ISSN:0975-9646

Association Rule Mining for Accident Record Data in Mines

Amber Hayat[1], Khustar Ansari[2], Praveen[3]

[1]*Assistant Professor, Department of Computer Engineering, Padmabhushan Vasantdada Patil Pratishthan's College of Engineering, Sion Mumbai, India*

[2]*Assistant Professor, Department of Computer Science and Engineering, Guru Gobind Singh Educational Society's Technical Campus, Bokaro Steel City, India*

[3]*Manager operation, Q-AB Tata Steel West Bokaro Division, India*

Abstract— **Association Rule Mining is one of the important and well accepted application areas in the field of Data Mining where rules are found between the data items which helps to determine the relationships between the data items. Thus helps to make the prediction for future in a better way. This study examined the group differences in accident susceptibility among underground coal mine workers accounting for different factors associated with them. Data was analysed with a bivariate modelling technique in the initial phase. Knowledge discovery in Database model- A Data mining technique was used for this purpose to take multiple parameters into account simultaneously. Data were collected from eight different underground coal mines for a period of ten years. The case study results revealed that different age of workers bear no significant differences in their accident susceptibility; however, the type of mines and designation show significant differences in their risk of injuries. It is inferred based on the Data Mining technique that among the various reasons of accidents, accident by haulage and conveyor, fall of person and fall of object possesses more risk than the traditional Roof and Side falls. There are other major conclusions which have been deduced from the work Later on various rules were also predicted from the dataset with the required rule accuracy and rule coverage enabling us to visualize and predict the general trend.**

Keywords— **Data Mining, Association Rule, Information Data Analyzer(ida), Support, Confidence, ESX(data mining tool).**

I. INTRODUCTION

The problem of finding association rules can be stated as follows[1] given a database of sales transactions, it is desirable to discover the important associations among items such that the presence of some items in a transaction will imply the presence of other items in the same transaction. An example of an association rule is: 30% of transactions that contain beer also contain diapers; 2% of all transactions contain both of these items. Here 30% is called the confidence of the rule, and 2% the support of the rule. The problem is to find all association rules that satisfy user specified minimum support and minimum confidence constraints [6]. The problem of mining association rules was first introduced in and the following formal dentition was proposed into[7] address the problem.

A. *Definition: Association Rules*

Let I ={I1, I2, … , Im} be a set of m distinct attributes, also called literals. Let D be a database, where each record (tuple) T has a unique identifier, and contains a set of items such that $T \subseteq I$. An association rule is an implication of the form $X \Rightarrow Y$, where X, $Y \subset I$, are sets of items called item sets, and X intersection $Y=\varphi$. Here, X is called antecedent, and Y consequent The problem of discovering all association rules can be decomposed into two sub problems[4]: a- Find all sets of items (item sets) that have transaction support above the minimum support. These are the frequent item sets. Other item set called infrequent item sets. b- Use the frequent item sets to generate the desired rules. There is a wide agreement among the literature that the first sub problem is the most important of the two. This is because it is more time consuming due to the huge search space (the power set of the set of all items) and the rule generation phase can be done in main memory in a straightforward manner once the frequent item sets are found[7]. That is the reason for the great attention researchers paid to this problem in the recent years. Another reason for such attention may be due to the dispute among the research community about the confidence as a measure of 23 rule importance or interestingness. Contrary to support, confidence received much criticism and many alternatives have been suggested with no global acceptance of one measure. [8,9].

Two important measures for association rules support (s) and confidence (α), can be defined as follows.

Support: The support (s) of an association rule is the ratio (in percent) of the records that contain $X \cap Y$(X and Y belong to item sets) to the total number of records in the database. Therefore, if we say that the support of a rule is 5% then it means that 5% of the total records contain $X \cap Y$. Support is the statistical significance of an association rule. Grocery store managers probably would not be concerned about how peanut butter and bread are related if less than 5% of store transactions have this combination of purchases. While a high support is often desirable for association rules, this is not always the case. For example, if we were using association rules to predict the failure of telecommunications switching nodes based on what set of events occur prior to failure, even if these events do not occur very frequently. Association rules showing this relationship would still be important.

$$\text{Support (s)} = \frac{x \cap y}{Total\ no.of\ records}$$

Where X and Y belong to item set, containing all items.

Confidence: For a given number of records, confidence (α)is the ratio (in percent) of the number of records that contain X∩Y to the number of records that contain X. where X,Y belong to item set containing all items.

$$\text{Confidence } (\alpha) = \frac{No.of\ records\ that\ contain x \cap y}{No.of\ records\ that\ contain\ x}$$

Where X and Y belong to item set, containing all items.

Thus, if we say that a rule has a confidence of 85%, it means that 85% of the records containing X also contain Y. The confidence of a rule indicates the degree of correlation in the 24 data set between X and Y. Confidence is a measure of a rules strength. Often a large confidence is required for association rules. If a set of events occur a small percentage of the time before a switch failure or if a product is purchased only very rarely with peanut butter, these relationships may not be of much use for management.

Mining of association rules from a database consists of finding all rules that meet the user specified threshold support and confidence.

B. Benefits and Applications

The most famous application of association rules is its use for market basket analysis [5]. Consider a supermarket setting where the database records items purchased by a customer at a single time as a transaction. The planning department may be interested in finding "associations" between sets of items with some minimum specified confidence. Such associations might be helpful in designing promotions and discounts or shelf organization and store layout.

However, association rules have many other fields in which it have been helpful. In [10], two successful examples for the application of association rules in the telecommunications and medical fields for performing partial classification is reported. Association rule mining has been also used on other types of data sets. It has been used to mine web servers log files to discover the patterns that access different resources consistently and occur together or the access of a particular place occurring at regular times. Other types of data include census data and text documents as in [17] for example. Other examples of applications of association rules include catalog design, customer segmentation based on buying patterns, fraudulent discovery and health insurance.

II. PROPOSED WORK

Indian Mining Industry makes a major contribution to the national economy and to the well-being of the society as a whole. The country's mineral industry represents a unique mix of very small to medium to large mines. Currently there are more than 600 coal mines and a large number of non-coal mines with more than 6000 being in record under The Mines Act (however the figure is estimated to be much larger). For the continuing viability and stability of the industry, it is important that full advantage be taken of developments in mining methods and procedures, modern machinery and equipment and advances in approaches to management of all mining activities, including health and safety.

While safety in any sphere of activities is important, it has special significance when the risk is greater. Unlike the major industries, mining has high potential risk of accidents. It has a dubious distinction of involving a very high actual hazard as the environment changes continually with the progress of work. It is, therefore, not possible for any external agency to ensure safety of any mine. The principal responsibility for the safety and good health of workers employed in mines rests with the management of that mine.

Considering the accident scenario in Indian Mining Industry, it has now become essential that risk assessment be undertaken of all hazardous operations, equipment and machinery, taking account of the procedures used, maintenance, supervision and management. Introduction of the risk management as a tool for the development of good health and safety management system is a breakthrough in the traditional strategy. The system is sure to be an effective tool for the improvement of the health and safety scenario in our mining industry. The risk assessment process will identify all the existing and probable hazards in the work environment and in all operations, assess the risk levels of those hazards in the order to priorities which needs immediate attention for redresser, where maintenance of on-going management will be sufficient and which are of very mild nature. Then for managing these risks, different Mechanisms (underlying causes) responsible for these hazards are identified and their control measures, set to timetable, are recorded pinpointing the responsibilities.

A. Collection of Data

The accident record data is collected from 10 mines randomly chosen from Eastern Coalfields Limited (ECL).The names of the mines are as follows:

(i) Bansra colliery(Kunstoria Area)
(ii) Chora Block Incline (Kenda Area)
(iii) Kumardhubi Colliery
(iv) Parsera 6&7 Incline
(v) Parbelia Colliery (Sodepur Area)
(vi) RatiBati Colliery (Satgram Area)
(vii) Sangramgarh Colliery
(viii) Mandmaan Colliery
(ix) Shyampur B Colliery
(x) Shyamsunderpur Colliery

The attributes of the accident record data of mines is as given below:

Table 1: Attributes of accident record data of mines

Sl.No.	Attributes
1.	Name of the mine/ area
2.	Type of the mine(mech/ semi-mech/ manual)
3.	Date of accident
4.	Name of the person
5.	Designation/ category
6.	Age(years)
7.	Type of injury(Fatal/ Serious/ Reportable)
8.	Place of accident
9.	Shift/ time of accident
10.	Period of absence/ loss of time
11.	Cause of accident
12.	Manpower/ Employment
13.	Production(TPD/ Annual)
14.	Source of information

B. Description of Data

The important terminologies involved in accident record database are as described below:

(i) **Name of Mine/Area:** It signifies the name of the mine or the area from where the data has been collected. i.e. Bansra colliery(Kunstoria Area),Chora Block Incline (Kenda Area), Kumardhubi Colliery, Parsera 6&7 Incline , Parbelia Colliery (Sodepur Area), RatiBati Colliery (Satgram Area), Sangramgarh Colliery, Mandmaan Colliery, Shyampur B Colliery, Shyamsunderpur Colliery.

(ii) Type of the mine: Mines could be of three types mechanical, semi mechanical and manual.

(iii) Date of accident: Date of accident signifies the date on which the accident took place.

(iv) Name of the person: It signifies the name of the person working in the mine.

(v) Designation/ category: It signifies the designation of the person working in mine. Designation of a person could be any one among the Loader, Mech Fitter, Trammer, Timberman, Fitter, Sirdar, Khalasi, Pump Khalasi, Mazdoor, Carrier,TimberMistry, Timber Helper, General Manager, Driller, Pit Clerk, Dumper Khalasi, Clipman, Repair Mistry, Haulage Operator,Electrician,Linemanetc

(vi) Age: Age defines the age of a person in years.

(vii) Type of Injury: Injuries are further divided in three types, namely Serious Bodily Injury, Reportable Injury, Minor Injury. They are described as follows:

a. **Serious Bodily Injury:** Serious Bodily Injury means any injury which involves the permanent loss of any part or section of the body or the permanent loss of sight or hearing.

b. **Reportable Injury:** Reportable Injury means any injury other than any serious bodily injury, which involves the enforced absence of injured person from work for a period of 72 hours or more.

c. **Minor Injury:** Minor Injury means any injury which results in enforced absence from work of the person exceeding 24hrs and less than 72 hours.

(viii) Period of Absence: It describes the number of days the person is absent.

C. Pre-Processing Of the Data

Information data analyzer (ida) software used contains a pre-processor. Before the data in a file is presented to ESX, the file is scanned for several types of errors, including illegal numeric values. Blank lines and missing terms. The pre-processor consists of several types of errors but doesn't attempt to fix the numerical data errors. The pre-processor outputs a data file ready for data mining as well as a document informing us about the nature and location of unsolved errors, if any. The data was available with 2 types of errors

- Duplicate record
- Missing values

Duplicate records were corrected by analyzing the summary of the domain statistics for categorical attribute and accordingly the main data base was modified. Missing values were represented by the open brackets by the pre-processor. Missing data items present a problem that can be dealt in several ways as discarding the records with the missing values, replacing the missing values with the class mean or replacing the missing attribute with the values found within other highly similar instances. The missing values here were corrected by the last technique of replacing the missing attribute with values found within other highly similar instances (Typicality).

In this accident record database our objective is to do the following

(i) Statistically analyze the data obtained from the mines.

(ii) Cross-tabulate the data and correlate the data amongst various parameters of accidents.

(iii) Rank the hazards based on the confidence and support of the rules generated.

(iv) Development of control measures for the prevention of the accidents and prioritize the resources for the hazard that are posing higher risk.

D. Accident In Mines And Their Analysis

After thorough analysis of the accident report of all the nine collieries.

(i) Bansra colliery(Kunstoria Area)
(ii) Chora Block Incline (Kenda Area)
(iii) Kumardhubi Colliery
(iv) Parsera 6&7 Incline
(v) Parbelia Colliery (Sodepur Area)
(vi) RatiBati Colliery (Satgram Area)
(vii) Sangramgarh Colliery
(viii) Mandmaan Colliery
(ix) Shyampur B Colliery

taken for the year 1996 to 2005 (10 years) on various parameters such as age, designation, type of accident, reason of accident, period of absence, year and shift many inferences can be easily concluded from the following simplified tables, charts and graphs.

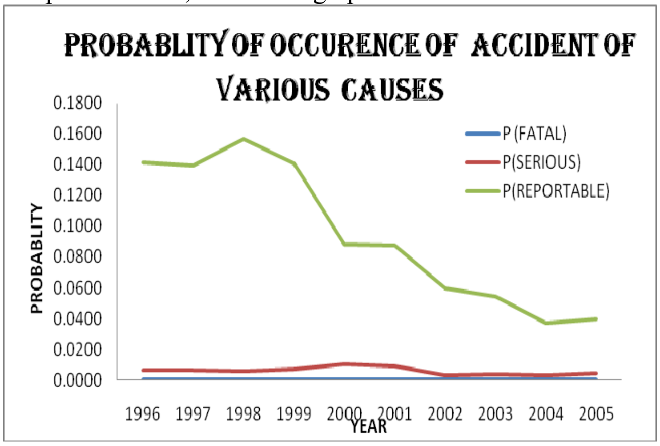

Figure 1: Probability of occurrence of accident of various causes

We observe that the probabilities of occurrence of reportable injuries are much higher than the serious injuries and further followed by the fatalities. This goes according to our expectation. Here as we see that the reportable injuries are having very high probability of occurrence then they will be increasing the risk (probability * consequence).All three P(fatal), P(serious) &P(reportable) are with decreasing trend as we go from 1996 to 2005 owing to the awareness about Risk Management. We can see from the charts below the accidents according to its type and percentage of occurrence in different years

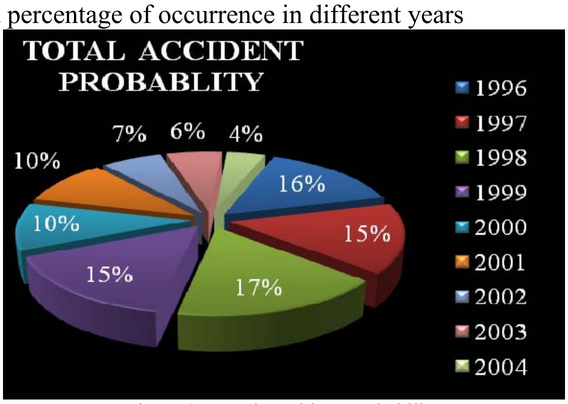

Figure 2: Total Accident Probability

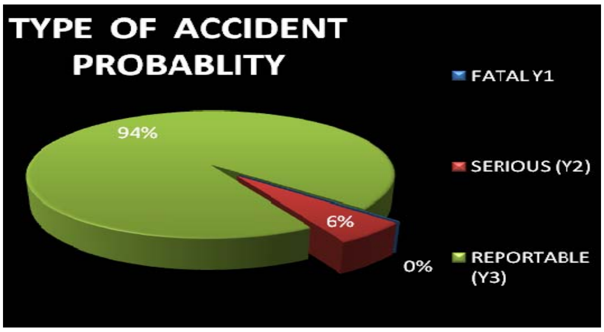

Figure 3: Type of accident probability

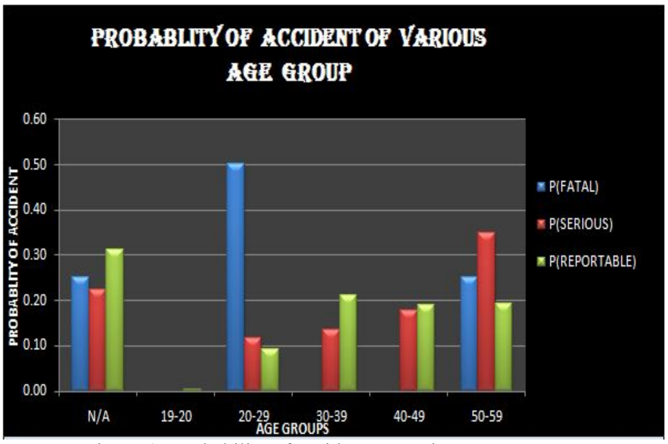

Figure 4: Probability of accidents to various age groups

The above graph of the probability of accidents to various age groups highlights an important point that lower age group are having higher rate of reportable injuries and the higher age group is having higher percentage of reportable injuries.

E. Comparing Reason of Accidents in Mines

Not many inferences can be drawn from the fatalities owing to very less frequency in the overall data.

Fatalities:-

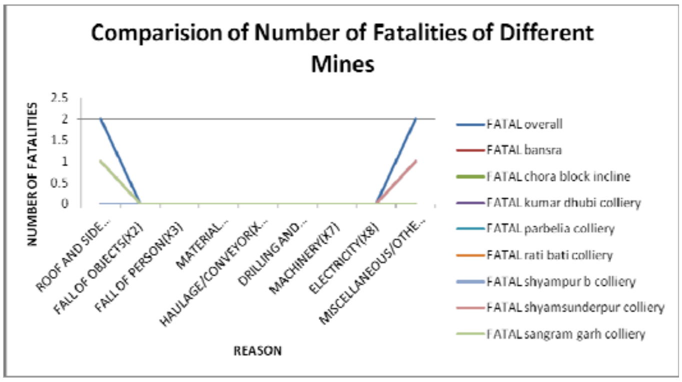

Figure 5: Comparison of no. of fatalities of diff. mines

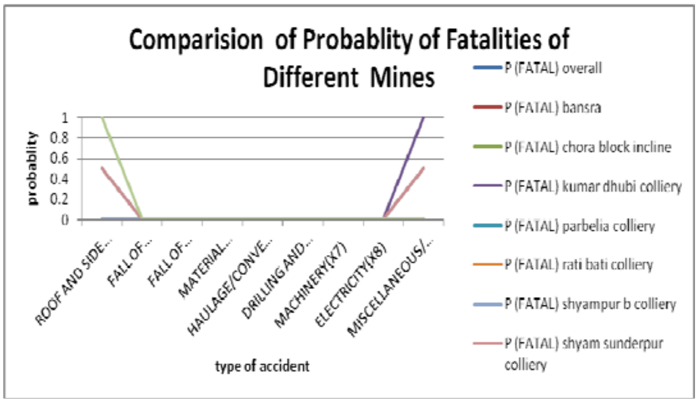

Figure 6: Comparisons of probability of fatalities of diff. mines

(i) Serious Accidents:-

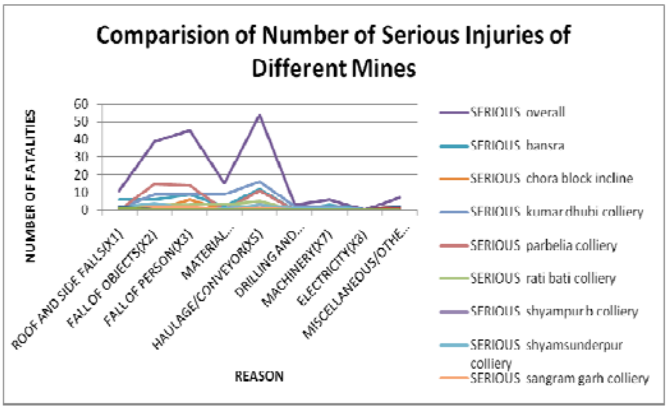

Figure 7: Comparisons of No. Of serious Injuries of diff. mines

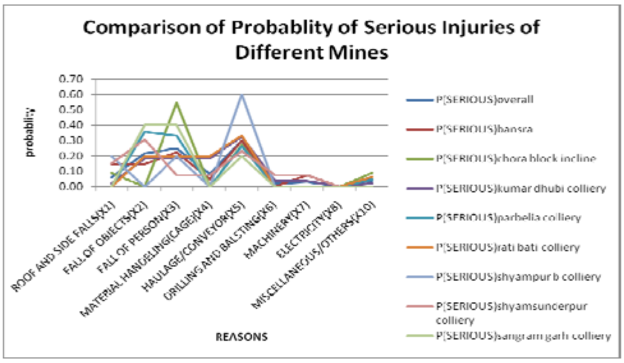

Figure 8: Comparisons of Probability Of serious Injuries of diff. mines

Chart in previous page clearly shows that nearly all the mines follows same trend with highest number if serious accidents in case of fall of object followed by fall of person and then haulage and conveyors and material handling. Rest all the reasons are negligible.

(ii) Reportable Accidents:-

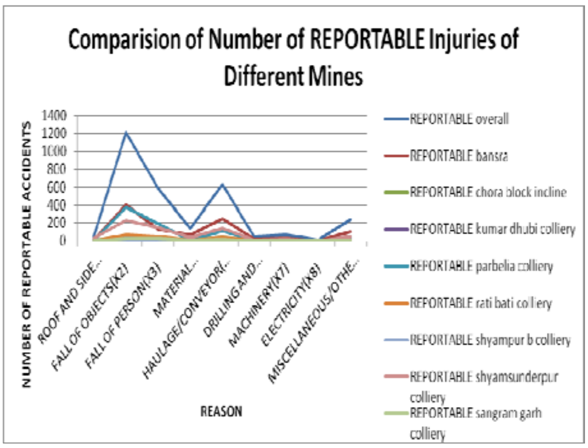

Figure 9: Comparison of No. Of Reportable Injuries of Diff. Mines

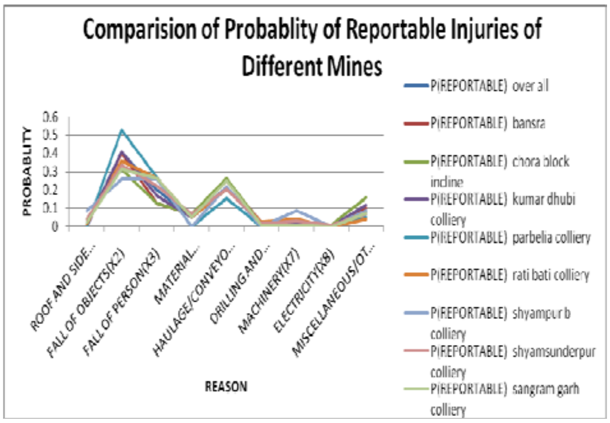

Figure 10: Comparison of Probability Of Reportable Injuries of Diff. Mines

The graphs shows that the reportable accidents also shows a general trend with highest number of injuries in case of fall of objects category, followed by fall of person and haulage and conveyors and then by material handling .Accidents due to rest all other categories are not of very much concern.

It is to be noted here that both the serious accidents and the reportable accidents follows the same trend of the reasons and hence this makes the greater concern about these causes. This also increases the risk due to these reasons i.e.

Below diagram shows the common decreasing trend of causes of both serious and reportable injuries

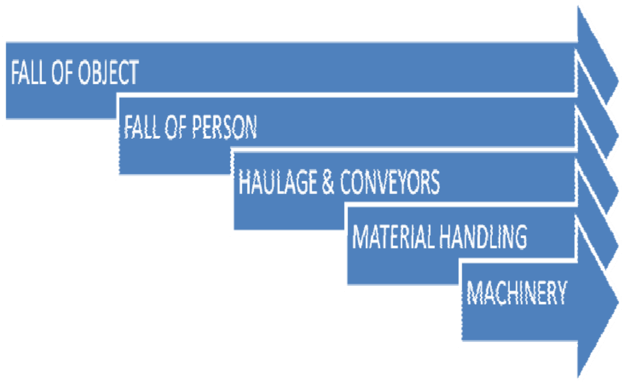

Figure 11: Decreasing trend of cause of both serious and reportable injuries

III. RESULTS

The different association rules which were found on the data described are as follows

OUTPUT:-

A. Generating Rules:

****

Rules for Class Reportable(R)

2972 instances

****

**Total Percent Coverage = 100.00%

RULE	RULE ACCURACY	RULE COVERAGE
Name Of Mines = Bansra Colliery	96.16%	33.71%
Name Of Mines = Parbellia Colliery	94.47%	24.16%
Name Of Mines = Shyamsunderpur Colliery	97.81%	22.54%
Type Of Mines = Mechanised	94.27%	93.51%
Designation = Loader	96.12%	52.46%
Designation = Trammer	91.67%	10.73%
10 <= Age <= 60	94.32%	100.00%
Shift = 1	93.33%	44.25%
Shift = 2	94.96%	27.89%
Shift = 3	94.74%	27.86%
Reason = Fall Of Object	96.87%	40.65%
Reason = Haulage And Conveyor	92.12%	21.23%
Reason = Fall Of Person	92.91%	19.85%
Name Of Mines = Shyamsunderpur Colliery And Designation = Loader	98.10%	12.52%
Name Of Mines = Shyamsunderpur Colliery And 10 <= Age <= 60	97.81%	22.54%
Name Of Mines = Shyamsunderpur Colliery And Type Of Mines = Mechanised	97.81%	22.54%
Reason = Fall Of Object And Name Of Mines = Bansra Colliery	98.53%	13.49%
Reason = Fall Of Object And Designation = Loader	97.58%	28.47%
Reason = Fall Of Object And Shift = 2	97.13%	11.37%
Reason = Fall Of Object And Shift = 3	97.01%	12.01%
Reason = Fall Of Object And Name Of Mines = Parbellia Colliery	96.19%	12.75%
Reason = Fall Of Object And 10 <= Age <= 60	97.03%	40.65%
Reason = Fall Of Object And Type Of Mines = Mechanised	96.94%	38.32%
Reason = Fall Of Object And Shift = 1	96.61%	17.26%
Name Of Mines = Bansra Colliery And Designation = Loader	97.48%	16.89%
Name Of Mines = Bansra Colliery And Shift = 3	97.76%	10.30%
Name Of Mines = Bansra Colliery And 10 <= Age <= 60	96.62%	33.71%
Name Of Mines = Bansra Colliery And Type Of Mines = Mechanised	96.16%	33.71%
Name Of Mines = Bansra Colliery And Shift = 1	95.06%	10.60%
Designation = Loader And Shift = 2	96.77%	16.10%
Designation = Loader And Shift = 3	96.14%	10.92%
Designation = Loader And Name Of Mines = Parbellia Colliery	96.45%	10.60%
Designation = Loader And 10 <= Age <= 60	96.29%	52.46%
Designation = Loader And Type Of Mines = Mechanised	96.10%	52.29%
Designation = Loader And Shift = 1	95.58%	20.39%
Designation = Loader And Reason = Fall Of Person	96.25%	11.24%
Shift = 2 And 10 <= Age <= 60	94.96%	27.89%
Shift = 2 And Type Of Mines = Mechanised	95.25%	25.64%
Shift = 3 And 10 <= Age <= 60	94.74%	27.86%
Shift = 3 And Type Of Mines = Mechanised	94.88%	26.21%
Name Of Mines = Parbellia Colliery And 10 <= Age <= 60	94.47%	24.16%
Name Of Mines = Parbellia Colliery And Type Of Mines = Mechanised	94.47%	24.16%
Name Of Mines = Parbellia Colliery And Shift = 1	94.28%	10.53%
10 <= Age <= 60 And Type Of Mines = Mechanised	94.43%	93.51%
10 <= Age <= 60 And Shift = 1	93.66%	44.25%
10 <= Age <= 60 And Reason = Fall Of Person	92.91%	19.85%
10 <= Age <= 60 And Reason = Haulage And Conveyor	92.52%	21.23%
10 <= Age <= 60 And Designation = Trammer	92.20%	10.73%
Type Of Mines = Mechanised And Shift = 1	93.29%	41.66%
Type Of Mines = Mechanised And Reason = Fall Of Person	92.77%	18.14%
Type Of Mines = Mechanised And Reason = Haulage And Conveyor	92.36%	19.92%
Type Of Mines = Mechanised And Designation = Trammer	91.79%	10.16%
Name Of Mines = Shyamsunderpur Colliery And Designation = Loader And 10 <= Age <= 60	98.10%	12.52%
Name Of Mines = Shyamsunderpur Colliery And Designation = Loader And Type Of Mines = Mechanised	98.10%	12.52%
Name Of Mines = Shyamsunderpur Colliery And 10 <= Age <= 60 And Type Of Mines = Mechanised	97.81%	22.54%
Reason = Fall Of Object And Name Of Mines = Bansra Colliery And Designation = Loader	98.35%	10.03%
Reason = Fall Of Object And Name Of Mines = Bansra Colliery And 10 <= Age <= 60	99.01%	13.49%
Reason = Fall Of Object And Name Of Mines = Bansra Colliery And Type Of Mines = Mechanised	98.53%	13.49%
Reason = Fall Of Object And Designation = Loader And 10 <= Age <= 60	97.80%	28.47%
Reason = Fall Of Object And Designation = Loader And Type Of Mines = Mechanised	97.57%	28.33%
Reason = Fall Of Object And Designation = Loader And Shift = 1	97.29%	10.87%
Reason = Fall Of Object And Shift = 2 And 10 <= Age <= 60	97.13%	11.37%
Reason = Fall Of Object And Shift = 2 And Type Of Mines = Mechanised	97.21%	10.57%
Reason = Fall Of Object And Shift = 3 And 10 <= Age <= 60	97.01%	12.01%
Reason = Fall Of Object And Shift = 3 And Type Of Mines = Mechanised	97.13%	11.41%
Reason = Fall Of Object And Name Of Mines = Parbellia Colliery And 10 <= Age <= 60	96.19%	12.75%
Reason = Fall Of Object And Name Of Mines = Parbellia Colliery And Type Of Mines = Mechanised	96.19%	12.75%
Reason = Fall Of Object And 10 <= Age <= 60 And Type Of Mines = Mechanised	97.10%	38.32%
Reason = Fall Of Object And 10 <= Age <= 60 And Shift = 1	96.98%	17.26%

RULE	RULE ACCURACY	RULE COVERAGE
Reason = Fall Of Object And Type Of Mines = Mechanised And Shift = 1	96.62%	16.35%
Name Of Mines = Bansra Colliery And Designation = Loader And 10 <= Age <= 60	98.05%	16.89%
Name Of Mines = Bansra Colliery And Designation = Loader And Type Of Mines = Mechanised	97.48%	16.89%
Name Of Mines = Bansra Colliery And Shift = 3 And 10 <= Age <= 60	97.76%	10.30%
Name Of Mines = Bansra Colliery And Shift = 3 And Type Of Mines = Mechanised	97.76%	10.30%
Name Of Mines = Bansra Colliery And 10 <= Age <= 60 And Type Of Mines = Mechanised	96.62%	33.71%
Name Of Mines = Bansra Colliery And 10 <= Age <= 60 And Shift = 1	96.05%	10.60%
Name Of Mines = Bansra Colliery And Type Of Mines = Mechanised And Shift = 1	95.06%	10.60%
Designation = Loader And Shift = 2 And 10 <= Age <= 60	96.77%	16.10%
Designation = Loader And Shift = 2 And Type Of Mines = Mechanised	96.76%	16.08%
Designation = Loader And Shift = 3 And 10 <= Age <= 60	96.14%	10.92%
Designation = Loader And Shift = 3 And Type Of Mines = Mechanised	96.12%	10.85%
Designation = Loader And Name Of Mines = Parbellia Colliery And 10 <= Age <= 60	96.45%	10.60%
Designation = Loader And Name Of Mines = Parbellia Colliery And Type Of Mines = Mechanised	96.45%	10.60%
Designation = Loader And 10 <= Age <= 60 And Type Of Mines = Mechanised	96.28%	52.29%
Designation = Loader And 10 <= Age <= 60 And Shift = 1	96.04%	20.39%
Designation = Loader And 10 <= Age <= 60 And Reason = Fall Of Person	96.25%	11.24%
Designation = Loader And Type Of Mines = Mechanised And Shift = 1	95.58%	20.36%
Designation = Loader And Type Of Mines = Mechanised And Reason = Fall Of Person	96.24%	11.20%
Shift = 2 And 10 <= Age <= 60 And Type Of Mines = Mechanised	95.25%	25.64%
Shift = 3 And 10 <= Age <= 60 And Type Of Mines = Mechanised	94.88%	26.21%
Name Of Mines = Parbellia Colliery And 10 <= Age <= 60 And Type Of Mines = Mechanised	94.47%	24.16%
Name Of Mines = Parbellia Colliery And 10 <= Age <= 60 And Shift = 1	94.28%	10.53%
Name Of Mines = Parbellia Colliery And Type Of Mines = Mechanised And Shift = 1	94.28%	10.53%
10 <= Age <= 60 And Type Of Mines = Mechanised And Shift = 1	93.65%	41.66%
10 <= Age <= 60 And Type Of Mines = Mechanised And Reason = Fall Of Person	92.77%	18.14%
10 <= Age <= 60 And Type Of Mines = Mechanised And Reason = Haulage And Conveyor	92.79%	19.92%
10 <= Age <= 60 And Type Of Mines = Mechanised And Designation = Trammer	92.35%	10.16%

****

Rules for Class Serious(S)

180 instances

****

**Total Percent Coverage = 0.00%

****

Rules for Class Fatal(F)

4 instances

****

**Total Percent Coverage = 0.00%

Rules which are having high value of rule accuracy and coverage can be used for prediction purposes of the type of accident. From the last rule we can predict that if there is an accident of a trammer aged between 10years to 60years in mechanised mine, 92.35 chances is that it is reportable accident.

IV. CONCLUSIONS

Using knowledge discovery from database process model we have done the multivariate analysis of the accident data. This multivariate analysis has lead us to take all the variable affecting the accidents to categories the data into the three class according to the type of accident. Using this categorization and the predictability scores which are provided as a summary we calculate the overall risk score due to an attribute (while considering all the attribute at a time) and prioritize it.

The multivariate analysis gives a much wider view than we get from the bivariate analysis and hence we are able to prioritize all the input attribute as shown in table. It is notable that the results do not violate the one which we get from the bivariate analysis rather it is the same both follow the same basics. The only difference between the bivariate analyses is that it only check the association of two variables at a time where as a multivariate uses all the input attribute for predicting the overall probability.

We also get a number of association rules which are discovered and can be used as a generalized rule for the different type of attribute to happen simultaneously. The findings matches with the actual analysis of the mines and this have been verified almost with the same degree of accuracy by the data of accidents of a mine (Mandman colliery) for the year 1996-2005.

Thus the work gives a view of the accident trend in ECL can be used as a model for Risk Assessment in any other mines.

REFERENCES

[1] Chen, M. S.; Han, J.; and Yu, P.S. Data Mining: An Overview from a Database Perspective. IEEE Transactions on Knowledge and Data Engineering, 8(6): 866-883, 1996.

[2] Bradley, P; Fayyad, U.; and Mangasarian, O. Data Mining:Overview and Optimization Opportunities. Microsoft Research Report MSR-TR-98-04, January 1998

[3] Agrawal, R. and Psaila, G. Active Data Mining. Proc. Ofthe 1st International Conference on Knowledge Discovery and Data Mining, Montreal, August 1995.

[4] Agrawal, R; Imielinski, T.; and Swami, A. Mining Associations between Sets of Items in Massive Databases. Proc. of the ACM

SIGMOD International Conference on Management of Data, pp. 207-216, Washington D.C., May 1993.
[5] Agrawal, R.; Imielinski, T.; and Swami, A. Database Mining: A Performance Perspective. IEEE Transactions on Knowledge and Data Engineering, Special issue on Learning and Discovery in Knowledge-Based Databases, 5(6): 914-925, December 1993
[6] Agrawal, R.; Arning, A.; Bollinger, T.; Mehta, M.; Shafer, J. and Srikant, R. The Quest Data Mining System.Proc.of the 2nd International Conference on Knowledge Discovery in Databases and Data Mining, Portland, Oregon, August 1996.
[7] Agrawal, R.;Srikant, R. Fast Algorithms for Mining Association Rules. Proc. of the 20th Int'l Conference on Very Large Databases,Santiago,Chile,Sept.1994
[8] Ahmed, Khalil M.; E1-Makki, Nagwa M.; and Taha, Yousry. A note on "Beyond Market Baskets, Generalizing association rules to correlations". In SIGKDD explorations: Newsletter of The Special Interest Group (SIG) on Knowledge Discovery & Data Mining, Vol. 1 Issue 2, January 2000.
[9] Brin, S.; Motwani, R; and Silverstein, C. Beyond Market Baskets: Generalizing Association Rules to Correlations. In Proc.Of the 1997 SIGMOD Conf. on the Management of Data, pp. 265-276.
[10] Ali, K; Manganaris, S.; and Srikant, R. Partial Classification using Association Rules. In Proc. of the 3rd International Conference on Knowledge Discovery in Databases and Data Mining, August 1997.
[11] MCKENNA, F.P. 'Accident proneness: A conceptual analysis.' Accident Analysis and Prevention, vol. 15, 1983. pp. 65-71