Ingénierie des Systèmes d'Information
Vol. 27, No. 1, February, 2022, pp. 81-91
Journal homepage: http://iieta.org/journals/isi

Analysis of E-Commerce Process in the Downstream Section of Supply Chain Management Based on Process and Data Mining

Ferra Arik Tridalestari[1,2*], Mustafid[1], Budi Warsito[3], Adi Wibowo[4], Hanung Nindito Prasetyo[5]

[1] Doctoral Program of Information System, School of Postgraduate Studies, Diponegoro University, Semarang 50241, Indonesia
[2] Department of Digital Business, Faculty of Business, Universitas PGRI Yogyakarta, DI Yogyakarta 55182, Indonesia
[3] Department of Statistics, Diponegoro University, Tembalang, Semarang 50275, Indonesia
[4] Department of Informatics, Diponegoro University, Tembalang, Semarang 50257, Indonesia
[5] Department of Information System Diploma, School of Applied Science, Telkom University, Bandung 40257, Indonesia

Corresponding Author Email: ferraarik@upy.ac.id

https://doi.org/10.18280/isi.270110

Received: 15 November 2021
Accepted: 12 February 2022

Keywords:
downstream supply chain management, E-commerce, process mining, data mining

ABSTRACT

Most businesses today use ecommerce stores and/or ecommerce platforms to carry out online marketing and sales activities. The rapid increase in the volume of E-commerce sales transactions normatively causes various problems that occur, especially in this case the buyer or consumer. Consumers expressed dissatisfaction in their e-commerce delivery experience. Customers often complain to sellers in the marketplace about the delay in sending the ordered package. This paper proposes a research model that is proposed in analyzing the datasets generated from the Downstream Supply Chain Management process, especially the process of selling and shipping E-Commerce goods to end customers. The mechanism used is collaborating process mining and data mining so that the resulting analysis becomes more powerful and better information is obtained compared to only analyzing separately. The results of the analysis in the case study of the E-commerce Costumer to Customer (C2C) marketplace show that process mining related to shipping goods can be explained by adding the results of data mining analysis from the datasets obtained, especially the processes in the Downstream Supply Chain Management Section.

1. INTRODUCTION

The Covid-19 Pandemic period has become one of the factors in accelerating digitalization at every line and level, starting from government, education, work behavior, and the most influential, namely in the trade sector [1]. In a trading company, Supply Chain Management has a very important function and role, because it relates to the company's business processes starting from the process of providing raw materials, processing raw material products into finished products, to how the process of distributing these products or services can be accepted by consumers [2]. To be able to increase the company's competitiveness is to make efficiency in terms of supply chains and apply the right information technology. Supply chain management that is assisted by using information technology will provide long-term benefits so that relevant parties in the supply chain can work together and be able to share information. With this Pandemic condition, the company must be sensitive in carrying out the supply chain management process. The need for demand, speed and customer satisfaction is very important for companies to always improve the service processes that are run [3]. No exception in this case are E-commerce companies in Indonesia. Currently, E-Commerce companies in Indonesia are very dominant in trading in national and international markets [4]. In the last decade, E-Commerce Startup companies have been born that are able to reach the Unicorn and Decacorn Levels in the Southeast Asia Region originating from Indonesia such as Tokopedia, Bukalapak and so on [5].

E-commerce is all buying and selling activities carried out through electronic media. Although the facilities include television and telephone, now e-commerce is more common through the internet [6]. In general, customers in C2C e-commerce only retail. The medium for such transactions to occur is known as the marketplace. Marketplace is one of the e-commerce models, where it serves as an intermediary between sellers and buyers [7]. Business models in E-commerce are very diverse, so in this paper the context used is the type of Customer to Customer (C2C) in the E-Commerce Marketplace. In this type of business model, where E-Commerce marketplace users can act as customers and sellers at the same time [8]. Users who sell on the marketplace only need to meet the needs of buyers. All other activities such as managing the website have been taken care of by the platform.

The increase in the growth of sellers in the marketplace is in line with the ever-increasing market demand [9]. The increasing needs of people's lifestyles make the volume of business transactions in the marketplace increase rapidly [10]. The very rapid increase in transaction volume has also caused various problems, especially for the buyer. Most of the fear when shopping online is shipping [11]. Will the ordered goods be delivered immediately? Or worst of all, what if it turns out that the order never arrived at the intended address? Customers often complain to sellers in the marketplace about the delay in the delivery of the ordered package. This raises the question who is at fault, is it the seller or the shipper?

Delivery Process is still the biggest challenge for e-commerce. In a survey conducted by iPrice Group and Parcel Perform, 36% of consumers expressed dissatisfaction with their e-commerce delivery experience. Of course, this complaint cannot be underestimated, especially considering that Indonesia's e-commerce growth is quite high, at 10.3% over the last five years. Consumer dissatisfaction with e-commerce shipments was revealed in a survey by iPrice Group and Parcel Perform, an independent parcel tracking service that tracks more than 600 logistics operators globally. For this survey alone, it involved more than 80,000 consumers in Malaysia, Singapore, Indonesia, Vietnam, and Thailand [12]. The results of the survey show 35% of consumers still consider shipping as the biggest problem in e-commerce. In addition, more than 90% of complaints and negative feedback from customers, usually related to delays in delivery or poor communication about delivery status [13].

There are many studies related to Supply Chain Management, especially in the field of Information Systems. But in the last 5 years not much has been discussed about process mining and data mining. Process mining research in the supply chain was conducted in 2015. In this paper, the researcher tries to investigate the feasibility of applying process mining to model supply chain processes through real case studies. The researcher determined that supply chain business process management as part of the implementation of Business Process Management (BPM) [14]. Then still by the same researcher, the research concept was improved by discussing CRM in 2018. In this publication, the researcher presents how the implementation of process mining is carried out to analyze the process of customer needs. This research is expected to provide a foundation for companies to develop standard procedures, establish service level agreements and propose potential planning processes to better meet customer demands based on CRM systems [15]. In 2018 there was another research publication that presented the application of process mining in the purchasing process from the heavy manufacturing industry. The Event Log of this process is extracted from the company's Supply Chain Management System related to Shipbuilding Processing plans, then filtered and analyzed using process mining techniques. The results of the analysis can be used by companies to optimize and improve processes at manufacturing companies of naval and ship spare parts [16]. There is another interesting research, namely research that tries to improve the manual mechanism and replace it with a process mining approach. The research conducted is to develop a methodology to apply process mining to the internal logistics department for a mixed model assembly line. This methodology combines multi-dimensional process mining (MDPM) techniques with proven production streamlining principles and VSM [17]. Another recent study related to process mining in the supply chain was published in 2021. In this research, the researcher proposes and evaluates the use of the Open Trip Model (OTM) for Process mining. Inspired by the industry's current use of OTM for reporting and business intelligence, researchers believe that the OTM data model can be used for unified storage, integration, interoperability, and querying of logistics event data. Therefore, the OTM model data is mapped to a generic event log structure to meet the minimum process mining requirements [18]. Still in the same year, another study tried to integrate the SIX Sigma concept with process mining. This research aims to describe the practical implementation of the problem solving structure provided by the Define, Measure, Analyze, Improve and Control (DMAIC) framework in combination with the analytical concepts provided by process mining, to improve the quality of the supply chain of health care providers. The mechanism is carried out to identify the existing framework by combining six sigma with process mining [19].

What about Data Mining? In 2017, there was a research publication proposing a food safety warning system, adopting the use of data mining with an association approach and Internet of Things technology to monitor all data in real-time from the entire supply chain and also automatically provide advance warnings. A case study of a dairy producer was conducted and the results show that the proposed early warning system can effectively identify safety risks and accurately determine whether a warning should be issued, depending on expert analysis when an abnormality is detected by the system [20]. In 2019, another study put forward the idea of customer clustering and dynamic customer clustering. A customer clustering dataset is provided to supply the production ranking order. The results show that the customer clustering idea effectively reduces the difficulty in solving modeling. The concept used is able to increase the flexibility of customer grouping [21]. Supply Chain research using data mining was also conducted in the same year. The research conducted focuses on the role of the use of information and proposes a relational model that considers the integration of information in the supply chain with a data mining approach. Complete integration is achieved when all supply chain members or actors share a common goal. This becomes the basis for the integration modeling carried out [22]. In the same year, there were studies that put forward proposals based on extreme distance calculation methods in the field of data mining technology, new supply chain distribution center selection methods combined with clustering algorithms and center of gravity selection methods. In theory, the location of the logistics distribution center involves many factors, and the selection of the distribution center can make a huge difference in the development of the industry [23]. The latest related research between supply chain and data mining is research conducted in 2020. In his research, a holistic approach integrates data mining and risk management activities in a unique framework for effective risk management. This study shows how Data Mining supports finding hidden and useful information from unstructured risk data to make smart risk management decisions [24].

Based on previous studies, there has never been a study that has tried a collaborative approach between process mining and data mining approaches. This paper proposes a process evaluation mechanism for E-Commerce transactions in the Downstream Supply Chain Management section based on process and Data mining. The use of Process mining in this analysis facilitates process analysis in E-Commerce more effectively. By integrating the analysis of Process mining with data mining, better information is obtained compared to only analyzing separately. The analysis of the downstream management supply chain process includes things such as the ordering process, packing and sending packages to end-customers involving sellers and third parties, namely expeditions.

The presentation in this paper is divided into several parts. Section 1 describes the background, problems, other related research, contributions and solutions provided. Section 2 describes the materials, literature, methods, and research paths used in this study, Section 3 shows the results and discussion of the analysis obtained from process mining and data mining from the data obtained as well as discussions related to the

results of the analysis. Section 4 contains recommendations to improve existing processes and the last is Section 5 contains conclusions and future research opportunities from this research.

2. MATERIAL & METHODS

2.1 Supply chain management

Supply Chain Management is a management activity from obtaining raw goods to finished goods to consumers. Supply Chain is divided into 3 main components, namely Upstream Supply Chain, Internal Supply Chain or called workflow, and Downstream Supply Chain [25].

2.2 Process mining & data mining

Process mining is a relatively new science in the last decade. Basically, process mining is a technique that helps organizations analyze and monitor internal processes to identify their activities and resources. In the past, organizations would manage this process through a series of interviews that produced an overall ideal summary of the business. Process mining allows companies to automatically monitor and track business processes in real-time without human intervention [26]. Another important aspect of process mining is that it enables business processes to identify process risks and bottlenecks, determine potential improvements that can improve efficiency, and monitor data to predict future events [27].

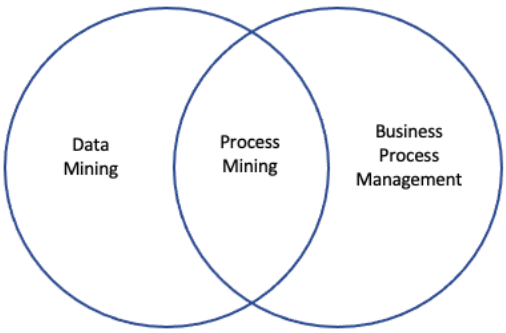

Figure 1. The relationship between process mining and data mining

While data mining is a process of extracting or collecting important information from a large data [28]. The data mining process often uses statistical, mathematical methods or utilizes artificial intelligence technology. Things that are done in data mining activities include data cleaning, data integration, data selection, transformation, data mining, pattern evaluation, and knowledge presentation. Data mining has many functions. There are two main functions, namely descriptive function and predictive function [29]. To see the relationship between process mining and data mining as shown in Figure 1.

Figure 1 shows that process mining and data mining are connected through the exploratory concept of business process management [30]. So, in this case the use of process mining and data mining can be done simultaneously to find a more powerful analysis instead of using only one. The Data Mining concept used in this study is K-Mean Clustering with Davies Bouldin Analysis.

2.3 Methodology

The methodology used in this paper is as described in the following steps:

(1) Identification of data needs

For the needs of this paper, data was collected from one of the E-commerce that serves online sales. The seller uses an E-Commerce system associated with various expedition parties. The system has provided a sales data import feature from the transaction process that occurs.

(2) Generate e-commerce Downstream data

At this stage, it takes data based on a certain period. Data retrieval is in accordance with the needs of the analysis so that the attributes or variables are adjusted to the method of analysis carried out. The mechanism is shown as Figure 2.

Figure 2 shows that after the data retrieval process, the data is then converted into excel data to then be adjusted to the needs of the analysis and tools used in this study.

(3) Pre-processing data

At this stage, data preprocessing is carried out. The data that has been obtained is divided into two parts namely Numeric data & Timestamp Data.

(4) Discovery model process & data mining process

At this stage perform data processing for both timestamp data and numeric data. The tools used in this process are ProM and DISCO tools for measurement mechanisms in process discovery and Excel and Rapidminer tools for data mining mechanisms such as regression and K-Mean Clustering.

(5) Result analysis

This stage analyzes the results from the previous stage to evaluate the Downstream E-Commerce process that occurs.

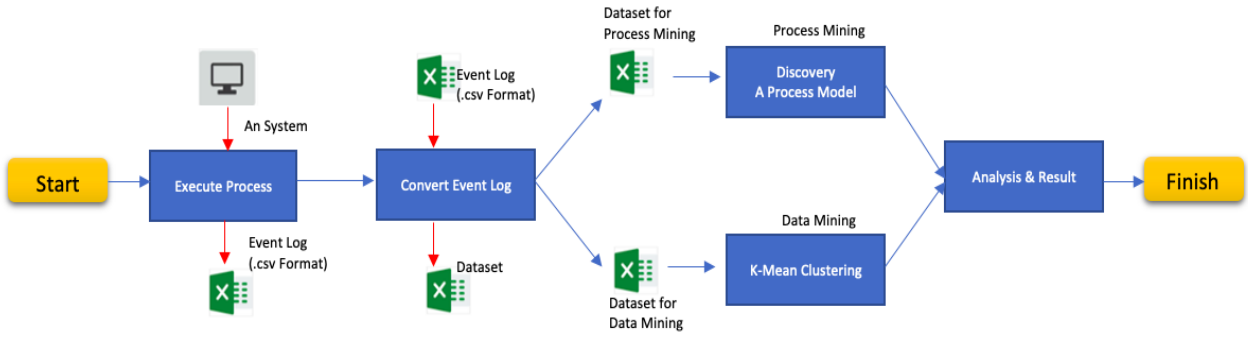

Figure 2. Steps of e-commerce Downstream data collaboration analysis

3. RESULT & DISCUSSION

3.1 Result

As the basis for the analysis is the Downstream SCM E-Commerce process as shown in Figure 3.

Figure 3 shows the processes that occur in Downstream Supply Chain Management. In the picture there are 3 processes involved, namely the ordering process from the customer, the approval process and packing the package from the seller and the shipping process by the expedition company.

The first mechanism is to obtain datasets taken from the running system according to the analysis needs to be used. The dataset is taken from three processes that occur as shown in Figure 3. The total attributes obtained are 24 with various data types such as numeric, character, category, date and time. The attributes obtained are as shown in Table 1.

The next step is to divide the table into 2 groups as shown in Table 2 and Table 3.

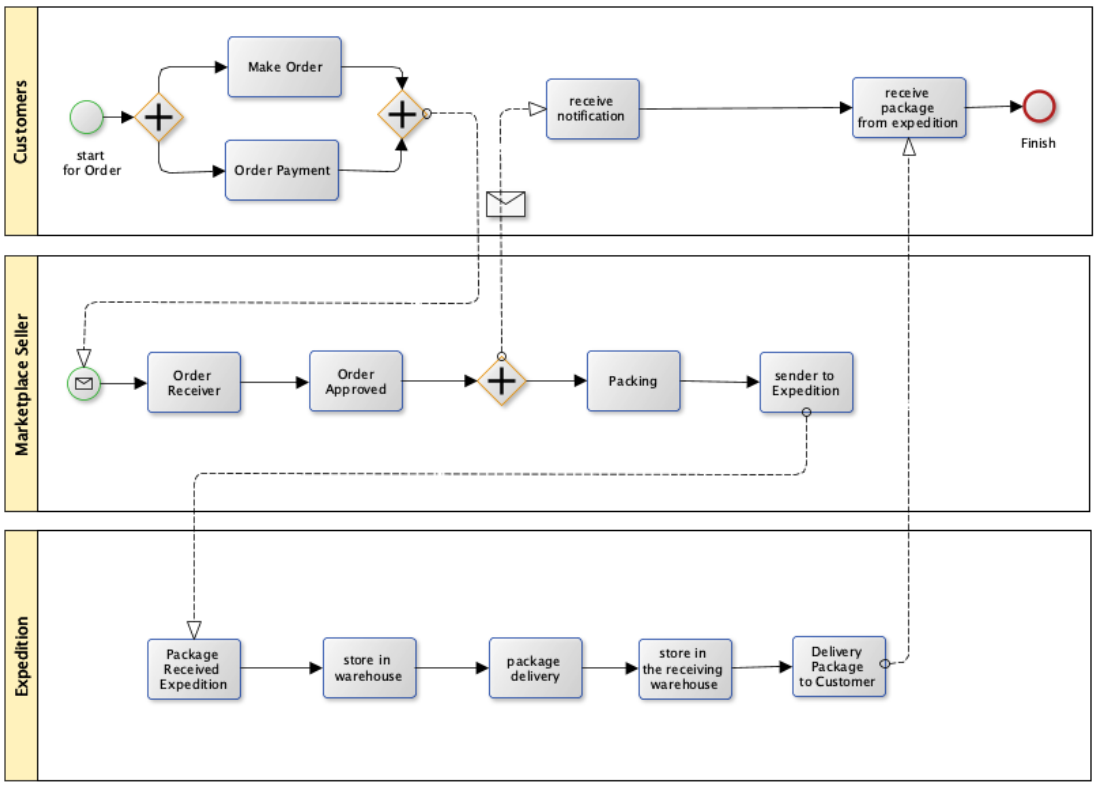

Figure 3. Process in Downstream supply chain management in B2C e-commerce

Table 1. Dataset attributes

Attribute	Data Type
ID Order	Character
Order Status	Category
Make order	Date & Time
No. receipt	Character
Product name	Category
Variation Name	Category
Initial Price	numeric
Amount	numeric
Total Product Price	numeric
Number of Products in Order	numeric
Total payment	numeric
Order Payment	Date & Time
Recipient's name	Character
Shipping address	Character
County town	Character
province	Category
Order Approved	Date & Time
Package Received Expedition	Date & Time
Shipping Warehouse	Date & Time
Package Shipping	Date & Time
Receiving warehouse	Date & Time
Package to Customer	Date & Time
Received by Customer	Date & Time
Done Approved	Date & Time

Table 2. Event Log dataset for data mining

Case ID	Events	Started	Finished	Duration (Minute)	Expedition Type	City
190710123661H8U	7	10.07.2019 11:36:00	13.07.2019 06:29:00	4013	Expedition A	City A
190720055566T5G	7	20.07.2019 04:55:00	24.07.2019 16:32:00	6457	Expedition A	City D
19072506336Q7QX	10	25.07.2019 05:33:00	27.07.2019 18:06:00	3633	Expedition B	City B
190725102165JA2	10	25.07.2019 09:21:00	29.07.2019 17:16:00	6235	Expedition B	City A
19073015476M26F	10	30.07.2019 14:47:00	02.08.2019 08:42:00	3955	Expedition B	City D
19073121496N31V	10	31.07.2019 20:49:00	02.08.2019 21:44:00	2935	Expedition C	City A
190803212265CA5	10	03.08.2019 20:22:00	08.08.2019 14:03:00	6821	Expedition B	City E
19081015286BQV6	10	10.08.2019 14:28:00	15.08.2019 17:16:00	7368	Expedition E	City A
19081119426KSPQ	10	11.08.2019 18:42:00	15.08.2019 21:52:00	5950	Expedition A	City C
19081311556NE7V	10	13.08.2019 10:55:00	20.08.2019 12:49:00	10194	Expedition B	City C
19082118596RDSF	10	21.08.2019 17:59:00	23.08.2019 06:05:00	2166	Expedition C	City A
19082219296HYY2	10	22.08.2019 18:29:00	24.08.2019 14:30:00	2641	Expedition B	City E
1908271408656K2	10	27.08.2019 13:08:00	28.08.2019 14:11:00	1503	Expedition B	City A
19083015266FP38	10	30.08.2019 14:26:00	01.09.2019 20:21:00	3235	Expedition B	City B
…	…	…	…	…		
200129VNVBVB5K	10	29.12.2019 15:26:00	02.02.2020 05:23:00	49797	Expedition B	City B

Table 3. Event log proses e-commerce Downstream supply chain management

ID Case	Activity	Time Stamp
190710123661H8U	Make order	2019-07-10 11:36
190710123661H8U	Order Payment	2019-07-10 11:36
190710123661H8U	Order Approved	2019-07-10 11:36
190710123661H8U	Package Received Expedition	2019-07-10 22:24
190710123661H8U	Package Shipping	2019-07-12 11:28
190710123661H8U	Received by Customer	2019-07-12 18:28
190710123661H8U	Done Approved	2019-07-13 06:29
190720055566T5G	Make order	2019-07-20 04:55
190720055566T5G	Order Payment	2019-07-20 04:58
190720055566T5G	Order Approved	2019-07-20 04:58
190720055566T5G	Package Received Expedition	2019-07-22 22:17
190720055566T5G	Package Shipping	2019-07-24 10:39
190720055566T5G	Received by Customer	2019-07-24 13:39
190720055566T5G	Done Approved	2019-07-24 16:32
19072506336Q7QX	Make order	2019-07-25 05:33
19072506336Q7QX	Order Payment	2019-07-25 05:34
19072506336Q7QX	Order Approved	2019-07-25 05:34
19072506336Q7QX	Package Received Expedition	2019-07-25 10:29
19072506336Q7QX	Shipping Warehouse	2019-07-25 16:27
19072506336Q7QX	Package Shipping	2019-07-25 19:18
19072506336Q7QX	Receiving warehouse	2019-07-27 00:30
…..	…..	…..

3.1.1 Process mining mechanism

In this case, after getting the Event Log, the next step is to model the process using the mining process mechanism using the Inductive Miner Algorithm. The use of the Inductive Miner Algorithm is based on many process mining algorithms. Inductive Miner Algorithm is the most stable algorithm in process modeling [31]. The data description of the event log is shown in Table 4.

Table 4. Event log description

Component	Score
Events	972
Cases	99
Activities	10
Mean Case Duration	4.7 day
Start	10-7-2019 11:36
End	02-02-2020 05:23

Table 4 shows that the total number of events in the Event Log Dataset reached 972 with 9 cases. The dataset period used was from 10-7-2019 to 2-2-2020 with an average duration of each case being 4,7 days which means the length of service for each transaction. For more details, the spread of transactions can be seen in Figure 4. Figure 4 shows that near the end of the year, the transaction volume experienced a significant increase.

Using the ProM Visual Machine Learning application and the Inductive Miner Algorithm, Petri net was obtained from the Downstream E-Commerce process from the existing dataset as shown in Figure 5.

Process performance shows that bottlenecks occur in the shipping process area by expedition as shown in Figure 6 for activities that are colored red. This is as shown in Table 5.

Table 5. Activities that are bottlenecked

Bottleneck Activities	Waiting Time
Sink Package Shipping to Receiving Warehouse	20.37 hours
Receiving Warehouse	21.01 hours
Received by Customer	7.35 hours
Sink package to customer to Receive by Customer	1.38 day
Sink Received by Customer to Done Approved	20.49 hours
Done Approved	23.32 hours

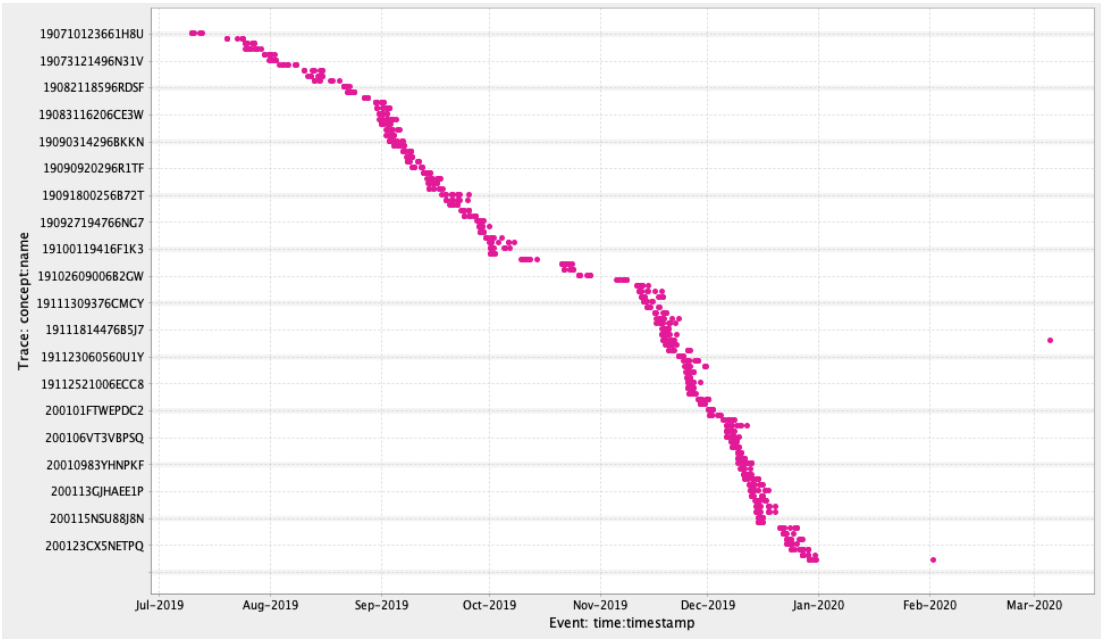

Figure 4. Distribution of transactions in period

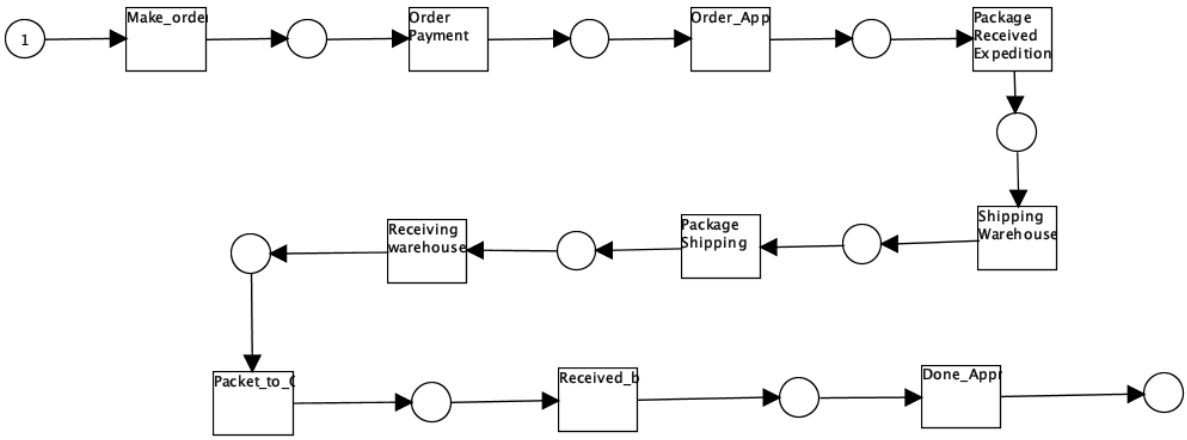

Figure 5. Petri net proses Downstream supply chain management e-commerce C2C

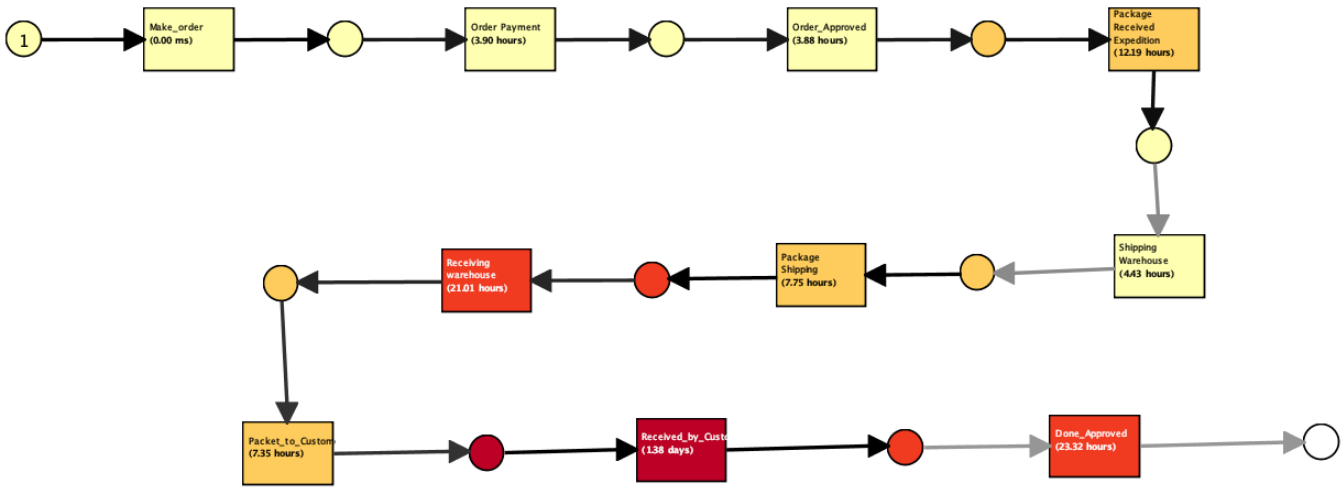

Figure 6. Measurement results of Downstream supply chain management process performance in C2C e-commerce

Table 5 shows the activities that experienced bottlenecks. Activities that experience bottlenecks are in the process of shipping goods carried out by expeditions. Activities that experienced a significant bottleneck were sending packages to the receiving warehouse, receiving packages at the warehouse, and sending packages to end customers. The activity that experienced the longest bottleneck condition was the activity of sending packages from the receiving warehouse to the end

customer which took up to 1.38 days. This activity is known to be a mechanism for distributing packages to end customers in a city by couriers.

3.1.2 Data mining mechanism

The next step is the data mining mechanism on the selected dataset. As previously explained, the dataset used is a special event log dataset for data mining in tracing the delivery process that occurs. The data mining mechanism carried out in this process is to use clustering. The algorithms used are K-Mean Clustering and Davies Bouldin Approach.

With the Davies Bouldin Approach, Cluster selection is carried out by performing 7 iterations, the first iteration (Clustering (k-2)) selecting 2 clusters and then obtaining the Davies Bouldin value, then the second iteration (Clustering (k-3)) selecting 3 clusters until the fifth iteration (Clustering (k-6)) with the number of selected clusters 6.

Determination of the number of clusters is based on Davies Bouldin values, based on data processing using Rapidminer, Davies Bouldin values are obtained as Table 6.

Table 6. Davies Bouldin's value in K-means clustering iteration

No	Number of K	Davies Bouldin's Value
1	2	0.016
2	3	0.031
3	4	0.270
4	5	0.327
5	6	0.361
6	7	0.383
7	8	0.327

Based on the Davies Bouldin value, the Davies Bouldin value drops back to K=8, meaning the number of selected clusters is 7 at the Davies Bouldin value of 0.383. The results of the cluster grouping are as shown in Table 7.

Obtained Cluster as Figure 7.

Table 7. Distribution of data on the 7 selected Clusters

Cluster	Items
0	33
1	1
2	1
3	9
4	33
5	15
6	17
7	23
Totally	251

3.1.3 Cluster division

Based on the data processing of the Cluster division based on indicators namely Duration, Expedition and Province, the order of the clusters is obtained, namely Cluster 0, Cluster 4, Cluster 6, Cluster 3, Cluster 5, Cluster 1, and Cluster 2. For the distribution of data in the cluster can be seen in the cluster grouping in Tables 8-14.

Table 8. Cluster 0

Average Duration	6310.70
Expedition	**Q**
Expedition A	21
Expedition B	4
Expedition C	3
Expedition D	4
Expedition E	1
Expedition F	0
Expedition G	0
Province	
Province A	7
Province B	8
Province C	8
Province D	3
Province E	1
Province F	3
Province G	2
Province H	1
Province I	0

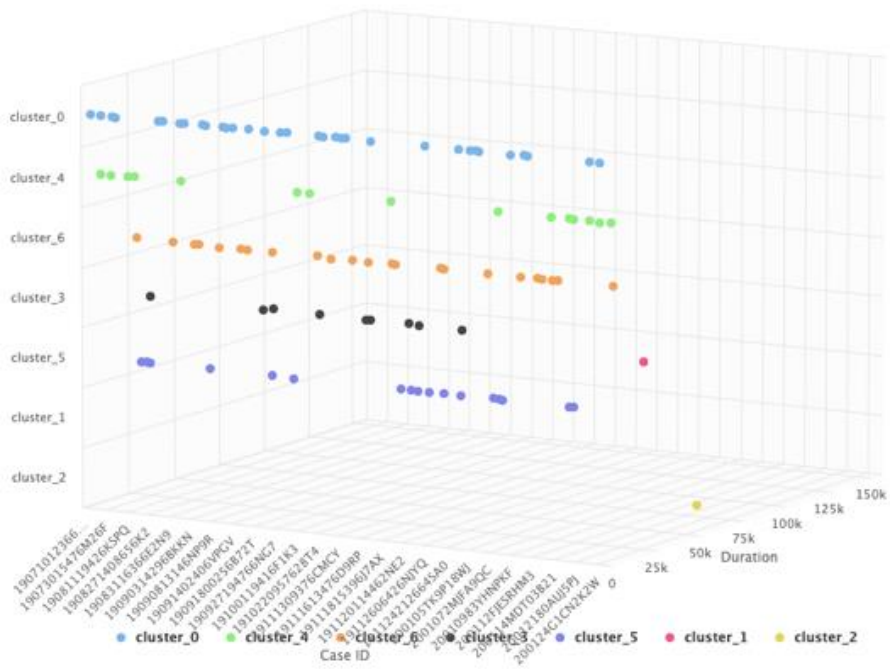

Figure 7. Graph of the determination of K-Mean clustering results on the dataset

Table 9. Cluster 4

Average Duration	**6340.70**
Expedition	**Q**
Expedition A	7
Expedition B	5
Expedition C	1
Expedition D	0
Expedition E	0
Expedition F	1
Expedition G	1
Province	
Province A	5
Province B	2
Province C	1
Province D	1
Province E	1
Province F	0
Province G	1
Province H	2
Province I	2

Table 10. Cluster 6

Average Duration	**6405.62**
Expedition	**Q**
Expedition A	18
Expedition B	4
Expedition C	1
Expedition D	0
Expedition E	0
Expedition F	0
Expedition G	0
Province	
Province A	4
Province B	3
Province C	5
Province D	1
Province E	4
Province F	1
Province G	3
Province H	1
Province I	1

Table 11. Cluster 3

Average Duration	**7280.87**
Expedition	**Q**
Expedition A	8
Expedition B	0
Expedition C	0
Expedition D	1
Expedition E	0
Expedition F	0
Expedition G	0
Province	
Province A	2
Province B	3
Province C	0
Province D	1
Province E	1
Province F	1
Province G	0
Province H	1
Province I	0

Table 12. Cluster 5

Average Duration	**6419.23**
Expedition	**Q**
Expedition A	12
Expedition B	4
Expedition C	0
Expedition D	0
Expedition E	1
Expedition F	0
Expedition G	0
Province	
Province A	5
Province B	4
Province C	4
Province D	2
Province E	1
Province F	1
Province G	0
Province H	0
Province I	0

Table 13. Cluster 1

Average Duration	**155245**
Expedition	**Q**
Expedition A	1
Expedition B	0
Expedition C	0
Expedition D	0
Expedition E	0
Expedition F	0
Expedition G	0
Province	
Province A	1
Province B	0
Province C	0
Province D	0
Province E	0
Province F	0
Province G	0
Province H	0
Province I	0

Table 14. Cluster 2

Average Duration	**49797**
Expedition	**Q**
Expedition A	1
Expedition B	0
Expedition C	0
Expedition D	0
Expedition E	0
Expedition F	0
Expedition G	0
Province	
Province A	0
Province B	0
Province C	0
Province D	1
Province E	0
Province F	0
Province G	0
Province H	0
Province I	0

3.2 Discussion

Based on the approach used in this study, namely the collaborative analysis of mining processes and data mining. The two approaches can reinforce each other in the results of the analysis. In the mining process can be seen the performance of the process and in data mining can be seen the pattern of data distribution. This relationship can be shown in Figure 8.

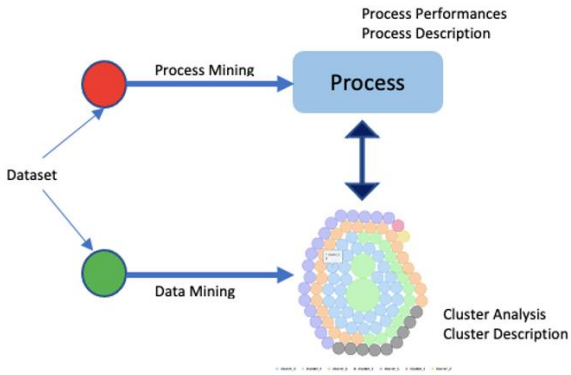

Figure 8. Collaboration point

The results of the analysis of the collaboration show that the problematic activities in the Downstream SCM process in the C2C pattern E-commerce are activities that are in the process of shipping goods by expeditions. Some processes experience bottlenecks as shown in the table. Activities that experienced a significant bottleneck were sending packages to the receiving warehouse, receiving packages at the warehouse, and sending packages to end customers. Then for the pattern of events of the process can be seen in the resulting cluster pattern. Expedition A as an entity that delivers goods carries out extreme delivery durations in clusters 2, 1, 3, and 5. Cities that can be of particular concern to Expedition A are provinces A, B, C, D, and E. Then the results of the regression analysis show that the number of activities in the process does not really affect the duration of delivery. The results of regression analysis were carried out to see the effect of the number of activities on the process on the duration of delivery. This is done because in the delivery process there are differences in the number of activities. There is a delivery process that does not go through the storage warehouse and is sent directly to the customer and there is a process that goes through the storage warehouse for delivery and the warehouse for receiving incoming goods.

4. RECOMMENDATION

Based on the results of the analysis and discussion, it is necessary to improve the process of shipping goods carried out by the expedition. The recommendations that are needed as notes in the evaluation are as follows.

(1) Receiving goods at the receiving warehouse should have a target time set by the shipping company.
(2) Delivery of goods to end-customers from the receiving warehouse should be added with a target time. If the target time is not achieved, then an evaluation of the personnel involved can be carried out. With a high volume of shipments, it can be considered to increase the number of courier personnel who distribute packages to end-customers.
(3) For provinces that are detected as experiencing bottlenecks in the process of sending goods or packages, the company can evaluate the processes that occur in that province.
(4) From the results of the regression analysis, it can be seen that the number of activities in the process does not really affect the duration of the delivery of packages or goods so that the number of activities can be ignored.
(5) Expedition A spread to almost all clusters because it dominated as the main expedition. However, Expedition A as the owner of the shipping process must at least improve and evaluate the duration of the package delivery process carried out by them, especially the delivery process carried out in provinces A, B, C, D, and E.
(6) With the results of the analysis, the delivery service company can evaluate the package delivery process as the company's core business. Delivery service companies have a very important role in online business because they are in direct contact with consumers. A big role is the issue of delivery time, security of goods, wider market reach, and another important thing is increasing customer loyalty.

5. CONCLUSION & FUTURE WORKS

The Downstream Process in Supply Chain Management is one of the important processes because it directly intersects with end-customers. In this study, using a collaborative approach between the mining process and data mining, it can be shown that in the three processes analyzed, namely the process of ordering goods, the process of packing goods and the process of shipping goods, bottleneck conditions occur in the delivery of goods, namely in the activity of sending packages to the receiving warehouse, receiving packages in warehouses, and sending packages to end customers. These activities have high waiting times. This is like the initial identification that customers often complain about the delay in the delivery of goods or packages. The process of delays in the delivery of goods does occur as a whole, but to find out which expedition or province is experiencing delays, using data mining, the results obtained are generally expedition an experienced delivery delays in several places, namely provinces A, B, C, D, and E. On this basis, the shipping company, especially Expedition A can improve and evaluate the ongoing process related to the delivery of goods or packages by taking into account the recommendations given in this study.

The collaborative approach in this research shows that in the future mining and data mining processes can strengthen each other in analysis by preparing dataset scenarios and the analysis to be carried out. Such an approach can open avenues of research in various fields that specialize in process analysis. Sometimes it is not enough to only perform process analysis but also to perform data mining to complement the resulting analysis. Although some studies sometimes separate process mining and data mining, it is very open to research that collaborates both. With the growth in the level of large volume of E-Commerce transactions, it can be associated with concepts such as Big Data and IoT technology that is currently developing very open to collaborating between process mining and data mining approaches.

ACKNOWLEDGMENT

This research was supported by the Doctoral Program in Information Systems at the Graduate School of Diponegoro University, Semarang. Without the support of the Supply Chain Management Laboratory in the Information Systems Doctoral program, access to international journals and applications owned by Diponegoro University, this research would not have happened.

REFERENCES

[1] Shoakhmedova, N., Hashimova, D., Belalova, G. (2021). Digitalization of the economy during a pandemic: accelerating the rate of development. Innovations in Economy, 4(3): 56-65. https://doi.org/10.26739/2181-9491-2021-3-8

[2] Sayin, A.A. (2020). The effects of information systems on supply chain and operating performance-analysis of the retail industry. MANAS Sosyal Araştırmalar Dergisi, 9(3): 1627-1639.

[3] Goshime, Y., Kitaw, D., Jilcha, K. (2019). Lean manufacturing as a vehicle for improving productivity and customer satisfaction: A literature review on metals and engineering industries. International Journal of Lean Six Sigma. 10(2): 691-714. https://doi.org/10.1108/IJLSS-06-2017-0063

[4] Ariansyah, K., Sirait, E.R.E., Nugroho, B.A., Suryanegara, M. (2021). Drivers of and barriers to e-commerce adoption in Indonesia: Individuals' perspectives and the implications. Telecommunications Policy, 45(8): 102219. https://doi.org/10.1016/j.telpol.2021.102219

[5] Esthi, R.B., Ekhsan, M. (2020). The effect of millennial intrinsic value toward employee outcomes with employee benefit as mediating variable for strengthening Indonesia's Startup business. Solid State Technology, 63(2s): 8856-8871.

[6] Cristobal-Fransi, E., Montegut-Salla, Y., Ferrer-Rosell, B., Daries, N. (2020). Rural cooperatives in the digital age: An analysis of the Internet presence and degree of maturity of agri-food cooperatives' e-commerce. Journal of Rural Studies, 74: 55-66. https://doi.org/10.1016/j.jrurstud.2019.11.011

[7] Wang, L., Chen, J., Song, H. (2021). Marketplace or reseller? Platform strategy in the presence of customer returns. Transportation Research Part E: Logistics and Transportation Review, 153: 102452. https://doi.org/10.1016/j.tre.2021.102452

[8] Abdul-Ghani, E., Hyde, K.F., Marshall, R. (2019). Conceptualising engagement in a consumer-to-consumer context. Australasian Marketing Journal (AMJ), 27(1): 2-13. https://doi.org/10.1016/j.ausmj.2018.06.004

[9] Xiao, D., Kuang, X., Chen, K. (2020). E-commerce supply chain decisions under platform digital empowerment-induced demand. Computers & Industrial Engineering, 150: 106876. https://doi.org/10.1016/j.cie.2020.106876

[10] Li, Y., Westlund, H., Liu, Y. (2019). Why some rural areas decline while some others not: An overview of rural evolution in the world. Journal of Rural Studies, 68: 135-143. https://doi.org/10.1016/j.jrurstud.2019.03.003

[11] Ma, S. (2017). Fast or free shipping options in online and Omni-channel retail? The mediating role of uncertainty on satisfaction and purchase intentions. The International Journal of Logistics Management, 28(4): 1099-1122. https://doi.org/10.1108/IJLM-05-2016-0130

[12] Team, I.R. (2019). Konsumen Indonesia menginginkan pengalaman pengiriman e-commerce yang lebih baik. IPrice. https://iprice.co.id/trend/insights/konsumen-indonesia-menginginkan-pengalaman-pengiriman-e-commerce-yang-lebih-baik-sebuah-survei-oleh-parcel-perform-membuktikan/.

[13] Sobandi, A., Somantri, B. (2020). Pengaruh Kepercayaan Konsumen Terhadap Keputusan Pembelian Secara Online. Winter Journal: Imwi Student Research Journal, 1(1): 41-52.

[14] Mahendrawathi, E.R., Astuti, H.M., Nastiti, A. (2015). Analysis of customer fulfilment with process mining: A case study in a telecommunication company. Procedia Computer Science, 72: 588-596. https://doi.org/10.1016/j.procs.2015.12.167

[15] Mahendrawathi, E.R., Arsad, N., Astuti, H.M., Kusumawardani, R.P., Utami, R.A. (2018). Analysis of production planning in a global manufacturing company with process mining. Journal of Enterprise Information Management, 31(2): 317-337. https://doi.org/10.1108/JEIM-01-2017-0003

[16] Rbigui, H., Cho, C. (2018). Purchasing process analysis with process mining of a heavy manufacturing industry. In 2018 International Conference on Information and Communication Technology Convergence (ICTC), 495-498. https://doi.org/10.1109/ICTC.2018.8539581

[17] Knoll, D., Reinhart, G., Prüglmeier, M. (2019). Enabling value stream mapping for internal logistics using multidimensional process mining. Expert Systems with Applications, 124: 130-142. https://doi.org/10.1016/j.eswa.2019.01.026

[18] Piest, J.P.S., Cutinha, J.A., Bemthuis, R.H., Bukhsh, F.A. (2021). Evaluating the use of the open trip model for process mining: An informal conceptual mapping study in logistics. In Proceedings of the 23rd International Conference on Enterprise Information Systems, 290-296. https://doi.org/10.5220/0010477702900296

[19] Ramires, F., Sampaio, P. (2021). Process mining and lean six sigma: A novel approach to analyze the supply chain quality of a hospital. International Journal of Lean Six Sigma. https://doi.org/10.1108/ijlss-12-2020-0226

[20] Wang, J., Yue, H. (2017). Food safety pre-warning system based on data mining for a sustainable food supply chain. Food Control, 73: 223-229. https://doi.org/10.1016/j.foodcont.2016.09.048

[21] Wu, M., Liu, K., Yang, H. (2019). Supply chain production and delivery scheduling based on data mining. Cluster Computing, 22(4): 8541-8552. https://doi.org/10.1007/s10586-018-1894-8

[22] Sener, A., Barut, M., Oztekin, A., Avcilar, M.Y., Yildirim, M.B. (2019). The role of information usage in a retail supply chain: A causal data mining and analytical modeling approach. Journal of Business Research, 99: 87-104. https://doi.org/10.1016/j.jbusres.2019.01.070

[23] Li, J. (2019). Optimal design of transportation distance in logistics supply chain model based on data mining algorithm. Cluster Computing, 22(2): 3943-3952. https://doi.org/10.1007/s10586-018-2544-x

[24] Kara, M.E., Fırat, S.Ü.O., Ghadge, A. (2020). A data mining-based framework for supply chain risk management. Computers & Industrial Engineering, 139: 105570. https://doi.org/10.1016/j.cie.2018.12.017

[25] Williamson, E.A., Harrison, D.K., Jordan, M. (2004). Information systems development within supply chain management. International Journal of Information Management, 24(5): 375-385. https://doi.org/10.1016/j.ijinfomgt.2004.06.002
[26] Aalst, W.V.D., Adriansyah, A., Medeiros, A.K.A.D., Arcieri, F., Baier, T., Blickle, T., Wynn, M. (2011). Process mining manifesto. In International Conference on Business Process Management, pp. 169-194. https://doi.org/10.1007/978-3-642-28108-2_19
[27] Caron, F., Vanthienen, J., Baesens, B. (2013). Comprehensive rule-based compliance checking and risk management with process mining. Decision Support Systems, 54(3): 1357-1369. https://doi.org/10.1016/j.dss.2012.12.012
[28] Hossain, M.Z., Akhtar, M.N., Ahmad, R.B., Rahman, M. (2019). A dynamic K-means clustering for data mining. Indonesian Journal of Electrical Engineering and Computer Science, 13(2): 521-526. http://doi.org/10.11591/ijeecs.v13.i2.pp521-526
[29] Bock, F.E., Aydin, R.C., Cyron, C.J., Huber, N., Kalidindi, S.R., Klusemann, B. (2019). A review of the application of machine learning and data mining approaches in continuum materials mechanics. Frontiers in Materials, 110. https://doi.org/10.3389/fmats.2019.00110.
[30] Bogarín, A., Cerezo, R., Romero, C. (2018). A survey on educational process mining. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(1): e1230. https://doi.org/10.1002/widm.1230
[31] Bogarín Vega, A., Cerezo Menéndez, R., Romero, C. (2018). Discovering learning processes using inductive miner: A case study with learning management systems (LMSs). Psicothema, 30(3): 322-329. https://doi.org/10.7334/psicothema2018.116