Knowledge Discovery And Virtual Tour To Support Tourism Promotion

Gusti Ngurah Mega Nata¹, Steven Anthony², Putu Pande Yudiastra³ Prodi Manajemen Informatika, Prodi Sistem Informasi, Prodi Sistem Komputer ITB STIKOM Kota Denpasar, Bali

e-mail: mega@stikom-bali.ac.id, steven.anthony96@gmail.com, yudiastra@stikom-bali.ac.id

To cite this document: Nata, G. N. M., Anthony, S., & Yudiastra, P. P. (2020). Knowledge Discovery And Virtual Tour To Support Tourism Promotion. IAIC Transactions on Sustainable Digital Innovation (ITSDI), 2(2), 94-106. DOI : https://doi.org/10.34306/itsdi.v2i2.387

Abstract

Planning a tourism trip is an important part for tourists so that their tour is satisfying. Service bureaus that have a function to help provide information and prepare tourist travel plans for tourists often provide random destination choices because they do not know the pattern of selecting tourist destinations. This will be detrimental to tourists when service bureaus make wrong tourism travel plans. Tourists also often find it difficult to determine which tourist destination to go to because they do not know the environmental conditions in tourist destinations. To overcome this problem, in this study, knowledge discovery and virtual tours are carried out to increase the promotion of tourism. Knowledge discovery is finding information or knowledge. Knowledge discovery uses data mining techniques to perform data analysis and find patterns. The data mining model that can be used is the frequent pattern by looking for Association Rule Mining from the data. Virtual tour is a technique that can provide 360 + 180 degree images. The virtual tour will be able to show the overall environmental conditions at the tourist destination. The results that have been obtained are in the form of a quick recommendation of tourist attractions in accordance with the country of origin of tourists based on the Association rule mining values. The virtual tour has presented a 360 degree panoramic photo view to inform the environment situation in the place commented by the system.

Keywords: Knowledge Discovery, Data Mining, Association Rule Mining, Apriori, Virtual Tour



Author Notification 22 November 2020 Final Revised 25 November 2020 Published 26 November 2020

1. Introduction

Tourism is a valuable asset for a country, therefore tourism needs to be increased to increase foreign exchange earnings, expand employment and introduce culture [1]. The choice of tourist destinations is one of the important factors that affect the satisfaction of a tour. The choice of tourist destinations is influenced by two factors, namely psychological and external (environmental) influences. Psychological factors include motivation, perceptions, attitudes and characteristics of tourists, while external factors include cultural values, social class, as well as situations and places of residence of tourists [2].

It is very important to understand well the pattern of tourists' choices in order to provide a tour that suits the interests of tourists. However, the fact is that in the field, the tourism industry still has difficulty using technology to provide information in accordance with the choice of tourists considering the large number of tourists who come from many countries [3], as well as displaying information about the environmental conditions of tourist destinations in a clear and interactive manner so this makes it difficult for tourists. to select a tourist destination. The industry gives tourists a choice of tourist destinations at random so they often offer choices that are not right to the tastes of tourists.

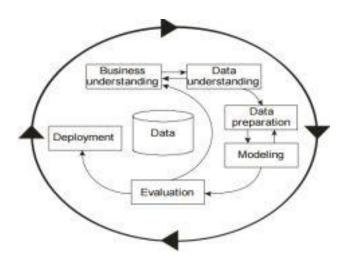
The tourism industry will be able to increase tourist satisfaction if it can provide information and also provide the right choices for tourists [2]. The tourism industry will be able to change the marketing strategy and determine the tour packages and discount coupons appropriately so that it will increase the sales and use of resources owned by the tourism industry to be more effective [2]. To be able to know the frequent patterns (Frequent Pattern) [10] of the choice of tourist destinations, the data mining method can be used to analyze data [5] tourism (Itemset) with the attributes of the countries of origin of tourists, packages and tourist objects.

The itemset used will form the Association model [6]. The stages to form an association data model between items are searching for frequent itemset by digging data (Data Mining) [10] in this study is tourism data. Frequent Itemset in the tourism sales data will show the pattern of tour locations that are often chosen by tourists from certain countries and the relationship between purchasing a tourism package with other tourism packages. Furthermore, the results from frequent itemset will be further processed using the Association Rules Mining Technique to find out how strong the rules for the relationship between items and other items. The result of using the Association Rules Mining technique is knowledge discovery in the form of tourist destination patterns based on the country of origin of the tourists are and by tourists from certain countries. From the analysis results also obtained knowledge discovery in the form of activity choice patterns with other activities.

To provide more accurate information about the environment in tourist destinations, a virtual tour technique can be used in the form of a panoramic view that displays 360 degrees [4] [8]. The 360 degree image presented is an image of the environment in a tourist destination. 360 degree images will be a technique that can provide visual information to visitors and potential buyers of tour packages [4]. With 360 degree images that are presented, visitors or potential buyers can interact by rotating the image to see the entire image. So that using virtual tour techniques will convince tourists about the destination and steps to overcome options that are not suitable for tourists [8].

2. Research Method

The engineering and analysis method uses the Crisp-DM method [9], namely the engineering steps as shown in Figure 1 below:



Gambar 1. Metode Crisp-DM

2.1 Business Understanding

The development of a Knowledge Discovery and Virtual Tour system to support tourism promotion is based on the needs of travel business in general. Based on these business needs in this study in collaboration with several travel agents. Data sources come from several Travel Agents in Bali.

2.2 Data Understanding

The characteristics of the data to be used are tourism package sales transaction data which contains the name of the tourist, the origin of the tourist, the number of purchases, the activity ordered, the order id, and the transaction id.

ia	guest_name	order_id	qTY	Nationality	guest_activity
1	Rasha Abeirabu	1	2	Saudi Arabia	Agung trekking
4	Rasha Abeirabu	1	2	Saudi Arabia	Bedugul Tour
2	Rasha Abelrabu	1	2	Saudi Arabia	Flying Fox
з	Rasha Abelrabu	1	2	Saudi Arabia	Refting
6	A Skeroft	2	1	Australia	Cangu Tour
5	A Ekeroft	2	1	Australia	Taman Safari Bali
7	Abbicox Wilk	з	2	Australia	Cooking Class
8	Abbicox Wilk	3	2	Australia	Rafting
10	Adam meilden	4	3	Australia	Fast Boat To Gili
9	Adam mellden	4	3	Australia	Rafting
12	Amanda Killer	5	4	Australia	Jimbaran Tour
11	Amanda Killer	5	4	Australia	Rafting
14	Amanyii	6	30	Australia	Dive In Nuss Lemborger
13	Amanyil	6	30	Australia	Nusa Dua Tour
15	Amy	7	2	Australia	Jimbaran Tour
15	Anny	7	2	Australia	Ratting
17	Anna Maria	8	4	Australia	Dive In Nusa Lemborga
18	Anna Maria	8	4	Australia	Rafting
19	Aude Pottier	9	4	Australia	Jimbaran Tour
20	Aude Pottler	9	4	Australia	Rafting
21	Barry Luine	10	4	Australia	Jimberan Tour
22	BarryLuine	10	4	Australia	Ubud Tour
24	Billiu Mallon	11	1	Australia	Batur Trekking
23	Billiu Mallon	11	1	Australia	Waterboom
25	Bred roper	12	2	Australia	Batur Trekking
25	Brad roger	12	2	Australia	Waterboom
14.45	A 1 AI			W. LEWIS	

Figure 2. Tourism Data 2.3 Preparation Data

2.1 Data Preparation

The data fields (itemset) that will be used to look for patterns of activity selection based on the country of origin of tourists are the origin of tourists, the number of purchases and activities ordered. Meanwhile, to look for patterns of activity selection based on the previously selected activity required data fields (itemset), namely order_id and ordered activities.

order_id	guest_activity	
1	Agung trekking	
1	Bedugul Tour	
1	Flying Fox	
1	Rafting	
2	Cangu Tour	
2	Taman Safari Bali	
3	Cooking Class	
3	Rafting	
4	Fast Boat To Gili	
4	Rafting	
5	Jimbaran Tour	
5	Rafting	
6	Dive In Nusa Lembongan	

Figure 3. Itemset Based on Activity Pattern

Nationality	qty	guest_activity
Saudi Arabia	2	Agung trekking
Australia	3	Agung trekking
Australia	4	agung trekking
Inited Kingdom of Great Britain and Northern Ireland	1	Agung trekking
Inited Kingdom of Great Britain and Northern Ireland	4	Agung trekking
USA	1	agung trekking
Australia	3	Atv
Australia	1	Atv
Australia	3	Atv
Australia	5	Atv
Australia	8	Atv
Australia	7	Atv
Australia	5	Atv
Australia	2	Atv
Australia	6	Atv
Australia	4	Atv
Australia	2	Atv
Australia	3	Atv
Australia	4	Atv
Australia	1	Atv
Australia	2	Atv
Australia	4	Atv
Australia	1	Atv
Australia	2	Atv
Australia	2	Atv
Australia	2	Atv
Canada	4	Atv
china	3	Atv

Figure 4. Itemset of Activity Patterns Based on the Country of Origin of the Tourist

3. Findings

Based on the results of business needs analysis and data analysis, the modeling and development that can be done are as follows:

3.1 Modelling

Modeling is the working stage of data analysis using data mining. The modeling used to process tourism data is the association rules model. To look for patterns with the association rules model, tourism data will be processed so that the Frequent Itemset can be identified. Then from the Frequent Itemset, an association rule will be formed by exploring the association rules (Association Rule Mining). Association rules will produce patterns of tourist selection with activities based on country of origin and patterns of relationships between one activity and other activities. To find Association Rule Mining, you can use the Apriori algorithm.

The Apriori algorithm uses 2 parameters to find associations between data (Frequent ItemSet), namely support and confidence. To find support and confidence, you need to calculate the number of itemset. Then support and confidence from an itemset will be selected based on minimum support.

The following is a discussion to look for patterns of activity selection based on the country of origin of tourists using the Apriori algorithm:

A. Looking for frequent itemset and calculating support and confidence. The formula for obtaining support for an item is as follows:

Support (A) = Jumlah Transaksi Mengandung A Total Transaksi

While the value of two item support is obtained from the following formula:

Support(A,B) = $(A \cap B)$ <u>Jumlah Transaksi Mengandung A dan B</u> Support(A,B) = Total Transaksi

The formula for obtaining the confidence of an item is as follows:

Confidence = P(B|A) = Jumlah Transaksi Mengandung A dan B Jumlah Transaksi Mengandung A

Tourism transaction data is data obtained from daily sales. The following is a representation of some tourism package sales data.

Nationality	Qty	Guest_Activity
Saudi Arabia	2	Agung Trekking
Saudi Arabia	2	Bedugul Tour
Saudi Arabia	4	Flying Fox
Saudi Arabia	2	Rafting

Table 1.	Examples of	f Sales	Transactions
----------	-------------	---------	--------------

Australia	1	Cangu Tour
Australia		Taman Safari Bali
Australia	2	Cooking Class
France	4	Agung Trekking
France	4	Bedugul Tour
France	2	Flying Fox
France 3		Rafting
Canada	3	Rafting

Then determine the support of the itemset: Support (Saudi Arabia) = $10/30 \times 100\% = 33.33\%$ Support (Australia) = $4/30 \times 100\% = 13.33\%$ Support (France) = $13/30 \times 100\% = 43.33\%$ Support (Canada) = $3/30 \times 100\% = 10\%$ Support (Agung trekking) = $6/30 \times 100\% = 20\%$ Support (Bedugul Tour) = $6/30 \times 100\% = 20\%$ Support (Flying Fox) = $6/30 \times 100\% = 20\%$ Support (Rafting) = $8/30 \times 100\% = 26.67\%$ Support (Canggu Tour) = $1/30 \times 100\% = 3.33\%$ Support (Bali Safari Park) = $1/30 \times 100\% = 3.33\%$ Support (Cooking Class) = $2/30 \times 100\% = 6.67\%$ With Minimum Support = 13%.

Table 2.Second iteration (Items Fulfilling Minimum Support)

Item	Support
Australia	13,33%
France	43,33%
Saudi Arabia	33,33%
Agung trekking	20%
Bedugul Tour	20%
Flying Fox	20%
Rafting	26,67%

Then in the third iteration, the items will be merged into 2-itemset.

2-Itemset	Suppor t	Persentase
Australia, Agung Trekking	0	0%
Australia, Bedugul Tour	0	0%
Australia, Flying Fox	0	0%

Australia, Rafting	0	0%
France, Agung Trekking	4	13,33%
France, Bedugul Tour	4	13,33%
France, Flying Fox	2	6,67%
France, Rafting	3	10%
Saudi Arabia Agung		
Trekking	2	6,67%
Saudi Arabia, Bedugul		
Tour	2	6,67%
Saudi Arabia, Flying Fox	4	13,33%
Saudi Arabia, Rafting	2	6,67%

Then in the fourth iteration, the confidence of the 2-itemset (country, activity) will be calculated.

Confidence = P(B|A) = Jumlah Transaksi Mengandung A dan B Jumlah Transaksi Mengandung A

Confidence = P(France|Agung Trekking) = 4 / 13 * 100 = 30,77%

Confidence = P(France|Bedugul Tour) = 4 / 13 * 100 = 30,77%

Confidence = P(Saudi Arabia|Flying Fox) = 4 / 10 = 40%

From the results of the 2-itemset association, a rule is formed, namely:

Aturan	Confidence	
If you are from France then buy	4/13	30,77
the Agung Trekking package		%
If you are from France then buy	4/13	30,77
the Bedugul Tour package		%
If you are from Saudi Arabia, you	4/10	40%
will buy the Flying Fox package		

The following is a discussion to look for patterns for selecting the next activity based on the selected activity using the Apriori algorithm:

Table 5.Examples of Sales Transactions (Order_id, and Itemset)

Ι	
d	Aktivitas
1	Atv, Cangu Tour, Rafting
2	Bedugul Tour, Jimbaran Tour, Water Sport
3	Atv, Cangu Tour, Rafting, Water Sport
4	Jimbaran Tour

Then, determine the activity support for the itemset:

	amoun	Support
Aktivitas	t	
atv	2	(2 / 4) * 100% = 50%
cangu tour	2	(2 / 4) * 100% = 50%
rafting	2	(2 / 4) * 100% = 50%
bedugul tour	1	(1 / 4) * 100% = 25%
jimbaran <i>tour</i>	2	(2 / 4) * 100% = 50%
water sport	2	(2 / 4) * 100% = 50%

Table 6. Support Table on 1-Itemset

With Minimum Support = 30%.

Table 7.Second iteration (Items Fulfilling Minimum Support)

	Jumla	Suppor
Aktvitas	h	t
jimbaran		50 %
tour	2	
water sport	2	50 %
atv	2	50 %
cangu tour	2	50 %
rafting	2	50 %

Then in the third iteration, the items will be merged into 2-itemset.

Table 8. Suppo	-	
Itomaat	Inmlah	Support

Itemset	Jumlah	Support
jimbaran tour, water sport	1	(1 / 4) * 100% = 25%
jimbaran <i>tour</i> , atv	0	0 %
jimbaran <i>tour</i> , cangu <i>tour</i>	0	0 %
jimbaran <i>tour</i> , rafting	0	0 %
water sport, atv	1	(1 / 4) * 100% = 25%
water sport, cangu tour	1	(1 / 4) * 100% = 25 %
water sport, rafting	1	(1 / 4) * 100% = 25%
atv, cangu <i>tour</i>	2	(2/4) * 100% = 50 %
		(2 / 4) * 100% = 50
atv, rafting	2	%
cangu <i>tour</i> , rafting	1	(1 / 4) * 100% = 25%

Table 9.Second iteration	(Items Fulfilling	Minimum Support)

	Jumla	Support
Itemset	h	
atv, cangu <i>tour</i>	2	(2 / 4) * 100% = 50%
atv, rafting	2	(2 / 4) * 100% = 50%

Then in the third iteration, the items will be merged into 3-itemset.

Itemset	Jumlah	Support
atv, cangu <i>tour</i> , rafting	2	(2 / 4) * 100% = 50 %

Then in the third iteration, the confidence calculation from the 3-itemset association will be carried out. From the 1 itemset obtained, there are 3 associations that can be formed, namely:

Confidence = P (atv, cangu tour | rafting) = 2/2 = 100%Confidence = P (atv, rafting | cangu tour) = 0/10 = 0%Confidence = P (cangu tour, rafting | Atv) = 0/10 = 0%

From the results of the calculation of confidence 3 associations formed from 1 itemset that fulfills support and calculation of confidence, a rule can be formed, namely: when a customer buys an atv and cangu tour activity package, he will buy a rafting activity package also with a 100% presentation.

3.2 System development

1. Context Diagram

Context diagram describes the system in general, which consists of 2 external entities that use this system, namely admin and users. As can be seen from the following context diagram.

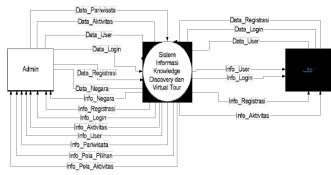


Figure 5. Context Diagram

4. System implementation results

The result of system development is a website with association rule mining algorithm and virtual tour. Here is an initial view that contains recommendations for visitors.

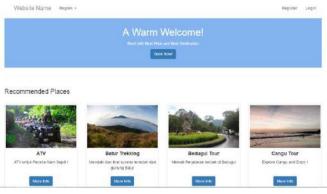


Figure 6. Main page

The page to find patterns of activity choices based on the country of origin of the traveler is a page where the frequent itemset search process and searches for association rules based on country and activity items

5. Result

1. Association Rule based on the tourist's country of origin for the activity

The result is a pattern of activity choices based on the country of origin of the tourists and also calculations of support and confidence. The results are processed using a website-based program. The results of the calculation are displayed in the form of tourist attractions recommendations using the location images shown in Figure 6 for tourists and in tabular form as in Figure 7 for the analysis process.

Negara	Aktivitas	Support	Confidence
Australia	Atv	1.98	9.04
france	Atv	1.01	15.11
italy	Atv	1.15	19.50
Singapore	Atv	1.09	14.02
Germany	Cangu Tour	1.03	15.22
france	Fast Boat To Lembongan	1.01	15.11
Germany	Fast Boat To Lembongan	1.33	19.57
Korea	Fast Boat To Lembongan	1.03	11.74
Australia	Flying Fox	1.18	5.40
Singapore	Flying Fox	1.03	13.26
Holland	Horse Riding	1.01	12.73
Australia	Junbaran Tour	1.12	5.13
Australia	Nusa Dua Tour	2.10	9.58
Australia	Rafting	4.82	22.00
Holland	Rafting	1.24	15.73

Tabel Asosiasi 2-Itemset Yang Memenuhi Minimum Support

Figure 7. Pattern of Activity Options based on Tourist Country of Origin

2. Association rule among tourist activities

The activity-based search activity selection page is a page where the frequent itemset search process and searches for association rules based on activity items. The result is a pattern of activity choices based on previously selected activities as well as calculations of support and confidence.

Aktivitas 1	Aktivitas 2	Support	Confidence
Atv	Cangu Tour	6	6.5217391304348
Atv	Cycling	14	15.217391304348
Atv	Elephant Riding	6	6.5217391304348
Atv	Fast Boat To Gili	17	18.478260869565
Atv	Fast Boat To Lembongan	11	11.95652173913
Atv	Flying Fox	23	25
Atv	Jimbaran Tour	9	9.7826086956522
Atv	Kintamani Tour	6	6.5217391304348
Atv	Nusa Dua Tour	9	9.7826086956522
Atv	Rafting	51	55.434782608696
Atv	Turtle And Glass Bottom Boat	6	6.5217391304348
Atv	Ubud Tour	7	7.6086956521739
Atv	Water Sport	20	21.739130434783
Atv	Waterboom	6	6.5217391304348
Bedugul Tour	Rafting	8	28.571428571429
Bedugul Tour	Ubud Tour	7	25
Cangu Tour	Fast Boat To Gili	10	33.3333333333333333

Figure 8. The Activity Selection Pattern Page Based on Previous Activities

The results of calculating support and confidence will produce a pattern. The explanation is as Next, when the activity chosen is ATV, tourists will choose Rafting activity with a percentage of 55.4%.

3. Virtual Tour Results

Virtual Reality Photography "(VRP)," Immersive Photography "or" Photo 360 "are techniques used to provide real-time information on the location of tourist destinations that are commented on by the system to potential tourists. The 360-degree panoramic photo used is capable of displaying borderless and seamless photos.

4. Conclusion

Based on the results of the analysis and discussion that has been carried out, several conclusions can be drawn as follows:

- 1. From the processing of tourism transaction data, it can produce a pattern of choice of tourist activities using a priori algorithm with 2 parameters, namely support and confidence.
- 2. The pattern of activity choices obtained is in the form of a pattern consisting of 2-Itemset, namely the pattern of the relationship between the country and the activity and the pattern of the relationship between the activity and the previously selected activity.
- 3. The Virtual Tour technique can provide visual information about tourist destinations with 360 + 180 degree images.

Acknowledgment

Thank you to ITB STIKOM Bali for providing space and facilities to researchers and also to travel tour owners who have provided travel data.

References

- [1] TAP Nomor : IV/MPR/1978 Tentang Garis Garis Besar Haluan Negara.
- [2] Imam Ahmad Adhi. Maria Goretti Wi Endang NP. Pravissi Shanti. Pengaruh Faktor Psikologis Terhadap Keputusan Berkunjung (Survei Pada Pengunjung Batu Secret Zoo Jawa Timur Park 2). 2016; Vol. 30 No.1 (Jurnal Administrasi Bisnis) : Halaman. 36.
- [3] Ni Made Ary Widiastini. Nyoman Dini Andiani. Trianasari. Strategi Pemasaran Pariwisata Di Kabupaten Buleleng, Bali. 2012; Vol. 1, No. 1 (Jurnal Ilmu Sosial dan Humaniora) : Halaman. 11.
- [4] NATA, I. Gusti Ngurah Mega. Aplikasi Virtual Tour Guide Sebagai Promosi Pariwisata Bali. Jurnal Sistem dan Informatika, 2017, 11.2:129987.
- [5] SAPTAWATI, Gusti Ayu Putri; NATA, Gusti Ngurah Mega. Knowledge discovery on drilling data to predict potential gold deposit. In: 2015 International Conference on Data and Software Engineering (ICoDSE). IEEE, 2015. p. 143-147.
- [6] NATA, Gusti Ngurah Mega. Association Rule Mining Pada Data Geokimia Pemboran. Eksplora Informatika, 2016, 5.2.
- [7] Kennedi Tampubolon, Hoga Saragih, Bobby Reza. Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat – Alat Kesehatan. 2013. Vol 1, No. 1 (Informasi dan Teknologi Ilmiah): Halaman. 95-99.
- [8] Hera Wulanratu Wulur, Steven Sentinuwo, Brave Sugiarso. Aplikasi Virtual Tour Tempat Wisata Alam Di Sulawesi Utara. 2015. Volume 6, No. 1 (E-Journal Teknik Informatika) : Halaman. 1-2.
- [9] Annisa Paramitha Fadillah. Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ). 2015. Volume 1, No. 3 (Jurnal Teknik Informatika dan Sistem Informasi) : Halaman. 261 -262
- [10] Agrawal, Rakesh. Srikant, Ramakrishnan. Fast Algorithm For Mining Association Rules. 1994. Proceedings of the 20th VLDB Conference Santiago.
- [11] Febriyanto, E., Naufal, R. S., & budiarty, frizca. (2019). Attitude Competency Assessment in the 2013 curriculum based on elementary school Prototyping methods. IAIC Transactions on Sustainable Digital Innovation (ITSDI), 1(1), 87-96. <u>https://doi.org/10.34306/itsdi.v1i1.6</u>
- [12] Sunarya, P. A., Aini, Q., Bein, A. S., & Nursaputri, P. (2019). The Implementation Of Viewboard Of The Head Of Department As A Media For Student Information Is Worth Doing Final Research. ITSDI Journal Edition Vol. 1 No. 1 October 2019, 18.
- [13] Febriyanto, E., & Naufal, R. S. (2019). Attitude Competency Assessment in the 2013 Curriculum Based On Elementary School Prototyping Methods. IAIC Transactions on Sustainable Digital Innovation, 1(1), 87-96.
- [14] Yusup, M., Aini, Q., Apriani, D., & Nursaputri, P. (2019, December). PEMANFAATAN TEKNOLOGI BLOCKCHAIN PADA PROGRAM SERTIFIKASI DOSEN. In SENSITIF: Seminar Nasional Sistem Informasi dan Teknologi Informasi (pp. 365-371)

Knowledge Discovery And Virtual...

- [15] Neelima, E., & Babu, M. P. (2020). Genome feature optimization and coronary artery disease prediction using cuckoo search. Computer Science and Information Technologies, 1(3), 106-115.
- [16] Ilamsyah, I., Wulandari, R., & Fahreza, R. R. (2020). Web Design Visual Comic Communication As A Visual Based Learning Media. IAIC Transactions on Sustainable Digital Innovation (ITSDI), 2(1), 1-10. https://doi.org/10.34306/itsdi.v2i1.22