Adaptive e-Learning Environment using Learning Style Recognition

George Abraham, Balasubramanian V., RA. K. Saravanaguru
School of Computer Science and Engineering, VIT University, Vellore, India

ABSTRACT
This paper aims to provide a survey of the major research works done in the domain of learning style recognition in an e-learning environment and proposes a Semantic Agent Framework for the e-Learning environment to detect individual differences existing among individuals, using their learning styles. Automatic detection of learning styles of an individual in an e-learning environment is an important problem that have been researched upon by many, as it proves beneficial to the learners to be provided with materials based on their individual preferences. To achieve this dynamic adaptability, we propose to use a mix of data-driven approach and literature-based approach. Out of the 71 models of learning styles that are described by different researchers, we consider the Felder-Silverman Learning Style Model (FSLSM) for our analysis.

Corresponding Author:
George Abraham,
c/o Prof. RA. K. Saravanaguru,
Assistant Professor (Selection Grade),
Member - Parallel and Distributed Computing Division,
School of Computing Science and Engineering, VIT University, Vellore – 632014,
Phone: 0416-2202811
Email: george.abraham1912@gmail.com

1. INTRODUCTION
The popularization of the Internet has changed the face of the education system with the introduction of e-Learning. There has been an increase in the demands of e-learning systems that cater to all needs in various fields of educations [1],[2]. One of the most desired characteristics of an e-learning system is personalization, as people with different skill sets use the system. Some people may be fast learners while some may be slow, some may need to practice more problems while others may need just example. These preferences are in general called the learning styles of an individual. The various preferences and requirements of an individual can be captured in a learner model that can be extracted from personality factors like learning styles, behavioural factors like user’s browsing history and knowledge factors like user’s prior knowledge [3]. Majority of the research work carried out are based on the learning styles as these are the most dynamic and give the best results if catered to properly [4].

The main challenge is the detection of the learning styles. Researchers have described various learning styles models like Myers-Briggs [5], Kolb [6], Honey & Mumford [7], Dunn & Dunn [8] and Felder-Silverman [9]. Research has proved that the Felder-Silverman Learning Style Model (FSLSM) is the most suited for the engineering students’ environment as it also considers the psychological aspects of a person [10]. The Index of Learning Styles [11] is a questionnaire-based approach for detection of learning styles based on the FSLSM. The problem with questionnaire-based approach is that it suffers from the “inaccurate self-conceptions of students” [12],[13] at a specific time. Moreover these questionnaires are incapable of tracking the changes in a learner’s learning style.

As a result of these problems, various researches have been conducted to come out with alternate automated solutions for learning style detection. These works can be broadly classified into two groups: data-driven approach and literature-based approach. Some of the noticeable works in the data-driven approach are by using
Bayesian Networks [14],[15], NBTree classifiers [3] and Genetic Algorithms [16]. Literature-based approach is a relatively new method with some of the noticeable works being done by Graf et al. [13], Dung and Florea [12] and Simsek et al. [18].

In this paper, we have taken a detail leveled comparision of all these major works in this domain, analysed them and have come to believe that literature-based and data-driven alone is not sufficient for predicting the learning style of an individual. There is also a lack of expressiveness in the representation of knowledge while analyzing the learning styles.

Personalization of the e-Learning system means to provide a system that adapts according to the learners’ learning process. In the next section, we will talk about adaptive web-based education followed by the process of automatic learning style recognition. Then we will study in detail many of the research work already done in this field. Finally we will summarise the complete works along with proposing a new solution.

2. ADAPTIVE WEB-BASED EDUCATION

The concept of an adaptive system was initially stressed by Bursilovsky and Peylo in 2003 [1]. They talk about improving the system of web-based education by providing an Adaptive and Intelligent Web-Based Educational System (AIWBES) as an alternative to the traditional systems. AIWBES adapts to the learners’ needs, knowledge and behaviour like a human teacher would do. An adaptive system modifies its solutions to a problem based on various factors, for instance the learners’ previous experience with the system whereas an intelligent system provides the same solution irrespective of the different needs of the learners. AIWBES is a mixture of adaptive hypermedia technologies and intelligent tutoring technologies. It also contains adaptive information filtering, intelligent monitoring and intelligent collaborative learning. Adaptive hypermedia mainly consists of adaptive presentation and adaptive navigation support while intelligent tutoring mainly consists of curriculum sequencing, problem solving support and intelligent solution analysis.

3. AUTOMATIC LEARNING STYLE RECOGNITION

Due to the various disadvantages of questionnaire-based learning style detection, the process has to be automated so that it can incorporate various aspects of the learner while modelling the learner. The process of automatic detection of learning styles consists of two phases: Identifying the relevant behaviour for each learning style and Inferring the learning style from the behaviour [19], as shown in Figure 1.

![Figure 1. Idea of Automatic Detection of Learning Style Preference](image)

The first step of identifying the relevant behaviour for each learning style consists of the following phases: Selecting the relevant features and patterns of behaviour, classifying the occurrence of the behaviour and defining the patterns for each dimension of the learning style [19], as shown in Figure 2. All this is performed by studying the various literatures of the respective learning model and other supportive research works that have already been done. The second step of inferring the learning style from the respective behaviour is where the approaches differ. But the
The initial step is to prepare the input data, which is common. This input data is prepared from the extracted information and is formulated in the form of matrices that correspond to each learning style. Then the calculation methodology can be data-driven or literature-based approach [19], as shown in Figure 3.

In our approach, we use the Felder-Silverman Learning Style Model and follow a mix of both the data-driven and literature-based approach by creating an ontological framework that can be then reasoned upon by a simple rule engine to detect the learning style.

Figure 2. Identifying the relevant behaviour for each learning style

Figure 3. Inferring learning styles from their respective behaviour

In our approach, we use the Felder-Silverman Learning Style Model and follow a mix of both the data-driven and literature-based approach by creating an ontological framework that can be then reasoned upon by a simple rule engine to detect the learning style.
4. RELATED WORK

A lot of research work has been undertaken in the field of automatic detection of learning styles and modelling of student behaviour for providing an adaptive personalized e-learning environment. Various techniques have been proposed and researched upon to automate the learning style detection process. All these techniques can be broadly classified into data-driven techniques and literature-based approaches. Some of the noticeable works in the data-driven techniques have been summarized below.

Garcia et al. [15] propose a Bayesian network based model that is used to infer the learning styles of the students according to the modelled behaviour, in order to provide adaptation and personalization in the Web-based education system. The proposed Bayesian model was evaluated for an Artificial Intelligence web-based course. They have opted for the Bayesian network as it helps them to model both the qualitative and quantitative information about a students’ behaviour. In the Bayesian network that they have modelled, the nodes represent the various student behaviours that determine their learning style while the arcs represent the relationships between the learning styles and these factors. The information that is used to model the Bayesian network is obtained from the student log file, maintained by the web-based system. This log file keeps track of various tasks carried out by the student like their activities in chat rooms and forums, the no. of test taken, marks obtained, etc. The authors have considered three dimensions of the Felder-Silverman Learning Styles Model namely perception, processing and understanding. They have omitted input and organization dimension as they are not currently considering the input methods of the lectures and also it has been demonstrated by Felder and Silverman themselves that most engineers are inductive learners. The value of the learning dimension is the one with the highest posterior probability. To verify the precision of their Bayesian model, they made the same set of students take the Index of Learning Styles questionnaire as well as made them attend the Web-based education system with the Bayesian model in place and they found out that they have obtained a precision of 77% in the perception dimension, 63% in the understanding dimension and 58% in the processing dimension.

In 2008, Garcia et al. [14] extends their previous work by providing suggestions based on the learning styles to the students. The eTeacher, i.e. the intelligent agent, assists students who are taking courses through an e-learning system called SAVER. SAVER has a predefined structure of how the individual course materials are structured; a well-defined hierarchical structure. Some of the recommendations that can be provided are for example to sequential learner, the agent can suggest to read a topic X before reading a topic Y whereas to an active learner, the agent can suggest to take part in a debate about the topic X in the forum. For analysis purpose, the option of accepting or rejecting a suggestion and even repeating a suggestion was provided to the student. This information is also vital to model the student behaviour. During analysis, they have found out that 83% of the total feedback received was positive. This showed that the system proved to be very promising.

Ozpolat and Akbar [3], propose an automated learner modelling based on diagnosing and classifying the learning styles by NBTree classification used in conjunction with Binary Relevance Classifier. The benefit of this learner model is that it uses only the data objects selected by the user for its modelling and is independent of the underlying LMS and other time-dependent learner behaviour. Hence this system can be integrated easily with any LMS or web-based search engines for education. The system will first find out the learner selected data objects (LSDOs) and then using a conversion unit, the learner profile table is obtained. Then a clustering unit based on the NBTree classification together with Binary Relevance classifier processes the profile table, by assigning labels to each row. These labels are based on the dimensions of the Felder-Silverman Learning Style Model. Finally, using the decision unit, the learning style is detected. In the profile table, the same combination of keywords can exist for learners of different learning dimension, hence multi-labelling are also allowed. Binary Relevance classifier is used for multi-label classification where it transforms the multi-label classification problem into one or more single label classification or regression problem. For experimental analysis, the authors used Personalized Search Tool for Teachers (PoSTech) as a front-end to the web-based search engines. For training the NBTree classifier, they used the ILS data from 10 graduate students to prepare the training data profile table. For the testing part, a group of 30 graduate students were used. This model yielded a success ratio of 70% in the processing dimension which is much more that even that given by Garcia et al. (2007). The perception and understanding dimension yielded an accuracy of 73.3% which is close enough to the other works. The accuracy in the input dimension suffered the most with only 53.3%.

Beragas-Suso et al. [20] designed two systems. The first system, iLessons, is a web browser-based system embedded within Microsoft Internet Explorer providing the teachers various features like content authoring by reusing the materials available on the World Wide Web (WWW) by drag and drop, creation of lessons like web pages, create special navigation zones for the students so that they don't lose focus on what they are doing and not wander into irrelevant websites and implementation of the lessons as a single file in a classroom selected by the teacher or to a specific group of students. This client-based system has been built by extending the functionality of the existing web browser Internet Explorer by using explorer bars, tool bands, browser help objects (BHOs) and asynchronous pluggable protocols (APPs). Subject-specific filtering using document categorization was the basis of the iLessons system. The second system extends the iLessons system by assessing the students' learning styles based
on FSLSM and recommending relevant pages to the students rather than restricting the usage of websites as a result of which students are turned towards a more research oriented approach of learning. For determining the patterns with which users with different learning styles make use of the Internet, the authors first analysed the learning styles of a group of people using the online available Index of Learning Styles (ILS). Finally, an accurate rule to predict the active-reflective dimension of learning style was determined taking into account the ratio between the images and text in a page combined with other parameters such as the average time spent on a page, the scroll distance and direction changes and the mouse movements. An approximate 24% increase in the accuracy of active-reflective dimension of learning style was achieved over the naïve prediction using the ILS.

In 2010, Beragasa-Suso et al. [17] extend their previous work and stress the need of a change in the prediction methodology as it was not efficient for the other dimensions. They introduced a new Unknown set for each pair of useful parameters, thereby creation of a more accurate rules to predict the Active or Reflective, Visual or Verbal and Sequential or Global dimensions of the learning styles, and they also worked out on rules for the Sequential/Global dimension. With the addition of the new Unknown set, users could now be categorized into, say, Active, Reflective or Unknown. This was required as most of the users do not fall crisply into any one set; rather tended to a particular dimension depending on the mood, circumstances or need. The accuracy in determining between Active or Reflective increased from 71% to 81% while that of Visual/Verbal increased from 71% to 82% and that of determining between Sequential or Global learners increased from 57% to 69%.

Deborah et al. [21] survey and outline the working of the existing learning style models and the various metrics used to identify them. They then suggest the use of FSLSM as the best suited for an e-learning system and suggest the use of fuzzy rules to handle certain uncertainty, mainly as an improvement to the work done by Bergasa-Suso and Sanders [17],[20]. For providing better classification, Deborah et al. extended the Unknown category to be further classified as reflective, medium reflective, active and medium active, by using a bell-shaped membership function for the fuzzy rules. They used the Sanders et al. model as base and studied the classification on the Computer Science and Engineering students of Anna University for the C-Programming Language course.

Chang et al. [16] propose a newer style of learning style detection by using an enhanced k-nearest neighbour (k-NN) combined with genetic algorithms (GA). The k-NN algorithm is enhanced using Pre-Contrast algorithm with Post-Comparison algorithm. The new algorithm was evaluated on a SCORM-compatible LMS by using social bookmarking software. The learners were made to interact with bookmarking sites like www.tagme1.com where the information of the links clicked is stored in the database. They evaluated the system for 25 learners with low number of links per learner and obtained a recognition accuracy of 72%. By changing the number of learners to 15 with high number of links per learner, a recognition accuracy of 86.66% was obtained.

In December 2011, Darwesh et al. [23], improved on the system by adding the concept of Learning Vector Quantization. They evaluated the system for different number of hidden neurons ranging from 20 to 100 and different learning rates in existence of varying epochs. It has been observed that for learning rate = 0.01, number of hidden neurons = 40 and epochs = 150, the recognition rate = 93.33% for 15 learners.

Montazer and Saberi [25] propose an Evolutionary Fuzzy Clustering (EFC) methodology with Genetic Algorithm (GA) for the recognition of learning styles of e-learners. Fuzzy clustering considers the real life uncertainty and classifies into more than one cluster. Standard clustering algorithm like Fuzzy C-means algorithm and K-means algorithm take into consideration only the compactness of a cluster and ignores their separation while the objective of EFC is to satisfy both. GA is used to optimize the useful objective function and find the centre of the clusters, hence evolutionary. It is observed that clustering based on EFC has more accuracy than the grouping the learners based on their behaviour logged in the LMS.

In 2012, Montazer and Saberi [25] again proposes a different methodology for the automatic detection of learning styles by using a three stage process in which the data was collected from the students using the ILS in the first stage. In the second stage, this data was evaluated using the Bayesian Networks. In the third stage, equivalent parameters for each question and for the dimension of the learning styles were formed based on the LMS. This system was evaluated on 40 M.Sc. Information Technology students studying four courses (IT, MIS, QPM and BPR).

Jyothi et al. [26] identify that most of the existing solutions lack accuracy in recognizing the learning styles when the number of data sets is below 150. As a result, they propose a recommender system that uses the learner information from the ILS and performs clustering. However prior information of the learner is necessary. The system was evaluated on a personalized e-learning system at C-DAC, Hyderabad R&D labs and tested on 105 student user profile data sets providing a good accuracy in the recognition.
<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Paper</th>
<th>Approach</th>
<th>Technology</th>
<th>Key Points</th>
<th>Assessment Methods</th>
<th>Precision /Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bergasa-Suso et al. (2005) [20]</td>
<td>Data-driven</td>
<td>Browser-based System with Rules</td>
<td>Processing dimension</td>
<td>67 students – ILS (Training) 7 students – iLessons</td>
<td>71% - Processing</td>
</tr>
<tr>
<td>2</td>
<td>García et al. (2007) [15]</td>
<td>Data-driven</td>
<td>Bayesian Networks</td>
<td>Detection only</td>
<td>27 Systems Engineering students – AI – SAVER</td>
<td>58% - Processing 77% - Perception 63% - Understanding</td>
</tr>
<tr>
<td>3</td>
<td>García et al. (2008) [14]</td>
<td>Data-driven</td>
<td>Bayesian Networks</td>
<td>Detection + suggestions</td>
<td>42 Systems Engineering students - AI - SAVER with eTeacher</td>
<td>83% feedback received was positive</td>
</tr>
<tr>
<td>5</td>
<td>Ozpolat and Akbar (2009) [3]</td>
<td>Data-driven</td>
<td>NBTree classification with Binary Relevance Classifier</td>
<td>Detection + suggestion Uses only data objects selected by the user LMS independent</td>
<td>10 graduate student (Training) 30 graduate students (Testing) – PoSTech</td>
<td>53.3% - Input 70% - Processing 73.3% - Perception and Understanding</td>
</tr>
<tr>
<td>6</td>
<td>Chang et al. (2009) [16]</td>
<td>Data-driven</td>
<td>Enhanced k-NN Clustering with GA</td>
<td>k-NN - Pre-Contrast and Post-Comparison Reduced no. of behavioural features</td>
<td>IRIS dataset by UCI 117 students - SCORM-compatible Java-based LMS - Windows XP</td>
<td>Increasing Accuracy</td>
</tr>
<tr>
<td>7</td>
<td>Bergasa-Suso and Sanders (2010) [17]</td>
<td>Data-driven</td>
<td>Browser-based System with Rules for Reasoning</td>
<td>More dimensions Improved rules Unknown category</td>
<td>67 students – ILS (Training) 7 students – same research task – iLessons</td>
<td>82% - Input 81% - Processing 69% - Understanding</td>
</tr>
<tr>
<td>9</td>
<td>Darwesh, Rashad and Hamada (2011) [22]</td>
<td>Data-driven</td>
<td>Analyzing web page content interactions using social bookmarking</td>
<td>No training methods used ILS filled using social bookmarking site</td>
<td>Study conducted on 25 and 15 learners participating in a bookmarking site such as <a href="http://www.tagme1.com">www.tagme1.com</a></td>
<td>For No. of learners = 25, Recognition = 72% For No. of learners = 15, Recognition = 86.66%</td>
</tr>
<tr>
<td>10</td>
<td>Darwesh, Rashad and Hamada (2011) [23]</td>
<td>Data-driven</td>
<td>Web pages tagging using social bookmarking and Learning Vector Quantization</td>
<td>No special effort from student side for collecting data LMS independent</td>
<td>By varying the Learning Rate Number of Hidden Neurons Values of Epochs</td>
<td>For Learning Rate = 0.01, Number of Hidden Neurons = 40 and Epochs = 150, Recognition Rate = 93.33% for 15 learners</td>
</tr>
<tr>
<td>11</td>
<td>Montazer and Ghorbani (2011) [24]</td>
<td>Data-driven</td>
<td>Evolutionary Fuzzy Clustering (EFC) method using Genetic Algorithm</td>
<td>People with similar learning styles in a cluster High computational and memory usage costs, so use Particle Swarm Optimization technique</td>
<td>98 undergraduate students – Fundamentals of Computer Networks course</td>
<td>EFC has more accuracy than grouping based on behaviour in log files of LMS</td>
</tr>
<tr>
<td>12</td>
<td>Deborah et al. (2012) [21]</td>
<td>Data-driven</td>
<td>Fuzzy Logic</td>
<td>Bell-shaped Membership function Better classification for “Unknown”</td>
<td>Comp. Sci. &amp; Engr. - Anna Univ., C-language</td>
<td>-NA-</td>
</tr>
<tr>
<td>13</td>
<td>Dung and Florea (2012) [12]</td>
<td>Literature-based</td>
<td>Simple rules on Matching Hints</td>
<td>LMS Independent Parameters - No. of visits and Time spent</td>
<td>44 UG students – Comp. Sci. – Politechnica Univ., Bucharest – AI course – Web-based LMS POLCA</td>
<td>70.15% - Input 72.73% - Processing 70.15% - Perception 65.91% - Understanding</td>
</tr>
<tr>
<td>14</td>
<td>Montazer and Saberi (2012) [25]</td>
<td>Data-driven</td>
<td>ILS + LMS logs + Bayesian Networks</td>
<td>Improved accuracy Decreased uncertainty</td>
<td>40 M.Sc. students on 4 different courses, done in three phases</td>
<td>-NA-</td>
</tr>
</tbody>
</table>
Literature-based approach is a new methodology that is being followed by researchers. This method is beneficial as it is LMS independent and also the data need not be present while modelling the students’ behaviour. Some of the noticeable works are those done by Graf [13], Dung and Florea [12] and Simsek [18]. These works differ in terms of the behavioural patterns that are considered for calculating the matching hints.

The work of Graf et al. [13] is an integral part of Graf’s Ph.D. thesis work [19], which first proposed the new methodology of literature-based approach for automatic detection of learning styles in LMSs. They studied the behaviours of 127 students during a course on Object Oriented Modelling in LMS Moodle. This method is works in a generic way and hence is independent of the LMS that we use. The main idea is to gather hints about a student’s learning style preferences from their behavioural patterns and then using a simple rule-based method calculating the learning styles from the number of matching hints. The behavioural patterns for the individual learning style dimensions are obtained from literatures as well as from the study of the model itself. Their occurrences and thresholds are obtained after studying various research works that have already been carried out. All this information is used to prepare the input data set. For the final calculation, based on the probability of occurrence of a pattern for a learning style, the sum of all the matching hints divided by the number of patterns available for that learning style gives the learning style of that person. Results show a higher precision in detecting the learning styles than data-driven approach.

Simsek et al. [18], in their paper, presents a method for prediction of learning styles of learners in a Learning Management System (LMS) by following a literature-based approach for automatic student modelling taking into consideration the learner interface interactions. They used the system, Moodle, to monitor a Mathematics course conducted for 27 learners and their learning styles were analysed with respect to active/reflexive dimension of the Felder Silverman Learning Styles Model and their approach gave them a precision of 79.6%. With the literature-based approach, first the behaviour patterns and their respective thresholds were determined. Then a simple rule-based method was used to calculate the learning styles from the number of matching hints. The six main features that were considered to study the behavioural pattern are videos, PDFs, forums, user profiles, quizzes and questionnaires. Their respective behavioural patterns with respect to the active/reflexive dimensions were determined and their thresholds assigned based on reviewing various literatures. The final comparison of the predicted value and that obtained from the Index of Learning Styles (ILS) was done by the formula developed by García et al. [15].

Dung and Florea [12] use the same literature-based approach proposed by Graf et al. [13] for automatic detection of learning style preference but consider the number of visits and time that the learner spends on learning objects as parameters. This method was evaluated on their own web-based LMS called POLCA by studying 44 under-graduate student over a course on Artificial Intelligence. Based on the characteristics of the FSLSM and the work carried out by Graf et al. [13],[19], the learning objects were properly labelled into each dimension. Then, for each learning style, the average of the ratios of time spent on each learning object to the expected time spent and number of learning objects visited to the total number of learning objects, is calculated. This average ratio is then used by a simple rule-base to decide the final learning style preference. The results obtained were approximate to that obtained by Graf et al. [13] as only two patterns were considered here.

5. CONCLUSION AND FUTURE WORK

Table 1 gives a summary of the complete literatures survey, mentioning the approach, technology, key points, assessment methods and the precision/accuracy obtained. Data-driven approach and literature-based approach both have their own benefits. Data-driven approach is more accurate as it is based on pre-collected data set, while literature-based approach has the freedom of LMS and other inherent systems. However in both the methodologies, the manner in which the knowledge is represented, like Bayesian Networks, or normal databases, etc. lacks the required level of expressiveness.

To provide more precision in the recognition of learning style, we propose a new methodology that incorporates both the approaches by creating an ontological framework for modelling the learner and using fuzzy reasoning engine. Simple rule-based reasoning can be performed on the ontology to extract the required content. The recognized learning style can be stored within the system in the learner database and can be used for further interactions with the learner so as to provide the learner with his/her relevant content. This makes the learning process similar to reinforcement learning in machines. The main advantage of using an ontological framework is that the level of expressiveness is very high when using OWL.

We are now working on creating a mathematical model for the complete recognition process based on the reinforcement learning algorithm SARSA so as to relate an adaptive e-learning environment to the approach of Reinforcement Learning, and also in creation of the ontological framework.
REFERENCES


ACKNOWLEDGEMENTS

The authors would like to thank the School of Computing Science and Engineering, VIT University, for giving them the opportunity to carry out this project and also for providing them with the requisite resources and infrastructure for carrying out the research.
BIOGRAPHIES OF AUTHORS

George Abraham received his B. E. degree in the year 2011 from the University of Mumbai, Mumbai, India. He is currently pursuing his M.Tech in Computer Science and Engineering from VIT University, Vellore, India. His areas of interests include Semantic Web, Web Services, Distributed Systems and Systems Security.

V. Balasubramanian is currently working as an Assistant Professor (Selection Grade) and Research Scholar in the School of Computing Science and Engineering at VIT University. His research interest is on Education Technology, Semantic Web and Artificial Intelligence in Education.

RA. K. Saravanaguru is currently working as an Assistant Professor (Selection Grade) and Research Scholar in the School of Computing Science and Engineering at VIT University. His research work focusses on Context Aware Systems, Middleware Development, Vehicular Safety and Web Services.