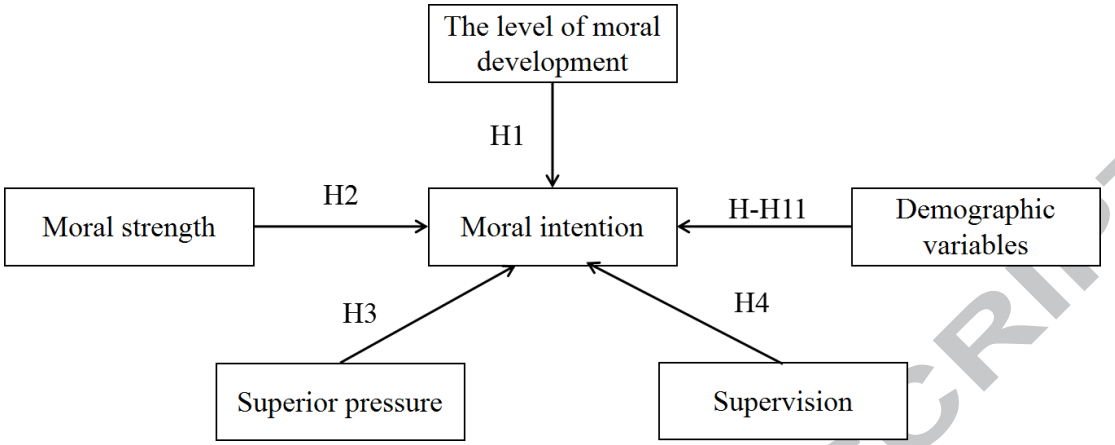

Fig.3 Moral behavior model of Accountants

3. Research on the evaluation of accounting risk based on support vector machine

As an important tool for classification, support vector machine is a new machine learning method, and has the ability to successfully solve the nonlinear and high dimension problems in the field of pattern recognition, which has attracted wide attention in the field of effectively improving the learning performance and generalization performance. The evaluation of accounting risk based on support vector machine has a good theoretical basis (Zhang X, et al. 2016) [15]. Support vector machine (SVM) method first appeared in the United States Bell Labs. As a new machine learning method, its theoretical basis is statistical learning (Khan S, et al. 2016) [16]. The SVM model can effectively solve the problems of small sample size, high dimensionality and nonlinearity in the field of identification. Under the principle of minimizing structural risk, SVM model can achieve a good balance between classification capacity and training error, and help different types of data samples to find the most suitable classification surface (Dixit S, et al. 2015) [17]. Compared with other classifiers, the most outstanding feature of this classifier is that it can work in the space of high dimension and even infinite dimension, and it is fast and accurate to judge the dimensionality and generalization performance in processing pattern recognition. Therefore, SVM has become the dominant machine learning standard in the fields of quadratic programming, relaxation research, hyperplane analysis and so on (Yao T T, et al. 2015) [18].

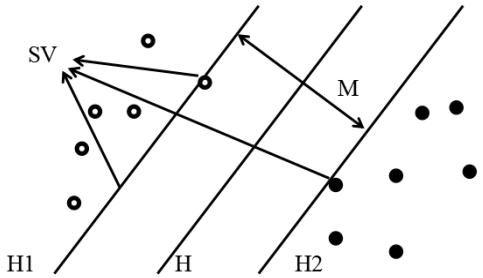

Fig.4 Flow chart of support vector machine classification

Support vector machine is a mathematical model for solving the optimal separating hyperplane based on the state of linear classification (Caraka B, et al. 2017) [19]. The optimal separating hyperplane here is used to distinguish two types of samples from the training sample and make the maximum interval between them (Gandoman S H, et al. 2017) [20]. Fig.4 shows the optimal separating hyperplane in a two-dimensional state in which the training samples can be classified linearly. The solid and open circles here represent two training samples, and the classification line H divides the training sample into two parts. H1 and H2 are the straight lines of the two types of samples that are

closest to H and are parallel to H. M is the distance between H1 and H2. The accurate identification of samples through the classification lines can effectively control the experience risk. When the distance between H1 and H2 reaches the maximum, the VC confidence value can be reduced to a smaller state. The classification line here is the optimal classification line. When the model is extended to the high dimensional space, the optimal classification line will become the optimal separating hyperplane. The relevant parameters of the training samples are set as follows: $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n), x \in R^d, y_i \in \{-1, +1\}$, and the expression formula of the classification line H in d dimensional space is shown as formula (4). Among them, w is a normal vector of the classification plane, and b is a deviation value. This formula for solving the optimal classification line can be transformed into quadratic programming formula, as shown in formula (5).

$$(w, x) + b = 0 \tag{4}$$

$$\begin{aligned} &\Phi(w, b) = \frac{1}{2}\|w\|^2 \\ &y_i(w \cdot x_i + b) - 1 \geq 0, \qquad i = 1, \ldots, n \end{aligned} \tag{5}$$

If there are many eigenvalues of samples collected and the distribution is complicated, some nonlinear separable problems can't be solved by adding slack variables and penalty parameters only. The SVM method is to map a sample to a high-dimensional feature space and construct a linear classifier in a high-dimensional feature space. Fig. 5 shows the mapping of the original space to the feature space.

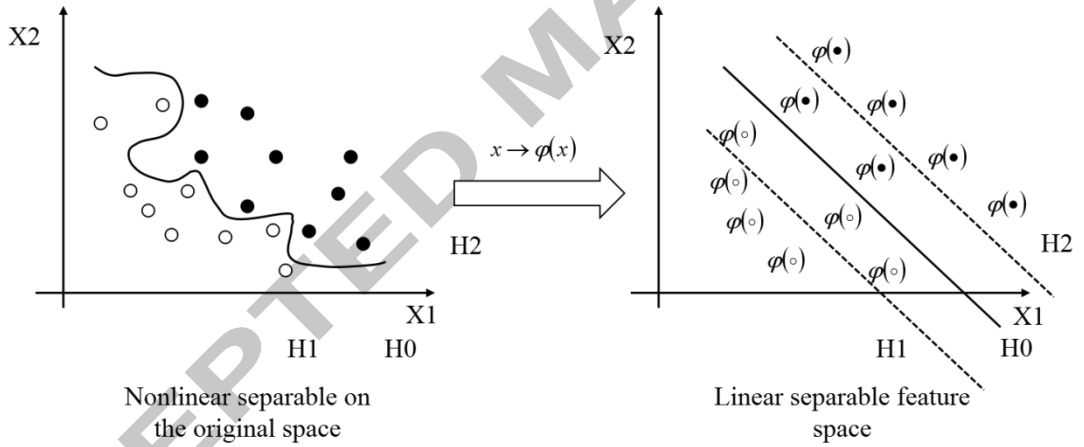

Fig.5 Mapping from the original space to the characteristic space

The process of classifying data samples based on SVM model is to intercept data samples with optimal separating hyperplane, and determine which side of the sample is in the hyperplane. When the data sample is composed of A and B, the data discriminant function is as shown in formula (6). Under the condition of f (x)> 0, the data sample is divided into A. On the contrary, the data sample is divided into B. The result of this formula is the distance from the sample to the classification hyperplane, which is also an important factor affecting the prediction accuracy.

$$f(x) = \operatorname{sgn}\left[\sum_{X_i \in SV} y_i \alpha_i^* (x_i, x) + b^*\right] \tag{6}$$

As can be seen from Fig. 6, the solid line represents the optimal hyperplane for the support vectors φ (x) and φ (x). If the mapping point φ(x) of the sample in the feature space is far from the hyperplane, it is easier to classify the sample, and if the distance is too close, the error will be caused, because the optimal separating hyperplane in the SVM algorithm is determined by the support vector contained in the training sample.

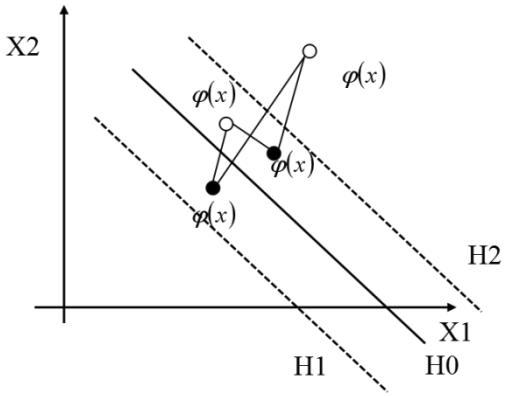

Fig.6 The distance diagram of the test sample and the classified hyperplane

In order to solve the problem that decision surfaces of support vector machines may bring down its generalization ability because of too much complexity in the case of two types of overlapped samples, K nearest neighbor algorithm is introduced to optimize it. K nearest neighbor algorithm is a non-parametric method in statistical methods. In the sample space, the sample x that needs classification is classified. The nearest neighbor training sample near K is found. The classification of every neighbor sample point is counted, and X is classified into the category with the largest number of statistics. The support vector machine uses the weighted K nearest neighbor algorithm to discriminate and optimize the sample points. It is necessary to introduce the distance between the sample point to be classified and its neighbors in the high-order feature space mapped by support vector machine. The distance formula is as shown in formula (7). Here K is a kernel function in the training process of support vector machine.

$$d_i = \left\| \phi(x_i) - \phi(x) \right\|^2 = K(x,x) - 2K(x,x_i) + K(x_i,x_i) \qquad (7)$$

After that, the training process of support vector machine is implemented as shown in Fig. 7. According to the criteria of category judgment, the sample points to be classified are judged, and the result of judgment is used to predict and judge scientifically.

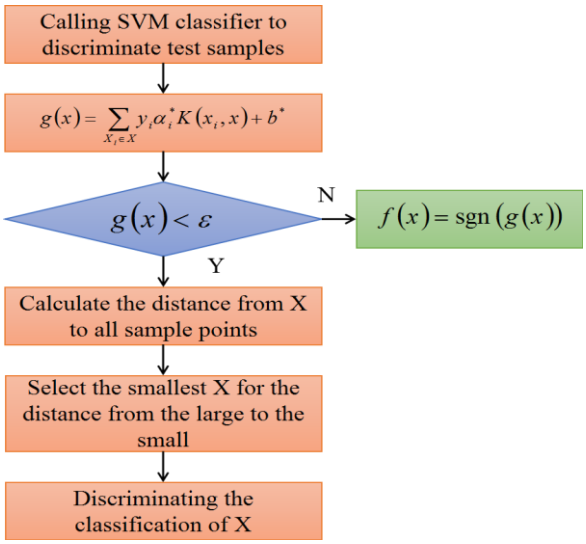

Fig.7 Test sample discriminant method improved flow chart

4. Experiment and result analysis

In order to verify the impact of this article's professional ethics on the government accounting

optimization, in this paper, the simulation experiments of the government accounting department of the provincial capital city of M province were carried out. The experiment was divided into two parts, and one was to experiment with the ethical behavior of accounting personnel. By changing the value or level of a variable, whether the manipulation of the variable affected the dependent variable was determined. If the variable was changed, and there was no other influential variable in the situation, it was considered that manipulating variables brought the change of dependent variables, so as to deduce the correlation and causality between them. The other part was to use the optimized support vector machine algorithm to research and analyze the government financial risk data, and establish a financial credit risk prediction model, so as to provide scientific decision-making basis for the government to improve risk prevention. The data in the ethical behavior simulation test was from 286 accounting staff of the government financial system, including 39 males and 247 females. The age was 22 years old to 45 years old. 286 copies of the survey were released and 280 copies of valid papers were returned. Considering the multi-factorial influence characteristics of the experiment, the control of the independent variables in experiment was not achieved through experimental instruments or environment, but by careful arrangement of experimental design plan and adopting scientific and effective statistical methods.

Tab.1 Multifactor variance analysis table

influence factor	Type three square sum	df	mean square	F	P
PHL	678.017	1	678.017	74.128	<0.001
Moral strength	11.213	1	11.213	1.167	0.284
Superior pressure	39.421	1	39.421	20.782	<0.001
Supervision	617.729	1	617.729	201.946	<0.001
PHL* superior pressure	11.267	1	11.267	5.921	0.018
PHL* supervision	4.698	1	4.698	1.529	0.216
Superior pressure * supervision	14.662	1	14.662	17.648	<0.001
PHL* superior pressure * supervision	1.097	1	1.097	1.323	0.253

After the experimental overall design, pre-experiment, statistical effects, statistical power and other processes, the results of this experiment were drawn. The results of multivariate analysis of variance are shown in Table 1. It can be concluded from the table that among the four main effects, the personal moral development, the pressure of superiors and the supervision of P score bring significant differences to the moral intention of accountants. The average moral intention will increase with the increase of moral intensity, but the increase is not obvious, so moral intensity has no significant influence on the moral intention of accountants. It can be seen from the interaction effect that the interaction between the level of moral development and the superior pressure has a significant impact on the ethical intention of accountants, as shown in Fig. 8.

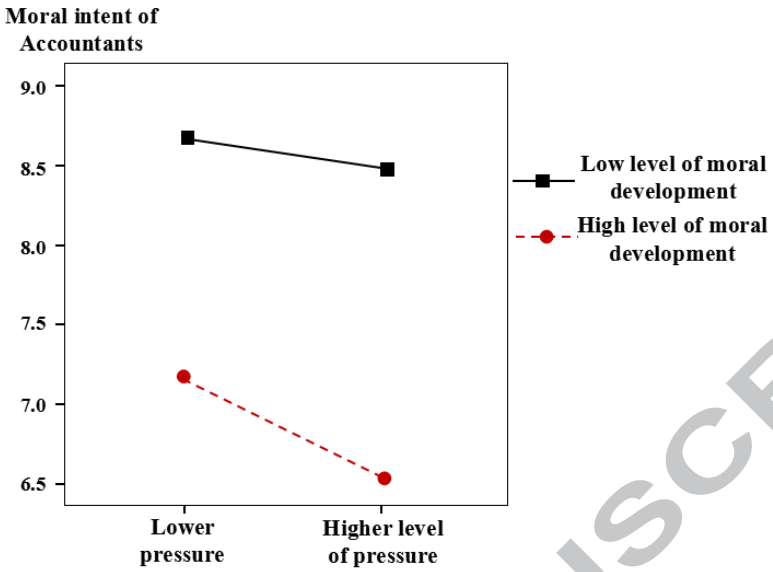

Fig.8 The interaction diagram of the level of moral development and the pressure of the superior

For the mining test of support vector machine in the financial risk, the experimental data came from the government financial management system. This data was collated and included 66 indicators in 5 aspects, such as short-term debt service, development ability and long-term debt service. In order to establish a precise data model, the data was pre-processed and the support vector machine algorithm was verified by a comparative experiment. The weighted K nearest neighbor algorithm introduced in this paper was used to optimize the support vector machine algorithm. The experimental results obtained are shown in Table 2. The results show that the weighted K nearest neighbor algorithm can optimize the discriminant method of support vector machine and improve the accuracy of the model.

Tab.2 The discriminant result of the weighted K- nearest neighbor algorithm

Discriminant sample	Actual category	g(x)	Pre discrimination category	Re discriminant category
x1	1	0.08765	1	1
x2	1	0.22517	1	1
x3	1	-0.00132	-1	1
x4	-1	0.00659	1	-1
x5	-1	0.03814	1	-1
x6	-1	-0.17219	-1	-1
x7	-1	0.050912	1	-1

5. Conclusions

Preventing financial risks is an important part of optimizing government accounting office work. The professional moral behavior of accountants will also have a direct impact on the quality of accounting work. Therefore, in this paper, cognitive psychology was introduced to analyze the ethical behavior of accountants, and the computer aided data mining technology was adopted, and the assessment and management of government accounting risk of financial accounting personnel were studied. Based on an analysis of the ethical behavior of accountants, quarter division method was used and the cost-benefit algorithm for accountants who encounter ethical dilemmas in compliance with professional ethics was proposed to maximize utility. The support vector machine algorithm of data mining technology was used, and the financial risk prediction model was studied. The basic structure and flow of support vector machine algorithm were discussed, and the optimal discriminant method of the weighted K nearest neighbor algorithm was proposed and adopted to effectively improve the accuracy of the model. Finally, the simulation experiment and comparative verification of the proposed model were carried out. The experimental results show that taking full account of the interaction between supervisory pressure and supervision can effectively improve the level of ethical development of accountants. The optimized data mining technology can make the financial risk prediction more accurate and effective. Although this research has achieved good results, the next step is to further analyze the risk of the accounting industry and adopt a more advanced prediction model to improve the level and quality of financial management.

Acknowledgements:

The study was supported by national social science fund of Chinese Government. Title of the project: Research on Government System of Comprehensive Financial Reporting. Grant No. 14FJY004.

References

[1]Skavronskaya L, Scott N, Moyle B, et al. Cognitive psychology and tourism research: state of the art. Tourism Review, 2017, 72(2).

[2]Johnson M K. Cognitive Neuroscience: Applied Cognitive Psychology. Journal of Applied Research in Memory & Cognition, 2016, 5(2):110-120.

[3]Pappas T N. Cognitive psychology meets art: exploring creativity, language, and emotion through live musical improvisation in film and theatre. Proceedings of SPIE - The International Society for Optical Engineering, 2015, 9394:939403-939403-15.

[4]Quinlan P. The Oxford handbook of cognitive psychology. Quarterly Journal of Experimental Psychology, 2015, 68(11):1-4.

[5]Wright D B. Learning From Others in an Educational Context: Findings From Cognitive Psychology. Journal of Cognitive Education & Psychology, 2016, 15(1):4–439.

[6]Turker M, Koc-San D. Building extraction from high-resolution optical spaceborne images using the integration of support vector machine (SVM) classification, Hough transformation and perceptual grouping. International Journal of Applied Earth Observation & Geoinformation, 2015, 34(5):58-69.

[7]Kong Y, Qu D, Chen X, et al. Self-Organizing Map (SOM) and Support Vector Machine (SVM) Models for the Prediction of Human Epidermal Growth Factor Receptor (EGFR/ ErbB-1) Inhibitors. Combinatorial Chemistry & High Throughput Screening, 2016, 19(5):32.

[8]Abd A M, Abd S M. Modelling the strength of lightweight foamed concrete using support vector machine (SVM). Case Studies in Construction Materials, 2017, 6:8-15.

[9]Li G, You J, Liu X. Support Vector Machine (SVM) based prestack AVO inversion and its applications. Journal of Applied Geophysics, 2015, 120:60-68.

[10]Yong Y. Spatial Choice Modeling using the Support Vector Machine (SVM): Characterization and Prediction. Studies in Computational Intelligence, 2017, 42(12):561.

[11]Proctor R W, Urcuioli P J. Functional relations and cognitive psychology: Lessons from human performance and animal research. International Journal of Psychology, 2016, 51(1):58.

[12]Smith D J. Why Cognitive Psychologists Should Know Comparative Psychology; Why

Comparative Psychologists Should Know Cognitive Psychology. International Journal of Comparative Psychology, 2015, 16(1):44-44.

[13]Houwer J D, Hughes S, Barnes-Holmes D. Bridging the Divide Between Functional and Cognitive Psychology. Journal of Applied Research in Memory & Cognition, 2017, 6(1):47-50.

[14]Ifenthaler D, Sampson D G, Spector J M. Interactions Between Cognitive Psychology, Educational Technology, and Computing in the Digital Age. Technology Knowledge & Learning, 2015, 20(2):129-131.

[15]Zhang X, Amin E A. Highly predictive support vector machine (SVM) models for anthrax toxin lethal factor (LF) inhibitors. Journal of Molecular Graphics & Modelling, 2016, 63:22-28.

[16]Khan S, Ullah R, Khan A, et al. Analysis of dengue infection based on Raman spectroscopy and support vector machine (SVM). Biomedical Optics Express, 2016, 7(6):2249.

[17]Dixit S, Gupta R K. Layered Approach to Classify Web Pages using Firefly Feature Selection by Support Vector Machine (SVM). International Journal of u- and e- Service, Science and Technology, 2015, 8.(1):125.

[18]Yao T T, Cheng J L, Xu B R, et al. Support Vector Machine (SVM) Classification Model Based Rational Design of Novel Tetronic Acid Derivatives as Potent Insecticidal and Acaricidal Agents. Rsc Advances, 2015, 5(61):49195-49203.

[19]Caraka B, Sumbodo B A A, Candradewi I. Klasifikasi Sel Darah Putih Menggunakan Metode Support Vector Machine (SVM) Berbasis Pengolahan Citra Digital. 2017, 7(1):25.

[20]Gandoman S H, Kiamehr N, Hemetfar M. Forecasting Initial Public Offering Pricing Using Particle Swarm Optimization (PSO) Algorithm and Support Vector Machine (SVM) In Iran. Business & Economic Research, 2017, 7(7 1).