IPSJ Transactions on Computers and Education Vol.3 No.2 20–27 (June 2017)

Short Paper

Intuitive Analysis by Visualizing Context Relevant E-learning Data

Pitoyo Hartono[1,†1,a)] Kayo Ogawa[2,b)]

Received: July 17, 2016, Revised: November 30, 2016,
Accepted: March 4, 2017

Abstract: In the last few years learning management systems have been widely introduced in many educational institutions with the primary objectives of supporting students with more flexible learning environments and also importantly acquiring learning pattern data from students and extracting meaningful contents from the data to be used to improve the learning quality. However, often due to the complexity and the multidimensionality of the data, the extraction of meaningful information from them is difficult. So far many methods for mining useful information from complex data have been proposed, and one of the most powerful is visualization that allows intuitive understanding on the underlying properties of the data. In this paper, visualization of E-learning data using a newly introduced context-oriented self-organizing map is introduced and compared against some traditional visualization methods.

Keywords: visualization, visual analysis, dimensional reductions, neural networks, Self-Organizing Maps

1. Introduction

Over the last two decades, many educational institutions have been adopting Learning Management Systems (LMSs) with mainly two primaries objectives. The first one is to support their students' flexibilities in learning by allowing them to learn outside their class rooms using information technologies [1], [2], [3], [4]. The second one is to automatically accumulate vast amount of data from students' learning activities and utilize the data to improve students' learning experience and quality. Because of the advances in network, sensor, and storage technologies, it is becoming easier to acquire and store massive amount of data, but analyzing the data is increasingly becoming an overwhelming challenge. In many situations, instead of becoming analytical tools for providing new insights from the data, many LMS are becoming expensive data dumps. The difficulty in discovering useful information from the data is primarily due to the volume, multidimensionality, and complexity of the data, as well as the lack of clear objective in the data analysis method itself. While, in recent years, many methods for discovering knowledge from data have been proposed, one of the most efficient means to discover underlying information from raw data is through visualization [5], [6], [7], [8]. Visual analysis of complex data combines algorithms to transfer complex data into a visualizable format, usually in 2-D or 3-D, and the insights of the human analyst to interpret the visualization results. Here, the flexibility, experience, intelligence and clear objective of the human analyst are crucial in discovering interesting hidden structure from the complex data. The human factor is also important in filtering noise included in the visualization results due to the imperfection in the data themselves or the algorithms to visualize them.

There are rich collections of algorithms to visualize complex data, but this paper focuses on dimensionality reduction. As meaningful data are likely to be multidimensional, naturally, one of the most logical choices is to utilize dimensionality reduction algorithms to visualize them. The challenge in reducing the dimension of high dimension data is to preserve the underlying structure or inherently embedded information of the data, as well as possible, in their low dimensional representations so that the structure and information can to some extent be visualized and intuitively understood. So far, dimensionality reduction is one of the most actively studied fields in data analysis. One of the most traditional dimensionality reduction techniques is Principle Component Analysis (PCA) [9], [10]. PCA is an elegant orthogonal transformation procedure to transfer multidimensional variables into linearly uncorrelated variables called the Principal Components (PCs). Expressing each data point using the first two or three PCs allows the visualization of the whole data set. However, the linearity of PCA often prevents it from disclosing the nonlinear nature of the data. PCA also does not access the categorical information embedded in the data, and so the context of the data does not have any role in their low dimensional representation. Another conventional dimensionality reduction method, Linear Discriminant Analysis (LDA) [11], is also an orthogonal transformation of high dimensional variables into low dimensional space, but unlike PCA it accesses the categorical information of the data, so that the data points belonging to same categories are mapped into close clusters while the distances be-

1 School of Engineering, Chukyo University, Nagoya, Aichi 466–8666, Japan
2 Faculty of Science, Japan Women's University, Bunkyo, Tokyo 112–8681, Japan
†1 Presently with Humanoid Research Institute, Waseda University
a) hartono@ieee.org
b) kogawa@fc.jwu.ac.jp

tween centroids of different categorical groups are maximized. Here, while the context or categories of the data play an important role in forming the low dimensional representation, similar to PCA, it is linearly constrained. At the same time, Self-Organizing Map (SOM) [12], [13], which is one of the most widely used visualization techniques, is unconstrained by the linearity of the data. SOM transfers high dimensional data into low dimensional representations while preserving their topological characteristics, such that data points that are located close to each other in their original high dimensional space are also collocated in their low dimension representation, and originally dissimilar data are mapped far from each other in the low dimensional representation space. However, similar to PCA, it does not access the categorical information of the data. More recently, new collections of dimensional reduction methods [14] have been proposed, such as Locally Linear Embedding (LLE) [15], Stochastic Neighborhood Embedding (SNE) [16] and its variant t-SNE [17], Neighborhood Component Analysis (NCA) [18], ISOMAP [19] and S-Isomap [20]. While these dimensionality reduction methods are built upon elegant mathematical foundations, some of them are difficult to implement, and the resulting low dimensional spaces are not easy to interpret.

In this study, a newly proposed neural network called Restricted Radial Basis Function (rRBF) Network [21] is applied as a visual analytical tool against real world e-learning data. The significant differences between rRBF and other dimension reduction methods are that rRBF produces not only 2-D representation of high dimensional data but also classifies the data. Hence, it is not only a dimension reduction method but also a classifier. Further, although the rRBF produces a kind of 2-D topographical maps, it significantly differs from SOM, in that the maps reflect the underlying topological structure of the high dimensional data in relevance to the context of the data. Hence, the map is called Context-Relevance Self-Organizing Maps (CRSOM). This is an interesting property of the dimension reduction function of rRBF that offers more flexibility for visual analysis. This property allows CRSOM to generate maps with different appearances for identical data with different contexts. For example, same e-learning data can be interpreted in different contexts, such that, graded differently. Unlike the conventional SOM, CRSOM is able to display the data with different contexts in different ways, opening the possibility for more flexible analysis. In this study, the rRBF was applied to some e-learning data with the objective of producing a 2-D representation of the data that can be visually analyzed to acquire intuitive information about the intrinsic properties of the data. The efficiency of the low dimensional representation produced by the rRBF, CRSOM, is compared against some traditional dimension reduction methods. This study is a significant improvement over the previous study [22] with new comparison experiments, classification experiments and explanation on the visual analysis from the resulting CRSOM.

The rest of this paper is structured as follows. In Section 2, an overview of rRBF and the resulting CRSOM are explained. Section 3 explains the experiments in producing 2-D maps from multidimensional data and their visual analysis. The classification performances of rRBF are also explained in this chapter. The final chapter provides the conclusion and future works.

2. Overview on Radial Basis Function Networks

The structure of the rRBF Network is illustrated in **Fig. 1**. The rRBF is a multilayered neural network inspired by the conventional Radial Basis Function (RBF) Networks [23]. The main difference between rRBF and RBF is that the activation function in rRBF is topologically restricted, which in effect generates an internal representation in the hidden layer that reflects not only the high dimensional inputs' topological structure but also their context. Hence, the resulting internal layer is called the Context-Relevant Self-Organizing Map (CRSOM). Since the CRSOM is two dimensional, it can be considered as a low dimensional representation of high dimensional data, hence rRBF is a dimension reduction algorithm. Furthermore, the rRBF is a supervised neural network that is trained to transfer high dimensional input into its category or label, hence it is also a classifier. As rRBF inherits the activation characteristics of the conventional RBF, for a given input only a part of the neurons that represent similar reference vectors with the input are intensely activated as illustrated with red neurons in Fig. 1. This activation characteristic is responsible for the topological-preservation nature of the rRBF.

The mathematical properties of the learning behavior of the rRBF have been explained in details in Ref. [21] but will be briefly outlined as follows.

At time t, observing high input, $\boldsymbol{X}(t) \in \mathcal{R}^d$, the rRBF selects a winning neuron, win among the hidden neuron in its two dimensional hidden layer as follows. Here, d is the dimension of the input vector.

$$win = arg\min_j \left\|\boldsymbol{X}(t) - \boldsymbol{W}_j(t)\right\|^2 \tag{1}$$

In Eq. (1), $\boldsymbol{W}_j(t) \in \mathcal{R}^d$ is the reference vector associated with the j-th neuron in the internal layer at time t. The output of the i-th neuron in internal layer, $O_i(t)$, is a function of the difference between its reference vector $\boldsymbol{W}_i(t)$ and the input $\boldsymbol{X}(t)$ and also between the geometric distance of its position in the internal layer and the position of the winning neuron, win, as follows.

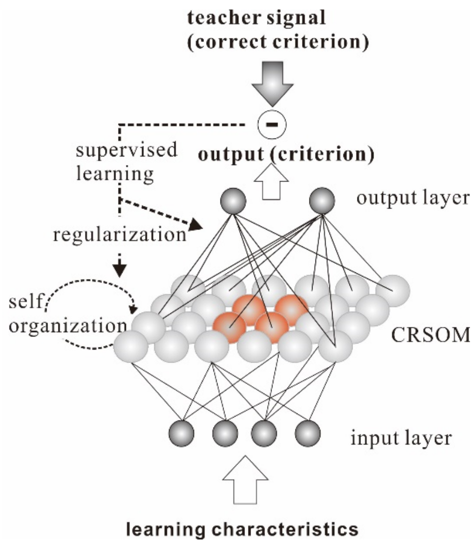


Fig. 1 Outline of rRBF and CRSOM.

$$O_i(t) = e^{-\|X(t)-W_i(t)\|^2}\sigma(win, i, t) \tag{2}$$

Here, $\sigma(win, i, t)$ is the neighborhood function defined as follows.

$$\sigma(win, i, t) = e^{-\frac{dist(win,i,t)}{S(t)}}$$
$$S(t) = S_{start}\left(\frac{S_{end}}{S_{start}}\right)^{\frac{t}{T}} \tag{3}$$

In Eq. (3), $dist(win, i, t)$ indicates the Euclidean distance between the i-th neuron and the winning neuron in the 2-D internal layer, while $S_{start} > S_{end}$ are positive constants and T is the maximum number of iterations. Hence, $S(t)$ is a constantly decreasing annealing function.

The output of the k-th neuron, $y_k(t)$, in the output layer can then be calculated as follows.

$$y_k(t) = f\left(\sum_i v_{ik}(t)O_i(t) + \theta_k(t)\right) \tag{4}$$

In Eq. (4) $v_{ik}(t)$ indicates the connection weight leading from the i-th internal neuron into the k-th output neuron at time t, $\theta_k(t)$ is the bias of the k-th output neuron while $f(x) = \frac{1}{1+e^{-x}}$ is a sigmoid function.

Here, the output of the network indicates the predicted context, for example grade of a student, given input $\boldsymbol{X}(t)$. The prediction error, $E(t)$ can be defined as follows.

$$E(t) = \sum_{k=1}^{N_{out}} (y_k(t) - T_k(t))^2 \tag{5}$$

Where T_k is the k-th teacher signal, the ideal answer that should be produced by the k-th output neuron while N_{out} is the number of the output neurons. In the same nature as Backpropagation in Multilayered Perceptron [24], [25], the prediction of the rRBF can be improved by modifying the connection weights between the internal and output layers and in this case the reference vectors associated with each internal neurons, as follows.

$$v_{ik}(t+1) = v_{ik}(t) - \eta\frac{\partial E(t)}{\partial v_{ik}(t)}$$
$$\boldsymbol{W}_i(t+1) = \boldsymbol{W}_i(t) - \eta\frac{\partial E(t)}{\partial \boldsymbol{W}_i(t)} \tag{6}$$

In Eq. (6), η is a positive learning rate.

The details of the derivation are elaborated in Ref. [21], and so here it is sufficient to explain that the modification of the reference vector associated with the i-th internal neuron can be expressed as follows.

$$\boldsymbol{W}_i(t+1) = \boldsymbol{W}_i(t) + \eta\delta_i(t)\sigma(win, i, t)(\boldsymbol{X}(t) - \boldsymbol{W}_i(t)) \tag{7}$$

Here $\delta_i(t)$ is a kind of regulatory signal from the output layer that transfers the output error into the modification of the reference vector. This internal layer generates a kind of reflection of high dimensional input into 2-D map, hence it can be readily visualized. Further, the regulatory signal $\delta_i(t)$ significantly distinguishes this 2-D internal layer from the conventional SOM, as it transfers the context of the inputs into the self-organizing process in the map, hence the name CRSOM.

The topographical preservation of SOM into 2-D maps allows users to visualize the underlying characteristics of the multidimensional data that are otherwise difficult to understand. However, the CRSOM differs from SOM in one important aspect. Whereas SOM only organizes a low dimensional map according to the topographical similarity of the input features, CRSOM also embeds the context of the data into the 2-D maps. The context-embedding in the 2-D allows users to visualize data in relevance with their context. This is an important property for visualization that is not sufficiently studied. For example consider data containing some features of students' learning patterns, where each data point is a multidimensional vector with the same dimension as the number of features. The multidimensionality naturally prevents us from visualizing the data, and thus it is difficult to understand the overall structures of the students' learning patterns. SOM maps those data while preserving their topological structures, in that data points that are similar in their original high dimension will be mapped close to each other in a low dimensional map, while dissimilar data points will be mapped far from each other. Now consider that those data points are associated with some contexts, for example the grades of the students. The introduction of the grades will not change the appearance of the SOM, because only the data's similarities are preserved. CRSOM provides novel visualization in that it also embeds the context of the data, in that it maps students with similar learning patterns that are associated to similar grades close to each other and generates more complex configuration for other cases. Further, CRSOM also generates different maps' appearances for the same data but different contexts. For example, the same students' learning data can be associated with different grades, like two steps pass-fail, or more detailed A-B-C grades. For these two different contexts, SOM will generate exactly the same 2-D maps, while CRSOM will generate two different maps. The context-relevance characteristic of CRSOM provides a new visual analytical tool, giving the users more information to intuitively understand the underlying structure of contextual data. Some examples of real-world e-learning data visualization will be explained in the next section.

3. Experiments with E-Learning Data

3.1 Visualization Experiments

In the first experiment, learning data were acquired from the LMS run by Japan Women's University in which 72 students participated to study physics through a web-based application. Over a semester, eight learning features (described in **Table 1**) were recorded. At the end of the semester the data were associated with the grades of the students. Hence, the grades become the context of the data. Naturally, there are various ways to grade students, for example, because each student either passed or failed the course, each data point can be associated with pass-fail labels in which 19 students passed while 53 students failed.

The challenge here is to transfer the eight dimensional data into two dimensional map and visually discover underlying structures of the data. Here, CRSOM is compared against widely used dimensional reduction and visualization methods, PCA, NCA and SOM.

Figures 2 and **3** show the 2-D representations of these data

Table 1 Features of Physics data.

dim	description
1	frequency of online contents accessed
2	number of different contents accessed
3	number of online physic simulations accessed
4	number of accesses one week before exam
5	average length in each access
6	frequency of access to practice problems
7	total length in solving practice problems
8	variance in access length

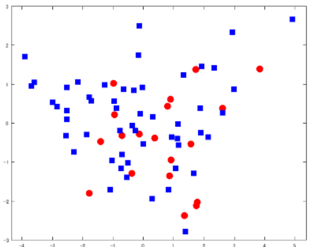


Fig. 2 PCA: Physics (pass-fail).

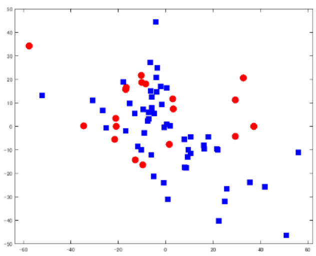


Fig. 3 NCA: Physics (pass-fail).

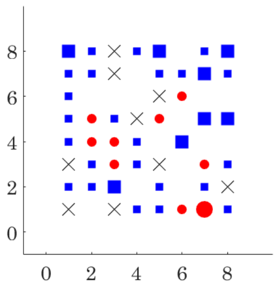


Fig. 4 SOM: Physics (pass-fail).

using PCA and NCA. In these figures ● shows a student that passed the exam and ■ shows a student who failed the exam. Visual observations of these figures do not indicate that there are distinctive underlying structures in these data.

Low dimensional projection of this data into a SOM is shown in **Fig. 4**. Here, ● shows the projections of the 8 dimensional learning patterns of one or more students that passed the course, ■ shows the projections of one or more students who failed while × shows the projection point that represents two students or a group of students having similar learning characteristics but with opposing results. The size of the marker indicates the number of data points that are projected into that marker. Although SOM, in Fig. 4, generates a clearer structure than PCA and NCA it is still difficult to use it for visual analysis. The CRSOM shown in **Fig. 5**

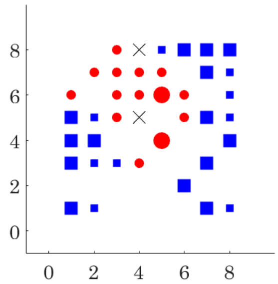


Fig. 5 CRSOM: Physics (pass-fail).

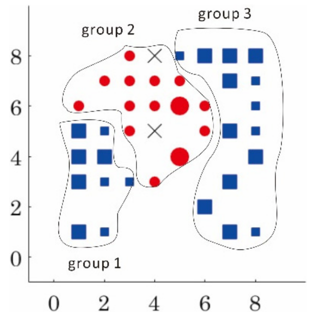


Fig. 6 Visual analysis.

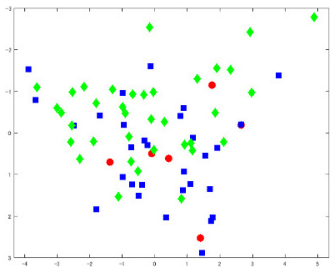


Fig. 7 PCA: (Physics-ABC).

generates more distinctive patterns.

The topographical-preservation nature, in which the positional proximity of two points in the 2-D map is related to their features similarity, of the CRSOM can be utilized for intuitive visual analysis. **Figure 6** is the enlargement of Fig. 5 shown with intuitive clusters of students. The first cluster is a relatively small group of students who failed, the second, in the middle is the group of students who passed, the third one is a large group of students who failed. Although the second group is rather large, there are two large circles indicating many students who shared similar learning behavior. Hence, it can be roughly understood that successful students have relatively distinctive learning patterns. The size of group 3 is an indication that there are many different learning behaviors that result in failure to pass this course. This is further clarified by the relatively large separation between groups 1 and 3, which is a strong indication that there are contrasting learning patterns resulting in failure. The past study in Ref. [26] indicates that many interesting insights can be discovered through this kind of visual analysis. To show the context-relevant properties of CRSOM, in the next experiment, the same data are associated with different contexts. Here, the data are associated not with pass-fail categories but more detail A-B-C levels. **Figures 7** and **8** show the 2-D representations using PCA and NCA, which is exactly the same as Figs. 2 and 3 aside from the color configurations and

Table 2 Features of Bookkeeping data.

dim	description	dim	description
1	num. attended classes	7	highest score (chap 4-6)
2	num. tests	8	num. practices (chap 4-6)
3	score of first test (chap. 2-3)	9	score of first test (chap. 7-9)
4	highest score (chap 2-3)	10	highest score (chap 7-9)
5	num. practices (chap 2-3)	11	num. practices (chap 7-9)
6	score of first test (chap. 4-6)		

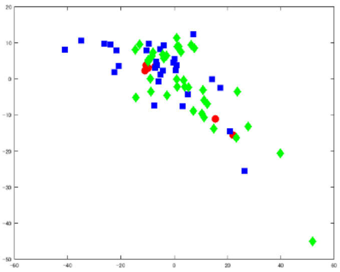


Fig. 8 NCA: (Physics-ABC).

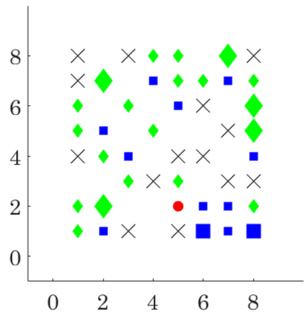


Fig. 9 SOM: (Physics-ABC).

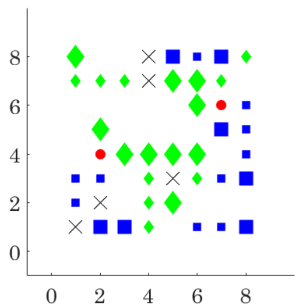


Fig. 10 CROM: (Physics-ABC).

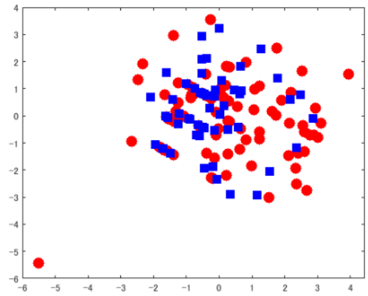


Fig. 11 PCA: (Bookkeeping).

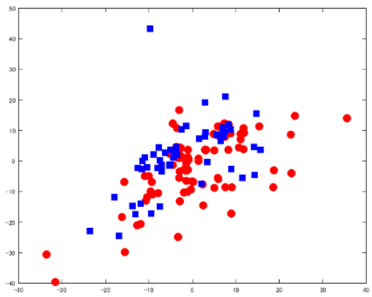


Fig. 12 NCA: (Bookkeeping).

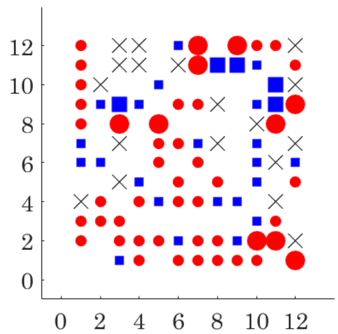


Fig. 13 SOM: (Bookkeeping).

marker types. In this experiment, there are 6 students with A grades, 30 students with B grades and 36 students with C-grades. A, B, C are represented by ●, ■ and ◆, respectively, in the following 2-D maps. The PCA and NCA 2-D representations do not offer any intuitive information about the data aside from the sizes of the distributions of A, B, and C students. SOM in **Fig. 9** is more informative in explaining the structure of the data but still lacks clarity. From CRSOM in **Fig. 10**, a few points about the underlying structure of the data can be learned. For example, the two A-students have relatively different learning styles, which is clear from their relatively large distance on the map. Some B students on the top right of the map are likely to share similar learning patterns with the C students in their surroundings. The closeness of A-students to other students on the map indicates that the grade excellence was not due to distinctive learning behaviors.

The data for the next experiment were acquired from a bookkeeping course run by a company with 145 participants, in which 11 learning features described in **Table 2** were recorded over several months. After the completion of the course, each student was graded and categorized as either above average (90 students) or below average (55 students). **Figures 11**, **12**, and **13** show the PCA, NCA and SOM representations of these data, where ● indicates an above-average student and ■ indicates a below-average student. For SOM in Fig. 13, × is a point in the 2-D map that is associated with two or more students from different categories. The PCA and NCA representations do not offer any clear pattern to learn from. SOM in Fig. 13 displays some clusters, but

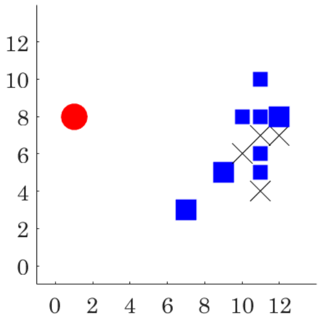


Fig. 14 CRSOM: (Bookkeeping).

Table 3 Features of English data.

dim	description
1	num. of access
2	total study time
3	num. lessons passed
4	num. lesson taken
5	check test average score

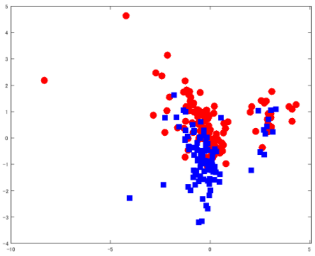


Fig. 15 PCA: (English).

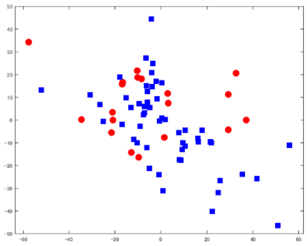


Fig. 16 NCA: (English).

it is still difficult to learn any underlying structure of the data. CRSOM in **Fig. 14** clearly displays some interesting structures of the data. The large ● indicates that the learning patterns of the above-average-students are somehow similar and the large distance of this point from the other shows that the learning patterns of these successful students are clearly distinguishable from other students' patterns. Some $\times$s indicate that there are some similar learning patterns with opposing results that are also similar to the learning patterns of the below-average students.

The data for the next experiment were acquired from an English course with 211 students, where the learning behavior of each student is characterized by five features described in **Table 3**. After the completion of the course, each student was graded and categorized into two groups, either above average or below average, indicated by ● or ■ in the following figures. **Figures 15**, **16**, and **17** respectively show the PCA, NCA, SOM representations of these data where the two categories are coarsely separated although the borders are not obvious. **Figure 18** shows

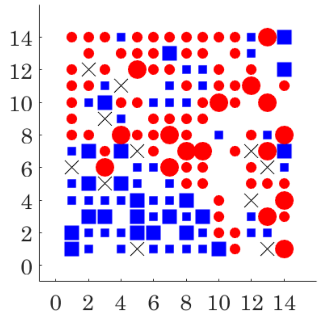


Fig. 17 SOM: (English).

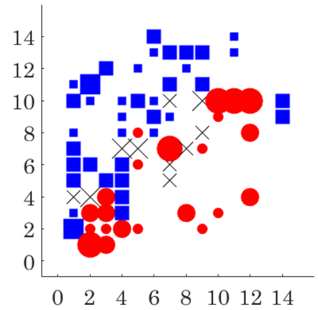


Fig. 18 CRSOM: (English).

the CRSOM representations of the same data. From this figure, it can be intuitively observed that there are a number of clusters where above-average students share similar learning patterns with below-average students and the $\times$s emphasize the borders. The few ●s at bottom right and ■s at the top indicate that there are distinctive learning patterns that lead to above-average as well as below-average results.

The experiments show that by embedding the context into topographic maps, CRSOM is able to visually present some underlying structures of the data that would otherwise be difficult to obtain from the traditional visualization methods.

3.2 Classification Experiments

Although the primary focus of this paper is to argue about the visualization properties of CRSOM, it should be noted that other than generating a visualizable map, rRBF is a neural network that can be readily utilized as a classifier. This can be a useful property to complement visual analysis, for example for predicting a student's expected grade based on the student's current learning behavior while also visualizing the student's relative position in the overall context-relevant 2-D map.

Here, the rRBF was compared against PCA, NCA, and K-Nearest Neighbors Classification [27] in the data's original high dimensional space and MLP. PCA and NCA are not classifiers hence, in the experiments the K-Nearest Neighbors classification was executed in the PCA and NCA representation spaces. K-Nearest Neighbor and MLP are not dimension reduction methods, and hence it is not possible to visualize their low dimensional representation of high dimensional data, but their classification performances can be used as relative benchmarks to evaluate the performance of the rRBF. In the experiment, K-fold cross validation tests were run for against the problems, and the error rates for all the classifiers were calculated. **Figure 19** shows the ex-

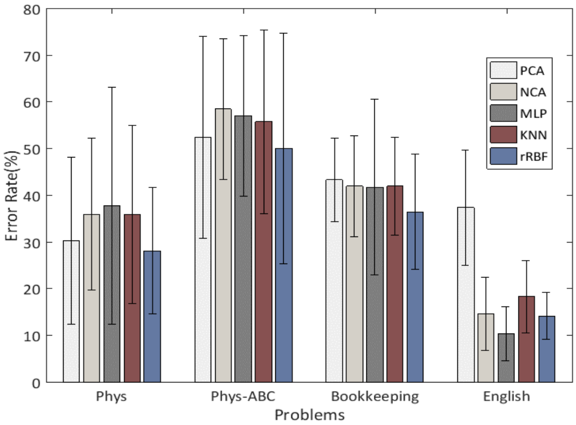


Fig. 19 Classification error comparison.

periment results, where it can be observed that rRBF performed nicely. For the first three problems the rRBF outperformed all other methods, even MLP and K-NN in the original high dimensional spaces of the data. This is surprising considering that as opposed to MLP and K-NN, which performed the original high dimensional space, rRBF must perform the classification in reduced dimensional spaces. For the final problem, rRBF did not outperform MLP, but it has the best performance among the three methods that ran classification in reduced dimensional-space.

4. Conclusion and Future Works

In this paper, it is shown that the rRBF can be potentially useful as an analytical tool for understanding learning data. Here, the objective is not to extract logical rules that govern the data but to intuitively discover the underlying structures of high dimensional data through visualization. Unlike many engineering system-generated data which are naturally bounded by logical or physical rules, learning data obtained from the learning behavior of students are often less logical, and so attempts to mine crisp logical rules from them often fail. In this light, this study attempts to give a different perspective for data understanding, by involving context-relevant visualization. The context-relevance of CRSOM is interesting for analytical purposes as it gives freedom for multi-perspective analysis by attaching different contexts to the same data.

This paper reports on preliminary attempts to use the rRBF and the resulting CRSOM for analytical means. The experiments show that intuitive understanding of high dimensional data through context-relevant visualization is promising. Further, rRBF does not only produce CRSOM for visualization but also predicts the context (in many cases the categories) of unknown inputs. The seamless combination of visualization and classification is one of the strengths of the rRBF as an analytical tool. In this paper, the rRBF was tested against small data. One of the immediate future objectives is to implement the rRBF for big data analysis. The overall objective is to build a user-friendly analytical system using the rRBF as an analytical engine. This tool will be seamlessly integrated into LMS, and will be opened to students as well as teachers to improve learning efficiency. For students, for example, the system will visualize a particular student's behavior relatively with the whole students group in some given context, and it also will predict the student's output based on past students' data. These functions hopefully help students to design their learning behaviors. Further, these functions may also improve the interface and reliability of the automatic recommendation systems, as unlike many current study-recommendation systems that often produce recommendations without any explanation, the visualization can offer intuitive justification for the contents of the recommendation. For teachers, multi-perspective visualizations offers better flexibility in analyzing students' learning data, which may help in designing better learning methods and environments.

It should be noted that CRSOM inherits the characteristics of the conventional SOM where the axes of the map cannot be easily interpreted. Unlike PCA or Factor Analysis, where the new axes are the linear combinations of the original ones, in CRSOM the dimensionality of the data are reduced in a non-linear manner while preserving their original topological structures. The linear transformation from a high dimensional feature space into a low dimensional one prevents us from easily understanding the meaning of the new axes. The interpretation of the axes of low dimensional representations is one of our future objectives. While in this preliminary study, the main objective was to generate intuitively visual maps for further analyzing and thus discovering the inherent information for the learning data, one of our immediate future objectives is to develop an efficient analytical tool for extracting common factors from clusters visualized on the map. For example, from a map like the one in Fig. 6, we can run variance analysis to extract common features that appear on a particular cluster. We believe that the idea to integrate the intuitive visual information and statistical analysis will potentially be a contribution for generating strong learning analytical systems.

Acknowledgments This work is supported by JSPS Grant-in-Aid for Scientific Research (C) Grant Number 25330419.

References

[1] McGill, T.J. and Klobas, J.E.: A task–technology fit view of learning management system impact, *Computers & Education*, Vol.52, No.2, pp.496–508 (2009).

[2] Coates, H., James, R. and Baldwin, G.: A critical examination of the effects of learning management systems on university teaching and learning, *Tertiary Education and Management*, Vol.11, pp.19–36 (2005).

[3] Lonn, S., Teasley, S. and Krumm, A.: Who needs to do what where?: Using learning management systems on residential vs commuter campuses, *Computer & Education*, Vol.56, No.3, pp.642–649 (2011).

[4] Lonn, S. and Teasley, S.: Saving Time or Innovating Practice: Investigating Perceptions and Uses of Learning Management Systems, *Computers & Education*, Vol.53, No.3, pp.686–694 (2009).

[5] Keim, D., Andrienko, G., Fekete, J.-D., Gorg, C., Kohlhammer, J. and Melancon, G.: Visual Analytics: Definition, Process, Challenges, in Information Visualization, *Lecture Notes in Computer Science*, Vol.4950, pp.154–175 (2008).

[6] Keim, D.: Information Visualization and Visual Data Mining, *IEEE Trans. Visualization and Computer Graphics*, Vol.8, No.1, pp.1–8 (2002).

[7] Kreuseler, M. and Schumann, H.: A Flexible Approach for Visual Data Mining, *IEEE Trans. Visualization and Computer Graphics*, Vol.8, No.1, pp.39–51 (2002).

[8] Costagliola, G., Fuccella, V., Giordana, M. and Polese, G.: Monitoring Online Tests through Data Visualization, *IEEE Trans. Knowledge and Data Engineering*, Vol.21, No.6, pp.773–784 (2009).

[9] Pearson, K.: On lines and planes of the closest fit to systems of points in space, *Philosophical Magazine*, Vol.2, pp.559–572 (1901).

[10] Jolliffe, I.: *Principal Component Analysis*, Springer-Verlag, New York (1986).

[11] Fisher, R.: The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, Vol.7, No.2, pp.179–188 (1936).

[12] Kohonen, T.: Essential of Self-Organizing Map, *Neural Networks*, Vol.37, pp.52–65 (2013).

[13] Kohonen, T.: Self-organized Formation of Topologically Correct Feature Maps, *Biological Cybernetics*, Vol.43, pp.59–69 (1982).

[14] van der Maaten, L.P.J., Postma, E.O. and van den Herik, H.J.: Dimensionality reduction: A comparative review, Technical Report TiCC-TR 2009-005, Tilburg University (2009).

[15] Roweis, S. and Saul, L.: Dimensionality reduction by locally linear embedding, *Science*, Vol.290, No.5500, pp.2323–2326 (2000).

[16] Hinton, G. and Roweis, S.: Stochastic neighbor embedding, *Advances in Neural Information Processing Systems*, Vol.15, pp.833–840 (2002).

[17] van der Maaten, L.P.J.: Visualizing high-dimensional data using t-sne, *Journal of Machine Learning Research*, Vol.9, pp.2579–2605 (2008).

[18] Goldberger, J., Roweis, S., Hinton, G. and Salakhutdinov, R.: Neighborhood Components Analysis, *Advances in Neural Information Processing Systems*, Vol.17, pp.513–520 (2004).

[19] Tenenbaum, J.B., de Silva, V. and Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction, *Science*, Vol.290, No.5500, pp.2319–2323 (2000).

[20] Geng, X., Zhan, D.-C. and Zhou, Z.-H.: Supervised Nonlinear Dimensionality Reduction for Visualization and Classification, *IEEE Trans. Syst. Man, and Cybernetics-Part B*, Vol.35, No.6, pp.1098–1107 (2005).

[21] Hartono, P., Hollensen, P. and Trappenberg, T.: Learning-Regulated Context Relevant Topographical Map, *IEEE Trans. Neural Networks and Learning Systems*, Vol.26, No.10, pp.2323–2335 (2015).

[22] Hartono, P. and Ogawa, K.: Visualizing Learning Management System Data using Context-Relevant Self-Organizing Map, *Proc. IEEE SMC 2014*, pp.3502–3506 (2014).

[23] Poggio, T. and Girosi, F.: Networks for approximation and learning, *Proc. IEEE*, Vol.87, pp.1484–1487 (1990).

[24] Rumelhart, D., Hinton, G. and William, R.: Learning representation by backpropagating errors, *Nature*, Vol.323, pp.533–536 (1986).

[25] Rumelhart, D., Hinton, G. and Williams, R.: Learning internal representations by error propagation, *Parallel Distributed Processing*, pp.318–362, MIT Press (1984).

[26] Merceron, L.A. and Yacef, K.: Clustering Students to Help Evaluate Learning, in Courtiat, J.-P. et al. (Eds), *Technology Enhnaced Learning*, Vol.171, pp.31–42 (2005).

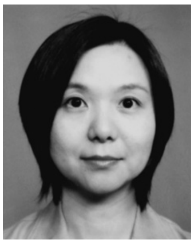
[27] Cover, T. and Hart, P.: Nearest neighbor pattern classifications, *IEEE Trans. Information Theory*, Vol.13, pp.21–27 (1967).

Editor's Recommendation

This paper applied the method called Restricted Radial Basis Function (rRBF) Network to the data obtained from LMS. rRBF produces the classification and 2D representation of high dimensional data as 2D topographical maps called Context-Relevance Self-Organizing Maps (CRSOM). Findings in this study show that very simple log data of students' behavior may correlate the learning performance of the students from the generated CRSOM. This would be the first step in showing the potential of rRBF applied to learning analytics and the efficiency of CRSOM. However, CRSOM itself generally has the difficulty in interpretation of generated CRSOMs. The future issue of this study, the interpretation of the axis of low dimensional representations, is also expected to be a good instance contribute to improve the method of rRBF and CRSOM.

(Michihiko Minoh, Academic Center for Computing and Media Studies, Kyoto University/An advisor of IPSJ Transactions on Computers and Education)

Pitoyo Hartono received his B.Eng., M.Eng., and Dr.Eng. degrees in applied physics from Waseda University, Tokyo, Japan, in 1993, 1995, and 2002, respectively. He was with Hitachi, Ltd., Tokyo, from 1995 to 1998. From 2001 to 2005, he was a Research Associate and Visiting Lecturer with Waseda University, and an Associate Professor with Future University Hakodate, Hakodate, Japan, from 2005 to 2010. He is currently a Professor with the School of Engineering, Chukyo University, Nagoya, Japan. His current research interests include computational intelligence and robotics.



Kayo Ogawa received her B.S. degree in 1993 from Japan Women's University, M.E. degree in 1995 from Waseda University, and completed her graduate school coursework in 1999 at Waseda University. She received her Doctor of Engineering degree in 2001 from Waseda University. She was a Research Associate at Waseda University from 1999 to 2001, and at Japan Women's University from 2001 to 2005. From 2005 to 2008, she was a Lecturer and from 2008 to 2014 she was an Associate Professor. She is currently a Professor at Japan Women's University. Her current research interests include learning analytics and e-portfolio systems.