Int. J. Advance Soft Compu. Appl, Vol. 11, No. 3, November 2019
ISSN 2074-8523

Temporal Learning Analytics Based on Triple-Factor Approach Using Self-Organizing Map

Kusuma Ayu Laksitowening, Denny, Zainal A. Hasibuan
Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
School of Computing, Telkom University, Bandung, Indonesia
e-mail: kusuma.ayu@ui.ac.id, ayu@telkomuniversity.ac.id
Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
e-mail: denny@cs.ui.ac.id
Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
e-mail: zhasibua@cs.ui.ac.id

Abstract

E-learning personalization aims to deliver learning activities and materials that suits to learners' needs. Therefore, the system must have the ability to analyze the profile and characteristics of each individual learner. Characteristics of learners, among others, can be identified from their behavior in using e-learning. Their most frequent learning resource accessed, their participation on discussions, and their assessment result are some of the variables from the activity logs that can describe their learning patterns. On the other side, learners' behavior may change over time. This research aims to capture and analyze the dynamic learning pattern throughout the semester. The learning analytics are conducted using temporal clustering approach to identify the learning style, motivation, and knowledge abilities. This research performs two-level clustering analysis to acquire learning patterns from activity logs from Moodle Learning Management System using Self-Organizing Map (SOM) and k-Means. SOM enables visualization high dimensional data by projection to lower dimensions. The proto-clusters of SOM are then clustered using k-Means. The temporal clustering results show that the learning patterns of learners are changing over time.

***Keywords**: clustering, e-learning, learning analytics, Self-Organizing Map, temporal*

1 Introduction

Traditional learning generally provides unvarying learning for all learners in the same classroom. Meanwhile, the ability of learners and the way they learn in a class may vary. Limitations of time and number of instructors in a class are the common reasons that personalization is still difficult to apply on traditional learning.

In contrast, personalization is highly possible to be applied on e- learning [1]. Technology enhancement enrich learning to be held without face-to-face learning in a classroom by using e-learning. Currently e-learning is generally delivered though the internet which enables learners to study without limitation of place and time. With e-learning, learning can be done individually. It means one's learning process would not interfere other learners' learning process.

Since the purpose of personalized learning is to deliver learning that accommodate learners' individual profile [2], the systems must first know the profile and characteristics of learners before deciding the suitable treatments to be delivered. This research analyses learners' learning patterns from activity logs in using e-learning. The activity logs show the frequent e-learning learning resource accessed by the learners, their learning patterns, their participation level and performance.

Our research considers the possibility that learning pattern may shift over time. If learning pattern is only analysed once, the achievement progress and dynamic behaviour of learners throughout the semester could not be described. To get the detailed view and development process, our analysis is performed using temporal approach and performed multiple times throughout the semester.

This research identifies students' learning patterns based on their learning style, motivation, and knowledge abilities factors. The factors were already defined in the Triple-Factor Approach [3]. Our research conducted the temporal analysis using two-level SOM clustering [4]. Two-level clustering is used to handle large amounts of data. First, the clustering is performed using SOM that reduce data dimensionality which enables visualization the data [5]. In the second stage, the prototype vectors of SOM are clustered using k-means. Then, the clustering results are analysed to understand various learning patterns based on the three factors.

2 Related Work

2.1 E-Learning Personalization

E-learning personalization has been applied according to analysis based on the learners' activities. It requires understanding of individual learning pattern. The analysis on learning pattern should include information related to learners that needed by the system to present personalization, among others: frequent page accessed [6] [7], time of access [6], preferred type of content [8], and participation on forum discussion [6] [7] [9]. There are at least two ways on analyzing learner [3]: question based and activity log based.

The question-based approach in understanding individual learning pattern requires learner to answer several questions that are prepared for learner profiling. Previous studies, among others [10] conducted learning style identification using Felder-Silverman Learning Style Model (FSLSM) questionnaires. This approach takes time for learner to fill out the questionnaire and the result strongly depends on the sincerity of the learner in answering the questions. It is not practical to conduct multiple data gathering using questionnaires.

On the other hand, the interaction between learners and e-learning system that are stored in the activity logs can be analysed by utilizing learning analytic techniques. With this approach, learners are not consciously being analysed. Thus, the result is considered more natural. Learning analytics enables more collected information and allows the detection of a wider variety of learning pattern.

2.2 Learning Analytics

Learning analytics (LA) is defined as a series of measuring, collecting, analyzing, and reporting data and knowledge related to learners. It aims for improved understanding of the learning process and education in general [11]. Learning analytics utilizes methods, techniques, and approaches, especially data mining and statistics [12]. The methods are developed to collect data and analyze the behavior and performance of learners with the purpose of enhancing learning process [13].

The methods that evolve in LA include prediction, association mining, and structural detection [14]. The prediction model is built to infer one variable data from some combination of other variables. Similar to our research, [6] used e-learning interaction data to analyze/predict students' failure using time series and temporal decomposition.

Association mining aims is to discover the relationship among variables in the data set. In general, relationship mining includes, among others, association rule, correlation, sequential pattern, and causal data mining [14]. [15] explored the correlation between time required by students on answering multiple choice quiz with their score.

Structural detection methods try to obtain structures in the data without first having an idea of what to look for. In this method, the structure is tried to be explored naturally from the data. Common techniques in structural detection include factor analysis, social network analysis, and clustering. The purpose of clustering is to find structures in which data are naturally divided into groups[14]. [16] observed patterns in the online environment using *k*-means and expectation-maximization (EM) algorithm.

Learning analytics have already became an important part in e-learning personalization. [17] used genetic algorithm to arrange the recommendations regarding the course's structure and students' learning styles. Neural networks was utilized to classify learning styles in [18], the result was used for personalizing the user interface. Our current research utilizes temporal clustering techniques to obtain learning patterns based on the factors explained in the following section.

2.3 The Triple-Factor Approach

This research uses a holistic analysis of learning patterns based on the Triple-Factor Approach. The approach [3] mentioned that e-learning process is influenced by learning style, motivation, and knowledge ability factors. Learning styles can be defined as personal characteristic on how a learner perceives, interacts with, and responds to the learning environment [19]. The approach identifies learning style from learners' actions in using e-learning [3].

Motivation can be measured through various ways, including: self-report measurement and participation on online discussion [20]. The online discussions activity is reliant on learners' motivational development [21]. In line with [20] and [21], the Triple-Factor approach measured learners' motivation level from discussion forum. The more often a learner doing the activity of the discussion forum shows the higher the learning motivation [3].

Knowledge ability level indicates the knowledge and ability of learners in understanding an information obtained through a learning or education [3]. Knowledge ability of learner can be identified from the evaluation of learning process. The common form of evaluation is by conducting series of assessments. In Triple-Factor Approach, knowledge ability can be measured from assessment such as quizzes.

2.4 Self-Organizing Map

A Self-Organizing Map (SOM) is a single layer Artificial Neural Network (ANN) in the form of low-dimensional grid (mostly in two-dimensional and rectangular or hexagonal grid) as the projection of high-dimensional data. The SOM can change their structure and function due to external stimulants [5]. In other words, SOM performs projection while following the distribution of the original datasets.

The basic principles of SOM are competitive and cooperative learning. It is called competitive since the neurons compete to each other for possession on each input. The neuron that is most alike to the input wins. The winner is called the best matching unit (BMU) [5] [22]. However, it is also a cooperative process. The winner neuron (BMU) modify their weights to adapt to the input. Neurons among this neighborhood then cooperate by also adjusting their weights as a response to the input [5] [23].

In general, SOM algorithm can be illustrated as follows. First, the weights of each neuron are initialized. For each input vector that is randomly chosen from the dataset and fed to the network. The most similar neuron (BMU) to the input vector is computed by comparing their distance/similarity. The weights of the BMU and its neighbors are adjusted towards the input vector.

Our research utilizes SOM to obtain learning patterns based on Triple Factor Approach. The clustering is run on several times to acquire time-varying analysis regarding the evolving behaviour of learners. Several previous researches already discussed the time-varying structure on clustering.

[24] proposed dynamic clustering framework to analyse the evolution of structure and its cohesion. Meanwhile, [25] discussed on structure and topology of the SOM that changes over time. Our research focused on changing behavior of each data instance (learner) and the overall structure (learning pattern) obtained from SOM on time-varying manners.

3 Method

The objective of this research is to conduct temporal pattern analysis based on e-learning factors previously described in Section 2.3. The analysis is conducted on activity logs to obtain learning pattern and observe the change that may occurred on the pattern over time. This research is conducted through several step as depicted on Fig. 1.

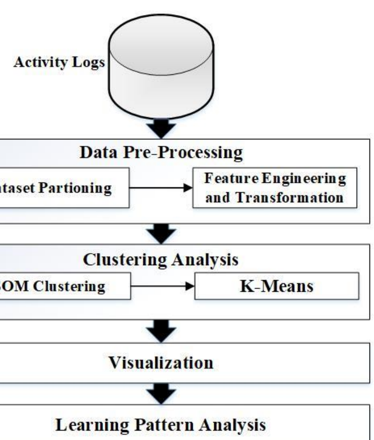


Fig. 1. Research steps

3.1 Data Pre-Processing

Data pre-processing step begin by partitioning dataset based on time. The partition is based on factor and the temporal label. The temporal label was added to the log as a note that each record belongs to which temporal period. Table 1 explains the detailed of the datasets.

The temporal data were collected from the activity log at four periods of time. The dataset column lists that each factor is analysed based on four temporal datasets (LS for learning style, M for Motivation, and KA for knowledge ability). Therefore, there will be 12 datasets to be analysed.

Table 1: Data partition

Period	Dataset	Time Retrieved	Period of Log Used
1	LS1, M1, KA1	End of 4th week	Week 1 - week 4
2	LS2, M2, KA2	End of 8th week	Week 1 - week 8
3	LS3, M3, KA3	End of 12th week	Week 1 - week12
4	LS4, M4, KA4	End of 16th week	Week 1 - week16

Each factor contains some features retrieved from activity log. The selection of the features were based on basic principles of Triple-Factor Approach [3], related studies such as [6][7][9], and the available log of the case study. The data were retrieved from 304 learners' activity log in using Moodle LMS during a course in Faculty of Computer Science in Universitas Indonesia. After retrieving LMS log, the next step is transforming the log using Pentaho Data Integration to get the value of each feature of the factor explained in Table 2.

Learning style is recognized by the features describing activities on accessing LMS and following the course. Learners' activities in forum discussion identify their motivation level. Meanwhile, the average and maximum grade of the learners in quizzes define the learners' knowledge ability.

Higher value of these features indicates a better condition. In contrast to other features, the less value of interval between learner's first quiz attempt and the first time the quiz is open indicates the better condition. Therefore, this feature is multiplied by -1. The value of each feature then be normalized by the z-score formula. The clustering process was later performed on the normalized data.

Table 2. Triple factor features

Factors	Features
Learning Style	Quantity of assignment pages visit (`ls_assignmentview`) Quantity of course pages visit (`ls_courseview`) Quantity of quiz attempt (`ls_quizattempt`) Average of attempt per quiz (`ls_quizmeanattempt`) Average of time difference between last time submitting quiz and the quiz deadline (`ls_quizmeanfinish`) Average of time difference between the time quiz is available online and the first attempt (`ls_quizmeanstart`) Quantity of quiz pages visit (`ls_quizview`) Quantity of resource pages visit (`ls_resourceview`)
Motivation	Quantity of discussion post (`m_addpost`) Quantity of discussion pages visit (`m_discussionview`) Quantity of discussion initiation (`m_discussionadd`) Quantity of forum pages visit (`m _forumview`)
Knowledge Ability	Maximum grade per quiz (`ka_quizmaxgrade`) Average grade per quiz (`ka_quizmeangrade`)

3.2 Clustering Analysis

The analysis process was conducted by using two-level clustering. At the first level, the data was processed with Self-Organizing Map (SOM) algorithm to create prototype vector. For each dataset, a map with 7 x 12 (84 prototype vectors) with hexagonal lattice structure is trained. Furthermore, the prototype vector was used as input on the next cluster process. The second level used *k*-means to produce the final clusters. The obtained clustering result are analysed to acquire the learning pattern based on the temporal data on each factor.

3.3 Visualization and Learning Pattern Analysis

The trained map and the clustering results are then visualized. The visualization can describe the distribution of each attribute and the final cluster on the map. It helps the learning pattern analysis.

The obtained clustering result are analysed to discover unique structures that may appear. Our research analyses structure of each factor using component plane visualizations of all four temporal datasets. The following section describe the analysis result and the obtained findings.

4 Result and Analysis

4.1 Component Plane Visualization

The characteristics of each cluster can be analysed by observing the component plane visualizations. These visualizations show the distribution of each attribute value in the maps. The colour of the node shows the average value of the attribute on the node. Darker colours show a higher value, while bright colours indicate a lower value. A set of nodes that have similar patterns, tend to be grouped on the same cluster, due to topological preservation property of SOM. In this research, the analysis is conducted by observing the distribution of each feature in the resulting maps.

4.1.1 Learning Style Factor Cluster

Although the distribution of attribute values generated by the temporal datasets is not exactly the same, the four temporal dataset maps show a similar pattern in general. The first pattern shows learners who access learning activities very rarely through e-learning. Figs. 2-5 show that the nodes located in the upper right area of the map show the brightest colors on all their attributes. This means that learners who are mapped to this region have characteristics of low course view, resource view, assignment view, and so on. Therefore, it can be concluded that these learners have inadequate learning style.

These visualizations also discover a region that contains learners who are frequently access the course page and the resource. This is shown by the darker node on `ls_courseview` and `ls_resourceview` attributes. This type of learners is in the lower left area of the maps (see Figs. 2-5). The maps also show that the attributes related to the quiz activity in the lower left area are also slightly darker. It means that these learners are frequent in accessing course and resource are also slightly active on quiz activities.

Another interesting cluster are learners who are very enthusiastic on quiz activity. The darkest node on `ls_quizattempt`, `ls_quizmeanattempt`, and `ls_quizview` attributes describe this kind of learners in Figs. 2-5. The attributes associated with quiz on the lower right area are a little darker than the lower left area. However, the `ls_courseview` and `ls_resourceview` attributes in the lower right area are relatively brighter than in the lower left area. It clearly visible on all dataset that learners who are very enthusiastic on quiz activity are on the lower right area of the maps.

Our analysis discovers another a group of learners who frequently access assignment page but are less enthusiastic on quiz activity. The component plane visualizations (Figs. 2-5) indicate that the most frequent learners accessing the assignment page are in the upper left nodes. Moreover, it appears that this region has the attributes associated with quiz are relatively bright. In contrast, learners who are enthusiastic on quiz activity (lower right region) are less frequent on visiting assignment page.

The last cluster consists of learners who perform all activities but in moderate frequency. This cluster is located in the middle area which, in all attributes, show a slightly brighter color. This indicates that the learner mapped in the area tends to occasionally visit course page, assignment page, and resource page, and not too excited about quiz activities.

From the identifiable patterns of component plane visualizations, the number of clusters selected is five. They are: learners that very rarely access learning activities; learners that are very enthusiastic on quiz activities; learners that frequently access course and resource, and active on quiz activities too; learners that are very frequent accessing the assignment page, but less enthusiastic on quiz activities; and learners that perform all activities but in moderate frequency.

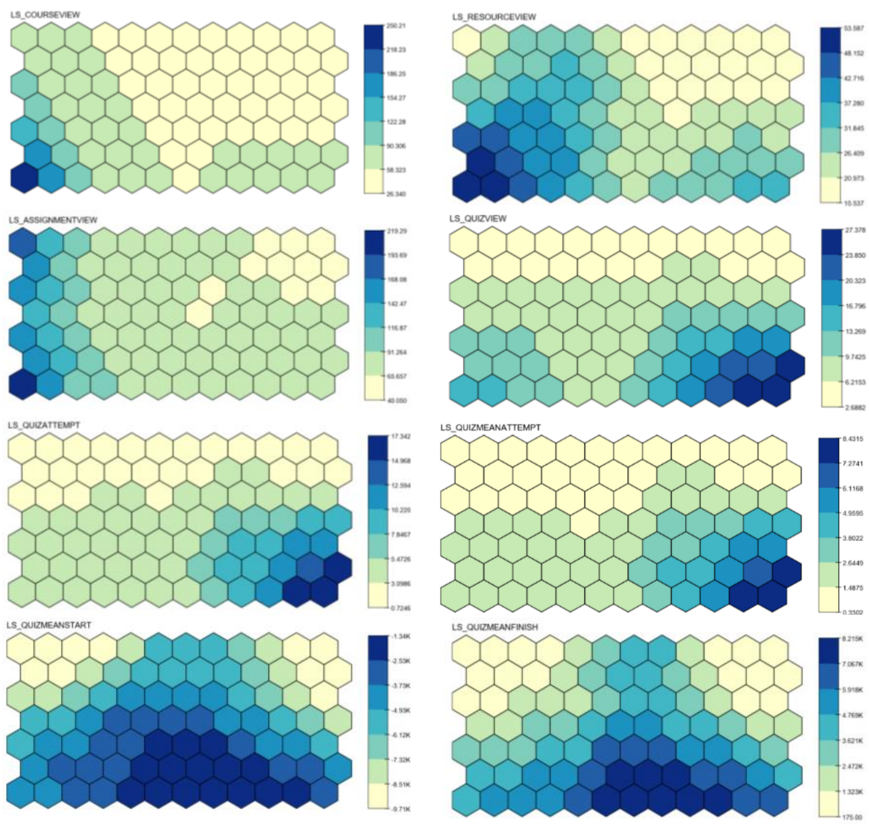


Fig. 2. Component plane visualization for Dataset LS1

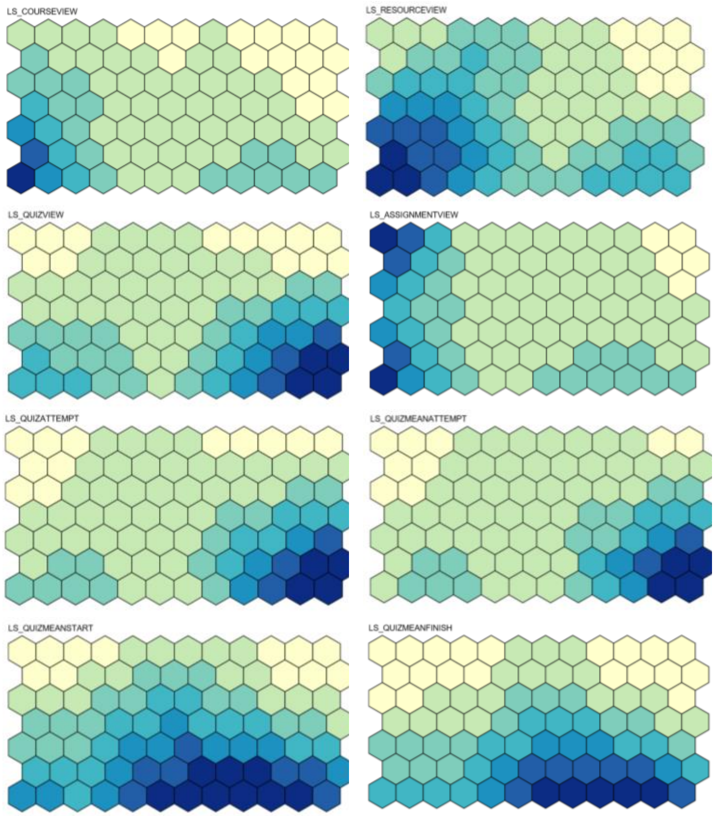


Fig. 3. Component plane visualization for Dataset LS2

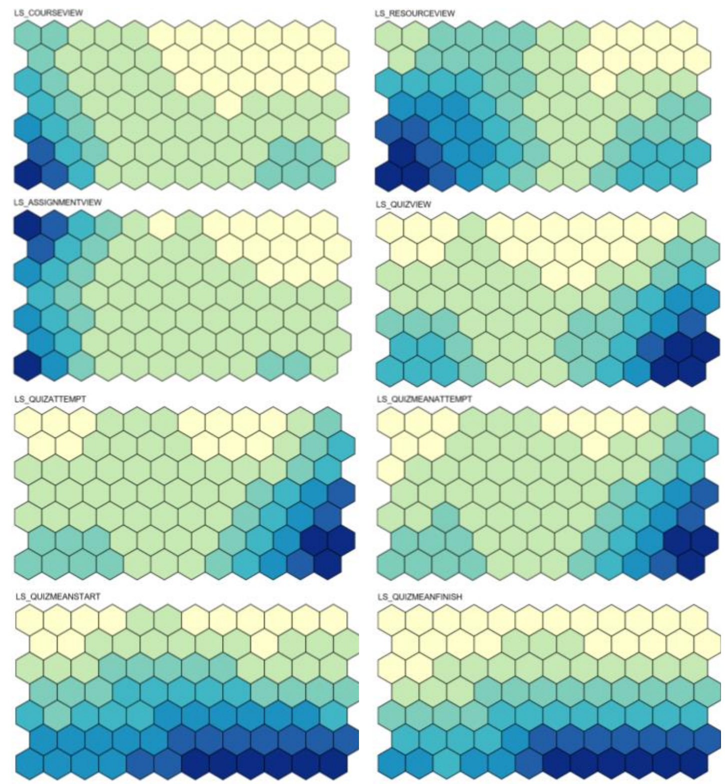


Fig. 4. Component plane visualization for Dataset LS3

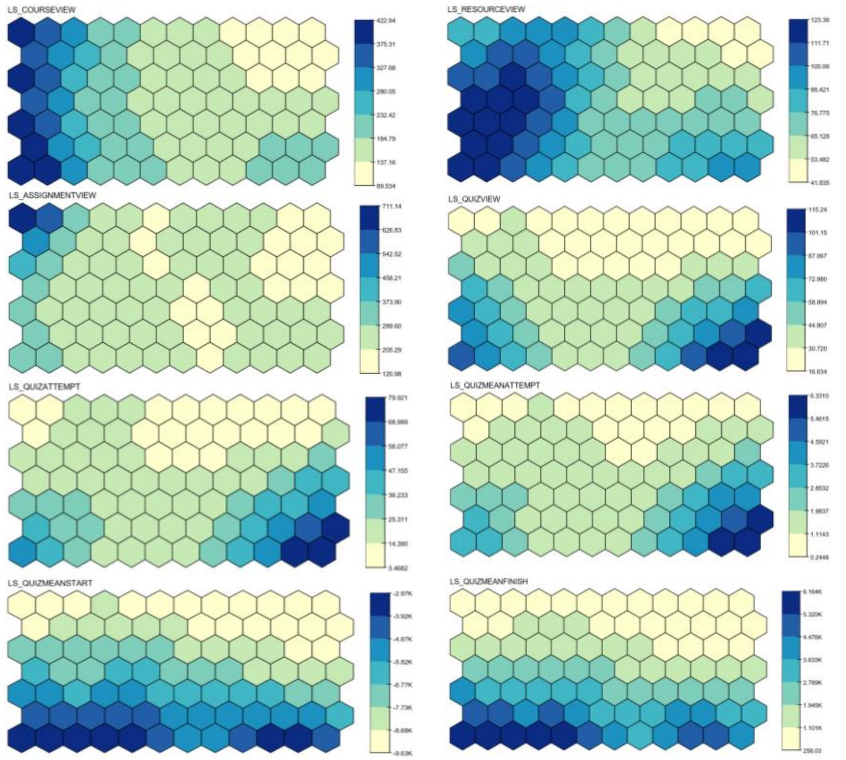


Fig. 5. Component plane visualization for Dataset LS4

4.1.2 Motivation Factor Cluster

The method used in analysis of learning style is also applied to analyze the motivational factor. Figs. 6-9 show the distribution of each motivation factor attributes in the resulting map of all four periods.

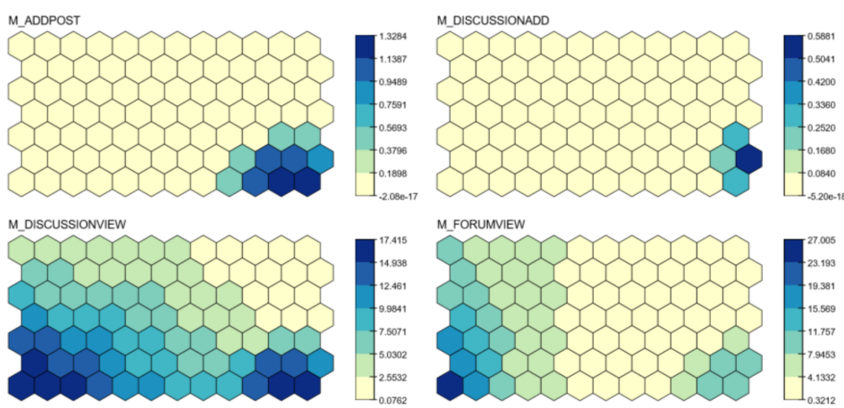


Fig. 6. Component plane visualization for Dataset M1

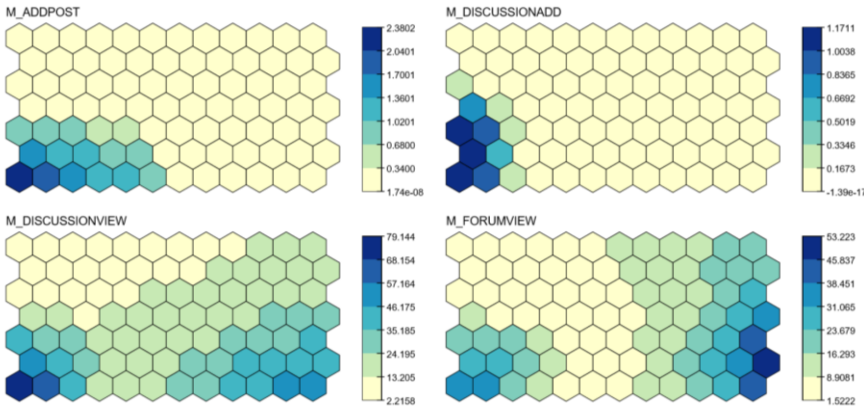


Fig. 7. Component plane visualization for Dataset M2

The Component plane visualizations on M1 and M2 show different positioning compared to the visualizations on M3 and M4. But the four maps tend to show the same pattern in their attribute information. The maps generally show three clusters of learner patterns identified by the area on the map.

The first cluster contains learners who rarely access discussion forums. This is visible from the top area of the map. Nearly all nodes in the upper area show bright colors in all attributes of this factor.

The second cluster consists of active learners in the discussion forums. In other words, the learners with this pattern participate actively in the discussions. This is indicated by areas with dark color on `m_addpost` and `m_discussionadd` attributes. The last cluster comprises learners who are quite frequent in accessing the discussion forum in passive way. This pattern is revealed in the area where the `m_discussionview` and `m_forumview` attributes are dark but the `m_addpost` and `m_discussionadd` attributes are light-colored. It means that the learners access the discussion forum regularly, but not active in posting or replying the message.

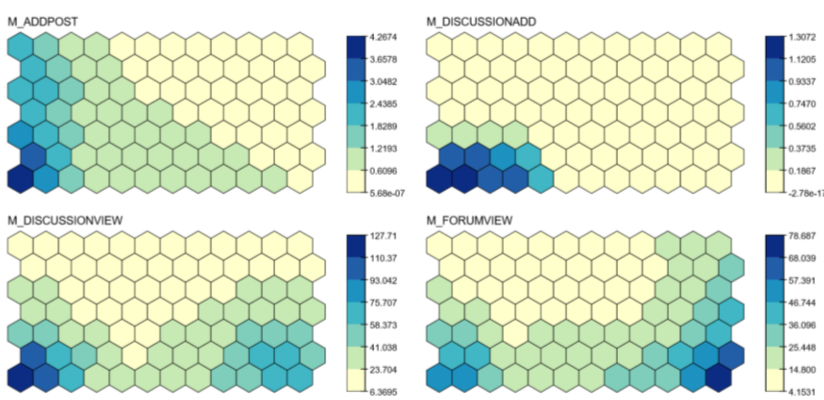


Fig. 8. Component plane visualization for Dataset M3

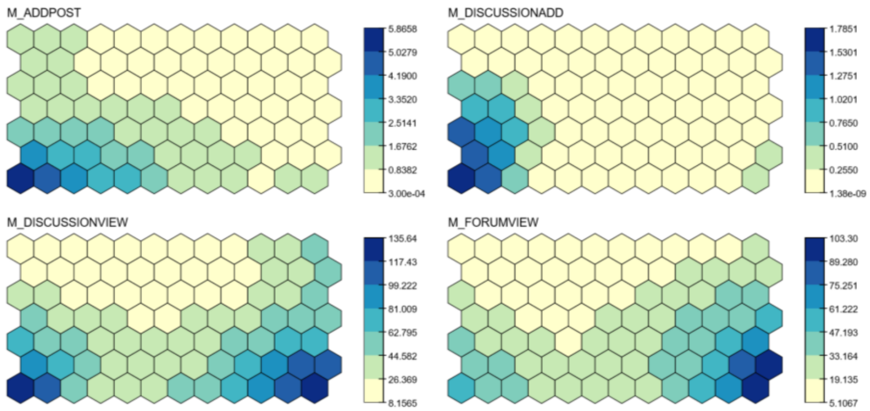


Fig. 9. Component plane visualization for Dataset M4

4.1.3 Knowledge Ability Factor Cluster

Fig. 10-13 show the distribution of the knowledge ability attributes on the SOMs of each period. The Component plane visualizations on KA1 and KA2 show different distribution of values compared to the visualizations on KA3 and KA4. Nonetheless the four maps tend to show the same pattern in their attribute information. The maps indicate four learner patterns shown by the area on the map.

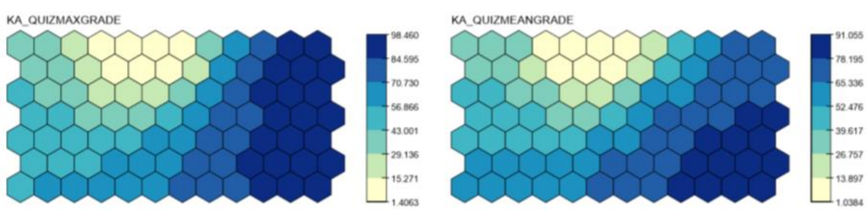


Fig. 10. Component plane visualization for Dataset KA1

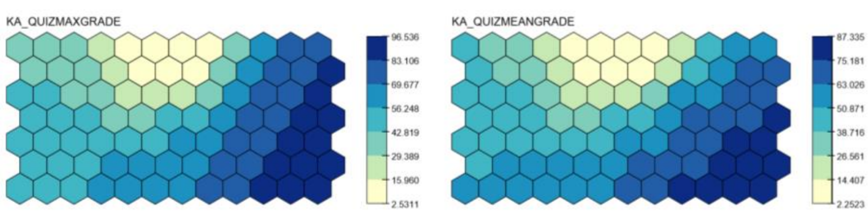


Fig. 11. Component plane visualization for Dataset KA2

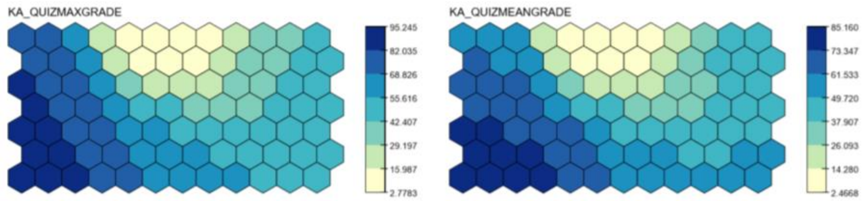


Fig. 12. Component plane visualization for Dataset KA3

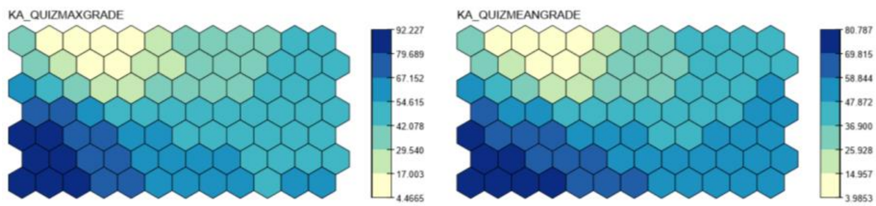


Fig. 13. Component plane visualization for Dataset KA4

There are two extreme patterns identified from the generated map on all four datasets. The first pattern is the learners that have very low quiz grades. This is shown in the upper middle area. In the all four component plane visualizations, the upper middle area is generally the brightest than any other area on the map.

In contrast, another remarkable cluster comprises learners with very high quiz grades. This is indicated by dark-colored nodes in both `ka_quizmaxgrade` and `ka_quizmeangrade` attributes. These clusters are located in the lower right area of the KA1 and KA2 maps, and in the lower left area of the KA3 and KA4 maps.

In addition to these two extreme patterns, there are two areas whose characteristics are relatively similar but have different values. Both patterns can be identified as learners whose quiz grades are low and learners who have fair quiz grades. Learner with fair quiz grades is quite dark colored but not as dark as learner node that has very high quiz grades. Whereas, learner with low quiz grades is indicated by relatively light node (green). From Component plane visualization analysis, the selected number of cluster for knowledge ability factor is 4 (four) clusters.

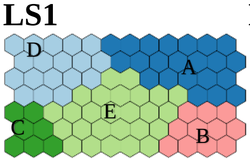
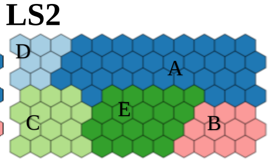
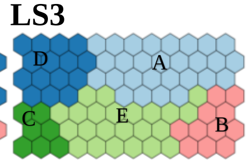
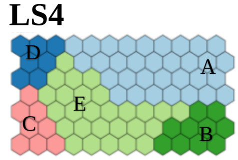
4.2 Clustering Results

Different datasets obtain different maps. Nonetheless, all the datasets produce relatively similar patterns. The final clustering result in Motivation, Learning Style, and Knowledge Ability factor respectively are depicted in Tables 3-5. In the Learning Style factor, the four temporal datasets display the similar clustering structure. The maps consistently present five clusters of learners (see Table 3). The component plane visualizations show the unique pattern of the five obtained clusters. This cluster patterns consistently appear in all datasets.

All four datasets produce three dominant clusters on Motivation Factor. The distinguishing characteristics can be seen on component plane visualizations: a cluster containing learner that rarely accesses forum discussions, a cluster containing passive user in forum discussions, and a cluster containing learner that show an active role in the forum discussions. Similarly, all datasets also produce same pattern in Knowledge Ability factor. The Knowledge Ability factor produces four clusters based on the quiz grades of the learners.

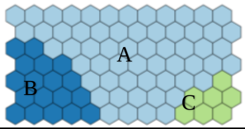
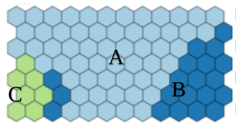
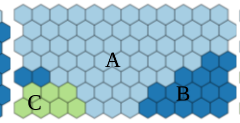
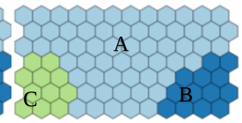
The difference between the cluster results between the four datasets is the number of nodes that belong to cluster. Likewise, in the learner grouped on the clusters. There is alteration in the number of learners incorporated in each cluster. Cluster changes on individual level are discussed in the following sub section.

Table 3. Clustering result on learning style factor

	Dataset			
	LS1	**LS2**	**LS3**	**LS4**
				
Cluster **LS_A**	22 nodes 101 learners	33 nodes 131 learners	28 nodes 117 learners	32 nodes 145 learners
LS_B	11 nodes 60 learners	11 nodes 57 learners	11 nodes 54 learners	10 nodes 43 learners
LS_C	7 nodes 23 learners	14 nodes 39 learners	7 nodes 16 learners	8 nodes 23 learners
LS_D	18 nodes 53 learners	7 nodes 22 learners	15 nodes 44 learners	7 nodes 17 learners
LS_E	26 nodes 67 learners	19 nodes 55 learners	23 nodes 73 learners	27 nodes 76 learners

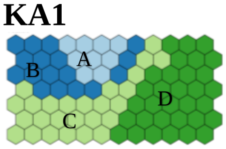
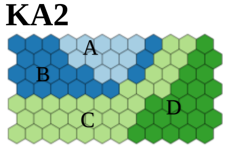
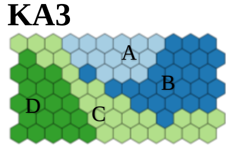
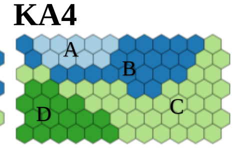
Cluster Description
LS_A: Learner that very rarely access learning activities
LS_B: Learner that are very enthusiastic on quiz activities
LS_C: Learner that frequently access course and resource, and active on quiz activities too
LS_D: Learner that are very frequent accessing the assignment page, but less enthusiastic on quiz activities
LS_E: Learner that perform all activities but in moderate frequency (not too often)

Table 4: Clustering result on motivation factor

		Dataset			
		M1	**M2**	**M3**	**M4**
					
Cluster	**M_A**	59 nodes	57 nodes	60 nodes	60 nodes
		241 learners	238 learners	230 learners	234 learners
	M_B	18 nodes	20 nodes	17 nodes	13 nodes
		52 learners	51 learners	46 learners	39 learners
	M_C	7 nodes	7 nodes	7 nodes	11 nodes
		11 learners	15 learners	28 learners	31 learners

Cluster Description
M_A: Learners who rarely access discussion forums
M_B: Learners who are quite frequent in accessing the discussion forum in passive way
M_C: Learner who are active on the discussion forum

Table 5. Clustering result on knowledge ability factor

		Dataset			
		KA1	**KA2**	**KA3**	**KA4**
					
Cluster	**KA_A**	9 nodes	11 nodes	14 nodes	9 nodes
		44 learners	46 learners	51 learners	52 learners
	KA_B	17 nodes	20 nodes	25 nodes	22 nodes
		18 learners	38 learners	42 learners	46 learners
	KA_C	26 nodes	31 nodes	25 nodes	36 nodes
		56 learners	71 learners	66 learners	55 learners
	KA_D	32 nodes	22 nodes	20 nodes	17 nodes
		186 learners	149 learners	145 learners	151 learners

Cluster Description
KA_A: Learner with very low quiz grades
KA_B: Learner with low quiz grades
KA_C: Learner with fair quiz grades
KA_D: Learner with very high quiz grades

4.3 Cluster Evaluation

Clustering is an unsupervised learning, so it does not have a comparable label to measure the quality of the obtained structure. However, there are several parameters that can be used for cluster evaluation, among others, Davies Boudin index and the Silhouette index. Both Davies-Boudin Index [26] and Silhouette Index [27] combine intra-cluster distance and inter-cluster distance for each cluster.

In the Davies-Boudin Index, the best structure is indicated by lower index value. In contrast, higher value in the Silhouette Index shows a better cluster structure. This study uses both parameters as depicted on Table 6. **Bold** writing indicates the best value, while *italic* writing marks the worst value amongst different k (number of cluster).

Table 6. Cluster evaluation

Factor	**Dataset**	**Davies-Boudin Index**			**Silhouette Index**		
		Number of Clusters (k)			**Number of Clusters (k)**		
		3	**4**	**5**	**3**	**4**	**5**
Learning Style	LS1	*1,262*	**0,988**	1,091	0,292	*0,139*	**0,367**
	LS2	*1,267*	1,049	**0,928**	0,339	**0,365**	*0,15*
	LS3	*1,267*	1,001	**0,977**	0,305	*0,204*	**0,378**
	LS4	*1,13*	1,047	**0,94**	*0,217*	**0,441**	0,282
Motivation	M1	0,848	**0,797**	*0,9*	**0,463**	*0,324*	0,446
	M2	*0,99*	0,915	**0,771**	**0,413**	0,395	*0,285*
	M3	*1,072*	**0,837**	0,907	0,158	*0,056*	**0,475**
	M4	0,853	**0,792**	*0,924*	0,309	*0,212*	**0,41**
Knowledge Ability	KA1	0,59	**0,572**	*0,636*	0,682	**0,683**	*0,605*
	KA2	0,599	*0,627*	**0,597**	**0,728**	*0,639*	0,658
	KA3	**0,571**	*0,626*	0,594	0,691	**0,7**	*0,646*
	KA4	*0,641*	**0,583**	0,586	0,641	**0,683**	*0,59*

Second stage of clustering analysis (*k*-means) was experimented with varied k value, from 3 to 5 clusters. Table 6 shows that the values of the Davies-Boudin Index and the Silhouette Index on all clusters tend to fluctuate. The best and worst values are not always obtained by a certain k value. Accordingly, this research also analyzes the distribution of the value of the attributes in determining the number of clusters (see Subsection 4.1 and Subsection 4.2), while still considering Davies-Boudin Index and Silhouette Index values.

4.4 Changes in Cluster Assignment

After the clusters were generated on each dataset, we analysed the cluster change by comparing the cluster result on each learner among datasets. For example, if a learner is mapped as member of a cluster on dataset 1 and identified as member of another cluster on dataset 2, then the learner is identified as experiencing cluster change. Table 7 shows the cluster mapping changes between datasets. The column describes the dataset being compared. For example, column 1-2 refers to the number of learner experience cluster change from dataset 1 to dataset 2. This research analysed the change between two sequential dataset and the change learner may experience at least once through the semester.

Table 7. Cluster mapping changes between datasets

Comparison Between Dataset	**1-2**	**2-3**	**3-4**	**1-2-3-4**
Learning style	86 28%	60 20%	73 24%	142 47%
Motivation	40 13%	27 9%	32 11%	73 24%
Knowledge ability	83 27%	38 13%	42 14%	115 38%
Change on at least one factor	163 54%	108 36%	118 39%	227 75%

The obtained results may indicate that the cluster maps on one dataset compare to other datasets were very likely to change. The changes occurred generally in the overall activity within the online class. The cluster changes occurred in all factors and all datasets, although with varying percentages.

From the experiment, up to 75% of learner experience changes in learning patterns at least once in a semester on at least one factor. It indicates the high possibility of learner to have behaviour change over time. Therefore, it confirms that the learning pattern should be analysed periodically to apprehend the behavioural changes as well as the progress of learning process.

5 Conclusion

This paper has shown that Triple Factor Approach using Self-Organizing Map (SOM) method can discover learning patterns from activity logs. The Triple-Factor Approach allows understanding of learner behaviour based on different factors, namely learning style, motivation, and knowledge ability. SOM reduces data dimensionality which enables data visualization.

While all four temporal datasets produced similar clusters' structure, the visualizations of the component plane show the existence of unique cluster structures. Our methods can indicate that distribution of nodes to each cluster change over time. Moreover, the number of learners assigned in each cluster is also dynamic. Therefore, it is capable to discovers the learner's behavioural change. The result confirmed that their learning patterns are dynamic over time. The cluster changes occur on all factors and all dataset comparisons.

References

[1] K. A. Laksitowening, H. B. Santoso, and Z. A. Hasibuan. (2017). E-Learning Personalization Using Triple-Factor Approach in Standard-Based Education. *IOP Conf. Ser. J. Phys. Conf. Ser. 801 012027*.

[2] S. Ghallabi, F. Essalmi, M. Jemni, and Kinshuk. (2013). Toward the reuse of E-Learning personalization systems. In *IEEE 2013 Fourth International Conference on Information and Communication Technology and Accessibility (ICTA)*. pp. 3–5. IEEE.

[3] Sfenrianto. (2014). Pendekatan Tipe Belajar Triple-Factor dalam Proses E-Learning Sebagai Basis Rekomendasi dan Personalisasi Pembelajaran (An Approach of Triple-Factor Learning Types in e-Learning Process as a Basis for Learning Recommendation and Personalization.). Universitas Indonesia.

[4] Denny, G. J. Williams, and P. Christen. (2010). Visualizing temporal cluster changes using Relative Density Self-Organizing Maps, *Knowl. Inf. Syst.*, vol. 25, no. 2, pp. 281–302.

[5] D. Miljković. (2017). Brief Review of Self-Organizing Maps, pp. 1252–1257.

[6] C. G. Nespereira, A. F. Vilas, and R. P. D. Redondo. (2015). Am I failing this course?. *Proc. 3rd Int. Conf. Technol. Ecosyst. Enhancing Multicult. - TEEM '15*, pp. 271–276.

[7] D. Vasic, A. Pinjuh, M. Kundid, and L. Seric. (2015, May). Predicting student's learning outcome from Learning Management system logs. *Proc. 2nd Int. Conf. Learn. Anal. Knowl. - LAK '12*, vol. 15, pp. 1–8.

[8] W. Chen, Z. Niu, X. Zhao, and Y. Li. (2014). A hybrid recommendation algorithm adapted in e-learning environments. *Springer - World Wide Web*, vol. 17, pp. 271–284.

[9] S. Goggins and W. Xing. (2016). Building models explaining student participation behavior in asynchronous online discussion. *Comput. Educ.*, vol. 94, pp. 241–251.

[10] K. A. Laksitowening, A. P. Yanuarifiani, and Y. F. A. Wibowo. (2016). Enhancing e-learning system to support learning style based personalization. In *Proceeding - 2016 2nd International Conference on Science in Information Technology, ICSITech 2016: Information Science for Green Society and Environment*, pp. 329–333. IEEE.

[11] Q. Liu and G. Fan (2014). Using Learning Analytics Technologies to Find Learning Structures from Online Examination System.

[12] T. Elias. (2011). *Learning Analytics : Definitions , Processes and Potential*.

[13] M. Pechenizkiy and D. Gasevic. (2015). Introduction into Sparks of the Learning Analytics Future. *J. Learn. Anal.*, vol. 1, no. 3, pp. 145–149.

[14] R. S. J. D. Baker and G. Siemens. (2012). Educational Data Mining and Learning Analytics. *Proc. 2nd Int. Conf. Learn. Anal. Knowl. - LAK '12*, p. 252.

[15] Z. Papamitsiou, V. Terzis, and A. A. Economides. (2014). Temporal learning analytics for computer based testing. In *Proc. Fourth Int. Conf. Learn. Anal. Knowl. - LAK '14*, pp. 31–35.

[16] R. Cerezo, M. Sánchez-Santillán, M. P. Paule-Ruiz, and J. C. Núñez. (2016) Students' LMS interaction patterns and their relationship with achievement: A case study in higher education. *Comput. Educ.*, vol. 96, pp. 42–54.

[17] M. M. El-Bishouty, T.-W. Chang, S. Graf, Kinshuk, and N.-S. Chen. (2014). Smart e-course recommender based on learning styles. *J. Comput. Educ.*, vol. 1, no. 1, pp. 99–111.

[18] W. Gong and W. Wang. (2011). Application research of support vector machine in E-Learning for personality. In *2011 IEEE Int. Conf. Cloud Comput. Intell. Syst.*, pp. 638–642. IEEE.

[19] R. M. Felder. (1996). "MATTERS OF STYLE," *ASEE Prism*, vol. 6, no. December, pp. 18–23.

[20] M. Hartnett. (2012). Relationships between online motivation, participation, and achievement: More complex than you might think. *J. Open, Flex. Distance Learn.*, vol. 16, no. 1, p. 28.

[21] K. Xie, V. Durrington, and L. L. Yen. (2011). Relationship between students' motivation and their participation in asynchronous online discussions. *J. Online Learn. Teach.*, vol. 7, no. 1, pp. 17–29.

[22] H. Yin. (2008). The self-organizing maps: Background, theories, extensions and applications. *Stud. Comput. Intell.*, vol. 115, pp. 715–762.

[23] MarcM. Van Hulle. (1995). Self-Organizing Maps. pp. 1–45.

[24] P. Caravelli, Y. Wei, D. Subak, L. Singh, and J. Mann. (2013). Understanding evolving group structures in time-varying networks. In *Proc. 2013 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min. - ASONAM '13*, pp. 142–148. IEEE/ACM.

[25] A. F. R. Araujo and R. L. M. E. Rego. (2013). Self-organizing maps with a time-varying structure. *ACM Comput. Surv.*, vol. 46, no. 1, pp. 1–38.

[26] D. L. Davies and D. W. Bouldin. (1979). A Cluster Separation Measure. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227.

[27] P. J. Rousseeuw. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, vol. 20, no. C, pp. 53–65.