

## Predictive Analysis For Credit Card Transaction Management System Based On Process Mining Techniques

<sup>1</sup>C. Rohith Bhat, <sup>2</sup>Dr. S. Ramamoorthy,

<sup>1</sup>Research Scholar, Dr.MGR Educational and Research Institute, Maduravoyal, Chennai.

Email: [rohithbhat2000@gmail.com](mailto:rohithbhat2000@gmail.com)

<sup>2</sup>Professor, Department of CSE, Dr.MGR Educational and Research Institute, Madurovoyal, Chennai.

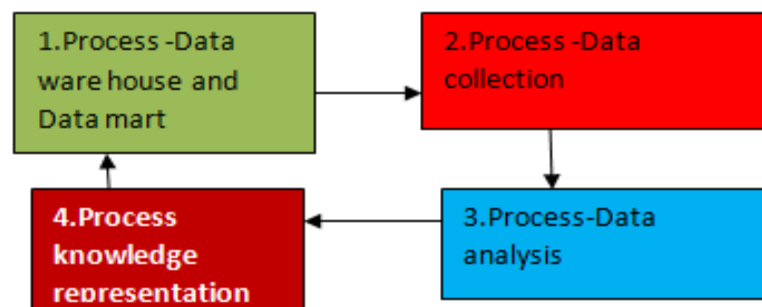
Email: [srm24071959@yahoo.com](mailto:srm24071959@yahoo.com)

**Article History:** Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 16 April 2021

**Abstract:** To attract the customers, business partners and organization, we need to execute processes with high consistency for achieving predictable quality. Now a day's many companies have been focused to improve Business process automation technologies .It can be used to improve the efficiency of both internal processes and customer services. Business process should be performed more effectively with good design process. Most of the organizations are selling their products ,services, pay, purchase of materials and submit their tax through information systems. These systems are used to record information about process execution on event logs .Data mining techniques are used to describe cash payments based on transactions. in this paper for In this paper process mining techniques are used to improve credit card transaction process. Model of the credit card transaction process is described with design constraints and events. In this paper proposed algorithm is introduced to get optimal transaction of credit card transaction process.

### 1. Introduction

Models are used to describe life-cycle of process instance and system configuration can be used to support business processes. Extract the useful information and produce the knowledge for improving the business process of organizations through the analysis process. The main objective of Process mining is to analyze the data and create useful information for improving business process of companies through the data mining tools and techniques[1]. It is placed in between analysis, process modelling and data mining . The term process mining can be used for distilling a structured process method from real-time data sets executions. Stages of architecture of process with the combination of process mining is given below[2].



**Fig 1: Stages of architecture of process based on Process mining**

Stage 1: In this stage understand the business and collect the relevant information about business.

Stage 2: Collect relevant data sets.

Stage 3:Data mining tasks can be applied.( Clustering,Association rules etc)

Stage 4:Extraction of knowledge.

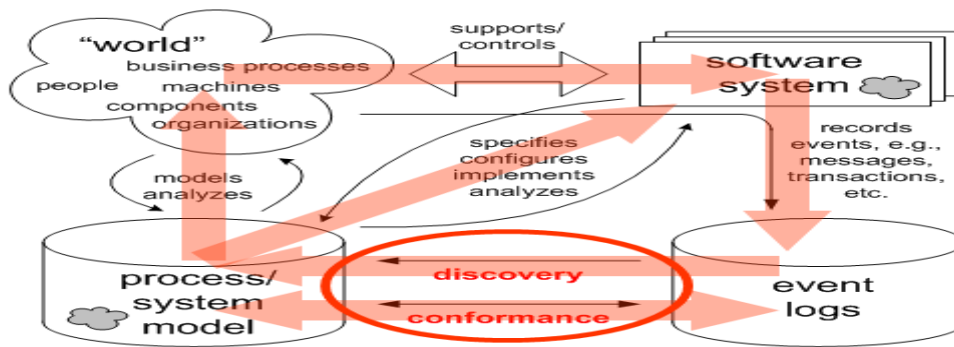


Fig 2. Process mining components[2]

Process mining components are world, software system, process or system model and event log[2]. Business processes, people ,machine components, organizations are included. Our approach is used to execute process logs or event logs by using data mining and data warehousing techniques . Business process automation systems can be applied at the start and completion time of each activity by recording the events. We can extract knowledge about the circumstances in which an exception occurred in the past, and use process mining to identify the causes of its occurrence and to predict future occurrences within running process instances. Related work is presented in section 2.In section 3 describes proposed work .Section 4 presents credit card transaction process with process mining and section 5 provide design constraints to achieve optimal process. section 6 gives predictive analysis of credit card transaction process with simulation model. Conclusions are given in remaining section.

## 2.Related work:

M. Pospíšil.etal provide detail information about “Process Mining in manufacturing Company[1],” W. M. P. Vander Aalst .etal [2] discussed about process mining techniques based on past executions and analyzes how to take decisions on data attributes . [3] [5] A. Rozinat.etal provide discussion on discovering of simulation and give “Workflow simulation for operational decision support”, [6].Maruster [13]proposes a redesign methodology based on process mining and simulation techniques. M. Song.etal proposes, “Time prediction analysis based on process mining”,.[9]

M. Pospisil.etal proposes“Business Process Simulation for Predictions.Shan.etal proposed Improvement of business process quality through exception understanding based prediction methods .Swami and agrawal proposed Association rule for large datasets.Srikant.etal proposed fast algorithms for association rules in large data bases.Business process models and work flow management models are described by vander.Agrawal.etal proposed mining process models from work flow.

## 3.Proposed work:

In this paper we propose a model consists of different stages .It can be used for solving the business problems and improving the business process[4]. Working stages of proposed work are expressed below.

- Understand the Business process
- Collect and prepare the Data
- Prepare a Model based on business process
- Evaluate the model
- Deployment

In this model Process mining can be applied with business techniques. It analyzes the operational data of process and improve the business performance.

### 3.1 Working stages of the proposed model

Data mining techniques are applied to extract the relevant knowledge about business process and analyse the process with dataset information. At first, dataset can be analysed to monitor the architecture of model[5].

#### Stage 1: Understanding’ the business process:

In an organisation, each process consists of group of activities. Each business processes consist of inputs and outputs. Their tasks can be completed with their special suppliers and customer.

The scope of the business problem is related to includes with input parameters, process activities, output parameters, suppliers, customer requirements.

Datasets are taken as input for the proposed solution with Data mining techniques to achieve target values of business process.

**Stage 2: Collect and prepare the Data:** This stage consists of the following attributes.

- 1 Business procedure
- 2 Suppliers
- 3 input parameters
- 4 process activity
- 5 output values
- 6 Customers

In data preparation phase attributes and records are removed with unique values. Here values are merged with nominal attributes and transforming the customer values in to business process of attributes. Clustering and feature selection are used for pre-processing on dataset.

**Stage 3: Prepare a Model based on business process:**

Here datamining techniques are used to describe a model. Classification, association rules, clustering techniques are used for predicting the model[19].

3.1 Technique 1: Clustering analysis: Business processes can be segmented in the form of clusters by using clustering techniques. It includes with process activities of input, supplier, process, output, customer. K-means clustering technique can be applied by taking the following parameters.

- a. Euclidean distance is taken as Distance function
- b. Maximum number of iteration for this algorithm: 60
- c. Number of clusters are taken :3
- d. To perform calculations on datasets process in training process by considering target attributes.

The results are given below:

**Table 1:Data processing by using clustering algorithm**

| Sl.no | Cluster squared errors | Cluster instances |
|-------|------------------------|-------------------|
| A     | 98                     | 45                |
| B     | 64                     | 10                |

**Extract the knowledge for process:**

After applying the clustering algorithm, Cluster instances are 45 and 10. respectively. Cluster squared errors are 98 ,64.Each process activity is separated in each cluster and do transactions are in each cluster .

**Summarisation technique:**

These observations are performed on each process activities of dataset. Process distribution of data values in attributes are taken for doing summarisation and operate the data pre-processing method.

**Technique 4: Association rules**

Association rules are used for providing link between input values, process activities, output values, supplier and customer by using if and then condition rules[16].

Association rule is used to express the formal equation is  $P \rightarrow Q$ , where P,Q are disjoint item sets.  $P \cap Q = \emptyset$ .

Normally Support and Confidence can be taken for determining the strength of association rule . Support provides applicability of rule with dataset. Confidence determines how the items will be Q appear transaction in P.Formal way of the Support is given below.

$$\text{Support, } S(P \rightarrow Q) = \frac{\sigma(P \cup Q)}{N}$$

$$\text{Confidence, } C(P \rightarrow Q) = \frac{\sigma(P \cup Q)}{\sigma(P)}$$

P is antecedent and Q is consequent parts of the rule respectively. Apriori algorithm can be used to create association rules .

Parameters of algorithm are as follows:

minimum support at lower level of boundary :0.1

nature of metric :Confidence

lowest range of Confidence:0.95

no of rules:10

Minimum support at upper level of boundary conditions: 1

The below table provides extraction of knowledge by using apriori algorithm under the form of if-then rules.

**Table 2: Extraction of knowledge by using apriori algorithm**

| S no | Antecedent                       | Consequent                | Confidence |
|------|----------------------------------|---------------------------|------------|
| 1    | Output=6,Business process=BP(10) | Input=12(10)              | 1          |
| 2    | Input=10,Business process=BP(10) | Output=08(10)             | 1          |
| 3    | Input=12,Output=08               | Business process=BP3      | 1          |
| 4    | Business process=BP 13           | Customer=e(9)             | 1          |
| 5    | Customer=d(9)                    | Business process=BP 12(8) | 1          |
| 6    | Business process=BP 5 and (9)    | Customer =a(8)            | 1          |
| 7    | Output=06,Customer=g(9)          | Input=10(8)               | 1          |
| 8    | Input=11,Customer=g(9)           | Output=06(9)              | 1          |
| 9    | Input=11,Customer=g(9)           | Business process=BP3(8)   | 1          |
| 10   | Output=06,Customer=g(9)          | Business process=BP3(9)   | 1          |

Ten important rules are presented in the above table.

For instance, At first rule, output is 6, business process with input of Here ten cases of process activities, the output and business process of have significant association with their inputs.

**Feature selection method:**

In this technique, identify the attributes, and explain hidden knowledge of process dataset.It express the link between the business process. Feature selection can be used in pre-processing. The Feature selection algorithm used to evaluate the function with 'information gain' and search the rank selection method is used to select and rank attributes. In this method to calculate the target attribute . Evaluation method can be used in this algorithm and get 10-fold cross validation[17].

The extracted process knowledge: Attributes of FS algorithm are Customer, output results, input parameters and suppliers.

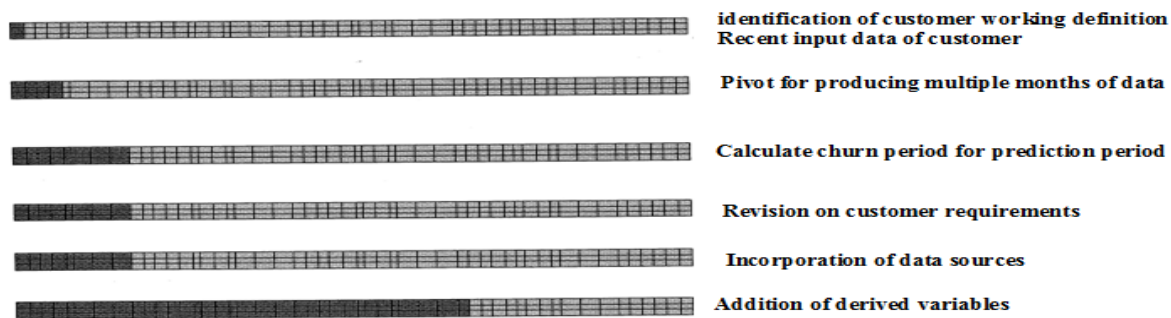
**4.Case study: Credit card transaction processing system.**

Predictive modeling of credit card process consists of following steps.

**Step 1: Format of Data preparation:**

Each row has entity ,customer values which are relevant to business. Particularly, identify the target column and remove all synonymous columns.

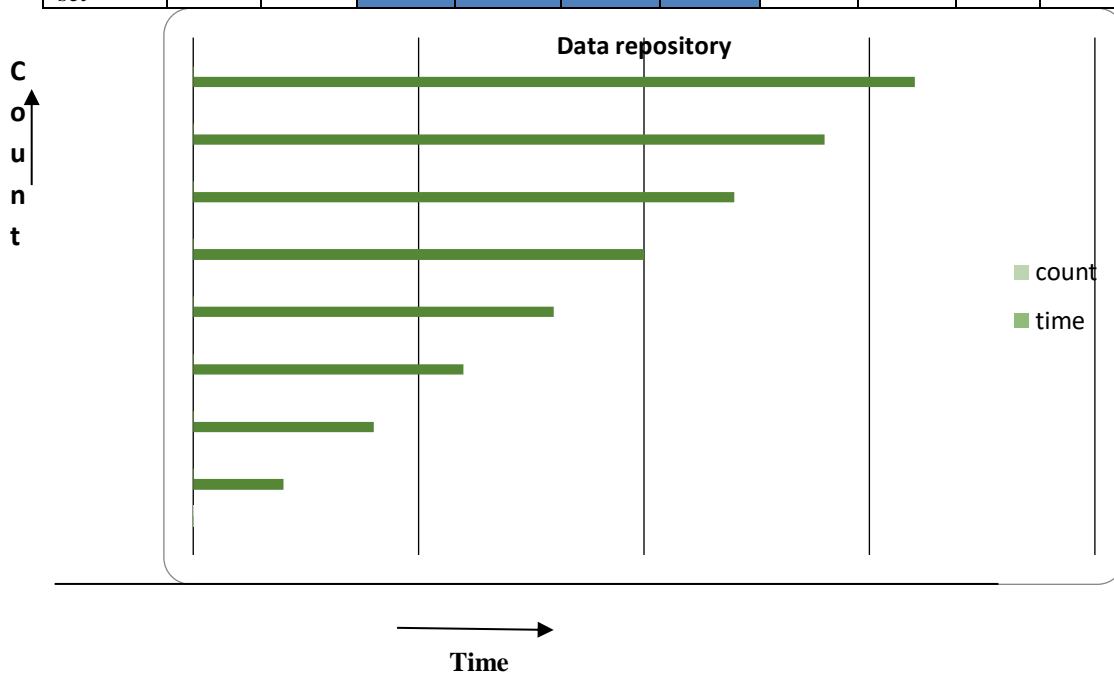
**Second step:** We use customer's signatures for understanding products, zipcodes and counties.



**Fig 3 :Data Catalog**

Time chart for model of data repository for 18 months of customer deposits to achieve target column. Data repository consists of three files. One is Unit\_masterfile,Trans\_master file,Bill\_master file. Time chart for raw data as shown above.

|           | Jan | Feb | Mar | April | May | June | July | Aug | Sep | Oct | Nov |
|-----------|-----|-----|-----|-------|-----|------|------|-----|-----|-----|-----|
| Score     |     |     |     |       |     | 4    | 3    | 2   | 1   |     |     |
| Model set |     |     |     |       | 4   | 3    | 2    | 1   |     |     |     |
| Model set |     |     | 4   | 3     | 2   | 1    |      |     |     |     |     |



**FIG 4 :Data repository**

**SQL Query for data base repository:**

SQL database for look up values in table with outer join .It uses no customer loss in rows.

```
SELECT C.*,1. value
FROM(left customer c in outer join lookup on C .code=1.code)
```

**5.Credit card model:** Model can be built up with system behaviour. This paper describes usage of credit card holders[20].

**Billing process of credit card:** Credit card companies are sending billing amount for every credit card holders once in a month.

**5.1 Design constrains for credit card model:**

Constraint 1: Calculate Balance Payable is the amount of credit card purchases, interest charges-payments and initial balance payments.

Constraint 2: Cash purchases amount is pay check amount and interest on balance amount is calculated in dollars/Month . Credit card purchases is calculated with spending fraction\*available credit Units is in dollars/Month. Credit limit is the maximum value of the "Balance Payable."

Constraint 3: final time = 72 Units: Month The final time for the simulation. Constraint 4: Consider initial balance is '0' and initial time is 0 Units The initial time for the simulation. interest charges is calculated with interest on balance .Calculate payments based on interest charges with respective balance amount.

Quality of life is the amount of cash purchases and credit card purchases.The value of spending fraction=0.1 and time is frequency of output.

**Model of credit card system:**

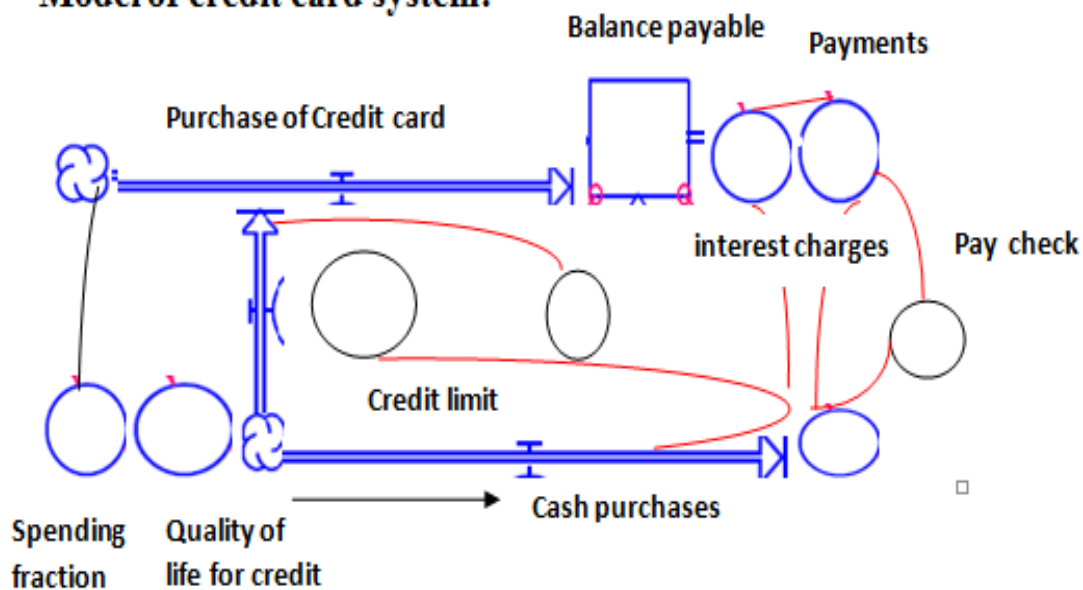


Fig 5 . Model of the credit card system

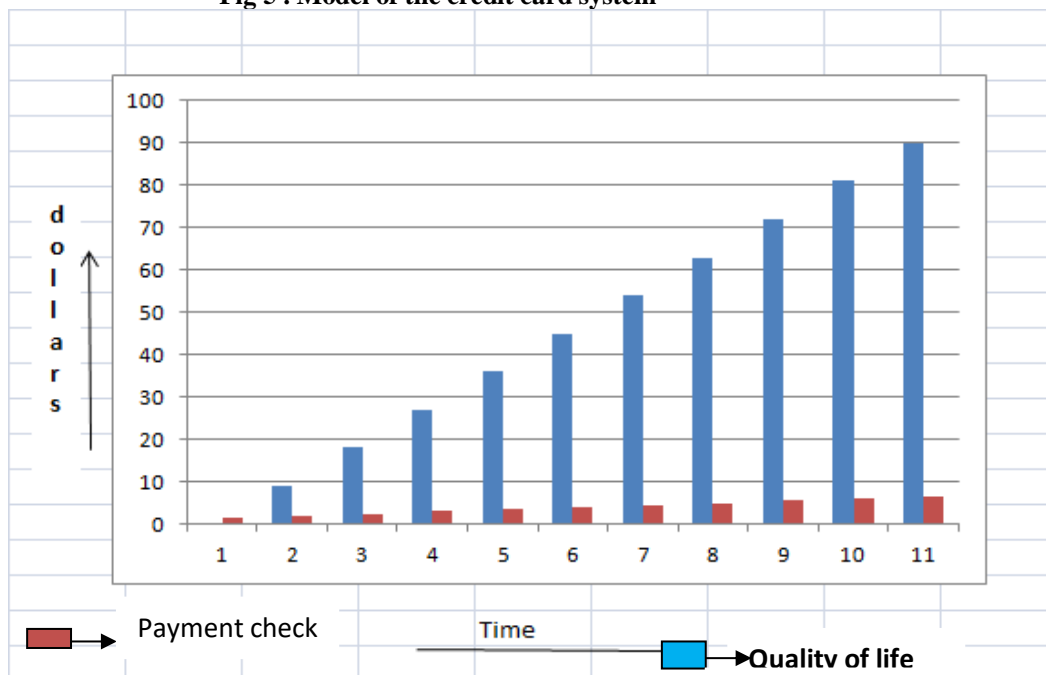


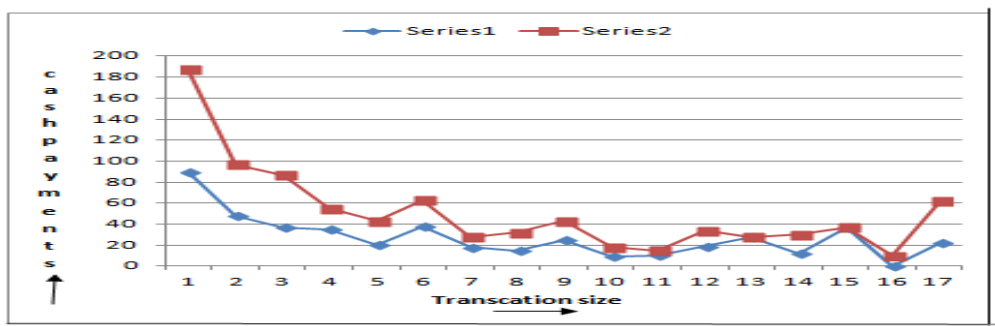
Fig 6. Payment check and quality o

Normally target attribute is taken as the main difference in between our process mining task and the process description[11] .because we will get prediction of based on processing time in data mining task. Regarding this , we will make two decisions – choose the discretization method and the size of intervals created from the target attribute values .

**7. Results and Discussion:**

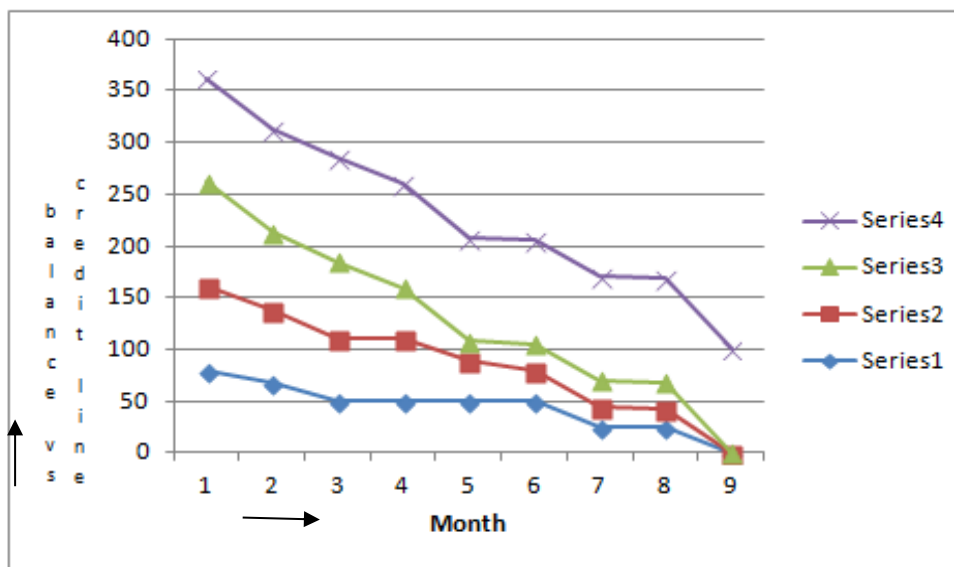
In company ,predictions are used to get lower execution time variance in internal logistics planning process. Precisions is needed to improve the production. Simulation environment is used for results for credit card transaction process. Figure 8 shows the cash payments on different transactions for Canada country and France country. It shows the performance of transaction process with respective cash payments. We produce results in the simulation environment . [21].

In this section procedure and cash transaction size based on economies. The below graph shows the differences in cash payments characteristics and transaction size in Canada and France.



Series 1= Canada region                      Series 2= France region

**Fig 8. Cash payments Vs Transaction size in Canada and France countries**



Here Series 1= Convenience 1 Series 2= Convenience2  
Series 3= ideal convenience                      Series 4= ideal transaction

**Fig 9: Credit card transaction from highest balance to lowest balance amount.**

The above graph shows the Credit card transaction from highest balance to lowest balance amount. The results are produced in simulation environment.

**8. Conclusions:**

In this paper proposed algorithm for improving the business process in banking sectors. process mining techniques can be used to get optimal solution of work flow management problems. Data preparation for transaction process is presented with data mining techniques. In this paper credit card model is proposed with events and describe credit card transaction process in simulation environment.

**References:**

1. M. Pospíšil, V. Mates, and T. Hruška, "Process Mining in Manufacturing Company," in The Fifth International Conference on Information, Process, and Knowledge Management, Nice, France, IARIA, 2013, pp. 143-148, ISBN 978-1-61208-254-7
2. W. M. P. van der Aalst, and A. J. M. M. Weijters, "Process mining: a research agenda", Computers in Industry, Volume 53, Issue 3, Process / Workflow Mining, April 2004, pp. 231- 244, ISSN 0166-3615, DOI: 10.1016/j.compind.2003.10.001
3. W. M. P. van der Aalst, B. F. van Dongen, J. Herbst, L. Maruster, G. Schimm, and A. J. M. M. Weijters, "Workflow mining: A survey of issues and approaches," Data & Knowledge Engineering, Volume 47, Issue 2, November 2003, pp. 237-267, ISSN 0169-023X, DOI: 10.1016/S0169- 023X(03)00066-1.
4. Vinoth Kumar V, Ramamoorthy S, Dhilip Kumar V, Prabu M, Balajee J.M., "Design and Evaluation of Wi-Fi Offloading Mechanism in Heterogeneous Network", International Journal of e-Collaboration (IJeC), IGI Global, 17(1), 2021.
5. W. M. P. van der Aalst, "Process Mining," Berlin, Heidelberg 2011, ISBN 978-3-642-19344-6
6. A. Rozinat, R.S. Mans, M. Song, and W.M.P. van der Aalst, "Discovering simulation models," Information Systems, Volume 34, Issue 3, May 2009, pp. 305-327, ISSN 030
7. A. Rozinat, M.T. Wynn, W.M.P. van der Aalst, A.H.M. terHofstede, and C.J. Fidge, "Workflow simulation for operational decision support", Data & Knowledge Engineering, Volume 68, Issue 9, Sixth International Conference on Business Process Management (BPM 2008)
8. Five selected and extended papers, September 2009, pp. 834-850, ISSN 0169-023X, DOI: 10.1016/j.datak.2009.02.014.
9. W.M.P. van der Aalst, M.H. Schonenberg, and M. Song, "Time prediction based on process mining", Information Systems, Volume 36, Issue 2, Special Issue: Semantic Integration of Data, Multimedia, and Services, April 2011, pp. 450-475, ISSN 0306-4379, DOI:10.1016/j.is.2010.09.001.
10. M. Pospisil., T. Hruška, "Business Process Simulation for Predictions," in BUSTECH 2012: The Second International Conference on Business Intelligence and Technology, Nice,France, IARIA, 2012, pp. 14-18, ISBN 978-1-61208-223-3
11. J. Nakatumba, A. Rozinat, and N. Russell, "Business Process Simulation: How to get it right," Springer-Verlag, 2010, doi:10.1.1.151.834
12. D. Grigori, F. Casati, U. Dayal, and M.C. Shan, "Improving Business Process Quality through Exception Understanding, Prediction, and Prevention," in Proceedings of the 27th VLDB Conference, Roma, Italy, 2001, 1-55860-804-4
13. R. Agrawal, T. Imielinski., A. Swami, "Mining Association Rules Between Sets of Items in Large Databases", Proceedings of the ACM SIGMOD Conference on Management of Data, Washington, USA, 1993, pp. 207-216.
14. R. Agrawal, R. Srikant, "Fast Algorithms for Mining Association Rules in Large Databases," in VLDB '94: Proceedings of the 20th International Conference on Very Large Data Bases, San Francisco, USA, 1994, pp. 487—499.
15. J. Han, J. J. Pei, Y. Yin, "Mining Frequent Patterns without M. L. Antonie, and O. Zaiane, "Text Document Categorization by Term Association," in Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM'02), Maebashi City, Japan, 2002, pp. 19-26.
16. V. Bartík, "Association Based Classification for Relational Data and Its Use in Web Mining," in IEEE Symposium on Computational Intelligence and Data Mining, Nashville, USA, 2009, pp. 252-258
17. W.M.P. van der Aalst, J. Desel, and A. Oberweis, editors. Business Process Management: Models, Techniques, and Empirical Studies, volume 1806 of Lecture Notes in Computer Science. Springer-Verlag, Berlin, 2000.
18. W.M.P. van der Aalst and K.M. van Hee. *Workflow Management: Models, Methods, and Systems*. MIT press, Cambridge, MA, 2002.
19. W.M.P. van der Aalst, A.J.M.M. Weijters, and L. Maruster. *Workflow Mining: Which Processes can be Rediscovered?* BETA Working Paper Series, WP 74, Eindhoven University of Technology, Eindhoven, 2002.
20. R. Agrawal, D. Gunopulos, and F. Leymann. Mining Process Models from Workflow Logs. In *Sixth International Conference on Extending Database Technology*, pages 469–483, 1998.
21. D. Angluin and C.H. Smith. Inductive Inference: Theory and Methods. *Computing Surveys*, 15(3):237–269, 1983.