The Second International Conference "Problems of Cybernetics and Informatics"
September 10-12, 2008, Baku, Azerbaijan. Section #2 "Intellectual Systems"
www.pci2008.science.az/2/25.pdf

APPLICATION OF ANALYTICAL METHODS IN RISK MANAGEMENT

Uzeyir Gurbanli

ANAS Institute of Information Technologies, Baku, Azerbaijan, *uzeyir@sinam.net*

Risk management process, which is one of the main parts of management process, is characterized in different ways for different activity areas. In general, it is impossible to create a global risk management system which covers different activity areas. But for main parts of risk management process it may be possible to generalize models and methods for different areas. Analysis, classification, parsing and evaluation are the very first steps of risk management for making correct desicions and minimizing risks. For risk analysis and classification, huge amount of statistical information must be processed. In Risk analysis and evaluation of different activity areas statistical and mathematical methods are widely used in parsing of information for previously mentioned activity areas. It is known that, more statistical information must be collected to obtain more correct results. In turn, it results in collection, analysis, preservation and other problems. Thus, creation of information systems for risk management process (RMIS) becomes very important . In creation process of RMIS different methods are used for parsing of given data ;

- Probability-statistical methods
- Fuzzy logic theory methods
- Games theory methods
- Artificial intelligence methods (neural networks, genetical algorithms)
- Decision trees and etc.

There are speaking about the role of Data Mining technology which is containing probability-statistical and airtificial intelligence methods, and its application in risk management in article. In the article Data Mining technology, which includes probability-statistics and artificial intelligence and its role in risk management process is discussed. Data Mining is a technology that contains methods for complex analysis of huge given data. These methods contain statistical models, mathematical algorithms, machine learning algorithms (neural networks, decision trees, bayesian method). Data Mining is not only used in data collection and management, but also successfully applied in analysis and prediction problems. There are different applications of Data Mining in different management systems for management of risks. Additionally, there are some software packages that include mentioned algorithms and are used for different systems. "Oracle", "SAS", "Base Group" are some of the companies that produce these kinds of software packages.

The classification and evaluation of risk objects, determination how risky these are and prevention them are basic systematical stages in risk management process. It is possible to solve these questions using statistical information. But, when information amount is as much as needed, investigation and evaluation of risk objects takes too much time and precision degree becomes too low. And sometimes it is completely impossible. Data mining helps with these problems regardless how huge the information amount is. To make analysis and predictions Data Mining helps with the solution of some important problems that are described below, using some well known methods and algorithms [1, 2].

- Classification – systematically distribution of objects according to type, shape, sex or any other property. In a classification problem, you have a number of cases (examples) and wish to predict which of several classes each case belongs to. Each case consists of multiple attributes, each of which takes on one of several possible values. The attributes consist of multiple predictor attributes (independent variables) and one target attribute (dependent variable). Each of the target attribute's possible values is a class to be predicted on the basis of that case's predictor attribute values.
- Regression - defines dependencies between input and output parameters that belong to object set. Regression is the oldest and most well-known statistical technique that the

data mining community utilizes. Basically, regression takes a numerical dataset and develops a mathematical formula that fits the data. When you're ready to use the results to predict future behavior, you simply take your new data, plug it into the developed formula and you've got a prediction. The major limitation of this technique is that it only works well with continuous quantitative data (like weight, speed or age). If you're working with categorical data where order is not significant (like color, name or gender) you're better off choosing another technique.

- Clustering - grouping of objects according to their properties. Clustering is a technique useful for exploring data. It is particularly useful where there are many cases and no obvious natural groupings. Here, clustering data mining algorithms can be used to find whatever natural groupings may exist. Data modeling puts clustering in a historical perspective rooted in mathematics, statistics, and numerical analysis. From a machine learning perspective clusters correspond to hidden patterns, the search for clusters is unsupervised learning, and the resulting system represents a data concept. From a practical perspective clustering plays an outstanding role in data mining applications such as scientific data exploration, information retrieval and text mining, spatial database applications, Web analysis, CRM, risk management, marketing, medical diagnostics, computational biology, and many others.

To make risk analysis Data Mining classification methods can be used to find out objects with risk and to define relations between risk objects. As a classification method bayesian method, decision tree, artificial neural networks and others can be used.

Those risks that are against the government interests during the international business processes belong to real risks and are very important during the customs registration. During the customs registration decision tree method of Data Mining is very useful for finding out the risks. Such as, the information of past years for risk objects included by risk profiles increases the application of these methods and chance of correct prediction during evaluation of customs risks. Let's look at the creation of decision table for risk objects with the example below;

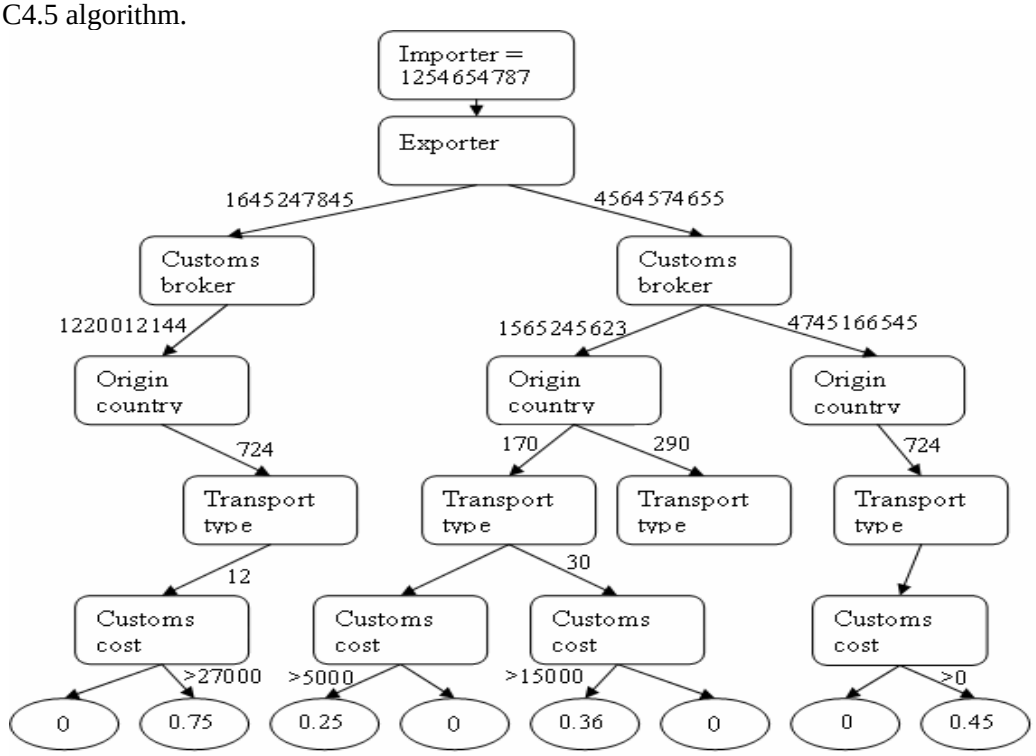
For example we take a class with 6 attributes. In this example each attribute defines an object that may have a risk during goods circulation process. Let's imagine a company which makes import and has been given a code of 1254654787 when it was first registered in the information system of customs office of goods customs declaration (GCD). Let's also assume that risk degree of that company can be defined with the table below (picture 1) according to the given attributes.

Importer	**Exporter**	**Customs Broker**	**Origin country**	**Transport type**	**Customs cost**	**Risk**
1254654787	1645247845	1220012144	724	12	>27000	0.75
1254654787	4564574655	1565245623	170	0	>5000	0.25
1254654787	4564574655	4745166545	724	0	>0	0.45
1254654787	1645247845	0	0	12,20,30	>3000	0.56
1254654787	4564574655	1565245623	290	30	>15000	0.36

Picture 1

As it can be seen from the table above, to estimate the risk of the importer company the mutual relation of mentioned company with some number of other objects and the result of this relation is defined. The decision tree for this table is illustrated in picture 2. The tree in the picture is composed of 6 levels. The statistical code of importer in customs declaration is verified by risk checking module of customs information system during customs process. The last step is the decision making step. After this step the risk degree that, object belongs to is defined. There are

several algorithms that are used in building process of decision tree: CART, C4.5, CHAID and others [2]. The decision that was built in the example above was built using the CART method. This method is used in construction of binary decision trees. In binary tree, points that are in different levels are separated into two branches. The first (left one) branch shows that the rule is applied and the second (right) branch shows that the rule is not applied. The tree in this example can also be constructed using the C4.5 algorithm. C4.5 method differs from CART method in that, each point can be separated to any number of points and branches. Each branch shows that some rules are applied. It must be kept in mind that CART algorithm works much faster than the C4.5 algorithm.



Picture 2

As a result, to make analysis, predictions and to make decisions for some activity areas which, has huge amounts of information, decision trees can be very effective for experts and analytics. This method differs from other methods with the simple structure of its classification models, with the precision degree of its predictions and etc. During the identification and evaluation of risk objects for discovery of situations with risk other methods of Data Mining such as using neural networks and Bayesian networks can also be used. Each of these methods, differ from others according to its properties. For example decision tree method is not as precision, as the neural networks, but it is faster. Decision tree method can successfully be applied to risk management process in bank business, production, medicine and some other areas.

Literature

1. K. Cios, W. Pedrycz, R. Swiniarski. Data Mining, Methods for Knowledge Discovery, 2000, 490 s.
2. M. Bramer. Principles of Data Mining(In English), 2007, 340 s.
3. Australian/ New Zealand Risk Management Standards AS/NZS 4360:2004.