# Research on Early Warning Methods of Economic Crimes Based on Data Mining

# Wang Hongning<sup>\*</sup>

### Abstract

This paper presents an early warning method of economic crime based on data mining. First, analyze the historical data of economic crime cases for the early warning targets, perform preliminary data processing on economic crime data, and use the maximum information coefficient method and the chi-square test method in the effective influencing factor screening method to calculate the attribute weights that affect various economic crime cases. Value, ranks the degree of influence of the attribute, retains the attribute value with high degree of influence on economic crime, and removes redundant attributes. Secondly, analyze and compare a variety of data mining algorithms. According to the characteristics of the economic crime data structure, the decision tree method that can mine the early warning rules is selected, and the defects of the traditional C4.5 algorithm in the decision tree are analyzed. The deficiencies in the mining of early warning rules are improved, and an economic crime early warning model is established on this basis, and the confidence of the model is verified: Finally, an economic crime early warning system is established based on the economic crime early warning model, and the excavated characteristic laws are used as the investigation The early warning rules of cases use feature matching methods to warn suspects and high-risk followers under investigation to provide auxiliary decision-making for the daily work of public security organs. At the same time, the reliability of the model is verified based on the sample data of a public security agency, and the efficiency of using this method to deal with economic crime cases is improved, which verifies that the system developed based on the economic crime early warning model has considerable practical application value.

**Keywords:** Economic crime, data mining, decision tree method, early warning model, auxiliary decision

### 1 Introduction

Economic crimes continue to develop with the emergence of economic activities. The economy is closely related to people's lives. While people enjoy the huge benefits brought by economic activities, criminals have motives to seek higher economic benefits. At present, the research of foreign economic crime mainly stays at the level of government economic management. It is believed that economic crime activities are closely related to society, and social conditions different different have characteristics of economic crimes, and with the improvement of social and economic levels, economic crimes The characteristics show a diversified trend, and the characteristics of various economic criminal activities can be analyzed through the study of economic management theory. Greenchet believes that the essence of organization is "a system that consciously coordinates the activities or forces of more than two people." In other words, human activity and effectiveness constitute the organizational system, and human activities directly affect the existence and development of the organizational system . The government organization system should organize and coordinate the management of economic activities of various units and departments, use effective analysis methods,

School of Law, Jilin University, Changchun, 130000, China \*Correspondence Email: 10144511@qq.com

Wang Hongning

clarify the characteristics of economic criminal activities in different situations, and severely crack down on economic crimes. Contingency Theory points out that there are internal and external factors in the analysis process of organizations, and these two factors are different, which together affect economic crimes. Economic crime activities are variability, complexity and multi-factor influence (02). Therefore, it is impossible to select effective management methods and case analysis methods, resulting in unclear characteristics of economic crimes, which is the main reason for the ineffective operation of management criminal activities, resulting in the continuous breeding of economic crimes, and the external environment is constantly changing and not fixed. Analyze the characteristics of economic crimes in a comprehensive and universal management method. The analysis of economic crime should combine the actual situation of economic activities and the actual situation of economic crime, appropriately coordinate units, and give play to the management role of each unit. Management should not be cumbersome, and new technical means should be used for analytical management. [1-15]

From a global perspective, the research on police work based on the dynamics of economic crimes has taken a relatively short time. Many countries are in the stage of continuous improvement in the field of economic crime early warning, and there is no clear, universal and applicable systematic method. Some developed countries represented by Europe and the United States have achieved a certain degree of research results in the field of economic crime dynamic early warning. Since the end of the 18th century, under the influence of the capital economic system, when dealing with and analyzing various and complex economic crime cases, police officers in Western countries such as Europe and the United States gradually developed the idea of economic crime early warning. They used the data and information of economic crime cases in the archive to integrate statistics. Data processing methods depended on the process of investigating economic crimes until the end of the early 19th century. From the early 19th century to the late 19th century, the police in Western countries discovered the convenience of the statistical analysis theory of mathematical statistics in the process of handling cases. They gradually started investigating the data and information involved in the case, and used statistical methods to combat various economic crimes. In the early 1990s, Western countries represented by the British Police Department focused on the analysis of the high socioeconomic crime rate and the low efficiency of police staff in handling cases in daily life. For the first time, they proposed the "intelligence-oriented police service ( ILP) management model", at the same time, the U.S. government has also proposed a prevention method for economic crime analysis based on case data and information using high-tech means. Since then, countries with stronger comprehensive strength represented by France, New Zealand and Germany have respectively formulated new police crime prevention models suitable for their own development. In the context of dynamic development of social crime, it is necessary to continuously collect and analyze economic crime information, and on the basis of effective analysis and control, interpret the possibility of economic crime in a specific location, and determine the location where crime may occur. Take effective control. At the same time, early warning technology emerged in the 1990s, gradually applied to the field of crime, and was recognized by various police agencies during the development process, becoming an analysis of crime, prevention of crime[16-20]

Necessary application of sin. At present, the methods of dealing with economic crimes still adopt methods such as on-the-spot investigation, investigation of relevant personnel, and questioning of experienced police officers, etc. to carry out analysis one by one, which is time-consuming and labor-intensive, and the effect is not good. Moreover, because different police officers have different levels of experience in detecting cases, the case detection cannot be unified, and there are differences in opinions. At the same time, with the increasing number of economic crime cases, it is difficult for the public security department to track suspects in a timely manner, and cannot proceed.[21-23]

Real-time and accurate analysis and early warning. Furthermore, there are many potential factors that affect the criminal motives of the people involved in economic crime cases. In the process of obtaining the effective influence and criminal characteristics of economic cases, it is found that they have strong concealment and ambiguity. In view of the above situation, this article proposes a data mining-based early warning method for economic crimes. The application of data mining technology to economic crimes can effectively find out the potential characteristics of economic crimes, and make coping strategies and necessary preparations for the

situation of the case. To the greatest extent, it helps the public security organs to quickly detect economic crimes, realize targeted investigations and prevent economic crimes. Therefore, in order to deal with the complex and changeable economic crime cases in the current society, the establishment of a sound economic case early warning system has far-reaching theoretical and practical significance for the detection of various types of economic cases.

# 2. Overall framework design of economic crime early warning model

### 2.1 Business analysis of economic crime cases

The subject of economic crime cases is people, and the prevention and stability control of the case should start from the human dimension. The detection process of the case is the process of identifying the suspects from the criminal suspects. The successful detection of economic crime cases can be regarded as a process. Bring the subjects involved in economic crime cases to justice. At the same time, timely and effective control of the targets involved in the case at the beginning of the case and tracking and grasping can effectively reduce the incidence of economic crimes. Therefore, the analysis of economic crime cases is mainly to analyze the characteristics of suspects involved in the case, to investigate suspects with strong characteristics of involvement in the case, and to track and control potential sources of crimes with characteristics of strong involvement. Through the analysis of the business of economic crime cases, three characteristics can be summarized.

1. Concealment. The persons involved in economic crime cases have different characteristics under different conditions. The staff should fully analyze the intensity of the case under various conditions, and look for persons involved in the current case and high-risk groups hidden in the case.

2. Diversified sources of knowledge. The analysis of various economic crime cases mainly relies on the subjective analysis of the staff. Different staff have different levels of understanding of the case, and different analysis results need to be combined to produce a scientific and effective analysis basis.

3. Measures to supplement education. After the economic crime case was filed, the staff launched investigations and took post-event remedial measures. The staff gradually looked for ways to stabilize and control high-risk sources at the beginning of the case in a timely and effective manner, reduce risks, and switch from post-remediation to pre-prevention.

## 2.2 Analysis of early warning applications of economic crimes

Based on the analysis of the business situation of economic crime cases, the early warning application analysis of economic crime cases can clarify the significance of data mining technology to the detection of economic crime cases. It mainly includes two parts: application feasibility analysis and technical advantages.

1. Application feasibility analysis

Data mining early warning technology is the process of using data for intelligent analysis, searching for the potential characteristics of the data, and using the rules to achieve the purpose of assisting decisionmaking. Generally, data mining early warning technology has the advantages of strong reliability, fast business processing speed and strong scalability. The practical application of data mining early warning technology in the field of economic crime early warning is mainly manifested in three aspects.

data processing: The data processed in the early warning process has the characteristics of large amount of data, long-term accumulation of historical data, and the diversity of data in the database; based on the data preprocessing technology in data mining, using different methods to solve data processing problems can improve data processing At the same time, it provides data that is convenient for pattern mining.

Pattern mining: According to the discrepancy law exhibited by data attribute characteristics, different data mining technologies are used to mutually restrict each other to generate rule patterns, and continuously improve the ability to solve actual cases in early warning work. According to different early warning goals, design pattern mining method frameworks; improve data While the accuracy of the mining results is ensured, the data mining method used is inherently scalable.

Pattern reasoning: The use of multi-mode crossover methods is mainly to improve the accuracy and practicability of mode reasoning in early warning work: Therefore, in the selection of reasoning mechanism, different technologies and methods are selected, and they are reasonably integrated, which is data mining early warning The application of technology is an important manifestation of the pattern reasoning process.

2. Advantages of early warning technology.

Through the application research of data mining technology in economic crime early warning, it can be found that the use of data mining early warning

Wang Hongning

technology has five advantages compared with traditional statistical analysis methods. The accuracy of early warning is high. For different types of early warning services, different mining methods can effectively improve the accuracy of early warning. In the process of pattern reasoning and matching, the method of data attribute association rules is used to save the field data and the integrity of the mining knowledge. Therefore, data mining technology can greatly improve the accuracy of early warning.

The early warning execution efficiency is high. Various technologies based on data mining can effectively increase the speed of data extraction and processing, and provide convenience for early warning. While using data mining technology, the fusion matching reasoning method speeds up the pattern search and calculation speed, which is conducive to improving the efficiency of early warning execution.

Strong extension ability. Using data mining technology to select reasonable mining modes and design the mining method framework of economic crime early warning modes, not only can realize the combination of various early warning technologies, but also make the pattern mining process diversified and expandable, and continuously meet various types of early warning.

Strong flexibility. Economic crime early warning is a data-based rule-based reasoning and mining process. The overall process of early warning is based on economic crime data. There is an independent business logic relationship between each component and data mining processing technology. Therefore, it is easy to change different technical methods and easy to operate. Save resource costs. Due to the low cost of data mining early warning technology, it can reduce method research cost and save resource space.

# 2.1.3 Establishment and description of objectives and tasks

Combining the research ideas of this article and the business needs of economic crime early warning, based on economic crime data information, business information as auxiliary support, combined with early warning algorithms in data mining technology, mining the potential characteristics of economic crime cases, and establishing economic crime early warning Model, early warning of economic crimes in investigating cases, focus on people and potentially high-risk groups, so as to provide the public security staff with the effect of analyzing economic crimes to assist decision-making, and improve the efficiency of economic crimes detection and prevention methods. On the basis of determining the above goals, using data mining ideas can divide the target task into three stages: preparation stage, feature mining stage and early warning stage. The preparation stage is mainly to preprocess and process the original economic crime data

Into a data format that satisfies feature mining. The feature mining stage is mainly to use data mining algorithms to dig out the characteristics of potential economic crime cases. The early warning stage is mainly to apply the economic crime early warning model to the detection process of actual economic crime cases, and use the characteristics of the case to match the suspects in the investigation case, and the high-risk stability control personnel who meet the characteristics of multiple cases. The specific structure is shown in Table 1.

Stage	Task	Description	
	Complete raw data on economic crime	Redundancy analysis of economic crime	
Preparation stage	Economic crime business data collection	Clear business crime knowledge	
r teparation stage	Data preprocessing         Economic crime data attribute extraction           Selection of early warning algorithm for         Analyze and compared to preprocessing	Economic crime data form conversion, effective attribute extraction	
Feature mining stage	Selection of early warning algorithm for data mining	Analyze and compare the advantages and disadvantages of multiple early warning algorithms for data mining	
	Selection of early warning algorithm	Select an algorithm that matches the data processing results and business knowledge	
	Establishment of Early Warning Model of Economic Crime	Use selected algorithms to build early warning models	
Early warning stage	The practical application of economic crime early warning model	Establishing an economic crime early warning system using an economic crime early warning	

Table 1 Target task structure table

# 2.1.4 Overall framework structure of economic crime early warning model

This paper establishes an economic crime early warning model based on data mining. The main ideas are: First, determine the economic crime early warning research goal; second, use data mining technology to preprocess the economic crime early warning data to convert the data into a pattern mining Data type, combining the results of data preprocessing with pre-business knowledge of

economic crimes, and selecting the most suitable data mining algorithm for economic crime early warning: Finally, a data mining early warning model is established to realize the mining of the characteristics of economic crimes. As can be seen from Figure 1, the economic crime early warning model mainly includes two parts: the selection of effective factors of economic crime and the design of economic crime early warning algorithm.



### Figure 1 Early warning model of economic crime

Aiming at the design of economic crime early warning model, two aspects are briefly summarized.

1. Screening of effective factors affecting economic crimes. First, carry out preliminary data processing of economic crime data and convert it into a form that can be exploited. Second, use feature screening methods to analyze the degree of influence of each involved attribute on various economic crime cases. Finally, complete the effective influencing factors of various economic crimes. filter.

2. The design of economic crime early warning algorithm. First, according to the preliminary data processing results and the screening results of effective factors affecting economic crimes, combined with a variety of data mining early warning algorithms, select the most suitable data mining early warning algorithms; secondly, establish an economic crime early warning mining model based on the early warning algorithms: Finally, The excavated characteristics of the case are stored in the pattern expression database, and the feature matching method is used to warn the suspects under investigation and the high-risk stability control personnel with multiple characteristics of the case.

# 2.2 Detailed frame design of economic crime early warning model

Based on the overall framework design of the economic crime early warning model, the detailed

framework of each part is outlined. It mainly includes two parts: the frame design of the effective factors of economic crime and the frame design of the economic crime early warning algorithm.

# 2.2.1 Screening framework for effective factors affecting economic crime

Obtaining effective factors affecting various economic crimes is the prerequisite for mining economic crime early warning rules. First, analyze the raw data of economic crimes and analyze the business data related to economic crime early warning. Second, conduct preliminary processing of economic crime data, using noise processing methods and data format conversion methods to convert economic crime data into usable mining Numerical data is a prerequisite for accurately mining the law of characteristics involved in economic crimes: Finally, as the economic crime data presents a state of categorization, the method for extracting features of categorized data is studied, and the maximum information coefficient method and chi-square test method are combined The method is to find out the degree of influence of each attribute of economic crime on the target economic crime case, screen the effective influencing factors of various economic crime cases, and eliminate irrelevant influencing factors. The frame design of effective factors affecting economic crimes is shown in Figure 2.



Figure 2 Screening framework for effective factors affecting economic crime

2020, Vol. XXIX, N°4, 621-633 REVISTA ARGENTINA **DE CLÍNICA PSICOLÓGICA** 

625

# 2.2.2 Economic crime early warning algorithm framework

On the basis of obtaining the effective influencing factors of various economic crime cases, the framework of economic crime early warning algorithm is designed. The selection of the algorithm is the core content of mining the characteristic law of economic crime. First, conduct research on economic crime early warning algorithms, analyze the characteristics of various data early warning mining algorithms, and select mining algorithms that are consistent with economic crime business and economic crime data processing results; second, use the selected data mining early warning algorithms to build the economy The crime early warning information mining model verifies the confidence of the early warning mining model: Finally, the excavated characteristics of economic crimes are stored. The structure diagram of the economic crime early warning algorithm framework is shown in Figure 3.



Figure 3 Structure diagram of economic crime early warning algorithm

# 3. Implementation method of economic crime early warning model

On the basis of completing the design of economic crime early warning framework, the realization method of economic crime early warning model is studied. It mainly includes two parts: the overall composition of economic crime early warning and the detailed composition of economic crime early warning.

# 3.1 Overall composition of economic crime early warning

The key to economic crime early warning is to use data mining technology to conduct in-depth analysis and detection of economic crime history data, summarize and refine the potential models for economic crime early warning data mining in a large amount of data, and describe the hidden relationship model between the data as much as possible. The actual detection of economic crime cases provides a strong theoretical support for economic crime early warning.

The overall economic crime early warning is based on historical data to dig out the characteristic law information available in the investigation of economic crimes. It mainly includes early warning information mining and early warning information matching to realize economic crime early warning. The overall composition of economic crime early warning is shown in Figure 4.



Figure 4 Overall composition of early warning

# 3.2 Detailed composition of economic crime warning

1. Early warning information mining The main components of early warning information mining include historical data, data preprocessing, pattern mining and early warning knowledge base. Through the understanding of historical data of economic crimes, the use of data

mining technology to realize the preprocessing of economic crime data, the processed data using data mining algorithms to mine potential rules involved in the case, and the use of rule notation to store the characteristics of the case in the early warning knowledge base.

2. Early warning information matching

Regarding the criminal suspects currently entered in the investigation case, it matches the characteristics of the cases involved with the use of data mining technology. If the characteristics of the criminal suspects in the current investigation cases and the characteristics of the mining characteristics are more consistent, an early warning will be triggered and a notification will be triggered. manager. Assuming that the early warning target is P, the different feature attribute data set of the early warning target is selected as L to form a matching pattern, so as to determine the feature matching situation as z, denoted as z=1, if at this time under the attributes of the suspect in the investigation of economic crimes The data and the pattern feature mining have strong consistency in the law of the attributes involved in the case, and the matching is

successful. Continue to match the attribute characteristics of Z=2 with the excavated characteristic rules. If the matching is successful, repeat the matching step until the data under the final attribute matches successfully, triggering an early warning. At the same time, the feature matching is repeated again to determine whether the characteristics of the suspects currently under investigation meet multiple early warning features, and trigger the focus on early warning. If in the process of feature matching, a certain attribute of the suspect does not match successfully, return to the target starting point and start the next round of matching process. According to the above-mentioned feature matching rules, an early warning plan is formed and applied to the actual work of economic crimes, which can effectively help police officers provide online analysis decision-making ideas, and timely deal with economic crime cases at work to detect potential problems and high-risk stability control problems . Figure 5 shows that the attributes of the suspects in the investigation case match with the unearthed regular features flow chart.



Figure 5 The flow chart of matching between investigation case data and mining features

# 4. Economic crime early warning mining model

The establishment of the model is the key to economic crime early warning. Based on the analysis of related early warning model methods, this article uses the decision tree method to establish an economic crime early warning model, and uses the C4.5 algorithm to generate readable characteristics rules and decision trees of the involved person, and use the generated The decision tree analyzes the characteristic data of economic crime suspects in investigative cases. The essence of model establishment is the process of categorizing the characteristic data of persons involved in economic crimes through a series of rules. It is the process of deriving the classification rules of the decision tree representation form from a set of unordered and irregular data. The process of establishment is adopted It is a recursive method, recursively from top to bottom, continuously compares the value of each attribute of economic crimes on the internal nodes of the decision tree, and judges the downward subbranch category of the node according to different economic crime attribute values, and finally The leaf nodes of the decision tree draw conclusions, and gradually improve and optimize the conclusions until the mining needs are met. It can be seen that the economic crime early warning model is mainly divided

#### Wang Hongning

into two stages: the generation of the early warning model and the optimization of the early warning model.

### 4.1 Generation of Early Warning Mining Model

The generation process of the early warning model is mainly the process of economic crime feature decision-making growth. The root node representing the entire economic crime feature data set is the starting node, and each given economic crime feature attribute is tested multiple times, and the data is divided by recursion. Become a smaller subset until the "purity" of the subset reaches expectations. At the same time, it can also be seen that when the purity of the sample on a certain tree branch is strong, it loses the meaning of continuing to group the sample. At this time, the division of the branch should be stopped immediately, when the sample collection of all economic crime characteristics At the same time, when the meaning of re-division is lost, the growth process of the decision tree with the characteristics of the person involved in economic crimes ends, and the establishment of the proper division standard is the core content of the decision tree to be solved in the growth process. The generation of the decision tree is shown in Figure 6.



# Figure 6 Decision tree generation diagram 4.2 Optimization of early warning mining model

In the process of finally generating the early warning decision tree model of economic crime characteristics, the phenomenon of decision tree overfitting often occurs. The reason for this phenomenon is that the process of generating the decision tree of the characteristics of economic crime involved is destroyed by the mixed values in the training sample set data. At this time, the decision tree rules obtained by the model cannot accurately classify the sample data. Therefore, in order to prevent the occurrence of decision tree over-fitting, decision tree pruning technology is needed. Prepruning technology and post-pruning technology are commonly used methods in the pruning process of decision trees.

The pre-pruning technology is the maximum growth of the predetermined feature decision tree. It can also be seen as the number of samples that can be accommodated by the predetermined branches of each economic crime attribute node to prevent the occurrence of some economic crime feature tree nodes. The sample size is too small The above two measures can solve the problem of too small sample size of some tree nodes. Therefore, before determining the appropriate decision tree depth or sample threshold, it is necessary to perform multiple adjustment checks on each parameter. The entire inspection process depends more on the user's level of understanding of the variable distribution law.

The post-pruning technology is to remove the scattered branches in the decision tree that have nothing to do with the characteristics of the economic crime involved in the classification effect after the mining feature decision tree completes the growth process. This is the pruning decision tree and the inspection decision tree. Steps that need to be implemented simultaneously. In the pruning process, the degree of influence of the current decision tree on the output value is calculated. Before building a decision tree, you need to set the maximum error rate range that the decision tree can accommodate. If it exceeds the error rate range that the decision tree can accommodate, pruning must stop immediately, if it continues to be lower than the error rate that the decision tree can accommodate The range should continue to be trimmed. The post-pruning technique is also based on the process of regular feature test sample set data. When the error rate of regular feature test set data becomes significantly larger, the current decision tree pruning needs to be ended.

The data mining method used in this paper is the C4.5 algorithm in the decision tree, and the pruning method that matches the C4.5 algorithm is the pessimistic pruning method. At the same time, the pessimistic pruning method is usually regarded as a method with higher accuracy in the post-decision tree pruning method. The purpose of this method is to improve the problems of the independent pruning

data set. It does not need to separate the pruning data set. Therefore, this article adopts the pessimistic pruning method to optimize the decision tree of the characteristics of the person involved in economic crimes, and it has the advantages of high accuracy and fast processing efficiency in the actual application process. In the pessimistic pruning method, in order to improve the accuracy of the classification of the characteristics of the involved person, the pruning method fully considers the influence of the error value to verify the continuity of the pruning. The basic principle of the pessimistic pruning method is that if e() < e(T) + S(e(T)) holds, then T should be pruned. As shown in formula 5.11, formula 5.12 and formula 5.13.

$$e'(t) = \left\lfloor e(t) + \frac{1}{2} \right\rfloor \quad (5.11)$$

$$e'(T_t) = \sum e(i) + \frac{N_t}{2} \quad (5.12)$$

$$S_e(e'(T_t)) = \frac{1}{\left[e'(T_t)\frac{n(t) - e'(T_t)}{n(t)}\right]^2} \quad (5.13)$$

In the above formula: e(t) is the error value existing at the decision tree node t; i is the branch node of the decision tree that occludes the subtree T: N is the number of nodes in the branch of the subtree T: n(t) is The number of training examples that exist at node t. The core idea of the pessimistic pruning method used in this article is to start with the root node of the generated characteristic decision tree, start searching for each node in the tree, calculate the relevant analysis value of all internal nodes in the decision tree, and focus on calculating each node The probability of being pruned, that is, the expected error rate value of each node, the error estimate in the pessimistic case is regarded as the upper limit of the confidence interval, and the confidence is set, denoted as c. The probability calculation is shown in formula 5.14.

$$c = P \left\lfloor \frac{f - q}{\sqrt{\frac{q(1 - q)}{n}}} \right\rfloor > u_{1-c} \quad (5.14)$$

In the above formula:Let N be the total value of the instance of the clipped subtree: E is the total value of instances where errors occur after the subtree is cut, P=EEN: f is the actual observed error value: let z=u1\_e, and the upper limit of the confidence interval is regarded as the expected error estimate of the current node. Through in-depth analysis of the above formula, the error rate calculation method under each node is shown in formula 5.15.

$$q = \frac{f + \frac{Z^2}{2N} + Z\sqrt{\frac{f}{N} - \frac{f^2}{N} + \frac{Z^2}{4N^2}}}{1 + \frac{Z^2}{N}} \quad (5.15)$$

In the above formula, q represents the error rate value estimated by the node. When q is positive, the value is the upper limit of confidence. The confidence value in the C4.5 algorithm is 0.25, at this time z=0.68. Let c be the highest value of the expected error rate. At this point, the node A is cut off. If the error rate q of the node after the cut is higher than the threshold c, all the subtrees under the decision tree node A of the economic crime involved feature decision tree are retained: if the node A is cut off, it will lead to After trimming, the error rate q of the node is lower than the threshold c, and the subtree under node A of the decision tree with the characteristics of economic crime involved is trimmed.

### 4.3 Confidence of pattern mining results

When the results of the early warning mode of a certain early warning type are obtained, the confidence test of the early warning mode is required to judge the reliability of the early warning model. Confidence is a measure of the accuracy of the results of pattern mining. It can reveal whether the state of involvement of various economic crime cases appears when the characteristics of the case appear. Set X as the characteristic of the case involved, and Y as the type of the case involved, and the confidence can be expressed as the proportion of cases of this type when X and Y occur simultaneously.

SUPPORT 
$$(X,Y) = \frac{P(X,Y)}{P(ALL)}$$
 (5.16)

In the above formula, SUPPORT (X, Y) represents the confidence level, and P (ALL) represents the total number of cases of the current type. When X and Y account for more than half of this type of cases, the model results are considered correct and credible.

# 4.4 Analysis of experimental results

The early warning model of economic crimes based on intellectual property rights is used as the early warning target. Since the establishment of the economic crime early warning model is based on the decision tree method, the analysis of the early warning model mainly includes the improvement of the C4.5 decision tree algorithm and the generation of early warning model has two parts.

### 4.4.1 Improvement of C4.5 Decision Tree Algorithm

The experimental data in this article uses 13 intellectual property economic crime data sets in the database for data analysis, and uses the intellectual property economic crime cases in the Saskatchewan region from January 2014 to January 2015 as the data

Wang Hongning

set to verify the algorithm But the result. The data set is divided into two categories, namely, training sample set and test sample set. 90% of the data is used as training sample and 10% is used as test sample. The improved C4.5 algorithm is named K-C4.5 algorithm. In this paper, under the same experimental data, the traditional C4.5 algorithm and the improved K-C4.5 algorithm are compared and analyzed, and the

improved algorithm is verified. The true validity of the algorithm. In the process of experimental analysis, the accuracy and time of the traditional C4.5 algorithm and the improved K-C4.5 algorithm were recorded. The algorithm efficiency statistics are shown in Table 2.

### Table 2 Algorithm efficiency statistics table

	C4.5 algorithm		K-C4.5 algorithm		
Data set	Accuracy (%)	Time (s)	Accuracy (%)	Time (s)	
AJ1	83.72	247	83.72	212	
AJ2	72.53	99	76.64	80	
AJ3	81.16	24	80.03	21	
AJ4	77.61	49	76.95	44	
AJ5	83.31	15	84.21	12	
AJ6	72.91	12	76.90	10	
AJ7	79.03	17	79.83	11	
AJ8	82.06	2354	82.07	2039	
AJ9	81.12	81	82.38	68	
AJ10	79.91	162	79.42	121	
AJ11	75.94	1526	75.51	1315	
AJ12	77.23	122	77.58	103	
AJ13	81.71	10	82.04	8	

In order to display the experimental results of the algorithm intuitively, the graphical representation is used to visually display some data test results, and the accuracy and execution speed of the original C4.5 algorithm and the improved K-C4.5 algorithm are compared and analyzed.

# 4.4.2 Generation of early warning model

Using the decision tree C4.5 algorithm in data mining, an early warning model can be constructed, with economic crime types as categories, and different decision trees established respectively, and finally determine whether the different characteristics of the case under different types of economic crimes meet the status of the case. Based on the extraction of effective features and influencing factors of similar economic crime cases, pattern mining is carried out. The generated description of some of the characteristic rules involved in the case is shown in Figure 7, and the excavated part of the rules involved in the case are displayed.



# Figure 7 Rules of characteristics involved in the case

Detailed description of the rule decision tree:Taking the root node NMSTHCS as an example,

this attribute is the number of unfamiliar calls per year; when the value is greater than 6443, it continuously

branches to the state of crime, and when the value of the attribute is less than 6443, it branches to the state of not satisfying the crime, but in the process of branching its child nodes There are also cases of branching to the criminal state: each time the attribute is split, the value with the higher information gain rate is selected as the current split attribute, which is represented by gini: samples represents the number of samples that the training data is assigned to the node: class represents the number of The node category status, kind represents the number of samples under the current class category.

Rule1: if (NMSRTHCS<6443) and (NMSDXL>1246) and (CARDNUMBER<4)

and( DEGREE>1.5 ) and ( YQDK> 25486 ) and (NDXZZTS<31 15 )an....and( SFCFZ<1.5)

then (state=yes);

Rule2: if (NMSRTHCS>6443) and (NMSDXL<1 246) and (NDXZZTS>3115) and

(DKED<100) and (DEGREE>1.5) a....a.n.. (YQDK<25486) then (state=YES);

Rule3: if(NMSRTHCS <6443 ) and(XB<1.5) and (AGE<18.7 )and( NMSDXL> 1246) .

and( CARDNUMBER>4 ) and (DKED>1 00) and (DEGREE>1 .5 )an....an.SFCF.>1.5)5

then (state=yes);

Rule4: if ( NMSRTHCS>6443) and ( NMSDXL>1246) and ( DEGREE>1.5) and

(SFCFZ<1.5) and (XB>1 .5) a..... and (CARDNUMBER<4) then (state=YES);

Rule5: if ( NMSRTHCS>6443) and (XB>1.5) and (AGE<40.76) and (DKED> 100)

and (DEGREE>2) an....a.nd.. (SFCFZ>1.5) then (state= yes) ....

The decision tree pruning is based on the confidence level of 0.25, then z=0.68, from which non-leaf nodes below XB<1.5 and AGE<18.76 are extracted for calculation and analysis. At this time, the total value of the error instances that occurred after the subtree is cut is 5, and the total value of the instances of the cut off is 50, so the actual error probability is 1110. According to the error rate calculation formula, q=0.1324, since q<0.25 at this time, this non-leaf node should be cut off for pruning operation. Extract the node XB>1.5, AGE<40.76 for analysis, the total value of the error instances that occurred after the subtree is cut is 10, and the total value of the instances of the cut off is 20, then the actual error probability is 112. According to the formula Get g=0.5752. Since g>0.25 at this time, this non-leaf node should be kept.

According to the characteristic decision tree of economic crime involved, the suspect characteristic matching of intellectual property economic crime cases is carried out. The characteristic matching information of some cases is shown in Table 3

Case name	Suspect	Early warning followers	Accuracy
Zhang intellectual property fraud	10	4	0.750
8.12 Patent fraud in Sa district	6	2	0.500
Property right Infringement in Sa district	5	3	0.667
9.2 Monograph fraud	10	4	0.600
Wang fraud case	7	3	0.556
6.3 Case of property rights encroachment in Sa district	3	0	0.000
Zhang property rights fraud	5	1	0.333
Zhao patent fraud	18	6	0.889
Zhang patent theft cases	7	4	0.667
10.11 Intellectual property cases	8	4	0.556

#### Table 3 Case feature matching table

Use the test data of the public security organs to analyze and mine the accuracy of the model, from the suspects involved in the early warning of the focus of attention, compare it with the actual case involved, and find out the number of the early warning focus of the person involved in the actual case. The accuracy of early warning. Among them, the number of people concerned in the early warning of the 6.3 property right infringement case in Saskatchewan is 0 because the suspect data entry is missing and cannot achieve

accurate feature matching, so the early warning is invalid.

### **5** Conclusion

According to the economic crime early warning model designed in this paper, an economic crime early warning system is established. Apply the researched economic crime early warning method to the daily case detection process of public security organs, use the economic crime early warning system to analyze

and verify the applicability of the model, and provide auxiliary decision-making effects for the daily work of public security organs. Through experimental analysis, it can be known that the early warning and prevention system developed based on the economic crime early warning model has good use value. At present, the research work of this subject has reached the expected goal, but there are still two problems to be optimized in the application of economic crime early warning model. -- In the process of mining effective factors of economic crime, there are a small number of economic crime data with complex types and high difficulty in mining that are not fully utilized, and further analysis of the data is needed. Second, due to the uncertain characteristics of the factors involved in economic crimes, the accuracy of early warning of economic crimes needs to be continuously tested, and a large amount of economic crime case data is needed to verify the accuracy of the early warning model.

### References

- Qin Qiaoqi, Zong Shuaibin. A preliminary discussion on the transformation of economic crime investigation in the era of big data [J]. Heilongjiang Province Political and Law Management Cadre College Journal, 2019(02):139-143.
- [2]Sun Jingjing. In the era of big data, the innovation of information statistics and investigation mechanism for public-related economic crimes [J] Observation, 201 9(01) 47-50.
- [3] Ouyang Aihui Thinking and Research on the Legal Guidance of Data Mining Technology in the Investigation of Chinese Cases [J]. Guangxi Socialist Science Journal of Yi College, 2015, 26(04): 102-106.
- [4] Xie Taichuan. Research and Application of Data Mining Technology in the Investigation of Daily Economic Crimes [J]. Information Network Security Full, 2012(12):36-38.
- [5] Kang Na. On the Light Penalty of Economic Crimes[D]. Shanxi University, 2010.
- [6] Cheng Xiaobai. On the basic work of economic crime investigation [J] Journal of Jiangxi Public Security College 2004(02): 5-8.
- [7] Agarwal,R. Home,M.Davies.M. Slevin,Dexin Ding.The Job Satisfaction of Economic Crime in the Police Directorates of Maribor and Some Reas ons for their Fluctuation[J]. Organizcija,2008,41(1).
- [8] Obiora C. Okafor, Benson Olugbuo. The Economic and Financial Crimes Commission and the

Accountability of Corrupt Foreign Actors in Nigeria[J]. The Law and Development Review,2012,4(3).

- [9] Anne Alves alo, Steve Tombs. The Emergence of a "War" on Economic Crime: The Case of Finland[J]. Business and Politics ,201 1,3(3).
- [10] Diana Ziegleder. Business and Self-regulation: Results from a Comparative Study on the Prevention of Economic Crime[J]. Zeitschrift fur Rechtssoziologie 2007,28(2).
- [11] Shipkova.Economics of crime as applied to regions (by the example of Siberian Federal Ditrict)[P]. Technology, 2004. KORUS 2004. Proceedings. The 8th Russian- Korean International Symposium on,2004.
- [12] Haitao Lv, Ruimin Hu, Bin Huang, Tao Lu, Jingjing Dai. Invetment decision model of crime prevention system based on expected monetary values of the economic loss caused byrisks[P].Computer Research and Development (ICCRD),2011 International Conference on,201 1.
- [13] Lan Yi,Nan Hu,Hongxiang Mu,Jing Sun,Jie Yin,Keren Dai, Fanghui Xu,Nanyang Yang,Dexin Ding. Identification of Coflin-1 and Destrin as Potential Early-warning Biomarkers for Gamma Radiation in Mouse Liver Tiss ues[J]. Health Physics,2019.
- [14] Peng Shaohui. The legal regulation of underground finance suspected of illegal fundraising [D] East China University of Political Science and Law, 2012.
- [15] Huang Rongkang's punishment of potential fraud crimes involving foreign financial credit [D]. People's Justice 2001 (06): 29-30.
- [16] Liang Yunpeng Research on Countermeasures of Economic Crime Investigation and Prevention [D]. Southwest University of Political Science and Law 2013.
- [17] Yue Mingqi. On the current situation, trends and countermeasures of economic crimes [J]. New Silk Road (Late), 2015 (08): 64-66.
- [18] Domingo Cartbonel, Review on sinkhole monitoring and performance trrestrial laser scanner in the salt karst of the Ebro Vally, Spain[J]. Engineering Geology, 2019, 248.
- [19] Shu-Rong Wang.Xiaom iao Chen, Shuang Ling ,Rong -zhen Ni,Huining Guo,Jin-Wen Xu. MicroRNA expres sion, targeting, release dynamics and early -waming biomarkers in rats[D]. Biomedicine & Pharmacotherapy,2019.111.
- [20] Yao Liun, Shengrong Lu. Erly-waming of socioeconomic-traffic resource environment

compound system carrying capacity in Yangtze River Economic Belt[J]. IOP Conference Series: Earth and Environmental Science ,2019,227(5).

- [21] Matthew Chinn, Mirko Wutkewicz. Artificial intelligence drives early-warning alerts to prevent ESP disruptions[J] World Oil, 2019.
- [2] Jian-fei ZHANG, Wen-sheng CUI, Ming LU, Shuo WANG. Research on Financial Risk Intelligent Early-warning Method for Colleges and Universities Based on Bayesian Classification and Delphi Method [P] DEStech Transactions on Computer Science and Engineering, 2018.
- [23] You-Chun Zhang, Zeng-Liang Liu, Jun Wei, Fang Ma. Early warning and tracking technology against large- scaleflow attack in information network for ITS[P].Machine Lerning and Cybernetics, 2009 Intermational Conference on,2009.

2020, Vol. XXIX, N°4, 621-633 REVISTA ARGENTINA **DE CLÍNICA PSICOLÓGICA** 

633