Generation of Process Models for Dyeing Process using Association Rule Mining Algorithms and Weka Library

Saravanan M. S.[#], Rama Sree R. J.[*]

[#] *R & D Centre, Bharathiar University, Coimbatore, Tamil Nadu, INDIA.*
Dept. of I.T in Vel Tech Dr. RR & Dr. SR Technical University, Chennai, INDIA.
[*] *Department of Computer Science, Rashtriya Sanskrit Vidyapeetha,*
Tirupati, Andhra Pradesh, INDIA.

[1]saranenadu@yahoo.co.in, [2]rjramasree@yahoo.com

***Abstract*— In recent study, we have identified that the process mining algorithms is not sufficient for the dyeing process, because of its dynamic nature. Hence the better process models need to be generated; so, the alternative process model is generated using association rule mining algorithms and Weka Library. Therefore, this paper aims at evaluating and analyzing the usefulness and applications of the association rule mining algorithms and as it was implemented to obtain simpler process models for the dyeing domain. The Emerald and Jayabala dyeing processes were used in implement the Apriori, FPGrowth, H-Mine LRM association rule mining algorithms and Weka library to gain some insights in its processes. Hence, this paper focus on these algorithms to contribute analysis of dyeing process to generate process models for the two dyeing units.**

***Keywords* - Process Model, Dyeing process, Association rule mining, Weka Library.**

I. Introduction

The Process Modeling is widely used within organizations as a method to increase awareness and knowledge of business processes, and to deconstruct organizational complexity [1]. Process models in general server two main purposes [2]. On the one hand, business process models are used for scoping the project, and capturing and discussing business requirement and process improvement initiatives with subject matter experts, best example of this is Event driven Process Chain (EPC) [3]. On the other hand, technical process models can also be used for process automation, which requires their conversion into executable specifications, best example of this is Petri nets [4], Business Process Execution Language for Web Services (BPEL4WS) [5].

The Process modeling does not decrease the complexity of describing the processes under consideration hence it helps people in making the problem at hand more insightful. Therefore the process modeling is widely used within organizations as a method to increase awareness and knowledge of business processes and to deconstruct organizational complexity [1]. In day to day the process discovery is conducted using the process mining. Hence the process mining research, so far has mainly focused on issues related to control flow mining, that is behavioral and operational perspective, different algorithms and advanced mining techniques have been developed and implemented in this context.

Dyeing is a traditional method of imbuing a cloth with color. To produce colored cloth, it necessary to do the coloring process for the cotton yarn. Then that yarn needs to send to different types of machineries to make a cloth. However, the color of dyed yarn is a function of the physical properties of the dye and the fabric. In other words, dyeing process can be completely controlled by the dyer, a person he always works for the dyeing process. Predicting the actual result of dyeing using graphics would be useful for designing the process model with accuracy, especially since the dyeing process is a troublesome task in reality [6]. More people would be interested in the dyeing process if they could better predict the final product.

The dyeing unit has colour processing task to colour the yarn. This process is not simple and static; hence there is a need to control the process completely by any technology and method. So, the process mining is an area which has process model constructs to simplify the process in a better way. In day-to-day the dyeing process conducted using cabinet machineries. Which is a machined used to mix colour, pre treatment, post treatment etc. The process, which is used by the dyeing master in the dyeing unit is not fulfill the need of colour processing, so the process mining can contribute more to help dyeing masters to process their task effectively.

In recent study, we identified that the process mining algorithms is not sufficient for the dyeing process, because of its dynamic in nature. Hence the better process models need to be generated; hence, the alternative process model is generated using association rule mining algorithms and Weka library.

This paper follows the description and explanation of the LinkRuleMiner (LRM) and experimental results in the previous chapter. Hence, this chapter aims at evaluating and analyzing the usefulness and applications of the LRM as it was implemented to obtain simpler process models for the dyeing domain as opposed to the complex models generated by the HM. The Emerald and Jayabala dyeing process are used in this chapter to compare the Apriori, FPGrowth, H-Mine and the newly proposed LRM association mining algorithms to gain some insights in its processes [7].

In this chapter, first the overview of the Emerald and Jayabala dyeing process is described in the Section 2 and illustrated these experiments using association rule mining algorithms and Weka library in the section 3 and concluded in the Section 4.

II. Overview of Emerald and Jayabala Dyeing Process

The data were collected from the Emerald dyeing unit and Jayabala dyeing units to implement the association rule mining algorithms and Weka Library. The database contains records of shades one, two, three and thirteen different shades of Emerald dyeing unit and twenty four, twenty eight, fifty two and hundred different shades of Jayabala dyeing unit colours.

Information is recorded from the dyeing unit for different shades which affect most of the time by problems. The shade data collected for which affected by problems are identified after the 6 hours from the shade problem. After the first 6 hours, the shade is considered to be in the sub-acute phase. The structure of the Emerald and Jayabala dyeing units dyeing process can be seen in the Entity-Relationship diagram given in Figure 1 and 2 respectively. These diagrams presents following facts:

- The shade's data being recorded, the dyeing history is also recorded. It means the treatments that a shade suffered from in the past, treatments and testing measurements prescribed to dyeing master are also kept in the records. The PHtests conducted on the shades as a part of the previous treatments are also recorded.
- Whether a shade is selected to the dyeing while shade is in acute phase or sub-acute phase, this information is recorded accordingly.
- All the measurements (PH tests etc.) and shade parameters (like pre treatment, post treatment etc.) are also recorded.
- All the treatments and testing's conducted on a shade during its treatment process phase are stored in the database.
- After the colouring process is completed the post treatment phase begins and data is then later on recorded for the shades follow up steps due to the shade difference of the dyeing process.

From the database it was also inferred that the number of different ATEs or events per case in Emerald dyeing unit is smaller compared to the number of events per case in Jayabala dyeing unit. This may be due to the fact that the Emerald dyeing process refers to the primary shades where a shade has limited number of treatments were used, whereas the Jayabala dyeing unit refers numerous number of shades, hence the number of different events is more and also the number of cases is less i.e. 600. The data is stored in the MS-Access database. Therefore before it can be used for experiments with plug-ins in ProM, it was converted to the MXML format using the MS-Access import algorithm [8] [9]. Logs for experiments were made for 1) various testing's (dyeing, shade test and PH test, etc.) given to the shades and 2) various measurements done on shades for testing and treatment purposes. The next section illustrates the experiments done on Emerald and Jayabala dyeing units dyeing process to gain insights into the underlying dyeing processes related to various shades.

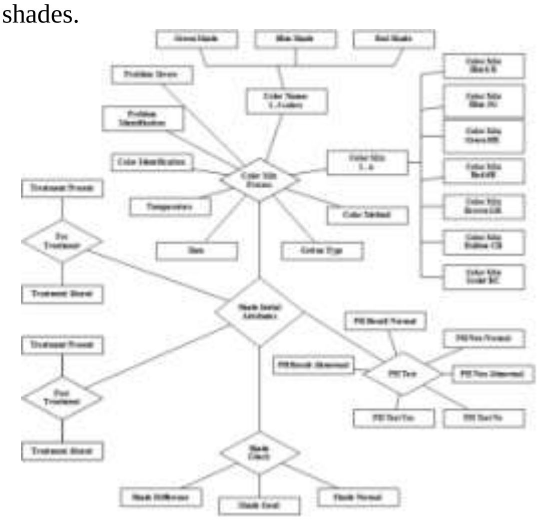

Fig.1 Structure (ER Diagram) of Emerald dyeing process

III. Experimental results of Dyeing processes

To experiment the dyeing process of Emerald dyeing unit and Jayabala dyeing unit, various mining algorithms and association rule mining algorithms are used. The Emerald dyeing unit has four implementations, they are one shade, two shades, three shades and thirteen shades and the Jayabala dyeing unit has four implementations, they are twenty four shades, twenty eight shades, forty eight shades and hundred shades dyeing processes.

Fig.2 Structure (ER Diagram) of Jayabala dyeing process

A. Implementation and study of one shade in Emerald Dyeing Unit

The log used in this experiment consists of information about various measurements (tests) done on Emerald dyeing process shades. This involves three types of measurements viz., *Pre PH test, Post PH test* and *Shade Check* also shown in Figure 1. The log has 203 PIs, 61 different ATEs and 3654 total number of ATEs. These PIs refer to shades that undergo these measurements. Each PI represents one single shade. The various ATEs in the log refer to the 3 types of measurements. Before applying the Apriori algorithm to this log, first discover the process model for the process registered in the log. Figure 3 gives the process model from the HM algorithm.

The process model is not easy to understand and shows that each shade undergoes testing during dyeing process. The shade is first measured for the colour and then the level of matching is recorded (shade check). The model also shows close interrelationship between various shades. For example, *blue shade* may be prescribed to a shade because of the colour depth, so the level of depth recorded during the shade check. This may be due to the fact that depth may affect a shade's normal processing to the extent that shade is unable to match and preprocess etc. by the dyeing expert.

Now the Apriori is applied to this log in order to gain some more insights into this process of measurements. The results of applying the Apriori algorithm with a population size =3654 can be seen in the Figure 5, obtained 10 rules. These rules depict the implicit associations between various measurement activities in the log. For example, Rule 1, 2, 3 and 4 indicates that a shade undergoing C_Name_GreenShade always undergoes Scolet RC, Post_Treat_Absent and PH_Res_abnorm with the confidence 1. The confidence of the rule is 1 indicating that the shade Green undergoes the C_Name_GreenShade also undergoes the Scolet RC, Post_Treat_Absent and PH_Res_abnorm such implicit information is not reflected in the process model in Figure 2.

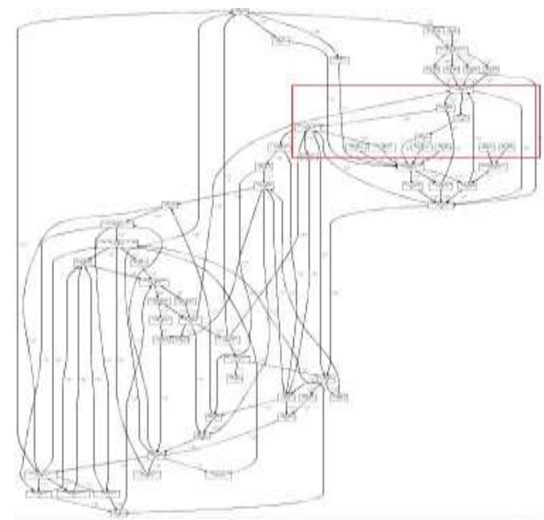

Fig.3 Process Model of one shade dyeing process for Emerald dyeing unit with 203 PIs, 61 different ATEs and 3654 total number of ATEs

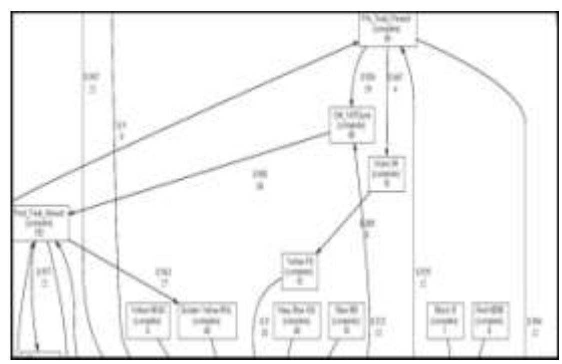

Fig.4 Selected area enlarged of the Fig.3

Rule 8 indicates if a shade undergoes the treatment Post_Treat_Absent, shade also undergoes Scolet RC and C_Name_GreenShade. The confidence of the rule is 0.95 indicating that most of the shades that undergo the Post_Treat_Absent, also undergo the Scolet RC and C_Name_GreenShade. Such implicit information is not reflected in the process model in Figure 2. If this kind of information is available with the stakeholders i.e. the dyeing staff then they can be well prepared with the skills and equipments needed to attend to such situations. This prior information will help them to improve the quality of their dyeing processes by saving their effort, time and cost of the services. Besides this information, some observations were also made from these association rules:

- The association rules also indicate a strong relationship between treatments like Post_Treat_Absent and Scolet RC. It should be noted that these rules do not involve the treatments: Pre_Treat_Absent and PH_Neu_normal. From the log statistics seen in ProM, it is found that the frequency of these measurements is the lowest. Therefore, when the association analysis is performed with lower value of support threshold (upper bound for minimum support is reduced from 1 to 0.85) rules involving these measurements are obtained.

- The presence of the activity Scolet RC in all the generated association rules signifies its importance as compared to other treatments. This is also verified from the log summary. Scolet RC is the most frequent treatment followed by the Post_Treat_Absent.
- For this process, the association rule algorithm captures all the events registered in the log. This may be an indication of the absence of noise or exceptional behavioural pattern in the log because it is quite possible that in case of noise or exceptional dyeing process cases some events would not be captured.

It is observed that the FPGrowth algorithm captures the low frequent activity: PH_Res_abnorm in the first 10 rules it displays. However, it does not capture the lowest frequent activity: PH_Neu_normal even when the number of rules is set quite high (30). This indicates a limitation of the algorithm because if a user gives less number of rules shade would not be able to find rules involving the activity PH_Neu_normal. It is shown in the figure 6. It is quite difficult to know what number of rules should be set so as the low frequent activities or most of the activities registered in the log can be captured by the association rules. In the next subsection, it is described that our experiment with a log pertaining to thirteen shades used for the Emerald dyeing process shades.

B. Implementation and study of two and three shades in Emerald Dyeing Unit

The log used in this experiment stores data about various testing's viz., Pre PH testing, Post PH testing, Shade check etc. It consists of 275 PIs, 62 different ATEs and 4950 total number of ATEs. The process also consists of activities dyeing process treatments indicating the treatments which the Emerald dyeing process shades receiving various testing's suffered from. This again indicates that the Emerald dyeing process treatment process is a complex procedure and requires long time.

Now it is experimented with the H-Mine to see what insights can be gained for the complex two shade dyeing process of Emerald dyeing unit. The result of applying the H-Mine algorithm with a confidence=0.5 and it is obtained 19 association rules with their confidence ranging from 0.98 to 0.81. Such a range of confidence values indicate that even rules with low confidence values indicate some correlation between the testing's or treatments involved. For instance, consider the following association rules:

1.[PH_Neu_normal=yes]:228==>[ShadeCheck_Good=yes]:224 <conf:(0.98)>
2.[ShadeCheck_Good=yes]:244==>[ScoletRC=yes,PH_Neu_normal=yes]:224 <conf:(0.92)>
3.[Scolet RC=yes]: 275 ==> [ShadeCheck_Good=yes]: 244 <conf:(0.89)>
4.[Scolet RC=yes]:275==>[ShadeCheck_Good=yes,PH_Neu_normal=yes]:224 <conf:(0.81)>

A rule 1 and 2 with confidence above 90% indicates that at least 90% of the times when a shade is given the testing are listed. This represents the strong correlation between these testing's. Similarly, considering the Rules 3 can be noted that 89% of the times when the shades undergo Scolet RC treatment and Rule 4 can be noted that 81% of the times, the PH test is conducted, hence this rule undergo 224 events out of 275. For these rules though the confidence value is lower than the confidence for the first two rules but however the latter rules also show a strong "implies" relationship between the various treatments they associate. It indicates that the treatments given to the shades are very much interrelated. This was also confirmed when the H-Mine algorithm was applied to this log. The rules from H-Mine algorithm signify that the treatment process of Emerald dyeing process is a process involving various tasks (treatments) at the same time and a shade may be required to give many testing's in the course of the treatment.

The clustering concept can also be achieved through the process mining algorithms such as HM and DWS. To implement the clustering feature in to association rule mining algorithms, it is necessary to convert the process mining log i.e. MXML file format to ARFF file format. The cluster log is derived from the whole log of the dyeing process using the DWS process mining algorithm. Then the output of DWS is fed in to the HM algorithm. The HM algorithm will produce the process model for the given input. Then the conversion of the process mining process model to association rule mining process model is done using the ProM and Weka Library tools.

To demonstrate the clustering feature of this algorithm, the three shades dyeing process is considered. The three shades dyeing process has 683 PIs, 70 different ATEs and 12,294 total number of ATEs. The above process mining model shown in figure 4 converted into association rule process model using H-Mine algorithm and it is shown in the Figure 7.

C. Implementation and study of thirteen shades in Emerald Dyeing Unit

The newly proposed association rule mining algorithm LRM also provides the functionality to cluster log traces. Clustering can help the user to get smaller process models which represent either an association rule or a frequent itemset (as selected by the user). Below it is shown that how clustering is based on an association rule helped in obtaining a simpler process model. The log has 205 process instances, 75 different ATEs and total number of ATEs of 3690 is used to illustrate this. When the log is mined with the default parameter settings of the LRM algorithm, 10 rules are obtained as shown in Figure 8.

Rule 3 is chosen to cluster the log in two parts: first part which contains all PIs satisfying this rule and the second part which do not satisfy this rule. Forty-seven PIs were found to satisfy this rule. Hence, the differences in the process models before and after clustering are analyzed. Figure 9 gives the process model of the whole log.

It is apparent that the structure of the process in Figure 8 is simple as compared to the complex structure of the entire log in Figure 9. Therefore, simpler models obtained through clustering can be used for gaining insights into the process. It can also provide the support count of the antecedents and consequent itemsets in a rule. In Figure 8, it is seen that one of the frequent itemsets is the *Scolet RC*, *PH_Neu_normal* and the number of process instances in which this itemset occurs in the event log.

This gives us the process instances which satisfy this itemset as well their count. The number of PIs satisfying this cluster is 12 rules. Similarly, the support count for the task *dyeing* process of *PH test* is found to be 100.

Hence, the obtained process model for the whole log used for this experiment, it was seen that it is difficult to trace out a control flow path for the similar “characteristic” shades i.e. the shades undergoing the same care flow path in the process. Therefore the resultant cluster describing an association rule and use it to mine a process model, the resulting process model is a specific and clean model which depicts homogeneity (cf. Figure 8).

This is extremely useful in case of the dyeing domain because it is characterized by less-structured processes. These processes are also not unique as every shade represents a unique case and may or may not follow the same care path as followed by some other shade suffering from the same complication or taking up the same treatment or same test. Such heterogeneity of the cases makes it difficult to find one clear and understandable process model. This is where the LRM and the clustering technique find their importance. After analyzing the LRM on Jayabala dyeing unit, it is established that the association analysis has the potential to gain insights into less structured processes like dyeing.

D. Implementation and study of twenty four and twenty eight shades in Jayabala Dyeing Unit

The Apriori is used for Emerald dyeing process to analyze the performance of the algorithm, it is necessary to execute the algorithm using Weka command line interface, i.e. CLI. Hence, the log from Jayabala dyeing units dyeing process with twenty four shades is taken into the account to measure execution time of the Apriori algorithm. Therefore, this log has more insights into this process of measurements. The twenty four shades dyeing process log has 144 process instances, 102 different ATEs and 2448 total number of ATEs. The experimental result can be seen in the Figure 10, it shows the obtained 10 rules. This log has different types of treatments such as C_Name_ParrotGreen, Pre_Treat_Present, Post_Treat _Absent and PH_Res_abnorm with the confidence 1. The confidence of the rule is 1 indicating that the shade ParrotGreen undergoes the C_Name_ ParrotGreen also undergoes the Scolet RC, Post_Treat_Absent and PH_Res_abnorm such implicit information is not reflected in the process model in Figure 10.

As already mentioned, the Apriori provides the functionality to cluster log traces. Clustering can help the user to get smaller process models which represent either an association rule or a frequent itemset (as selected by the user). Below it is shown that how clustering is based on an association rule helped in obtaining a simpler process model. Hence to demonstrate this algorithm considers the twenty eight dyeing process log from Jayabala dyeing unit. The log has 168 process instances, 84 different ATEs and total number of ATEs of 2856 is used to illustrate this. When the log is mined with the default parameter settings of the FPGrowth algorithm, 10 rules are obtained as shown in Figure 11.

E. Implementation and study of forty eight shades in Jayabala Dyeing Unit

To demonstrate the clustering feature of H-Mine algorithm for Jayabala dyeing unit with forty eight shades dyeing process is considered. The forty eight shades dyeing process has 288 PIs, 155 different ATEs and total number of ATEs is 4896. Figure 12 shows the whole log and the cluster derived from the DWS process mining algorithm is shown in the Figure 13. It is identified that the whole log is has less frequent items and less confidence of each rule than the clustering log.

```
Apriori
=======
Best rules found:

 1. Scolet RC=yes 193 ==> C_Name_GreenShade=yes 193    <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
 2. Post_Treat_Absent=yes 192 ==> C_Name_GreenShade=yes 192    <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
 3. Scolet RC=yes Post_Treat_Absent=yes 182 ==> C_Name_GreenShade=yes 182    <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
 4. PH_Res_abnorm=yes 181 ==> C_Name_GreenShade=yes 181    <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
 5. C_Name_GreenShade=yes 203 ==> Scolet RC=yes 193    <conf:(0.95)> lift:(1) lev:(0) [0] conv:(0.91)
 6. Post_Treat_Absent=yes 192 ==> Scolet RC=yes 182    <conf:(0.95)> lift:(1) lev:(0) [0] conv:(0.86)
 7. Post_Treat_Absent=yes C_Name_GreenShade=yes 192 ==> Scolet RC=yes 182    <conf:(0.95)> lift:(1) lev:(0) [0] conv:(0.86)
 8. Post_Treat_Absent=yes 192 ==> Scolet RC=yes C_Name_GreenShade=yes 182    <conf:(0.95)> lift:(1) lev:(0) [0] conv:(0.86)
 9. C_Name_GreenShade=yes 203 ==> Post_Treat_Absent=yes 192    <conf:(0.95)> lift:(1) lev:(0) [0] conv:(0.92)
10. Scolet RC=yes 193 ==> Post_Treat_Absent=yes 182    <conf:(0.94)> lift:(1) lev:(0) [0] conv:(0.87)
```

Fig.5 Association rules using Apriori algorithm for various testing’s conducted on the Emerald dyeing process with 203 PIs, 61 different ATEs and 3654 total number of ATEs.

```
FPGrowth found 12 rules (displaying top 10)

 1. [Scolet RC=yes]: 193 ==> [C_Name_GreenShade=yes]: 193   <conf:(1)> lift:(1) lev:(0) conv:(0)
 2. [Post_Treat_Absent=yes]: 192 ==> [C_Name_GreenShade=yes]: 192   <conf:(1)> lift:(1) lev:(0) conv:(0)
 3. [PH_Res_abnorm=yes]: 181 ==> [C_Name_GreenShade=yes]: 181   <conf:(1)> lift:(1) lev:(0) conv:(0)
 4. [Scolet RC=yes, Post_Treat_Absent=yes]: 182 ==> [C_Name_GreenShade=yes]: 182   <conf:(1)> lift:(1) lev:(0) conv:(0)
 5. [C_Name_GreenShade=yes]: 203 ==> [Scolet RC=yes]: 193   <conf:(0.95)> lift:(1) lev:(0) conv:(0.91)
 6. [Post_Treat_Absent=yes]: 192 ==> [Scolet RC=yes]: 182   <conf:(0.95)> lift:(1) lev:(0) conv:(0.86)
 7. [Post_Treat_Absent=yes]: 192 ==> [C_Name_GreenShade=yes, Scolet RC=yes]: 182   <conf:(0.95)> lift:(1) lev:(0) conv:(0.86)
 8. [C_Name_GreenShade=yes, Post_Treat_Absent=yes]: 192 ==> [Scolet RC=yes]: 182   <conf:(0.95)> lift:(1) lev:(0) conv:(0.86)
 9. [C_Name_GreenShade=yes]: 203 ==> [Post_Treat_Absent=yes]: 192   <conf:(0.95)> lift:(1) lev:(0) conv:(0.92)
10. [Scolet RC=yes]: 193 ==> [Post_Treat_Absent=yes]: 182   <conf:(0.94)> lift:(1) lev:(0) conv:(0.87)
```

Fig.6 Association rules using FPGrowth algorithm for various testing’s conducted on the Emerald dyeing process with 203 PIs, 61 different ATEs and 3654 total number of ATEs.

```
HMine found 50 rules (displaying top 10)

 1. [Scolet RC=yes]: 58 ==> [Pre_Treat_Present=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
 2. [Pre_Treat_Present=yes]: 58 ==> [Scolet RC=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
 3. [Scolet RC=yes]: 58 ==> [Post_Treat_Absent=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
 4. [Post_Treat_Absent=yes]: 58 ==> [Scolet RC=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
 5. [Scolet RC=yes]: 58 ==> [PH_Res_abnorm=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
 6. [PH_Res_abnorm=yes]: 58 ==> [Scolet RC=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
 7. [Pre_Treat_Present=yes]: 58 ==> [Post_Treat_Absent=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
 8. [Post_Treat_Absent=yes]: 58 ==> [Pre_Treat_Present=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
 9. [Pre_Treat_Present=yes]: 58 ==> [PH_Res_abnorm=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
10. [PH_Res_abnorm=yes]: 58 ==> [Pre_Treat_Present=yes]: 58   <conf:(1)> lift:(1) lev:(0) conv:(0)
```

Fig.7 Association rule mining Process model using H-Mine algorithm of three shades dyeing process for Emerald dyeing unit of Cluster shown in figure 4

```
LinkRuleMiner found 12 rules (displaying top 10)

 1. [Scolet RC=yes]: 51 ==> [PH_Res_abnorm=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
 2. [PH_Res_abnorm=yes]: 51 ==> [Scolet RC=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
 3. [Scolet RC=yes]: 51 ==> [PH_Neu_normal=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
 4. [PH_Neu_normal=yes]: 51 ==> [Scolet RC=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
 5. [PH_Res_abnorm=yes]: 51 ==> [PH_Neu_normal=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
 6. [PH_Neu_normal=yes]: 51 ==> [PH_Res_abnorm=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
 7. [Scolet RC=yes]: 51 ==> [PH_Res_abnorm=yes, PH_Neu_normal=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
 8. [PH_Res_abnorm=yes]: 51 ==> [Scolet RC=yes, PH_Neu_normal=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
 9. [Scolet RC=yes, PH_Res_abnorm=yes]: 51 ==> [PH_Neu_normal=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
10. [PH_Neu_normal=yes]: 51 ==> [Scolet RC=yes, PH_Res_abnorm=yes]: 51   <conf:(1)> lift:(1) lev:(0) conv:(0)
```

Fig.8 Association rule mining Process model using LRM algorithm of thirteen shades dyeing process for Emerald dyeing unit of Cluster log with 205 PIs, 75 different ATEs and 3690 total numbers of ATEs.

```
LinkRuleMiner found 12 rules (displaying top 10)

 1. [CT_highcont=yes]: 136 ==> [Scolet RC=yes]: 136   <conf:(1)> lift:(1.05) lev:(0.03) conv:(6.63)
 2. [Temp_medium=yes]: 132 ==> [Scolet RC=yes]: 128   <conf:(0.97)> lift:(1.02) lev:(0.01) conv:(1.29)
 3. [PHTest_yes=yes]: 134 ==> [Scolet RC=yes]: 129   <conf:(0.96)> lift:(1.01) lev:(0.01) conv:(1.09)
 4. [PH_Res_abnorm=yes]: 178 ==> [Scolet RC=yes]: 168   <conf:(0.94)> lift:(0.99) lev:(-0.01) conv:(0.79)
 5. [ShadeCheck_Good=yes]: 175 ==> [Scolet RC=yes]: 165   <conf:(0.94)> lift:(0.99) lev:(-0.01) conv:(0.78)
 6. [Blue BR=yes]: 166 ==> [Scolet RC=yes]: 156   <conf:(0.94)> lift:(0.99) lev:(-0.01) conv:(0.74)
 7. [Post_Treat_Absent=yes]: 158 ==> [Scolet RC=yes]: 148   <conf:(0.94)> lift:(0.98) lev:(-0.01) conv:(0.7)
 8. [PH_Res_abnorm=yes, Blue BR=yes]: 149 ==> [Scolet RC=yes]: 139   <conf:(0.93)> lift:(0.98) lev:(-0.01) conv:(0.66)
 9. [PHTest_yes=yes]: 134 ==> [ShadeCheck_Good=yes]: 125   <conf:(0.93)> lift:(1.09) lev:(0.05) conv:(1.96)
10. [PH_Res_abnorm=yes, ShadeCheck_Good=yes]: 148 ==> [Scolet RC=yes]: 138   <conf:(0.93)> lift:(0.98) lev:(-0.01) conv:(0.66)
```

Fig.9 Association rule mining Process model using LRM algorithm of thirteen shades dyeing process for Emerald dyeing unit of Whole log with 205 PIs, 75 different ATEs and 3690 total numbers of ATEs.

```
Apriori
=======

Best rules found:

 1. Pre_Treat_Present=yes 102 ==> Post_Treat_Absent=yes 102    <conf:(1)> lift:(1.13) lev:(0.08) [12] conv:(12.04)
 2. PH_Res_abnorm=yes Pre_Treat_Present=yes 89 ==> Post_Treat_Absent=yes 89    <conf:(1)> lift:(1.13) lev:(0.07) [10] conv:(10.51)
 3. PH_Neu_abnorm=yes 82 ==> Post_Treat_Absent=yes 82    <conf:(1)> lift:(1.13) lev:(0.07) [9] conv:(9.68)
 4. Temp_medium=yes 89 ==> PH_Res_abnorm=yes 88    <conf:(0.99)> lift:(1.09) lev:(0.05) [7] conv:(4.02)
 5. Sev_medium=yes 87 ==> ShadeCheck_Good=yes 85    <conf:(0.98)> lift:(1.19) lev:(0.1) [13] conv:(5.24)
 6. PH_Res_abnorm=yes Sev_medium=yes 81 ==> ShadeCheck_Good=yes 79    <conf:(0.98)> lift:(1.19) lev:(0.09) [12] conv:(4.88)
 7. Sev_medium=yes 87 ==> PH_Res_abnorm=yes 81    <conf:(0.93)> lift:(1.02) lev:(0.01) [1] conv:(1.12)
 8. ShadeCheck_Good=yes Sev_medium=yes 85 ==> PH_Res_abnorm=yes 79    <conf:(0.93)> lift:(1.02) lev:(0.01) [1] conv:(1.1)
 9. PHTest_yes=yes 109 ==> PH_Res_abnorm=yes 99    <conf:(0.91)> lift:(1) lev:(0) [0] conv:(0.89)
10. Sev_medium=yes 87 ==> PH_Res_abnorm=yes ShadeCheck_Good=yes 79    <conf:(0.91)> lift:(1.25) lev:(0.11) [15] conv:(2.62)

=== Evaluation ===

Elapsed time: 0.234s
```

Fig.10 Association rule mining Process model using Apriori algorithm of twenty four shades dyeing process for Jayabala dyeing unit of Whole log with 144 PIs, 102 different ATEs and 2448 total numbers of ATEs.

```
> java weka.associations.FPGrowth -t d:\shade28.arff

FPGrowth found 15 rules (displaying top 10)

 1. [Post_Treat_Absent=yes, Temp_medium=yes]: 41 ==> [ShadeCheck_Good=yes]: 41    <conf:(1)> lift:(1.2) lev:(0.07) conv:(6.83)
 2. [Post_Treat_Absent=yes, CT_highcont=yes, PH_Neu_normal=yes]: 41 ==> [ShadeCheck_Good=yes]: 41    <conf:(1)> lift:(1.2) lev:(0.07) conv:(6.83)
 3. [PH_Neu_normal=yes, Pre_Treat_Absent=yes]: 45 ==> [ShadeCheck_Good=yes]: 44    <conf:(0.98)> lift:(1.17) lev:(0.06) conv:(3.75)
 4. [Post_Treat_Absent=yes, PH_Neu_normal=yes]: 55 ==> [ShadeCheck_Good=yes]: 53    <conf:(0.96)> lift:(1.16) lev:(0.07) conv:(3.06)
 5. [PH_Neu_normal=yes, CM_ReactiveDyes=yes]: 48 ==> [ShadeCheck_Good=yes]: 46    <conf:(0.96)> lift:(1.15) lev:(0.06) conv:(2.67)
 6. [Post_Treat_Absent=yes, PHTest_yes=yes, PH_Neu_normal=yes]: 43 ==> [ShadeCheck_Good=yes]: 41    <conf:(0.95)> lift:(1.14) lev:(0.05) conv:(2.39)
 7. [Post_Treat_Absent=yes, PH_Neu_normal=yes, MIX_Others=yes]: 43 ==> [ShadeCheck_Good=yes]: 41    <conf:(0.95)> lift:(1.14) lev:(0.05) conv:(2.39)
 8. [CT_highcont=yes, PH_Neu_normal=yes]: 52 ==> [ShadeCheck_Good=yes]: 49    <conf:(0.94)> lift:(1.13) lev:(0.06) conv:(2.17)
 9. [ShadeCheck_Good=yes, Pre_Treat_Absent=yes]: 47 ==> [PH_Neu_normal=yes]: 44    <conf:(0.94)> lift:(1.29) lev:(0.1) conv:(3.23)
10. [ShadeCheck_Good=yes, CT_highcont=yes, PHTest_yes=yes]: 46 ==> [Post_Treat_Absent=yes]: 43    <conf:(0.93)> lift:(1.21) lev:(0.07) conv:(2.59)

=== Evaluation ===

Elapsed time: 0.062s
```

Fig.11 Association rule mining Process model using FPGrowth algorithm of twenty eight shades dyeing process for Jayabala dyeing unit of Whole log with 168 PIs, 84 different ATEs and total number of ATEs of 2856

```
HMine found 22 rules (displaying top 10)

 1. [Pre_Treat_Absent=yes]: 141 ==> [PH_Res_abnorm=yes]: 141    <conf:(1)> lift:(1.14) lev:(0.06) conv:(17.14)
 2. [CM_ReactiveDyes=yes]: 165 ==> [ShadeCheck_Good=yes]: 165    <conf:(1)> lift:(1.21) lev:(0.1) conv:(28.07)
 3. [PH_Res_abnorm=yes, CM_ReactiveDyes=yes]: 149 ==> [ShadeCheck_Good=yes]: 149    <conf:(1)> lift:(1.21) lev:(0.09) conv:(25.35)
 4. [ShadeCheck_Good=yes, Pre_Treat_Absent=yes]: 140 ==> [PH_Res_abnorm=yes]: 140    <conf:(1)> lift:(1.14) lev:(0.06) conv:(17.01)
 5. [Post_Treat_Absent=yes, CM_ReactiveDyes=yes]: 131 ==> [ShadeCheck_Good=yes]: 131    <conf:(1)> lift:(1.21) lev:(0.08) conv:(22.29)
 6. [PH_Neu_normal=yes, CM_ReactiveDyes=yes]: 130 ==> [ShadeCheck_Good=yes]: 130    <conf:(1)> lift:(1.21) lev:(0.08) conv:(22.12)
 7. [Post_Treat_Absent=yes, PH_Neu_normal=yes]: 144 ==> [ShadeCheck_Good=yes]: 143    <conf:(0.99)> lift:(1.2) lev:(0.08) conv:(12.25)
 8. [Pre_Treat_Absent=yes]: 141 ==> [ShadeCheck_Good=yes]: 140    <conf:(0.99)> lift:(1.2) lev:(0.08) conv:(11.99)
 9. [Pre_Treat_Absent=yes]: 141 ==> [PH_Res_abnorm=yes, ShadeCheck_Good=yes]: 140    <conf:(0.99)> lift:(1.39) lev:(0.14) conv:(20.32)
10. [PH_Res_abnorm=yes, Pre_Treat_Absent=yes]: 141 ==> [ShadeCheck_Good=yes]: 140    <conf:(0.99)> lift:(1.2) lev:(0.08) conv:(11.99)

=== Evaluation ===

Elapsed time: 0.156s
```

Fig.12 Association rule mining Process model using H-Mine algorithm of forty eight shades dyeing process for Jayabala dyeing unit of Whole log with 288 PIs, 155 different ATEs and total number of ATEs of 4896

```
> java weka.associations.HMine -t d:\Shade48cluster1.arff

HMine found 12 rules (displaying top 10)

 1. [Pre_Treat_Present=yes]: 30 ==> [Post_Treat_Absent=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)
 2. [Post_Treat_Absent=yes]: 30 ==> [Pre_Treat_Present=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)
 3. [Pre_Treat_Present=yes]: 30 ==> [PH_Res_normal=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)
 4. [PH_Res_normal=yes]: 30 ==> [Pre_Treat_Present=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)
 5. [Post_Treat_Absent=yes]: 30 ==> [PH_Res_normal=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)
 6. [PH_Res_normal=yes]: 30 ==> [Post_Treat_Absent=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)
 7. [Pre_Treat_Present=yes]: 30 ==> [Post_Treat_Absent=yes, PH_Res_normal=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)
 8. [Post_Treat_Absent=yes]: 30 ==> [Pre_Treat_Present=yes, PH_Res_normal=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)
 9. [Pre_Treat_Present=yes, Post_Treat_Absent=yes]: 30 ==> [PH_Res_normal=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)
10. [PH_Res_normal=yes]: 30 ==> [Pre_Treat_Present=yes, Post_Treat_Absent=yes]: 30    <conf:(1)> lift:(1) lev:(0) conv:(0)

=== Evaluation ===

Elapsed time: 0.0s
```

Fig.13 Association rule mining Process model using H-Mine algorithm of forty eight shades dyeing process for Jayabala dyeing unit of Cluster log with 288 PIs, 155 different ATEs and total number of ATEs of 4896

```
LinkRuleMiner found 20 rules (displaying top 10)

 1. [Post_Treat_Absent=yes, PH_Neu_normal=yes]: 301 ==> [ShadeCheck_Good=yes]: 297   <conf:(0.99)> lift:(1.17) lev:(0.07) conv:(9.23)
 2. [Pre_Treat_Absent=yes]: 305 ==> [PH_Res_abnorm=yes]: 299   <conf:(0.98)> lift:(1.11) lev:(0.05) conv:(4.54)
 3. [ShadeCheck_Good=yes, Pre_Treat_Absent=yes]: 294 ==> [PH_Res_abnorm=yes]: 288   <conf:(0.98)> lift:(1.1) lev:(0.05) conv:(4.78)
 4. [CT_highcont=yes, PH_Neu_normal=yes]: 287 ==> [ShadeCheck_Good=yes]: 281   <conf:(0.98)> lift:(1.16) lev:(0.06) conv:(6.29)
 5. [PH_Neu_normal=yes]: 387 ==> [ShadeCheck_Good=yes]: 376   <conf:(0.97)> lift:(1.15) lev:(0.08) conv:(4.95)
 6. [CM_ReactiveDyes=yes]: 339 ==> [ShadeCheck_Good=yes]: 329   <conf:(0.97)> lift:(1.15) lev:(0.07) conv:(4.73)
 7. [PH_Res_abnorm=yes, PH_Neu_normal=yes]: 348 ==> [ShadeCheck_Good=yes]: 337   <conf:(0.97)> lift:(1.14) lev:(0.07) conv:(4.45)
 8. [PHTest_yes=yes, PH_Neu_normal=yes]: 316 ==> [ShadeCheck_Good=yes]: 306   <conf:(0.97)> lift:(1.14) lev:(0.06) conv:(4.4)
 9. [PH_Res_abnorm=yes, CM_ReactiveDyes=yes]: 307 ==> [ShadeCheck_Good=yes]: 297   <conf:(0.97)> lift:(1.14) lev:(0.06) conv:(4.28)
10. [PH_Res_abnorm=yes, PHTest_yes=yes, PH_Neu_normal=yes]: 282 ==> [ShadeCheck_Good=yes]: 272   <conf:(0.96)> lift:(1.14) lev:(0.06) conv:(3.93)

=== Evaluation ===

Elapsed time: 0.124s
```

Fig.14 Association rule mining Process model using LRM algorithm of hundred shades dyeing process for Jayabala dyeing unit of Whole log with 600 PIs, 262 different ATEs and total number of ATEs of 10,200

F. Implementation and study of hundred shades in Jayabala Dyeing Unit

The hundred shades dyeing process of Jayabala dyeing unit computes the association rules in 0.124 seconds using LRM algorithm. This has the 10 different rules and each represents the different confidence, lift values. It is shown in the Figure 14. Hence the LRM algorithm has good performance than rest of the association process mining algorithms to generate process model.

IV. CONCLUSION

In this paper, we experiment the dyeing process of the Emerald and Jayabala dyeing unit's using the association rule mining algorithms and Weka Library. Hence, it is analyzed the association rules and understood the strengths of these rules indicated by confidence and predictive accuracy for the association mining algorithms such as Apriori, FPGrowth, H-Mine and LRM were analyzed in the view of performance. Therefore the outcome of these algorithms generates process models. The generated process models were analyzed and the importance of f these algorithms were discussed. In spite of the limitations of the association rule algorithms, the clustering approach can be conducted through Weka Library. The various dyeing processes were analyzed to generate the process model even it has the complex and less-structured processes. Also these algorithms were used to determine homogeneous shades, color mix processes with various treatments. Therefore, this paper contributes more on by implementing the dyeing process using association rule mining algorithms with Weka Library.

REFERENCES

[1] Wasana Bandara, "Bibliography of Process Modeling: An emerging Research Field", School of Information systems, Queensland University of Technology, 2005, pp. 11-12.

[2] Dehnert. J, van der Aalst. W.M.P, "Bridging the Gap between Business Models and Workflow Specifications", In International Journal of Cooperative Information Systems", 2004, pp. 289-332.

[3] Keller. G, Meinhardt. S, *SAP R/3 Analyzer: Business Process Reengineering based on the R/3 Reference Mode*, SAP North-America, Philadelphia, Pennsylvania, 1994.

[4] Murata. T, Petri Nets Properties, Analysis and Applications, In Proceeding of the IEEE, 1989, pp. 541-580.

[5] Andrews. T, Curbera.F, Dholakia.H, Goland.Y, Klein.J, Leymann.F, Liu. K, Roller.D, Smith.D, Thatte.S, Trickovic.I, Weerawarana.S, "Business Process Execution Language for Web Services", Version 1.1, 2003.

[6] Y. Morimoto, R. Tsuruno and K. Tomimatsu, "Liquid Penetration Modeling for Cloth Dyeing", Journal of Computational Aesthetics in Graphics, Visualization, and Imaging, Euro graphics Publication Board, 2007, Vol. No 2, pp. 121- 137.

[7] J. Herbst, *A Machine Learning Approach to Workflow Management*, In Proceedings 11th European Conference on Machine Learning, volume 1810 of Lecture Notes in Computer Science, Springer-Verlag, Berlin, 2000, pp. 183–194.

[8] R. Agrawal, D. Gunopulos, and F. Leymann, "Mining Process Models from Workflow Logs". In Sixth International Conference on Extending Database Technology, 1998, pp. 469–483.

[9] Dumas. M, van der Aalst. W.M.P. *Process Aware Information Systems: Bridging People and Software through Process Technology*, John Wiley & Sons, Hoboken, New Jersey, 2005.