# Rough Set Approach in Finding the Cause of Decline and Down Fall of Jute Industries and the Remedy

Sujogya Mishra
Research Scholar,
Utkal University
Bhubaneswar-751004, India

Shakti Prasad Mohanty
Department of Mathematics,
College of Engineering and Technology
Bhubaneswar-751003, India

Sateesh Kumar Pradhan
Department of Computer Science
Utkal University
Bhubaneswar-751003, India

Radhanath Hota
Department of Computer Science
College of Basic Science & Humanities,
OUAT, Bhubaneswar-751003, India

## ABSTRACT
The jute industries 10 to 15 years back especially in south Asian countries are bread and butter for the people belonging to middle and lower income group , now the scenario is completely different. In the present scenario jute is found to be expensive and not much useful as compared to other parallel packaging material available in the market due for this reason most of the jute mills suffered, from severe financial crisis which forced the jute mills owner to close down their unit .In this context we tries to find the cause of failure of jute industries in recent age and how to develop the jute industries in recent age,. For this purpose we develop an algorithm by using rough set concept on data which we gathered from different sources, develop algorithm is simple and user friendly then validate this concept by using statistical validation method in our paper we basically focused on issues which leads to sick jute industries. Initially we gathered 10000 samples for our purpose then applying correlation technique on the collected data the data set reduced to 20 which are dissimilar in nature . Once we have the data set by correlation technique we then apply rough set techniques on those data to generate an efficient algorithm . The entire paper is sub divided into three sections. Section 1 deal with literature review and last two section deals with the experimental result and statistical validation of our proposed algorithm.

## Keywords
*Rough Set Theory, Raw data regarding Jute industries , Granular computing, Data mining.*

## 1. INTRODUCTION
The increasing demand and the application of modern technology for the growth of business has produced huge data .The amount of data that has generated always make it difficult to extract inference for further application .This is a challenge for the analyst to find a definite data set and extract rules from those data set.

Application of rough set theory is a very useful tools in knowledge discovery findings from the data base .The ever growing field of knowledge discovery (KD) helps in the collection of hidden information from large database [3]. Data mining is also considered as an essential tool in knowledge discovery process which uses techniques from different disciplines ranging from machine learning, statistical information , database, visualization,([4]-[12]).Further, prediction of business failure needs a systematic and scientific study. The first approach to predict business failure started in 1995 by Zopounidis( [24]-[26]). The methods proposed are the "five C" methods, the "LAPP" method, and the "credit-men" method. Then, financial ratios methodology

Prediction of business failure , this idea can be extended to develop method which can predict business failure efficiently based upon multivariate statistical analysis. Altman([13]-[15]), Beaver[17], Courtis[18]). Frydman *et al*[19] first employed recursive portioning, while Gupta *et a*l[20] use mathematical programming as an alternative to multivariate discriminate analysis for business failure prediction problem. Other methods used were survival analysis by

Luoma, Laitinenl[21] which is a tool for company failure prediction, expert systems by Messier and Hansen[22] , neural network by Altman *et al*[16], multi-factor model by Vermeulen *et al*[23] are also other methods developed for business failure prediction. This paper presents a methodology for prediction of failure of jute industries. .

## 2. PRILIMINARIES
1. **Rough set Rough** set theory as introduced by Z. Pawlak [8] is an extension of conventional set theory that support approximations in decision making.

2.1.2. Approximation Space: An Approximation space is a pair (U, R) where U is a non empty finite set called the universe R is an equivalence relation defined on U.

2.1.3. Information System: An information system is a pair S = (U , A), where U is thenon-empty finite set called the universe, A is the non-empty finite set of attributes

2.1.4. Decision Table: A decision table is a special case of information systems S= (U , A= C U {d}), where d is not in C. Attributes in C are called conditional attributes and d is a designated attribute called the decision attribute

2.1.5. Approximations of Sets: Let S = (U, R) be an approximation space and X be a subset of U.The lower approximation of X by R in S is defined as $\underline{RX}$ = { e ε U | [e] ε X} and The upper approximation of X by R in S is defined as $\overline{RX}$ □{e□U/[e]□X □□}where [e] denotes the equivalence class containing e. A subset X of U is said to be R-definable in S if and only if $\overline{RX}$=$\underline{RX}$ .A set X is rough in S if its boundary set is nonempty.

## 2. Dependency of Attributes

Let C and D be subsets of A. We say that D depends on C in a degree k (0 ≦k ≦1) denoted by C→k D if K=y(C,D)= $\left|\frac{POS_C(D)}{|u|}\right|$ where $POS_C(D)$ =U $\underline{C}$(X), is called positive region of the partition U/D with respect to C where $x \in u/d$ , which is all elements of U that can be uniquely classified to the block of partition U/D. If k = 1 we say that D depends totally on C. If k < 1 we say that D depends partially (in a degree k) on C.

**3. Dispensable and Indispensable Attributes-** **Let** *S = (U, A = C υ D) be a decision table.* Let c be an attribute in C. Attribute c is dispensable in S if $POS_C$ (D)= $POS_{(C-\{c\})}$(D)otherwise, c is indispensable. A decision table S is independent if all attributes in C are indispensable.

*Let S = (U, A = C υ D) be a decision table.*

Rough Set Attribute Reduction (RSAR) provides a filter based tool by which knowledge may be extracted from a domain in a concise way; retaining the information content whilst reducing the amount of knowledge involved.

**4. Reduct and Core** Let S = (U, A=C U D) be a decision table. A subset R of C is a reduct of C, if $POS_{\mathbf{R}}$(D) = $POS_{\mathbf{C}}$(D) and S’ = (U, RUD) is independent, ie., all attributes in R are indispensible in S’. Core of C is the set of attributes shared by all reducts of C. CORE(C) = ∩RED(C) where, RED(C) is the set of all reducts of C. The reduct is often used in the attribute selection process to eliminate redundant attributes towards decision making.

**5. Correlation-** Correlation define as a mutual relationship or connection between two or more things .The quantity *r*, called the *linear correlation coefficient*, measures the strength and the direction of a linear relationship between two variables. The linear correlation coefficient is sometimes referred to as the *Pearson product moment correlation coefficient* in honor of its developer Karl Pearson. The mathematical formula for its coefficient given by the formula

$$r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\left(\sum x^2\right) - \left(\sum x\right)^2}\sqrt{n\left(\sum y^2\right) - \left(\sum y\right)^2}}$$

**6. Goodness of fit-**The **goodness of fit** of a statistical model describes how well it fits a set of observations. Measures of **goodness of fit** typically summarize the discrepancy between observed values and the values expected under the model in question.

**7. Chi squared distribution-** A **chi-squared test**, also referred to as **$\chi^2$ test**, is any statistical hypothesis test in which the sampling distribution of the test statistic is a chi squared distribution when the null hypothesis is true. Also considered a chi-squared test is a test in which this is *asymptotically* true, meaning that the sampling distribution (if the null hypothesis is true) can be made to approximate a chi-squared distribution as closely as desired by making the sample size large enough. The chi-square (I) test is used to determine whether there is a significant difference between the expected frequencies and the observed frequencies in one or more categories. Do the number of individuals or objects that fall in each category differ significantly from the number you would expect? Is this difference between the expected and observed due to sampling variation, or is it a real difference

**8. Further analysis of chi square test-** Basic properties of chi squared goodness fit is that it is non symmetric in nature .How ever if the degrees of freedom increased it appears to be to be more symmetrical .It is right tailed one sided test. All expectation in chi squared test is greater than 1 $E_I=np_i$ where n is the number samples considered $p_i$ is the probability of $i^{th}$ occurrence .Data selected at random there are two hypothesis null hypothesis and alternate hypothesis null hypothesis denoted by $H_0$ alternate hypothesis denoted by $H_1$. $H_0$ is the claim does follow the hypothesis and $H_1$ is the claim does not follow the hypothesis here $H_1$ is called the alternate hypothesis to $H_0$.If the test value found out to be K then K can be calculated by the formula $K=\sum(O_I-E_I)^2/E_I$. Choice of significance level always satisfies type 1 error.

### 2.9. Different types of error-

1) Type 1 error-Rejecting a hypothesis even though it is true
2) Type 2 error-Accepting the hypothesis when it is false
3) Type 3 error-Rejecting a hypothesis correctly for wrong reason

## 3. BASIC IDEA

The basic idea being conceived by looking at the present scenario of jute industries .Both employer and employees of the jute industries suffers in the recent past.. We intensively analyze and rigorously studies the different data’s which we collected about jute industries to extract certain inference about jute industries and upon applying rough set on those drawn inference we develop an algorithm. The generalize version of the algorithm is develop which we develop from the original one is simple and useful and helpful to find the cause for the failure of jute industries. In our research we find jute industries sick within a period of 10 to 12 years of time span ( 1996-2008). Initially we find the cause for the failure of jute industries as the failure of jute industries is due to lack of foresight of both mills owner and the government. Most of time mills proprietor, takes whimsical decision by disengaging the workers, Reduction of the salary of the existing staff, terminating all ministerial staff and disengagement of all women employees to counter the financial crisis. To develop the algorithm we modified the initial cause of failure to derive the conditional attributes by rigorously studies the situation about jute industries in recent past . Our modified attributes for the purpose as follows

1. Government policy towards jute industries, 2.Owner’s attitude towards the employee of the industries 3. Other parallel materials available which counter the materials

generated by jute industries (like poly thin , plastic etc)4.Lack of advertisement of jute industries in the national and international market 5.The number of products generate from jute consider as our conditional attributes and consider two decision attributes as success and failure .The values as accept , reject , not exact (intermediate state of reject and failure ). We rename the conditional attributes as $a_1$,$a_2$,$a_4$,$a_5$ and it's values are rename as $b_1$,$b_2$ and $b_3$ respectively. We consider the decision attribute true , false further rename as $c_1$ and $c_2$ respectively (that is significance and non significance).

## 4. DATA REDUCTION

As size of the data increased , it is very difficult to find which attributes are actually essential attributes and which are not essential for a particular problem . The aim of data reduction is to find the relevant attributes that have all essential information of the data set. The process is illustrated through tables for rough classification. In this particular problem we consider the data presented in the table -1

Table-1:

| E | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | d |
|---|---|---|---|---|---|---|
| $E_1$ | $b_2$ | $b_2$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_2$ | $b_2$ | $b_2$ | $b_1$ | $b_1$ | $b_1$ | $c_1$ |
| $E_3$ | $b_1$ | $b_2$ | $b_2$ | $b_3$ | $b_3$ | $c_1$ |
| $E_4$ | $b_1$ | $b_2$ | $b_2$ | $b_3$ | $b_1$ | $c_1$ |
| $E_5$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $b_2$ | $c_2$ |
| $E_6$ | $b_1$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $c_2$ |
| $E_7$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $c_1$ |
| $E_8$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_9$ | $b_1$ | $b_2$ | $b_2$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{10}$ | $b_1$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $c_2$ |
| $E_{11}$ | $b_2$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{12}$ | $b_1$ | $b_2$ | $b_3$ | $b_1$ | $b_2$ | $c_1$ |
| $E_{13}$ | $b_3$ | $b_2$ | $b_2$ | $b_2$ | $b_1$ | $c_2$ |
| $E_{14}$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{15}$ | $b_2$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{16}$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{17}$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{18}$ | $b_1$ | $b_2$ | $b_2$ | $b_3$ | $b_2$ | $c_2$ |
| $E_{19}$ | $b_1$ | $b_3$ | $b_1$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{20}$ | $b_1$ | $b_2$ | $b_1$ | $b_3$ | $b_3$ | $c_1$ |

The decision table -1 , takes the initial values before finding the reduct looking at the data table it is found that entities $E_1$, and $E_2$, ambiguous in nature and $E_{16}$, $E_{17}$ gives same result so both $E_1$,$E_2$ drop from the table because of it's ambiguity nature and from $E_{16}$, $E_{17}$ we keep one record that is either $E_{16}$ or $E_{17}$ for our purpose so the new table appears as table -2 is called reduce table now we apply the rough set concept on table -2 to find the strength[27]

Table-2

| E | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | d |
|---|---|---|---|---|---|---|
| $E_3$ | $b_1$ | $b_2$ | $b_2$ | $b_3$ | $b_3$ | $c_1$ |
| $E_4$ | $b_1$ | $b_2$ | $b_2$ | $b_3$ | $b_1$ | $c_1$ |
| $E_5$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $b_2$ | $c_2$ |
| $E_6$ | $b_1$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $c_2$ |
| $E_7$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $c_1$ |
| $E_8$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_9$ | $b_1$ | $b_2$ | $b_2$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{10}$ | $b_1$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $c_2$ |
| $E_{11}$ | $b_2$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{12}$ | $b_1$ | $b_2$ | $b_3$ | $b_1$ | $b_2$ | $c_1$ |
| $E_{13}$ | $b_3$ | $b_2$ | $b_2$ | $b_2$ | $b_1$ | $c_1$ |
| $E_{14}$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{15}$ | $b_2$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{16}$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{18}$ | $b_1$ | $b_2$ | $b_2$ | $b_3$ | $b_2$ | $c_2$ |
| $E_{19}$ | $b_1$ | $b_3$ | $b_1$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{20}$ | $b_1$ | $b_2$ | $b_1$ | $b_3$ | $b_3$ | $c_1$ |

### Indiscernibility relation:

Indiscernibility Relation is the relation between two or more objects where all the values are identical in relation to a subset of considered attributes.

### Approximation:

The starting point of rough set theory is the indiscernibility relation, generated by information concerning objects of interest. The indiscernibility relation is intended to express the fact that due to the lack of knowledge it is unable to discern some objects employing the available information Approximations is also other an important concept in Rough Sets Theory, being associated with the meaning of the approximations topological operations (Wu et al., 2004). The lower and the upper approximations of a set are interior and closure operations in a topology generated by the indiscernibility relation. Below is presented and described the types of approximations that are used in Rough Sets Theory.

## a. Lower approximations

Lower Approximation is a description of the domain objects that are known with certainty to belong to the subset of interest.The Lower Approximation Set of a set X, with regard to R is the set of all objects, which can be classified with X regarding R, that is denoted as $R_L$.

## b. Upper Approximation :

Upper Approximation is a description of the objects that possibly belong to the subset of interest. The Upper Approximation Set of a set X regarding R is the set of all of objects which can be possibly classified with X regarding R . Denoted as $R_U$

## c. Boundary Region (BR) :

Boundary Region is description of the objects that of a set X regarding R is the set of all the objects, which cannot be classified neither as X nor -X regarding R. If the boundary region X= ϕ then the set is considered "Crisp", that is, exact in relation to R; otherwise, if the boundary region is a set **X**≠ϕ the set X "Rough" is considered. In that the boundary region is BR = $R_U$-$R_L$.

The lower and the upper approximations of a set are interior and closure operations in a topology generated by a indiscernibility relation. In discernibility according to decision attributes in this case has divided in to two groups one group consist of positive case and another group consists of negative cases

$E_{true}$ ={$E_3$, $E_4$, $E_7$, $E_{11}$, $E_{12}$,$E_{13}$, $E_{14}$ ,$E_{20}$}…...(1)

$E_{false}$ ={$E_5$, $E_6$, $E_8$, $E_9$, $E_{10}$, $E_{15}$, $E_{16}$,$E_{18}$,$E_{19}$ }……(2)

E($a_1$)$_{accept}$={ $E_3$, $E_4$, $E_6$, $E_8$, $E_9$, $E_{10}$, $E_{12}$,$E_{16}$, $E_{18}$ $E_{19}$, $E_{20}$}…….(3)

E($a_1$)$_{reject}$ ={ $E_7$, $E_{11}$, $E_{15}$ }……………….……..(4)

E($a_1$)$_{nonexactt}$ ={ $E_5$, $E_{13}$, $E_{14}$ }………………..……..(5)

The above result when compared with the failure cases E($a_1$)$_{true}$ strength[27]

Found to be 4/11 about 36% where as for false cases of accepting E($a_1$) strength[27] is 7/11 about 63% similarly for non exact case success E($a_1$) strength[27] gives rise to be 1/3 about33% so we see that adopting case $a_1$ we have failure is 36% and success is about 63% and non exact $a_1$ we have a success about 33% from this analysis we have the following ,$a_1$ provide some significance , now analyzing $a_2$ we have the following result

E($a_2$)$_{accept}$={ $E_8$, $E_{15}$, $E_{16}$ }……………..…..(6)

E($a_2$)$_{reject}$={$E_3$,$E_4$,$E_6$,$E_7$,$E_9$,$E_{10}$,$E_{13}$,$E_{18}$, $E_{20}$}…………………..…..(7)

E($a_2$)$_{nonexact}$ ={ $E_5$,$E_{11}$,$E_{14}$, $E_{19}$}………..……..(8)

STtrength[27] for accepting $a_2$ is found to be nil and strength[27] of true by rejecting $a_2$ is 2/4 about 50% similarly non exact $a_2$ strength[27] will be 2/3 about 66% non exact means we cannot definite about our decision similarly

E($a_3$)$_{accept}$={ $E_8$,$E_{12}$ ,$E_{15}$,$E_{16}$ }..(9)

E ($a_3$)$_{reject}$={$E_3$, $E_4$, $E_6$, $E_7$, $E_9$, $E_{10}$ ,$E_{13}$} … (10)

E($a_3$)$_{nonexact}$={$E_5$,$E_{11}$,$E_{12}$,$E_{14}$}…………………….….(11

Now the strength[27] of accept truth $a_3$ ¼ about 25% where as failure is about ¾ about 75% So now up to this stage adopting attribute $a_1$ and $a_3$ no longer helpful so we drop both attribute from the decision table for next round analysis

E($a_4$)$_{accept}$ ={$E_8$,$E_{12}$,$E_{15}$,$E_{16}$} ……………….………(12)

E($a_4$)$_{reject}$ ={$E_6$,$E_7$,$E_{10}$,$E_{13}$} ………………(13)

E($a_4$)$_{nonexact}$ ={$E_3$,$E_4$,$E_5$,$E_9$,$E_{11}$,$E_{14'}$,$E_{20}$}…………………….(14)

Accept $a_4$ strength[27] for truth

case is ¼ about 25% rejecting $a_4$ the strength[27] of success case is 5/8 is about 62.5% so rejecting $a_4$ gives a significant result so we keep $a_4$ for further analysis now analyzing $a_5$findingE($a_5$)$_{accept}$={$E_4$,$E_8$,$E_9$,$E_{13}$,$E_{15}$,$E_{16}$}…(15)

E($a_5$)$_{reject}$={ $E_5$,$E_6$,$E_7$,$E_{10}$,$E_{12}$,$E_{18}$ }………………..(16)

E($a_5$)$_{nonexact}$ ={ $E_5$,$E_6$,$E_{10}$,$E_{12}$ }…………………….(17)

E($a_5$)$_{accept}$={ $E_4$,$E_8$,$E_{13}$,$E_{15}$,$E_{16}$}………………….(18)

Strength[27] accepting $a_5$ truth is 2/8 about 25% and non exact case Strength[27] 1/5 is about 20% so $a_5$ does not provide any significant result in the in this case of true, but provide some significance result on considering the false case .WE can get one conclusion that attribute $a_5$ has some amount of significance so we can drop the attribute $a_2$ from the table to get some significant result which is presented in table-3 now in the table -3 we have a set of values

Table-3

| E | $a_1$ | $a_3$ | $a_4$ | $a_5$ | d |
|---|---|---|---|---|---|
| $E_3$ | $b_1$ | $b_2$ | $b_3$ | $b_3$ | $c_1$ |
| $E_4$ | $b_1$ | $b_2$ | $b_3$ | $b_1$ | $c_1$ |
| $E_5$ | $b_3$ | $b_3$ | $b_3$ | $b_2$ | $c_2$ |
| $E_6$ | $b_1$ | $b_2$ | $b_2$ | $b_2$ | $c_2$ |
| $E_7$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $c_1$ |
| $E_8$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_9$ | $b_1$ | $b_2$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{10}$ | $b_1$ | $b_2$ | $b_2$ | $b_2$ | $c_2$ |
| $E_{11}$ | $b_2$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| | | | | | |
| $E_{12}$ | $b_1$ | $b_3$ | $b_1$ | $b_2$ | $c_1$ |
| $E_{13}$ | $b_3$ | $b_2$ | $b_2$ | $b_1$ | $c_1$ |

| | | | | | |
|---|---|---|---|---|---|
| $E_{14}$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{15}$ | $b_2$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{16}$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{18}$ | $b_1$ | $b_2$ | $b_3$ | $b_2$ | $c_2$ |
| $E_{19}$ | $b_1$ | $b_1$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{20}$ | $b_1$ | $b_1$ | $b_3$ | $b_3$ | $c_1$ |

(The table is the part of table-3)

As analyzing table -3 we find $E_6$, $E_{10}$ and $E_8$, $E_{16}$ forms a group so we keep one record from $E_6$, $E_{10}$ and $E_8$, $E_{11}$ respectively now the reduced table is represented Reduced table -4 that is as follows

from Table-4

| E | $a_1$ | $a_3$ | $a_4$ | $a_5$ | d |
|---|---|---|---|---|---|
| $E_3$ | $b_1$ | $b_2$ | $b_3$ | $b_3$ | $c_1$ |
| $E_4$ | $b_1$ | $b_2$ | $b_3$ | $b_1$ | $c_1$ |
| $E_5$ | $b_3$ | $b_3$ | $b_3$ | $b_2$ | $c_2$ |
| $E_6$ | $b_1$ | $b_2$ | $b_2$ | $b_2$ | $c_2$ |
| $E_7$ | $b_2$ | $b_2$ | $b_2$ | $b_2$ | $c_1$ |
| $E_8$ | $b_1$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_9$ | $b_1$ | $b_2$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{11}$ | $b_2$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{12}$ | $b_1$ | $b_3$ | $b_1$ | $b_2$ | $c_1$ |
| $E_{13}$ | $b_3$ | $b_2$ | $b_2$ | $b_1$ | $c_1$ |
| $E_{14}$ | $b_3$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{15}$ | $b_2$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{18}$ | $b_1$ | $b_2$ | $b_3$ | $b_2$ | $c_2$ |
| $E_{19}$ | $b_1$ | $b_1$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{20}$ | $b_1$ | $b_1$ | $b_3$ | $b_3$ | $c_1$ |

Upon analyzing table-4 we have a very peculiar result that is attribute $a_1$ we find a lot of ambiguity result for example with respect to attribute $a_1$ value $b_1$ is ambiguous for the group ($E_3$,$E_4$,$E_{12}$,$E_{20}$) ambiguous with the group ($E_6$,$E_8$,$E_9$,$E_{18}$,$E_{19}$) similarly with respect to value $b_2$ of attribute $a_1$ ($E_7$,$E_{11}$) group ambiguous with $E_{15}$ similarly if we consider the value $b_3$ for attribute $a_1$ is also giving ambiguous result that is $E_5$ ambiguous with the group ($E_{13}$,$E_{14}$) so we find this not important from the point of adopting it so we remove this attribute from the table as we are having lots of ambiguous result in this attribute so adopting the attribute is not in the success result now table reduced table -4 it provide the following result Reduced table -5 that is as follows from

Table-5

| E | $a_3$ | $a_4$ | $a_5$ | d |
|---|---|---|---|---|
| $E_3$ | $b_2$ | $b_3$ | $b_3$ | $c_1$ |
| $E_4$ | $b_2$ | $b_3$ | $b_1$ | $c_1$ |
| $E_5$ | $b_3$ | $b_3$ | $b_2$ | $c_2$ |
| $E_6$ | $b_2$ | $b_2$ | $b_2$ | $c_2$ |
| $E_7$ | $b_2$ | $b_2$ | $b_2$ | $c_1$ |
| $E_8$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_9$ | $b_2$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{11}$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{12}$ | $b_3$ | $b_1$ | $b_2$ | $c_1$ |
| $E_{13}$ | $b_2$ | $b_2$ | $b_1$ | $c_1$ |
| $E_{14}$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{15}$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{18}$ | $b_2$ | $b_3$ | $b_2$ | $c_2$ |
| $E_{19}$ | $b_1$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{20}$ | $b_1$ | $b_3$ | $b_3$ | $c_1$ |

Now analyzing the table -5 we have ($E_3$,$E_9$) and ($E_6$,$E_7$) are ambiguous in nature and ($E_8$,$E_{15}$) and ($E_{11}$,$E_{15}$) forms a group so we eliminate all records that is present in $E_3$ in $E_9$ and $E_6$ in $E_7$ respectively and keep single record from ($E_8$,$E_{15}$) and ($E_{11}$,$E_{15}$) so now the new table -6 we get from table-5 so we have the following result that is Reduced table -6 that is as follows from

Table-6

| E | $a_3$ | $a_4$ | $a_5$ | d |
|---|---|---|---|---|
| $E_4$ | $b_2$ | $b_3$ | $b_1$ | $c_1$ |
| $E_5$ | $b_3$ | $b_3$ | $b_2$ | $c_2$ |
| $E_8$ | $b_1$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{11}$ | $b_3$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{12}$ | $b_3$ | $b_1$ | $b_2$ | $c_1$ |
| $E_{13}$ | $b_2$ | $b_2$ | $b_1$ | $c_1$ |
| $E_{18}$ | $b_2$ | $b_3$ | $b_2$ | $c_2$ |
| $E_{19}$ | $b_1$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{20}$ | $b_1$ | $b_3$ | $b_3$ | $c_1$ |

(This is the part of table -6)

After analyzing the table -6 in particular attribute $a_3$ for value $b_1$ shows $E_8$ and $E_{20}$ ambiguous in nature similarly upon analyzing the the value $b_2$ for the attribute $a_3$ provide ambiguous result $E_{13}$ and $E_{18}$ similarly analyzing the attribute $a_3$ for the value $b_3$ also has an ambiguous result as we are getting lots of ambiguity in $a_3$ so we drop $a_3$ attribute from the table -6 to get new table-7

Reduced table -7 that is as follows from

Table-7

| E | $a_4$ | $a_5$ | d |
|---|---|---|---|
| $E_4$ | $b_3$ | $b_1$ | $c_1$ |
| $E_5$ | $b_3$ | $b_2$ | $c_2$ |
| $E_8$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{11}$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{12}$ | $b_1$ | $b_2$ | $c_1$ |
| $E_{13}$ | $b_2$ | $b_1$ | $c_1$ |
| $E_{18}$ | $b_3$ | $b_2$ | $c_2$ |
| $E_{19}$ | $b_3$ | $b_3$ | $c_2$ |
| $E_{20}$ | $b_3$ | $b_3$ | $c_1$ |

Now upon analyzing the table-7 we have ($E_5$,$E_{18}$) forms a group and ($E_{19}$,$E_{20}$) ambiguous result so we drop $E_{19}$,$E_{20}$ from the table and keep one record from $E_5$,$E_{18}$ now we have the reduced table-7 from table - 6 gives the following result that is new table given below

Reduced table -8 that is as follows from

Table-8

| E | $a_4$ | $a_5$ | d |
|---|---|---|---|
| $E_4$ | $b_3$ | $b_1$ | $c_1$ |
| $E_5$ | $b_3$ | $b_2$ | $c_2$ |
| $E_8$ | $b_1$ | $b_1$ | $c_2$ |
| $E_{11}$ | $b_3$ | $b_3$ | $c_1$ |
| $E_{12}$ | $b_1$ | $b_2$ | $c_1$ |
| $E_{13}$ | $b_2$ | $b_1$ | $c_1$ |

The paper actually find the down fall of jute industries and remedy to get success

From table -8 we have the following decision present as an algorithm state as follows that is
Step-1 Advertisement and the product generated from jute has taken little more care leads to success for jute industry
Step-2 Other parallel material which counter jute product should be checked ,That government should put a check in the use of other available parallel material (Along with jute)

Running time of this comparison will take $O(n^2)$ as every time we compare to find the reduct every record compare with rest n records then n-1 records till we reach 1 record so total running time will be n+(n-1)+ (n-2) +……..+1=n(n+1)/2 Of order $n^2$ . And further breaking the table will take O(nlgn) so total time complexity will take $O(nlgn+n^2)$

**Statistical validation-** For validate our findings we basically depends upon chi-square test for this purpose we consider we take a survey by taking data regarding the truth I,e the success case and we are not focused on one particular to collect our data to collect the data we approached several jute industries and the apply chi square test to validate our claim. . Chi square test-Expected15%,10%,15%,20%,30%,15% and the Observed samples are 25,14,34 45,62,20 so totaling these we have total of 200 samples so expected numbers of samples per each day as follows 30,20,30,40,60,30 . We then apply chi square distribution to verify our result assuming that $H_0$ is our hypothesis that is correct $H_1$ as alternate hypothesis that is not correct , Then we expect sample in six cases as
chi squared estimation formula is $\sum(O_i-E_i)^2/ E_i$ where i=0,1,2,3,4,5 so the calculated as follows

$X^2=(25-30)^2/20+(14-20)^2/20+(34-30)^2/30+(45-40)^2/40+(62-60)^2/60+(20-30)^2/30$

$X^2=25/20+36/20+16/30+25/40+4/60+100/30=7.60$ the tabular values we have with degree of freedom 5 we get result 11.04

As we find our result lies much below the critical values so this result is statistically validate

Future work- this work can be extended and applicable to different business house like film industry , software industries small and large scale industries .

## 5. REFERENCES

[1] S.K. Pal, A. Skowron, Rough Fuzzy Hybridization: A new trend in decision making, Berlin, Springer-Verlag, 1999

[2] Z. Pawlak, “Rough sets”, International Journal of Computer and Computer and Information Sciences, Vol. 11, 1982, pp.341–356

[3] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, System Theory, Knowledge Engineering and Problem Solving, Vol. 9, The Netherlands, Kluwer Academic Publishers, Dordrecht, 1991

[4] Han, Jiawei, Kamber, Micheline, Data Mining:Concepts and Techniques. San Franciso CA, USA, Morgan Kaufmann Publishers, 2001

[5] Ramakrishnan, Naren and Grama, Y. Ananth, “Data Mining: From Serendipity to Science”, IEEE Computer, 1999, pp. 34-37.

[6] Williams, J. Graham, Simoff, J. Simeon, Data Mining Theory, Methodology, Techniques, and Applications (Lecture Notes in Computer Science/ Lecture Notes in Artificial Intelligence), Springer, 2006.

[7] D.J. Hand, H. Mannila, P. Smyth, Principles of Data Mining. Cambridge, MA: MIT Press, 2001

[8] D.J. Hand, G.Blunt, M.G. Kelly, N.M.Adams, “Data mining for fun and profit”, Statistical Science, Vol.15, 2000, pp.111-131.

[9] C. Glymour, D. Madigan, D. Pregibon, P.Smyth, Statistical inference and data mining", Communications of the ACM, Vol. 39, No.11,1996,pp.35-41.

[10] T.Hastie, R.Tibshirani, J.H. Friedman, Elements of statistical learning: data mining, inference and prediction, New York: Springer Verlag, 2001

[11] H.Lee, H. Ong, "Visualization support for data Mining", IEEE Expert, Vol. 11, No. 5, 1996, pp. 69-75.

[12] H. Lu, R. Setiono, H. Liu,"Effective data Mining using neural networks", IEEE Transactions on Knowledge and Data Engineering, Vol. 8, No. 6, 1996, pp. 957-961.

[13] E.I Altman, "Financial ratios, discriminants analysis and prediction of corporate bankruptcy", The journal of finance, Vol. 23 , 1968, pp.589-609

[14] E.I.Altman, R.Avery, R.Eisenbeis, J. Stnkey, "Application of classification techniques in business, banking and finance. Contemporary studies in Economic and Financial Analysis", vol.3, Greenwich, JAI Press,1981.

[15] E.I Altman, "The success of business failure prediction models: An international surveys", Journal of Banking and Finance Vol. 8, no.2, 1984, pp.171-198

[16] E.I Altman, G. Marco, F. Varetto, "Corporate distress diagnosis: Comparison using discriminant analysis and neural networks", Journal of Banking and Finance, Vol.18, 1994, pp. 505-529

[17] W.H Beaver, "Financial ratios as predictors of failure Empirical Research in accounting : Selected studies", Journal of Accounting Research Supplement to Vol- 4, 1966, pp.71-111

[18] J.K Courtis, "Modelling a financial ratios categoric frame Work", Journal of Business Finance and Accounting, Vol. 5, No.4, 1978, pp71-111

[19] H.Frydman, E.I Altman ,D-lKao, "Introducing recursive partitioning for financial classification: the case of financial distress", The Journal of Finance, Vol.40, No. 1 1985, pp. 269-291.

[20] Y.P.Gupta, R.P.Rao, P.K. , Linear Goal programming as an alternative to multivariate discriminant analysis a note journal of business fiancé and accounting Vol.17, No.4, 1990, pp. 593-598[21] M. Louma, E, K. Laitinen, "Survival analysis as a tool for company failure prediction". Omega, Vol.19, No.6, 1991, pp. 673-678

[22] W.F. Messier, J.V. Hanseen, "Including rules for expert system development: an example using default and bankruptcy data", Management Science, Vol. 34, No.12, 1988, pp.1403-1415

[23] E.M. Vermeulen, J. Spronk, N. Van der Wijst., The application of Multifactor Model in the analysis of corporate failure. In: Zopounidis,C.(Ed), Operational corporate Tools in the Management of financial Risks, Kluwer Academic Publishers, Dordrecht, 1998, pp. 59-73

[24] C. Zopounidis, A.I. Dimitras, L. Le Rudulier, A multicriteria approach for the analysis and prediction of business failure in Greece. Cahier du LAMSADE, No.132, Universite de Paris Dauphine, 1995.

[25] C. Zopounidis, N.F. Matsatsinis, M. Doumpos, "Developing a multicriteria knowledge-based decision support system for the assessment of corporate performance and viability: The FINEVA system, "Fuzzy Economic Review, Vol. 1, No. 2, 1996, pp. 35-53.

[26] C. Zopounidis, M. Doumpos, N.F. Matsatsinis, "Application of the FINEVA multicriteria knowledge decision support systems to the assessment of corporate failure risk", Foundations of Computing and Decision Sciences, Vol. 21, No. 4, 1996, pp. 233-251

[27] 11 Renu Vashist Prof M.L Garg Rule Generation based on Reduct and Core :A rough set approach International Journal of Computer Application(0975-887) Vol 29 September -2011