

Research on Execution of Civil Servants and Professional Ethics based on Data Mining Technique and Joint Modeling Analysis of Multiple Factors under Big Data Environment

Yang Du^{1,2} and Hongwei Wang^{3,*1}

¹College of Humanities and Law, Northeast Agricultural University, No.59 Mucai Street Xiangfang District, Harbin, P.R. China 150030

²School of Marxism, Northeast Forestry University, No.26 Hexing Road Xiangfang District, Harbin, P.R. China 150040

³College of Humanities and Law, Northeast Forestry University, No.26 Hexing Road Xiangfang District, Harbin, P.R. China 150040

^{1,2}duyang973@126.com, ³weiwei129@sina.com

Abstract

With the continuous advancement of the data science and engineering, the combination of data analysis with other applications has been a trend. Data mining is in accordance with the established business goals from huge amounts of data to extract the effective potential and can be understood model of advanced treatment process. In this research paper, we conduct discussion on the execution of civil servants and professional ethics based on data analysis. Starting from the concept of data mining, this paper introduces the objects as the functions of data mining and mining process, mining algorithm combined with several kinds of common data mining, decision tree method, association rules method and neural network to its main ideas and improve the related description. We firstly review traditional data classification algorithms such as the Bayes classification algorithm to serve as the foundation. Later, we propose our method by using the modified SVM and deep neural network optimization. Then, we review the principles of civil servants and professional ethics with the combination of the prior discussion. The experimental result illustrates the feasibility and effectiveness of our method. We also discuss the future research plan in the final part.

Keywords: Professional Ethics, Data Mining Algorithm, Execution of Civil Servants, Joint Modeling Analysis, Big Data Environment, Experimental Simulation

1. Introduction

Data mining as a new multidisciplinary cross application field that is decision support from all walks of life activities play an increasingly important role. With the rapid development of the information technology, all walks of life have accumulated the massive amounts of the heterogeneous data. These data are often implies a variety of useful information, just rely on the database query retrieval mechanism and the statistical methods are difficult to obtain this information, is an urgent need to put these data into useful information and knowledge, so as to achieve the aim of service for decision making.

Data mining is a great deal of data storage, use of pattern recognition, statistical and mathematical techniques, screening of discovering new meaningful relationships, patterns and trends. It is the large, incomplete, noisy, fuzzy and random data, extract implicit in it, people don't know in advance, but it is potentially useful information and knowledge of

* Corresponding Author

the process. It mainly rely on artificial intelligence, machine learning and the statistical techniques, the data is inductive reasoning, to dig out the potential patterns, predict the future trend to provide support for decision making. At first, all kinds of business data are stored in database, only then gradually to the business data from a database query and then develop to instant traversal of the data.

Data as the main carrier of the information plays an important role in information society nowadays, fields everywhere from all walks of life is the presence of data, data provides us with a wealth of information. However, how to go from a large number of seemingly random data reveal the inherent law of implicit, discover useful information to guide people to make scientific inference and decision which also need to analyze these complicated data. In basic classification, the current data mining techniques could be categorized as three follows. (1) Knowledge discovery. Knowledge discovery class of statistics and data mining technology is a kind of data mining technology is completely different kind of mining technology. It can filter information from a large amount of data of data warehouse, looking for possible new operating mode in the market and find people did not know the facts. (2) Statistical analysis of class. Statistical analysis techniques used in the data mining model with linear analysis and the nonlinear analysis, regression analysis, logistic regression analysis, univariate analysis, multivariate analysis, time series analysis, nearest neighbor algorithm and clustering analysis technology. These techniques can be used to check the abnormal data then use a variety of statistical model and mathematical model to explain these data. (3) Other mining technology. Other data mining techniques include text data mining, Web data mining, classification system, visual system, spatial data mining and distributed data mining, *etc* In the following image one, we show the application scenario of the big data.

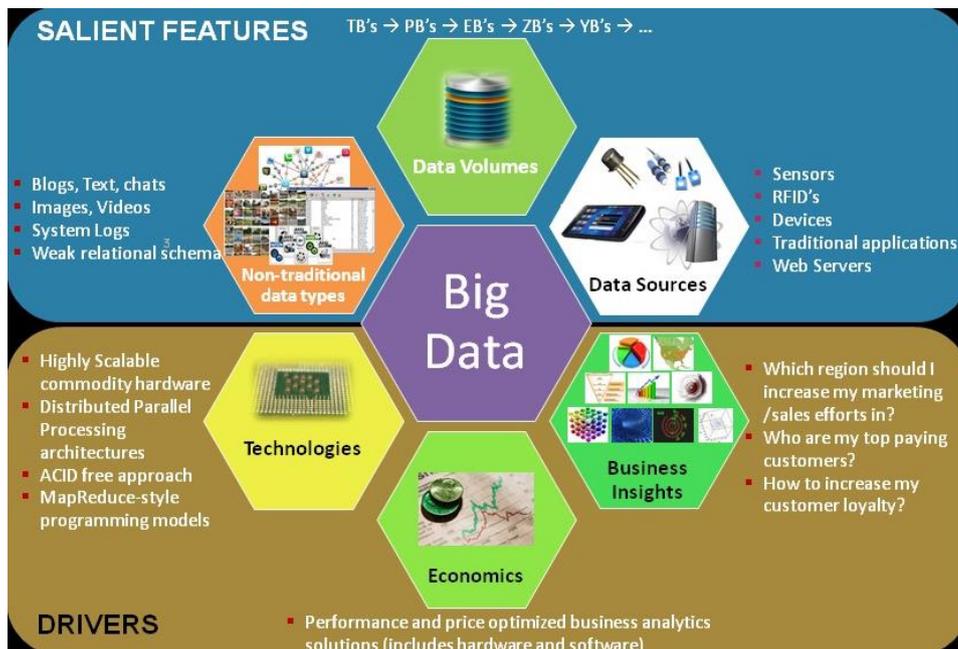


Figure 1. The Application Scenario of the Big Data Technique

As the main body of carrying out the policy of the civil servant is in the process of policy implementation is the most important part of the strength of the policy executive power of civil servants directly affected the policy can be smoothly carried out. At present, the weak part of the civil execution, subjective one-sided understanding of the policy, cannot correctly reflect the intention of the decision-making in the execution, the interruption or failure of the policy implementation. Therefore we achieve purpose of policy in real full implementation, to improve policy executive power of civil servants.

Ethics of civil servants is one of the social moral form, refers to the civil servants in the process of public administration must abide by the code of a common code of conduct and the sum of value concept and mode of thinking. The quality of the civil servants, namely behavior shows thinking character recognize such substantive reflects its ethical and moral nature. Ethics as a special social consciousness, in some cases, can play a role not by law. Especially in the big data and the complicated social background, the release of this effect is better and more necessary.

In this paper, we conduct research on execution of civil servants and professional ethics based on data mining technique and joint modeling analysis of multiple factors under big data environment. Pattern recognition techniques have been in the social life that obtained wide application in various fields. However in the past most classic classification algorithm is the premise of success requirements sample data is roughly balanced between classes, but with the development of technology, there are more and more new classification problem, does not meet this premise type can be seen as unbalanced data classification problem. To combine the data analysis technique with execution of civil servants and professional ethics we organize the reminding of the paper as the follows. In the Section 2, we analyze the traditional basic and the core data classification and the clustering methodologies with the theoretical analysis. In the Section 3, we propose our novel data classification methodology with the detailed and in-depth analysis with discussion. In the section 4, we discuss civil servants and professional ethics enhancement and evaluation model combined with prior discussion. In the Section 5, we numerically experiment the proposed methodology. As the summary and conclusion of the paper, we summarize our article and cast the prospect for further research.

2. Overview of Traditional Data Classification Algorithms

In this section, we conduct overview of the traditional data classification algorithms. We choose the algorithms based on following criteria. In practice, each class of misclassification cost should be different, but it needs prior knowledge, which makes the real cost is difficult to accurately estimate fault points. Classification is one of the important missions of the data mining and its goal is according to the data of the existing categories summing up the general description of each category. Classification of different type errors is often corresponded to the different price. When the sample of misclassification costs are not equal, based on the precision of the traditional classification algorithms usually cannot be used directly.

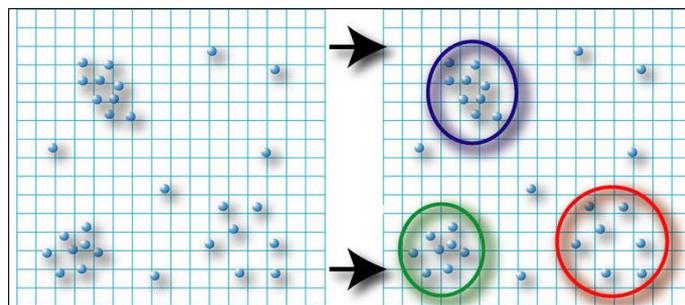


Figure 2. The Demonstration of the Data Classification Process

2.1. Rotation Forest Classification Algorithm

Integration algorithm is of the same object with the multiple rather than the single learning machine learning and prediction, and the predicted results of each learning machine, in order to get a more comprehensive and the reliable judgment. Rotation forest algorithm is a new integration algorithm, this algorithm chooses decision tree as its base classifier, and in the decision tree learning, due to the classifier is too complex, so too to

adapt to the noise, leading to a fitting problem. Integration algorithm is the difference between the base classifier and precision in two crucial factors affecting algorithm performance. That is to say, an integrated classifier to obtain satisfactory results, should be the base classifier has high precision, at the same time, the difference between the base classifier to as large as possible. Therefore, by improving the base classifier to improve the effect of classification of integration algorithm become the focus of attention. The following formulas denote the processes.

$$l_k = \sum_{i=1}^r \beta_i g(w_i \cdot x_k + b_i), \quad k = 1, 2 \dots N \quad (1)$$

$$J_{Optimized} = \sum_{i=1}^n \sum_{k=1}^c \mu_{ik}^m |p_i - v_k| \cdot \langle p_i^2 - v_k \rangle \cdot \langle p_i - v_k^2 \rangle \quad (2)$$

Algorithm firstly by the rotation of the core original forest algorithm to deal with the base classifier training set, in order to improve the differences of each base classifier. Then choose ELM algorithm as the base classifier, a fitting problems existing in the original algorithm, at the same time improve the classification accuracy. The table one shows the pseudo-code.

Table 1. The (RFC) Rotation Forest Classification Algorithm

Algorithm 1. Rotation Forest Classification Algorithm (RFC)

1. **Input:** Training data set, the instance of the corresponding class label the number of base classifiers in the integration algorithm.
 2. **Output:** A list of the clusters and partitions as the classification result.
 3. The results of the base classifier synthetically using the (1).
 4. The transformed sample class label as the output of the network (2).
 5. Update the parameters of the $\langle p_i^2 - v_k \rangle \cdot \langle p_i - v_k^2 \rangle$ as reversion.
 6. **While** (Do not meet the stopping criteria).
-

2.2. The Modified Naive Bayes Classification Algorithm

Price sensitive theory has been successfully applied in many of the fields. The cost of data sensitivity problem relies on the theory of cost sensitive applications, only will be the price of sensitive theory combining with the specific application field can effectively solve the actual problem. Price sensitive classification is a different price for the different types of the error distribution, pay attention to avoid the costly error when classification, in order to achieve the classification errors of the lowest total cost target. Using the algorithm model training is the key to reasonable training instances of misclassification cost factor. By adjusting the positive and the negative samples or decision threshold indirectly realize study price sensitive, so can't guarantee the price sensitive learning effect. Cost function is difficult to in the present rare class classification algorithm according to the specific data distribution problem of adaptive determine, introducing global cost matrix, the naive Bayesian classification is improved.

Table 2. The Modified Naive Bayes Classification Algorithm

Algorithm 2. The Modified Naive Bayes Classification Algorithm (MNB)

1. **Input:** Put the price of a few kinds of divided into most classes is generally higher than the price of the majority class is divided into a few classes.
2. **Output:** The optimized data classification partitions and the unique parts.
3. Conditional independence assumption based feature computing and analysis.

$$P(a_1, a_2, \dots, a_m | c_j) = \prod_{i=1}^m P(a_i | c_j)$$

4. Discriminant for the corresponding record expression for the records in the class.

$$P(u) = \sum_{i=1}^l P(u | c_i)P(c_i)$$

5. Increment $t=t+1$; repeat steps 2 and 3 until the convergence condition.

3. Our Proposed Data Classification Methodology

3.1. Kernel Based Support Vector Machine Principles

In the real world, most of the pattern classification problems in the original sample points within the sample space is linear inseparable, solve the problem of the general approach is to adopt method of nonlinear mapping, the original space samples mapped to high-dimensional feature space, make the samples in the high-dimensional space linearly separable. Determine whether a certain parameter value in the optimal algorithm, is essentially to select the proper parameter value to make the algorithm corresponding minimum the error rate. In the training sample lot and the condition of high proportion of support vector, the limit support vector machine (SVM) to a training sample subset, can reduce the number of support vector. In the formula three, we demonstrate the objective function for the SVM.

$$\min_{w, b, \xi} \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \quad s.t. \quad \xi \geq 0 \tag{3}$$

Kernel function description is two sample points in high dimensional feature space, inner product, so kernel function is reflected in the high dimensional feature of the space position relations between any two sample points expressed as the formula 4.

$$x \rightarrow \Theta(x) = (\Theta_1(x), \Theta_2(x), \dots, \Theta_l(x))^T \tag{4}$$

Need to know is the most important of the sample points in high dimensional feature space, inner product, namely, relation between the location of the relationship between each other if know the inner product, this is the particular form of the kernel function is also do not need to know, so we care about is eventually the inner product between the sample points in high dimensional feature space, the study of kernel function can be converted into the study of the relationship between inner product. The optimized objective function is shown below.

$$f(x) = \text{sgn}(w \cdot \Theta(x) + b) = \text{sgn}\left(\sum_{i=1}^l a_i y_i \Theta(x_i) \cdot \Theta(x) + b\right) \tag{5}$$

A time for working knowledge of the training samples, and fixed the training sample. The algorithm is the key to select optimal working set selection algorithm, and the random method is adopted in the selection of the working set, thus limiting the rate of convergence of the algorithm. We optimize is by using the following formula.

$$L(w, b, a) = \frac{1}{2} \|w\|^2 - a(y((w \cdot x) + b) - 1) \quad (6)$$

It is obvious that the nuclear matrix is symmetric matrix and for inner product matrix, said any two samples in high dimensional feature space, inner product, determines the relative location of sample points in high dimensional feature space. Concrete form, also don't have to know specific coordinates of samples in high dimensional feature space and only need to look for a suitable to meet the theorem of the nuclear matrix, and therefore can be converted into the problem of parameter selection of kernel function to find suitable matrix of the problem.

3.2. Optimized Deep Structure Assisted Model for Enhancement

Neural network is an important branch of machine learning and is a mainstream research direction of intelligent computing has long been the attention of many scientists and research. It is rooted in many disciplines, combination of mathematics, statistics, physics, computer science and engineering. Have found that it can solve some difficult to solve the problems in the traditional sense, also provides the solution of the problems with the new ideas. The basic idea is neural network algorithm. The positive learning process by signal transmission with error back propagation of two processes forward propagation, input samples from the input layer, after each processing step by step a hidden layer to output layer. If the actual output and desired output of output layer, into the error back propagation phase, in some form of error back propagation is the output error back propagation step by step through the hidden layer to the input layer, and the error to all elements of each layer to get error signal of each unit, the error signal as the basis of a fixed weights of each unit. This signal is spread and error back propagation process of each layer weights adjustment is carried out repeatedly. Weights are constantly adjusts process, that is, the network training process. This process has been to the network output error reduced to an acceptable level, or at the preset number of learning.

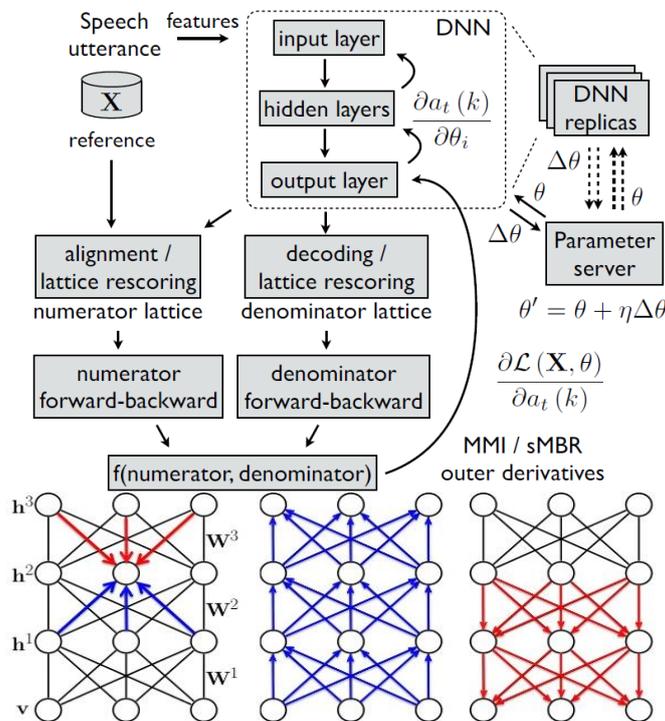


Figure 3. The Optimized Deep Structure Assisted Model

As demonstrated in the Figure 3, we show the optimized deep structure assisted model. There is a flat area error curved surface. Weight adjustment into the plain, the reason is that the neuron output into the transfer function of the saturated zone. If after the adjustment into flat area, managed to compress neuron net input, make its output out of the saturated zone transfer function, we can change the shape of the error function, so that the adjustment from the flat area. Implement the idea of the specific practice is introduced in the original transfer function a gradient factor. In practice, it is difficult to determine vector from beginning to end all the right best. Can be seen from the error of curved surface, the vector in plain area is too small will increase the number of training, thus hope to increase more.

3.3. The Optimization Techniques Introduction

When the dimension of the training sample set is too large, or data is high, its computation cost is higher. In order to improve the core efficiency of classification algorithms, mainly by dimension reduction and establish the efficient index structure, the narrowing the scope of the search and shorten the time of looking for the nearest neighbor. By looking in the original training focus for representative sample set to replace the training set. In most of unbalanced classification problem, rare class is the key of the classification. So wrong points of rare class samples need to pay a greater price and the price sensitive study by giving each category in the different fault points, to deal with unbalanced classification problem in two types of problems, for example, suppose that is scarce, and has a higher fault points, while training the classifier, right and wrong points are kind of samples for more punishment, forcing the final classifier for sample class has a higher recognition rate.

$$d(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (7)$$

The training data set is another effective method of training set balance. Its main idea is the kinds of samples were randomly divided into a series of disjoint subsets, the size of these subsets by sparse number of such sample set and learning sample distribution proportion of decision in advance. The optimized information entropy is defined as follows.

$$I(s_1, s_2, \dots, s_m) = -\sum_{k=1}^m p_k \log_2(p_k) \quad (8)$$

Resampling method is by increasing the rarest type of training samples were on sampling and reduce the categories of unbalanced sampling under a sample distribution of the sample become more balanced, so as to improve the recognition rate of classifier to the rare class. The following formula defines the optimization step.

$$E_{Optimized}(A) = I(s_1, s_2, \dots, s_m) \frac{\sum s_{\square}}{s} \quad (9)$$

4. Principles of Execution of Civil Servants and Professional Ethics

Improve execution of civil servants is very important to strengthen the construction of the organ administrative efficiency. This article in view of the current main problems of the civil servants at the grass-roots level on the execution, from the theory of armed, responsibility is not strong enough, the organization perform poorly, the respect such as lack of supervision and restriction are analyzed, and put forward from the breeding advanced execution concept, improve the execution of the fine of the quality, the efficient implementation of innovation mechanism, the strict constraints mechanism, build efficient execution ability and so on to strengthen the construction of authority administrative efficiency, improve execution of civil servants at grass-roots level. The grass-roots civil

Primary Variable	Secondary Variables
1.Loyal to the motherland and people	Firm belief and ideal
	Correct world outlook and on life and values
	Safeguard the unity of the country
	Close ties with the masses
2.Administration according to law	Strong legal sense
	Observe law and discipline
	Maintain strict discipline
	In performing their duties according to law
3.Be loyal to their duties and work hard	Dedication and sense of responsibility
	Love the job
	Do not dereliction of duty and work adversely
	Working efficiency of be particular about
4.Solidarity and collaboration	Despite the big picture
	Not going
	Solidarity and collaboration, dare to responsible
	To ensure that the order
5.Integrity, hard work and plain living	Strict requirements, set a good example
	Uphold justice
	Strict economy, industry and thrift
	Honest self-discipline
6.Honesty, seeking truth from facts	Aboveboard, match words with deeds
	Dare to tell the truth. Starting from the actual
	Against subjectivism and the formalism
	The thorough investigation and study, pragmatic
7.Study hard, proficient in business	To master modern office technology
	Improve the service ability
	Improve the innovation consciousness
	To strengthen their own quality training
8.Modest and prudent, and civility	Manners polite, talk properly
	Healthy life interest
	Personable, clothes clean
	The right attitude towards achievement

Figure 5. The Evaluation Criterion Parameter List

Along with the social transformation, the inherent integration between economic, political, cultural, social relations, prone to moral crisis for example, the influence of the traditional morality is becoming more and more weak as the new moral system has not yet been forming, the capital of the socialist ideological and ethical infiltration, feudal moral is quite serious, the influence of religious moral, very, folk backward traditional moral tradition has a strong vitality. As the mainstream of socialist moral system is not strong enough show in the figure 5, we analyze the issues and measure from the following perspectives. (1) Honesty, seeking truth from facts; (2) Loyal to the motherland and people; (3) Administration according to law; (4) Solidarity and collaboration; (5) Integrity, hard work and plain living; (6) Study hard, proficient in business; (7) Modest and prudent, and civility.

5. Experimental Analysis and Simulation

In order to test the robustness of our proposed model, we conduct simulation in this part. The process of the decision tree induction algorithm based tree, the tree growth process is a process of constant shard the data, controlling a corresponding to a problem, also corresponds to a node. Requirements for segmentation are divided into the group of

"difference" between the biggest. All kinds of decision tree algorithm are the main difference between the measures of the difference between "differences". In the following Table 3, we demonstrate numerical data of the different clustering algorithms and the related figure.

Table 3. The Numerical Data of the Different Clustering Algorithms

Number	Ours	Algorithm1	Algorithm2	Algorithm3	Algorithm4	Algorithm5
1	88.3	82.2	80.4	81.5	84.1	80.3
2	89.5	87.9	88.3	85.2	87.1	87.6
3	93.4	91.3	89.7	90.1	88.8	88.9
4	89.4	81.3	87.6	88.2	88.3	85.5
5	89.1	86.1	81.0	87.9	85.2	83.3
6	94.2	89.9	90.6	89.6	91.1	88.6
7	87.9	88.9	86.4	86.9	85.5	86.1
8	89.9	86.2	84.2	89.3	88.1	86.5

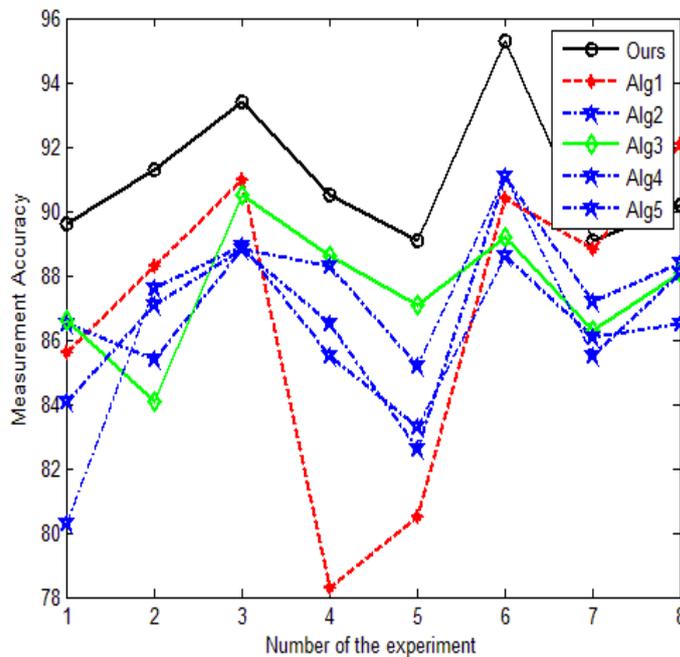


Figure 6. The Visual Comparison of the Different Clustering Algorithms

From this simulation, we record weight of different elements and the influence factors for execution of civil servants. Under the guidance of the theory, the policy execution should be strengthened to promote construction of incentive mechanism, stimulate the policy executors positive implementation, training executives of public spirit and administrative culture, enable it to provide better service for the policy object, such not only can improve policy executive power, also can improve the image of the government in the heart, to the construction of the government credibility. Because of the small group have the better execution, so should be streamlined and change the present situation of our local government bloated to improve the efficiency of policy implementation.

Table 4. The Experiment on Execution of Civil Servants and Professional Ethics

Number	Case1	Case2	Case3	Case4
1	12	23	18	21
2	13	22	19	15
3	12	25	17	19
4	8	19	22	19
5	12	15	18	9
6	9	19	15	13
7	6	19	16	13
8	9	22	15	13
9	18	21	17	19
10	15	17	13	19

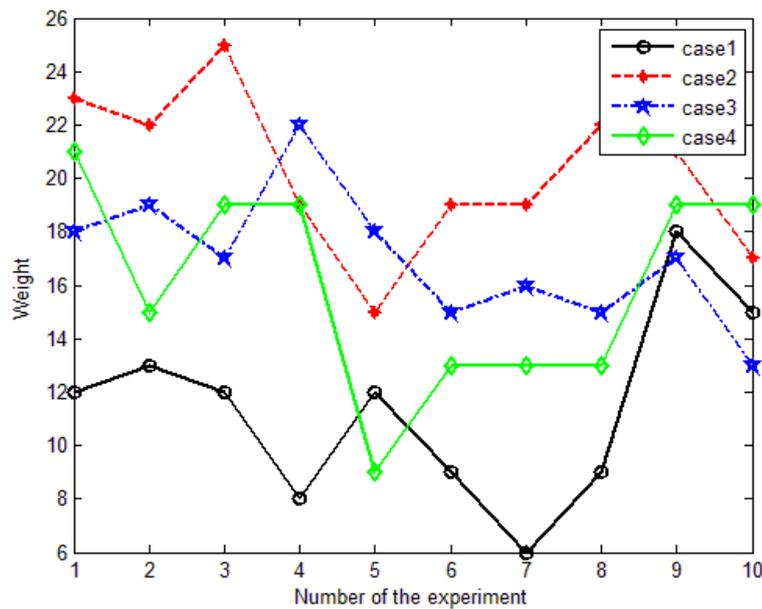


Figure 7. The Visual Comparison of Execution of Civil Servants and Professional Ethics

5. Conclusion and Summary

The birth of data mining technology is meaningful, so that we can from a large amount of data to extract useful information for the decision makers. Union rule mining is to discover a large number of items in the data set of interest between the correlation information. After more than ten years of the development association rule mining has become a more mature and in data mining technology is one of important ways. This system describes association rule mining the concepts involved, association rule mining algorithm and application fields such as association rules. Therefore, we conduct research on execution of civil servants and professional ethics based on data mining technique and joint modeling analysis of multiple factors under big data environment in this research manuscript. We modify the traditional of the evaluation algorithm to propose our revised one. The experiment proves the effectiveness and the feasibility of our methodology.

References

- [1] G. N.-Arribas, V.Torra, A. Erola and J. C.-Roca, "User k-anonymity for privacy preserving data mining of query logs", *Information Processing & Management*, vol. 48, no. 3, (2012), pp. 476-487.
- [2] A. Fernández, S. del Río, F. Herrera and J. M. Benítez, "An overview on the structure and applications for business intelligence and data mining in cloud computing", In *7th International Conference on Knowledge Management in Organizations: Service and Cloud Computing*, Springer Berlin Heidelberg, (2013), pp. 559-570.
- [3] M. A. Faisal, Z. Aung, J. R. Williams and A. Sanchez, "Securing advanced metering infrastructure using intrusion detection system with data stream mining", In *Intelligence and Security Informatics*, Springer Berlin Heidelberg, (2012), pp. 96-111.
- [4] T. C.-K. Huang, C.-C. Liu and D.-C. Chang, "An empirical investigation of factors influencing the adoption of data mining tools", *International Journal of Information Management*, vol. 32, no. 3, (2012), pp. 257-270.
- [5] Y. S. Cho, S. C.Moon and K. H. Ryu, "Mining association rules using RFM scoring method for personalized u-commerce recommendation system in emerging data", In *Computer Applications for Modeling, Simulation and Automobile*, Springer Berlin Heidelberg, (2012), pp. 190-198.
- [6] I. Lee, G. Cai, and K. Lee, "Exploration of geo-tagged photos through data mining approaches", *Expert Systems with Applications*, vol. 41, no. 2, (2014), pp. 397-405.
- [7] D. Delen, D. Cogdell and N. Kasap, "A comparative analysis of data mining methods in predicting NCAA bowl outcomes", *International Journal of Forecasting*, vol. 28, no. 2, (2012), pp. 543-552.
- [8] S. Gupta, D. Kumar and A. Sharma, "Data mining classification techniques applied for breast cancer diagnosis and prognosis", *Indian Journal of Computer Science and Engineering (IJCSE)*, vol. 2, no. 2, (2011), pp. 188-195.
- [9] F. Alonso, L. Martínez, A. Pérez and J. P. Valente, "Cooperation between expert knowledge and data mining discovered knowledge: Lessons learned", *Expert Systems with Applications*, vol. 39, no. 8, (2012), pp. 7524-7535.
- [10] P. Chandrasekar and M. Krishnamoorthi, "BHOHS: A Two Stage Novel Algorithm for Data Clustering", In *Intelligent Computing Applications (ICICA), 2014 International Conference*, IEEE, (2014), pp. 138-142.
- [11] G. Krishnasamy, A. J. Kulkarni and R. Paramesran, "A hybrid approach for data clustering based on modified cohort intelligence and K-means", *Expert Systems with Applications*, vol. 41, no. 13, (2014), pp. 6009-6016.
- [12] P.-L. Lin, P.-W. Huang, C.-H. Kuo and Y. H. Lai, "A size-insensitive integrity-based fuzzy c-means method for data clustering", *Pattern Recognition*, vol. 47, no. 5, (2014), pp. 2042-2056.
- [13] S. R. Kannan, S. Ramthilagam, R. Devi and Y.-M. Huang, "Novel Quadratic Fuzzy c-Means Algorithms for Effective Data Clustering Problems", *The Computer Journal*, bxs085, (2012).
- [14] Y. Guo and X.-Q. Lai, "Use of Partial Supervised Model of Fuzzy Clustering Iteration to Mid-and Long-Term Hydrological Forecasting", In *Fuzzy Information & Engineering and Operations Research & Management*, Springer Berlin Heidelberg, (2014), pp. 287-293.
- [15] Y. Jiao, H. Liu, P. Zhang, X. Wang and H. Wei, "Unsupervised performance evaluation strategy for bridge superstructure based on fuzzy clustering and field data", *The Scientific World Journal*, (2013).
- [16] S. Ramthilagam, R. Devi, T.-P. Hong and S. R. Kannan, "Robust fuzzy clustering techniques for analyzing complicated colon cancer database", *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology*, vol. 27, no. 5, (2014), pp. 2573-2595.
- [17] C. Bi, H. Wang and R. Bao, "SAR image change detection using regularized dictionary learning and fuzzy clustering", *Cloud Computing and Intelligence Systems (CCIS), 2014 IEEE 3rd International Conference*, IEEE, (2014), pp. 327-330.
- [18] L. Bulysheva and A. Bulyshev, "Segmentation modeling algorithm: a novel algorithm in data mining", *Information Technology and Management*, vol. 13, no. 4, (2012), pp. 263-271.
- [19] A. Janusz, H. S. Nguyen, D. Ślęzak, S. Stawicki and A. Krasuski, "JRS'2012 Data Mining Competition: Topical Classification of Biomedical Research Papers", In *Rough Sets and Current Trends in Computing*, Springer Berlin Heidelberg, (2012), pp. 422-431.
- [20] N. Abdelhamid, A. Ayesh and F. Thabtah, "Phishing detection based Associative Classification data mining", *Expert Systems with Applications*, vol. 41, no. 13, (2014), pp. 5948-5959.
- [21] S. K. Wajid and A. Hussain, "Local energy-based shape histogram feature extraction technique for breast cancer diagnosis", *Expert Systems with Applications*, (2015).
- [22] H. Wang and J. Wang, "An effective image representation method using kernel classification", In *Tools with Artificial Intelligence (ICTAI), 2014 IEEE 26th International Conference*, IEEE, (2014), pp. 853-858.
- [23] M. Hussain, S. K.Wajid, A. Elzaart and M. Berbar, "A comparison of SVM kernel functions for breast cancer detection", In *Computer Graphics, Imaging and Visualization (CGIV), 2011 Eighth International Conference*, IEEE, (2011), pp. 145-150.

- [24] J. Yousafzai, Z. Cvetkovic and P. Sollich, "Effects of domain-specific SVM kernel design on the robustness of automatic speech recognition", In Digital Signal Processing (DSP), 2013 18th International Conference on IEEE, (2013), pp. 1-7.
- [25] S. Han, C. Qubo and H. Meng, "Parameter selection in SVM with RBF kernel function", In World Automation Congress (WAC), IEEE, (2012), pp. 1-4.
- [26] M. Polk, "Achieving the promise of transdisciplinarity: a critical exploration of the relationship between transdisciplinary research and societal problem solving", Sustainability Science, vol. 9, no. 4, (2014), pp. 439-451.
- [27] R. Gómez and T. Verge, "Party patronage in Spain: Appointments for party government", Party Patronage and Party Government in European Democracies, (2012), pp. 316-334.
- [28] M. Portal, E. Lande, R. Jones and K. Lüder, "French revolution and German circumspection: reforming budgeting and accounting in national governments", Public Money & Management, vol. 32, no. 5, (2012), pp. 357-361.
- [29] C. Onoka, O. Onwujekwe, B. Uzochukwu and N. Ezumah, "Promoting universal financial protection: constraints and enabling factors in scaling-up coverage with social health insurance in Nigeria", Health Res Policy Syst, vol. 11, (2013), pp. 20.
- [30] C. Wen, D. Wu, H. Hu and W. Pan, "Pose estimation-dependent identification method for field moth images using deep learning architecture", Biosystems Engineering, vol. 136, (2015), pp. 117-128.

