Elman Neural Network Trained by using Artificial Bee Colony for the Classification of Learning Style based on Students Preferences

Nor Liyana Mohd Shuib¹, Ahmad Shukri Mohd Noor², Haruna Chiroma³,⁴ and Tutut Herawan¹,⁵,⁶,*

¹ Department of Information Systems, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia
² School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, Terengganu, Malaysia
³ Department of Information Systems, University of Malaya, 50603 Pantai Valley, Kuala Lumpur, Malaysia
⁴ Department of Computer Science, Federal college of Education (Technical), Gombe, Nigeria
⁵ Universitas Teknologi Yogyakarta, Kampus Jombor, Yogyakarta, Indonesia
⁶ AMCS Research Center, Yogyakarta, Indonesia

Received: 3 Apr. 2017, Revised: 11 May 2017, Accepted: 16 May 2017
Published online: 1 Sep. 2017

Abstract: Efforts have been made in recent times by educators and researchers to provide learners with appropriate learning objects (LO) based on their learning style (LS). Previous studies on the classification of LS, typically classify LS based on the description of the LS preference itself without giving attention to the students preferences. This study presents a new knowledge in classifying learning material based on learning style. In this paper, we propose Elman Neural Network (ENN) trained by using Artificial Bee Colony (ABC) (ABCENN) to create a classifier for the classification of LS (Diverging, Accommodating, Converging, and Assimilating) based on student preference of teaching strategies (TS) and LO. Our research extends on our previous work which considered only LO without TS. For the purpose of comparison, hybrid of ABC and backpropagation neural network (ABCBPNN) and ENN were applied to classify the LS of learners. Simulation results indicated that the propose ABCENN classifier outperforms ABCBPNN, and ENN classifiers with an accuracy of 97.12% and converges faster than the comparison methods. The propose ABCENN of this research can offer valuable information for educators, school administrators, and researchers to reach a decision on their respective students and to appropriately adapt their teaching methods. This in turn can significantly improve learners performance in understanding the subject matter.

Keywords: Learning Style, Learning Objects, Teaching Strategies, Kolb Model, Artificial Bee Colony, Elman Neural Network

1 Introduction

Recent studies show that there is a lot of effort by educators to provide students with Learning Object (LO) or Teaching Strategy (TS) based on their Learning Style (LS). This is as a result of studies shows that learning can be enhanced through the presentation of materials that are consistent with a students particular LS [1][2]. The LO is a tool for learning [3] whereas TS is the elements given to the students by the teachers to facilitate a deeper understanding of the information [4]. Educators tend to provide various types of LO and TS to cater for different type of LS preference. Researchers also proposed various educational tools such as e-learning to help students learn in their preferred ways [5][6]. Most of the researchers classify the LO and TS based on the description of LS preference itself. For example, Alharbi et al. [7] proposed an evaluation metric that can be used to categorize LO based on their compatibility with different LS. Franzoni & Assar [4] proposed taxonomy for combining and adapting TS with LS. They used Felder-Solomon LS. In their taxonomy, they categorize TS by linking the characteristics of TS with characteristic mentioned in the LS theory. Each LS can be accommodated with one or more TS. By using the taxonomy, educators can select the specific TS that is appropriate with students LS. Tulbure [8] categorized TS based on LS and different field of studies using Kolb LS Model. The study shows that students with different LS preference and field of study prefer different types of TS. Most researchers that
proposed the relationship between students LS used Kolbs LSI [9] and Soloman-Felder Index of LS (IFS)[4].

Researchers proposed a classification of LO and TS to help educators in choosing LO and TS that meet students requirements and preferences [7]. Churchill [10] proposed the classification of LO into six (6) types: presentation, practice, simulation, conceptual models, information, and contextual representation of objects whereas Wiley [11] classified LO into five (5) types: single type, combined-intact, combined-modifiable, generative-presentation and generative-instructional. However, LS was not considered in their classification.

Educators are facing a challenge in developing various LO and planning of TS to cover all the LS [12]. Our initial work Shuib et al. [13] has applied data mining technique to classify LS based on LO according to the students preference of use. This current research is to extend the work of Shuib et al. [13] which classifies LS based on LO according to the students preference of using the LO, whereas TS was not included in the study. Thus, the present study included both LS and TS to classify LS based on the LO and TS according to the students preference of using the LO and TS. Presently, data mining and educational systems are attracting unprecedented growing interest which makes educational data mining a novel and rising research community [14]. For that reason, data mining methodology is applied in this research. The commonly used technique in data mining is the back propagation neural network (BPNN) as reported in [15].

However, The BPNN is vulnerable to limitations: the possibility of being stuck in local minima, selection of the best BPNN structure, and training parameters is difficult to realize [16]. Elman Neural Network (ENN) is more effective in performance than the BPNN [17]. Karaboga and Basturk [18] reported that several comparative studies have shown that the Artificial Bee Colony (ABC) outperforms established population based algorithms such as genetic algorithm, Ant Colony, Differential Evolution, and particle swarm optimization. Thus, we propose to train ENN by using ABC (ABCENN) to build a classifier for the classification of LS based on LO and TS. To improve the performance of the ENN since hybrid ENN is more effective than the standard ENN [19]. The rest of this paper is organized as follows: Section 2 presents theoretical background of the study. Section 3 describes the theoretical background of the research work. Section 4 presents methodology, including the concept of ABC and ENN, data collection procedure and the LO and TS.

2 Theoretical Background

The use of LO and TS in the learning process is very crucial. Learning can be very effective if educators can use multiple options of LO and TS to accommodate various LS [20]. Differences between students LS and the way information is presented affects the students understanding of the subject and their attitude in class [21]. Students may lose their focus, become demoralized and assume that the course is too difficult for them, thus giving up on the subject [22].

2.1 Learning Object

The LO can be defined as any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning. LO is used to help students understand and apply the knowledge of the content. LO are typically design by instructional designers with inputs from an instructor or subject matter expert [23]. Faculty usually selects one type of learning object for a topic. However, they also provided multiple LO to help students explore a topic from different perspectives [3]. Learning object includes slide presentations, book, lecturer notes, educational games, video, audio-recorded, lecture, animation, instruction, real object, model, mind map, multimedia content, interactive tool, technology-supported learning, systems Intelligent and computer-aided instruction systems [7].

2.2 Teaching Strategies

Numerous studies had shown that learning process will be more effective if teaching strategies (TO) are in accordance with students LO [24]. Examples of TS are presentation, demonstration, quiz, simulation, exercise, exam, tutorial, lecture, problem solving, practice, laboratory, case study, discussion, question and answer, brainstorming, project based, practical training, educational visit, individual assignment, group assignment, self assessment and collaborative learning.

2.3 Learning Style

Each individual has a unique combination, or profile. This is influenced by individual characteristics, such as prior knowledge, education level, past experience, level of literacy, motivation, task confidence, aptitude, and LS [25]. These differences affect learning activities and outcomes. One of the most important factors that affect the outcomes of learning is LS [21]. The LS is defined as the preference (or predisposition) of an individual to perceive and process information in a particular way, or a combination of ways [26]. LS is accepted to be a students suggested way of learning [27].

2.3.1 The Learning Style Model

Several LS models exist in the literature. In this study, Kolbs LS Inventory is chosen because it is well
established in the literature and the commonly used as stated in the introduction. The LS Inventory is based on the Experiential Learning Theory [9]. Kolb's LS model is measured using LSs Inventory (LSI). The model outlined two levels. The first level is a four-stage cycle which is Concrete Experience (CE), Reflective Observation (RO), Abstract Conceptualization (AC) and Active Experimentation (AE). The ideal learning process engages all stages in response to situational demands. For learning to be effective, all stages must be incorporated. The second level represents the combinations of two preferred stages as illustrated in Fig. 1. The LSs are as follows:

a. **Diverging**: Diverges tend toward CE and RO. They are imaginative and are good at coming up with ideas and seeing things from different perspectives.

b. **Assimilating**: Assimilators are characterized by AC and RO. They are capable of creating theoretical models by means of inductive reasoning.

c. **Converging**: Converges are characterized by AC and AE. They are good at making practical applications of ideas and using deductive reasoning to solve problems.

d. **Accommodating**: Accommodators use CE and AE. They are good at actively engaging with the world and doing things instead of merely reading about the concept and studying them.

### 3 Proposed Method

#### 3.1 Artificial Bee Colony

In the ABC algorithm, the colony of artificial bees (AB) comprises of three (3) major steps, including: employed bees, onlookers, and the scout but the detail stages involved in ABC search is represented in Figure3. The first portion of the colony comprises of the employed AB, whereas the second portion contain onlookers. For each of the food source, only one employed bee exists. The employed bee of the discarded food source becomes the scout [18]. The pseudo-code of the ABC algorithm is as follows [33] (refer to Figure 1):
input neurons into the network in one direction to output neurons, and there is no feedback connection in such networks architecture. A network with recurrent connections has been described by Jordan. These recurrent networks associate static patterns with output patterns that are sequentially ordered. Hidden nodes visualize their own previous output to serve as a guide for subsequent behavior. Memory is provided in the network with the recurrent connections [35].

The ENN structure is present in Figure 4, where $\tau^{-1}$ represents time delay. $W_1, W_2, W_3$ are weight matrix between input and hidden layer, weight matrix between hidden and output layer, and weight matrix between context and hidden layer, respectively. Compositions of vectors at $s$th iteration are

$$x_i^{(s)} \in X, i = 1, 2, ..., n, y_j^{(s)} \in Y, j = 1, 2, ..., m, z_k^{(s)} = Z, k = 1, 2, ..., l, c_i^{(s)} = \mathcal{C}_i = j,$$

where $X$ is input layer vectors, $Y$ is output layer vector and $Z$ is hidden layer vector while $i, j, k$ and $i'$ are number of their respective nodes. However, the context units only interact with the hidden layer, not the external layer [17]. The network architecture of the ENN consists of three layers of neurons, namely internal and external input neurons in the input layer. The internal input neurons are also referred to as context units or memory. Internal input neurons receive input from their hidden nodes. Hidden neurons accept their inputs from external and internal neurons. Previous outputs of hidden neurons are stored in neurons in the context units [37].

### 3.3 Data Collection

#### 3.3.1 Sampling and Data Collection

To collect information about students preferential use of LO and TS, a five (5) points survey questionnaire was developed based on LOand TS reported in Tables 1 and 2 respectively, and LS described in section 2. The survey questionnaire was developed through previous literature and interview with experts in education. The interviews and literature provided a platform for the identification of unnecessary survey items and identified items required for inclusion in the survey questionnaire. A pilot test was conducted among five (5) students in a separate location from the target population to prevent interference. The pilot test allowed the identification and resolution of ambiguity and misunderstood wordings so as to present clear survey items to respondents.

The survey questionnaires used were printed and online for wider coverage. Filling of the survey questionnaire was voluntary and it was indicated in the questionnaire. The survey questionnaires were subsequently distributed to students in UiTM Arau, Perlis, Malaysia, through personal visits to the faculties, and online University mailing systems. The distribution was conducted by the principal investigator and the research assistance recruited purposely for the research project. The University selected for our research were chosen because of the willingness of the students to participate in the survey. The period for the data collection covered April 2013 to May 2013. A total of two hundred and fifty (250) survey questionnaires were distributed and two hundred and ten (210) marked questionnaires were successfully retrieved.

A total of two hundred (200) usable survey questionnaires were chosen for inclusion because they were appropriately filled by the respondents and ten (10) survey questionnaires were rejected due to inappropriate
response or carelessly answered which makes them unsuitable for data analysis. Finally, 80% of the questionnaires were found to be usable which are included in the final stage of the research.

### 3.3.2 Data Description and Partitioning

In the dataset, unnecessary rows of data were filtered out because they are not required in the modeling. Example of data that were removed includes gender, familiarity, awareness, among others. The data collected were tabulated which comprised of thirteen (39) columns. Thirteen (13) columns represent the LO1-LO13 (see Table 1), twenty two (22) columns representing TS14-TS35 (see Table 2), and the last four columns are LS (Diverging, Accommodating, Converging, and Assimilating). The last column is the classification of the LS based on LO and TS. The database contained seven hundred eight hundred (7800) observations. The performance of neural network depends on the data type presented to the network. Therefore, coding data into acceptable format plays a critical role in improving the neural networks performance. The coding of the data needed representation of the data in such a way that neural network can easily classify inputs based on their classes [38]. The ABCENN require numerical data to operate. As such, each column in the dataset was coded into nuimerial values for suitable operation by the ABCENN.

### 3.4 Experiment Setup of the proposed ABCENN

The ABCENN need training and test data for building the classifier. The data were partitioned for training and testing. Since there is no specific data partition ratio, we experiment with several partition ratios such 60% for training and 40% for testing, 70% for training and 30% for testing, 50% for training and 50% for testing, and lastly 80% for training and 20% for testing. In this study, we realized 80% : 20% to be more suitable than the other ratios. The data are presented to the ENN in the following order because the ENN is a dynamic NN: LO1-LO13, TS14-TS35. The parameters of the ENN require initialization for the ABC algorithms to run. We have thirteen (13) LO (see Table 1) and twenty two (22) TS (see Table 2) making a total of thirty five (35) independent research variables. Four (4) LS as dependent variables comprised of assimilation, diverging, accommodating and converging. The input neurons and output neurons were set to thirty (35) and four (4) respectively, because input neurons and output neurons depend on the number of independent variables and dependent variables [39]. The context layer neurons and hidden layer neurons are determined by the ABC search and the hidden layer was set to one (1), as argued by Pan and Wang [40] that one hidden layer is enough for approximation of any function. The activation functions at the hidden and output layer are tansig and purelin as recommended in [41].

The number of generations was set to a maximum of one thousand (1000). The objective function used for the ABC in the experiment is Mean Square Error (MSE). The ABC involves the searching of optimal weight parameters of the ENN. For Each Bee in the population represent the ENN for the minimization of MSE. The Bee requires setting of parameters associated with the successful implementation of ABC for searching the best solution.
The ABC performance is influenced by these parameters settings involving the number of bees recruited for the selection (NBRBS), the initial size of the patches (ISOP), the number of bees recruited for the best sites (NBRBS), number of sites chosen out of number of scout bees visited sites (NSNSB), a number of the elite sites out of selected sites (NESOSS), and number of scout bees (NCB).

Conversely, setting of these values has no systematic mechanism for automatically realizing the optimal values of the ABC algorithm. Typically, researchers employ trial-and-error to find the estimated settings of the parameters. We perform several experiments to realize the required parameter values. The modelling of the ABCENN continues for one thousand (1000) generations until learning curve with the minimum MSE is returned with the optimal ENN parameters as the best solution. For purpose of comparison, the ABCBPNN, ENN were also implemented to build a classifier for the classification of LS. The experiments were conducted on Personal computer (HP L1750, 4 GB RAM, 232.4 GB HDD, 32-bit OS, Intel Core 2 Duo CPU @ 3.00 GHz). The propose modeling process is depicted in Fig. 2.

## 4 Results and Discussion

### 4.1 Demography of the Respondents

Organization of the survey questionnaire respondents is a University. Among the respondents, 62.5% were found to be females, whereas 37.5% were males. The population of the respondents females is almost twice the population of males, which signify the number of females in the target population is significantly greater than the males. The respondents that indicated their familiarity with the LO and TS constitute 63.5%, whereas 36.5% indicate otherwise. This percentage shows that most of the respondents have familiarity with the LO and TS under study. With regard to awareness, 56% indicated that they are aware of the LO and TS. In the other hand, 44% of the total respondents revealed they are not aware of the LO and TS presented in the survey questionnaire.

### 4.2 Reliability and Validity Test

4.2. The reliability of the survey questionnaire measures was evaluated using Cronbachs coefficient alpha. The results indicated that the Cronbachs coefficient alpha value is 0.878 which is above the accepted benchmark of 0.6 [42]. Therefore, the survey questionnaire measures, show very high reliability. The content validity of the questionnaire was performed by experts, including the principal investigator.

### Table 3: Parameter settings of the ABCENN classifier

<table>
<thead>
<tr>
<th>Parameters of the ABCENN</th>
<th>Settings of the parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSNSB</td>
<td>29</td>
</tr>
<tr>
<td>NESOSS</td>
<td>4</td>
</tr>
<tr>
<td>NCB</td>
<td>134</td>
</tr>
<tr>
<td>NBRBS</td>
<td>60</td>
</tr>
<tr>
<td>NBRS</td>
<td>25</td>
</tr>
<tr>
<td>ISOP</td>
<td>0.8</td>
</tr>
<tr>
<td>ENN input neurons</td>
<td>35</td>
</tr>
<tr>
<td>ENN hidden neurons</td>
<td>8</td>
</tr>
<tr>
<td>ENN hidden layer Activation function</td>
<td>Tansig</td>
</tr>
<tr>
<td>NN output layer activation function</td>
<td>Purelin</td>
</tr>
<tr>
<td>ENN Number of hidden layers</td>
<td>1</td>
</tr>
<tr>
<td>ENN context layer</td>
<td>1</td>
</tr>
<tr>
<td>ENN context neurons</td>
<td>8</td>
</tr>
<tr>
<td>Number of output neurons</td>
<td>4</td>
</tr>
<tr>
<td>Generation</td>
<td>1000</td>
</tr>
</tbody>
</table>

### 4.3 The performance of ABCENN in classification of learning style based learning objects and teaching strategies

The methods described in the preceding sections can successfully be implemented using ABCENN and the empirical results are presented in this section. The parameters used for the modelling of the ABCENN are reported in Table 3, the optimal parameter settings are required for successful execution of the ABC to optimize the ENN.

The training performance of the ABCENN is depicted in Fig. 4 showing convergence performance, the curve has no oscillation and the straight line at the end of the curve indicate convergence. This shows successful convergence as argued by Nawi et al.[43]. Fig. 5, shows the performance of the convergence for ABCBP after training, the convergence is also successful. Comparing the search results of the ABCENN after training with that of ABCBP, it can clearly be seen that the ABCENN performs better than the ABCBP considering the values of the MSE which is the fitness function of the comparable models. Among the three (3) compared methods, only the two Figures (refers to Figures5 and 6) are shown because they have the best performance.

The convergence speed and accuracy on both training and test dataset for ABCENN, ABCBPNN, and ENN were observed and reports in Table 4. The comparative results on the same classification task shows the superiority of the ABCENN over ABCBPNN, and ENN in both accuracy and convergence speed which are the major performance indicators in machine learning, data mining, and soft computing. Jin [44] states that a model might introduce false optima, whereas it performs very well on the training dataset.
The performance of the ABCENN can be viewed as having a consistent performance since the classifier performs better than the comparative methods in both training and test dataset. For a performance to be consistent, two factors must have a consistent performance and the most significant is the accuracy on both the training dataset and the test dataset according to the opinion of Jin [44]. Thus, the performance of the ABCENN classifier is effective, reliable and robust. The possible reason for the performance exhibited by the ABCENN could be attributed to two possibilities: (1) the ability of the ABC to effectively deviate from the limitations of the ENN to improve its performance. (2). The LO and TS were collected from the students’ preferences whose salient attributes, possibly were detected easily by the ABCENN than the ABCBPNN, and ENN.

The propose ABCENN classifier can confidently be applied in the classification of student LS based on the student preference use of LO and TS. The LS is critical in a learning process. It can help students enhance their learning capabilities. For example, if educators know and understand students LS, they can propose the suitable LO or TS that is suited to students LS, which can assist the students in understanding the subject better and enhance their learning capabilities [20].

The classifier proposes in this research does not mean to replace the educators, but to serve as a decision support system for aiding the educators to reach an informed decision that can significantly improve student’s learning and easy understanding of a concept. Despite the high accuracy achieved by the classifier, yet, there some errors in the miss classification of the LS as can be seen in Table 4. The results in Table 4 are obtained after implementation of the experimental description presented in sub-section 3.4.

The educators and student need to be aware that the propose ABCENN classifier only assist in providing knowledge of the student LS based on LO and TS. For that reason, this does not in any way change the fact that any classification of student LS calls for immediate attention from the educators. The use of the classifier lies in supporting any decision made in regard to any kind of student. Shuib et al. [13] stated that the classifier propose in their study can ease the tedious manual method practice by the educators in the teaching and learning process for classifying students according to the student preference use of the LO. In this study, the classifier is more advanced for the classification since both LO and TS are considered unlike the study of Shuib et al. [13]. In the process of recommending LO and TS to students, researchers need to identify which LO and TS belong to which LS preference.

### 5 Conclusion

The main objective of the study is to the classify LS based on students preferential use of LO and TS. In this paper, we develop a classifier based on ABC and ENN using a dataset collected from students through the administration of questionnaires. The questionnaire was developed based on a LO and TS. The questionnaire provides a platform for the students to indicate their preference in the use of LO and TS. The ABCENN classifier build in our research classify student LS according to their LO and TS with an accuracy of 97.12%. Comparative performance analysis
indicated that the ABCENN outperforms ABCNN, and ENN in both training and testing dataset.

The finding of this research improve the result of our initial work. Showing that the performance of the propose ABCENN is effective, robust, and reliable in the classification of students LS based on LO and TS. The propose ABCENN classifier have the potentials of helping educators to select the appropriate LO and TS while considering student preference with very high level of accuracy. The appropriate selection of the students LS can help in enhancing student understanding of a particular subject matter and concept. This can improve the students performance significantly. The propose ABCENN of this research can offer valuable information for educators and researchers to reach a decision on their respective students and to appropriately adapt their teaching methods.

Acknowledgement.

This research is funded by University Malaya Research Fund Assistance (BK045-2013). The work Ahmad Shukri Mohd Noor is supported by Universiti Malaysia Terengganu Research Grant from Ministry of Higher Education Malaysia. The work of Tutut Herawan is partly supported by Universitas Teknologi Yogyakarta Research Grant Ref number O7/UTY-R /SK/O/X/2013.

References


recent research work focuses on proposing and improving the existing models in Distributed Computing and Artificial Intelligence. This includes research interests in the design, development and deployment of new algorithms, techniques and frameworks in Distributed System and Soft Computing. He has published more than 30 research papers on the distributed system at various referred journals, conferences, seminars and symposiums. In addition, he also reviewed more than 10 research papers. Furthermore, he has served various duties such as scientific committee, secretary, programme committee and session chair for various levels of conferences. He was also an invited speaker at International Conference on Engineering and ICT 2007 (ICEI 2007) organised by Universiti Teknikal Malaysia Melaka(UTeM). He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and IEEE Computer Society, Internet Society Chapter (ISOC-Malaysia Chapter), Malaysian Software Engineering Interest Group (MySIG) and International Association of Computer Science and Information Technology (IACSIT).

**Haruna Chiroma** is a faculty member in Federal College of Education (Technical), Gombe. He received B. Tech. and M.Sc both in Computer Science from Abubakar Tafawa Balewa University, Nigeria and Bayero University Kano, respectively. He is a PhD candidate in Artificial Intelligence department in University of Malaya, Malaysia. He has published articles relevance to his research interest in international referred journals, edited books, conference proceedings and local journals. He is currently serving on the Technical Program Committee of several international conferences. His main research interest includes metaheuristic algorithms in energy modeling, decision support systems, data mining, machine learning, soft computing, human computer interaction, social media in education, computer communications, software engineering, and information security. He is a member of the ACM, IEEE, NCS, INNS, and IAENG.

**Tutut Herawan** received PhD degree in computer science in 2010 from Universiti Tun Hussein Onn Malaysia. He is currently a senior lecturer at Department of Information System, FCSIT, University of Malaya. His research area includes rough and soft set theory, DMKDD, and decision support in information systems. He has successfully co-supervised two PhD students and published many articles in various book, book chapters, international journals and conference proceedings. He is an editorial board, guest editors, and act as a reviewer for various journals. He has also served as a program committee member and co-organizer for numerous international conferences/workshops.