This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

Automatic Student Modelling for Detection of Learning Styles and Affective States in Web based Learning Management Systems

FARMAN ALI KHAN[1], AWAIS AKBAR[1], MUHAMMAD ALTAF[2], SHUJAAT ALI KHAN TANOLI[1], AYAZ AHMED[2]

[1]COMSATS University Islamabad, Attock Campus, Attock, 43600 Pakistan
[2]COMSATS University Islamabad, Wah Campus, Wah, 47040 Pakistan

Corresponding author: Farman Ali Khan (e-mail: farman_marwat@cuiatk.edu.pk)

This work was supported by COMSATS University Islamabad, Attock Campus

ABSTRACT In traditional learning environments, it is easy for a teacher to get an accurate and deep understanding about how students are learning and undertaking tasks. This results in teacher understanding about each student's learning preferences and behavior, which exhibits, for instance, students' learning styles and affective states. On the other hand, identification of learning styles and affective states in web-based learning environments is quite challenging. The existing approaches for the identification of learning styles such as questionnaire is not without limitations. Similarly, affective states identification approaches mentioned in the literature indicate several shortcomings. This paper proposes an automatic approach for the identification of learning styles and affective states in web-based Learning Management Systems (LMSs). The unique feature of this approach is that it is generic in nature. Using this approach, the students learning styles and affective states are calculated automatically from their learning preferences and behavior within a course. Evaluation of this approach was performed by following a study with 81 students. The results of the study were then compared with the learning styles and affective states questionnaires, which demonstrate that the suggested approach is more appropriate for the identification of learning styles and affective states. Therefore, using this approach, a tool (AsLim) has been developed and can be used by the teachers for the identification of learning styles and affective states of their students.

INDEX TERMS Learning Management Systems (LMSs), Learning Styles, Affective States, Adaptivity, Personalization

I. INTRODUCTION

The adaptive systems that were developed earlier, are either based on the user's knowledge, interest, and background etc. On the other hand, recent research is emphasizing over developing learning styles and affective states based adaptive systems, due to their high significance towards learning. Adaptive systems that are based on either learning styles or affective states usually adopt a collaborative approach for student modelling by intimating students to fill out a concerned survey questionnaire. Regarding learning styles, questionnaires have been criticized due to its static nature, as it describes learning styles of students in a particular environment and at a specific point [1]. Furthermore, this type of measurement can be slanted due to its dependency on the judgment of students. Similarly, regarding affective states, recent research indicates that subjective assessment with questionnaires is far from ideal due to the difficulty that students have in identifying their own affective states. Feldman et al., [2] also highlighted similar issues with the use of questionnaires, and concluded that rather than asking students about their learning preferences and providing explicit information about their state of learning behavior, their actual preferences and state of learning behavior can be exploited as an efficient source for the identification of their learning styles and affective states. Research conducted so far focuses only over the identification of learning styles in various learning environments, but affective states are continuously ignored. Although learning styles and affective states are equally important and both contribute towards the

students' learning process. Therefore, this research introduces an automatic approach of student modelling for the detection of learning styles and affective states in the web-based learning environments such as LMSs.
A concept of students' tracking, for example, how they are learning and demonstrating in a LMS, and using these information to infer the learning preferences and behavior of students is used. This approach has the capability of overcoming the problems associated with the use of questionnaires for the identification of learning styles as well as verbal, nonverbal, and intrusive approaches used for the identification of affective states. In an automatic student modelling approach, extra effort at the students' end is not needed for obtaining information regarding their learning styles and affective states. They only need to interact with system, and the system, on the basis of their preferences, behavior and actions, infers their learning styles and affective states. The information obtained using this approach is more authentic and error free. [3] Kay & Kummerfeld and Liyanage et al., [4] also indicated that this approach (i.e. data obtained from students log files and records) has the potential to overcome the challenge of developing useful student model for effective adaptive technologies. Furthermore, Shute & Zapata-Rivera [5] described that the data obtained from student log files and records represents an untapped, wasted resource that could be used for developing rich student model.

Automatic approach for student modelling can be more reliable and beneficial due to the longer time span for information gathering. Moreover, owing to the feasibility of frequent collection and analysis of students' behavior and actions, the automatic approach for student modelling can update the information in student model by identifying this change. The two main features make the automatic student modelling more prominent as compared to the other approaches employed for identifying the learning styles and the affective states. The first feature deals with developing a student model regarding the learning styles and the affective states from scratch. The second feature deals with updating and improving the already existing student model.

This paper is organized as follows: an overview of the related research work, which deals with an automatic student modelling with respect to the learning styles and affective states in web-based learning environments, is provided in the next subsection. In Section III, an automatic approach of student modelling for detecting the learning styles and affective states in LMSs is presented. Within this section, first of all, generic issues dealing with the discovery of relevant behavioral patterns for the identification of learning styles and affective states are discussed. Afterwards, for inferring the learning styles and the affective states from determined preferences and behavior respectively, two approaches namely the literature-based approach and the data-driven approach, are introduced. Moreover, for the purpose of evaluating these two approaches of automatic student modelling, a study is presented in the section IV. Finally, the discussion and conclusion is presented in section VI.

II. RELATED WORK

This section presents an overview of the related research work, which deals with the automatic student modelling in relation to the learning styles and affective states.

A. LEARNING STYLES

Karagiannis & Satratzemi [6] presented an automatic approach for the detection of learning styles in Moodle with the objective of providing adaptive courses. This approach initially exploits a questionnaire such as Index of Learning Styles (ILS) for detecting the students' learning styles. Later on, students' interaction data is exploited by a data mining technique for updating the student model. Similarly, Liyanage et al. [4] applied a classification algorithm, namely J48 decision tree, for the prediction of students' learning styles by analyzing the Moodle log data.

Likewise, regarding identifying the students' learning styles, WELSA [7] adopted automatic student modelling approach. The WELSA data analysis tool interprets the preferences of students and consequently builds and updates the student model. Similarly, the EDUCE [8] employs a blended approach for students modelling. In this system, when learners log in for the first time, they are emphasized to fill out multiple intelligence inventory named MIDAS (Multiple Intelligences Developmental Assessment Scales). This information is then used to determine their Multiple Intelligences (MI) profile and highest ranking intelligence. Regarding the MI profile, EDUCE maintains a static profile and a dynamic profile. A static profile is used for the initialization of student model and is generated from the MIDAS filled in by the learner, and a dynamic profile is generated by the system based on the learner behavior and navigation.

Cha et al., [9] through the pattern extraction module collected the behavioral patterns of learners regarding each learning style dimension mentioned by Felder and Silverman, from the user interface. The data regarding the patterns was then analyzed using two different types of machine learning techniques such as Hidden Markov Models (HMM) [10] and Decision Trees (DT) [11]. In this regard, an experiment was conducted with 70 students using an online course in the intelligent learning environment on the basis of respective patterns. Initially,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

the students were required to fill out the ILS questionnaire in order to assess their learning preferences. However, in this experiment, only that data of ILS questionnaire was evaluated that indicates either a moderate or strong preference for the respective learning style dimension.

Then the respective students' behavior for the interface was observed for deriving the students' learning style preference from the events of interface with the use of HMM and DT. Since, the proposed approach only considers the data when students have a strong or a moderate preference, this makes proposed approach only capable of being applied for the identification of learning style preference of students when students have a strong or moderate preference on either one or other pole of a specific dimension. Therefore, it is required to further investigate the approach which is highly accurate and incorporates the balanced preference as well. Similarly, Graf et al., [12] exploited automatic student modelling approach. They investigated learning preferences of students when considering adaptivity on the basis of learning styles in the LMSs. This approach looks into only detecting learning styles and ignores the behavior of students with respect to detecting affective states in LMSs.

B. AFFECTIVE STATES

Rajendran [13] developed an approach for the identification of affective states using log file analysis. This approach analyses the student interactions in an intelligent tutoring system i.e. Mindspark. However, this research considered only one affective state i.e. frustration, out of the six affective states reported by Bakera et al., [14]. Similarly, Salmeron-Majadas et al., [15] proposed a set of indicators such as keystrokes, digraphs and trigraphs, and mouse clicks and movement, derived from the interaction of user with keyboard and mouse. The aim of this study was to evaluate their use in the identification of affective states through non-intrusive method in one of the e-learning platform. The learners participated in the experiment scored their affective states' level of valence (emotion's pleasantness) and arousal (emotion's strength) after each mathematical problem. For this purpose, Self-Assessment Manikin scale (SAM) was used. SAM is a standardized questionnaire that allows users in rating their valence as well as arousal. In addition, the participants also entered the details about their own feelings. The information supplied by those various indicators, processed from original interactions logs of users, was evaluated with the help of that affective marking. A total of 96 mouse and 42 keyboard indicators were computed.

Qu et al., [16] while interacting with Virtual Factory Teaching System (VFTS) [17] inferred students' affective states such as confidence, effort, and confusion from student's current task, focus of attention, and expected time required for performing the task. For inferring the learners' focus of attention, a Dynamic Bayesian Networks (DBNs) [18] was used for combining the evidence from learners' interface actions and eye gazes. Subsequently, for the detection of learners' confidence level, effort and confusion, combination of the focus of attention and the information regarding activities of learners was used with the inference of a plan recognizer.

De Vicente & Pain [19] conducted a study for the identification of affective states by interacting with the systems, such as MOODS, which is a working model of ITS having a facilitation of self-reporting with an additional motivation to learn Japanese numbers. The participants were requested for observing the recorded students' interaction with the MOODS, and inferring and commenting on the students' affective states during interaction. In order to identify the students' affective states, some rules were formulated on the basis of inferences. Their validated results give an indication that inferring affective states diagnosis on the basis of information provided by the interaction between the students and the computer, is feasible.

Hershkovitz & Nachmias [20] argued that affective states affect the students' learning process and also explains their individual differences. However, it is very difficult to evaluate without direct contact with the student. They presented an approach which makes it possible to learn about the online learners by means of automatically and continuously collected digital traces. Affective states identification was based upon the three dimensions such as engagement, energisation, and source and by the seven computable learning variables corresponding to these dimensions. The measurement of three dimensions were based on the data solely extracted from log files. Learning variables, which describe patterns of learner's behavior, were identified using the technique of analysing Learnogram [21].

Cocea & Weibelzahl [22] argued that an affective (motivation) diagnosis which is based only on the interaction between the students and the system is not complete in spite of the evident benefits of unnoticeable diagnosis. They presented an approach which is founded on Social Cognitive Theory (SCT) [23]. For an affective diagnosis, a two-step approach was proposed. First of all, the system monitored the learners unobtrusively and diagnosed their disconnection on the basis of log files. The aim was to distinguish the engaged learners from the disengaged ones. The purpose of this practices was to focus further on disengaged learners. To predict level of engagement of the students from log data, eight different data mining techniques such as Bayesian nets, Logistic regression etc. were applied. The results of study concerning the first step were presented and showed overall good prediction. The next step was proposed to engage the disengaged learners in a dialogue for assessing their self-regulation, self-efficacy, and other relevant concepts.

III. AUTOMATIC IDENTIFICATION APPROACH FOR LEARNING STYLES AND AFFECTIVE STATES

The approaches described in section II regarding the learning styles and affective states are formulated for particular learning environments and hence, are customized specifically for these learning environments with the use of only those features and patterns, which are incorporated in the respective learning environments. Our proposed approach includes commonly used features of the LMSs. Therefore, this approach is capable of being applied to the LMSs in general. Nevertheless, developing an approach, which is applicable to almost all LMSs, is more difficult than developing an approach which is applicable to just one particular system. It takes into consideration a variety of features supported by different LMSs as well as the available data stored in the LMSs' database with respect to each pattern.

The proposed approach of automatically identifying the learning styles and the affective states can be divided into the two parts. The first part focuses on determining the learners' relevant preferences and behavior. This part is normally achieved on the basis of literature regarding the relevant learning style model and also the affective states factors or models. This part also involves investigations regarding the included features and patterns, the relevant patterns related to each learning style dimension and affective state, and the thresholds for classification of data. The second part focuses on arrangement and collection of data regarding the preferences and behavior of students for inferring the learning styles and the affective states respectively from this data. The two different approaches namely the data driven-approach and the literature-based approach are applied in this part.

The data-driven approach, for the purpose of developing a model each for identifying the learning styles and the affective states, exploits sample data. For instance, Cha et al., [9] derived the learners' patterns of behavior regarding each learning style dimension from the literature and then the data regarding the patterns were analyzed with DTs and HMMs. Similarly, Cocea & Weibelzahl [22] derived the behavioral patterns of learners regarding affective states from the literature and then the data regarding the patterns were analyzed with eight different data mining techniques such as Bayesian nets, Logistic regression etc. Viola [24] described that due to the possibility to track users' actions during navigation in the Electronic Learning Environments (ELEs), the data obtained are fully authentic. Therefore, it is recommended to use data driven approaches to analyze such data. Moreover, such approaches have been applied inside different research communities dealing with data mining, user modelling and intelligent tutoring systems, and e-learning. These communities have investigated data driven approaches for different purposes, for example, for providing adaptivity [25].

The literature-based approach uses a research literature as a base for developing a model each for the identification of the learning styles and the affective states. According to the literature, learners having specific learning style dimensions demonstrate specific preferences. Similarly, learners with specific affective states demonstrate specific behavior. The philosophy of the literature-based approach is to get hints about the learning styles of students by using their learning preferences, and similarly, to get hints about their affective states using their learning behavior. Afterwards, a simple rule-based mechanism is applied to the gathered hints. This method calculates learning styles and affective states from the number of corresponding matching hints. This simple method regarding the learning styles is similar to the one employed for the calculation of learning styles in ILS questionnaire, and also regarding affective states is similar to the approach employed for the calculation of affective states such as academic confidence and independence in the ACS and independence survey questionnaire respectively. This approach is advantageous because of its generality and applicability to the data collected from any course in the LMSs.

The data-driven approach is developed on the basis of concerned learning styles and affective states questionnaire calculation and aim of the approach is to imitate it. On the other hand, in addition to the FSLSM [26], information from the literature is used as a basis for the literature-based approach regarding the learning styles; whereas, in case of literature-based approach regarding the affective states, only the information from literature is used as a base. The proposed approach for student modelling aims at the identification of learning styles and affective states on a three-item scale. Thus, the suggested approach regarding the calculation of learning styles distinguishes between an active, reflective or balanced learning style. Similarly, the suggested approach regarding calculating the affective states

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

distinguish between, a low, balanced, or high affective state level.

A detailed description about the proposed approach for student modelling is presented in the following subsections. In the next subsection, an introduction regarding included features and patterns, the relevant patterns related to the each learning style dimension and the affective state, and the thresholds for classification of data, are presented. Subsequently, a description regarding the two approaches, each for inferring the learning styles from students' preferences and the affective states from the behavior of students, is presented. In addition to this, a method for the preparation of input data is described. In the final subsection, an evaluation of our automatic student modelling based on literature based approach, is presented by making a comparison of the results with the data-driven approach.

A. DETERMINING RELEVANT PREFERRED WAY OF LEARNING AND THE LEARNING BEHAVIOR

The approaches for automatic student modelling aim at the identification of learning styles on the basis of preferred way of learning and affective states on the learning behavior in the LMSs. For making the approaches capable of being applied to the LMSs in general, it is important to identify which learning preferences are appropriate for the identification of learning style and which learning behavior properly demonstrates the affective states. Regarding the preferred way of learning and the learning behavior, the LMSs features and patterns selection depends upon two requirements. Firstly, the relevance of patterns is necessary for the identification of learning styles based on the FSLSM [26] and affective states based on P. Sander & Sanders [27] (reported also by Shaukat & Bashir [28]) and Singh & EMBI [29] etc. The second requirement deals with gathering information about the patterns in LMS. The probability of information gathering for patterns should be as high as possible. This entails that selected features are required to be available and integrated in the most LMSs, the selected patterns are tracked by most of the LMSs, and the features are required to be normally used by the course developers and the teachers.

For the purpose of accomplishing the first requirement, features and patterns regarding the learning styles were deduced from the literature reported by Felder & Silverman [26]. Similarly, features and patterns regarding affective states such as confidence, independence, effort, and confusion were derived from literature reported by P. Sander & Sanders [27], Singh & EMBI [29], Pintrich & DeGroot [30], Qu et al., [16] respectively. Regarding the second requirement, only those features and patterns were selected which are available and incorporated in the most LMSs and normally used by the course developers and the teachers.

The following subsection provides a detailed description about the integrated features and patterns regarding learning styles and affective states. Afterwards, the next subsection describes how to classify the data of the incorporated patterns regarding learning styles and affective states. This classification distinguishes between various existences of patterns data, such as fewer visits or a higher stay at a specific type of learning object. The next subsections introduce the relevant patterns regarding the learning style dimensions and also the affective states respectively.

1) FEATURES AND PATTERNS SELECTION

For this study, commonly used features of the LMSs were exploited such as exercises, content objects, self-assessment tests, discussion and peer rating forum relating to the content objects, discussion forum related to queries about assignments, examples illustrating concepts, assignments, outlines, and patterns dealing with the students' navigation behavior within the course. Similarly, the patterns against each feature were selected in terms of their generality in LMSs and with respect to their pertinence for the learning style dimensions and also the affective states.

For example, patterns (content_visit) and (content_stay) were selected against the content objects, which describe the total visits and the time spent by learners over the contents consecutively. Table 1 describes the features and associated patterns exploited for the identification of learning styles and affective states. In the subsequent section, a recommendation for the classification of different occurrences of patterns data regarding the learners' learning preferences and behavior is introduced. Afterwards, in addition to relevant patterns regarding the learning style dimensions and the affective states, the respective occurrence of learning preferences and behavior are discussed.

2) CLASSIFYING THE DIFFERENT OCCURRENCES OF PATTERN DATA

This section describes the classification of different occurrences of patterns (introduced in the previous section) data regarding the learners' learning preferences and behavior. The uniqueness of this approach is that it is generic and applicable to various courses having different attributes. Regarding this, a three (3) item scale is employed, which divides each, the learning preferences and the behavior into 3 groups. The three groups describe the high, low, and moderate occurrence. This classification

IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

employs generic threshold values instead of considering the average learning preferences and behavior, in a particular course. Use of generic threshold values is advantageous because, results in the form of detected learning styles and affective states are not dependent over learning preferences and behavior of other students. On the other hand, using the average learning preferences and behavior for getting threshold values will result in a distribution of learning styles and affective states which is defined in advance for each respective pattern, and it may not be applicable for medium and small size groups.

Therefore, generic thresholds were used to make the suggested approach suitable for the small and medium size groups. However, Alberer et al., [31] argued that general thresholds can vary, depending on subject, course structure, as well as students' experiences.

In the subsequent paragraphs, recommendations for thresholds from the literature (e.g. Graf et al., [12], Garcia et al., [32], Picciano [33], Cheng [34]) are used as a base, with the further consideration of the attributes of particular course, and are described in table 2.

The thresholds of 75% and 100 % regarding visiting content objects were used. Content objects are initially, delivered verbally through the lecture in a classroom, and later on uploaded through LMS by the teacher or course designer, where the students had the opportunity to visit and explore the learning material. Moreover, the learning material can be downloaded by the students either for taking a print or reading it offline. The content objects are helpful for students, not only for the course examination preparation, but for other purposes as well. For example, for looking up information about the questions presented in exercises and self-assessment tests. This demonstrates that by viewing the pattern regarding content objects visits, it can be concluded that whether the students used the content objects for getting knowledge about certain topic or they looked at examples etc. Similarly, regarding the students' stay at content objects i.e. students were expected to spend on this type of object, threshold values of 10% and 20 % of the predefined time estimate, were used. The reason behind introducing and using the low percentage threshold values is due to the fact that students had the opportunity to read the content objects online as well as they can download the learning material and read it offline.

Regarding the outlines visit, threshold values of 75% and 150 % were used. Moreover, regarding the outline stay i.e. the time students were expected to spend on this type of object, the threshold values of 50% and 75 % of the predefined time estimate, were used.

TABLE 1
PATTERNS FOR THE IDENTIFICATION OF LEARNING STYLES AND AFFECTIVE STATES

Features	Patterns	Description of behavioral patterns
Content Objects	content_visit	Number of visits to content objects
	content_stay	Total time spent on content objects
Outlines	outline_visit	Number of visits to outlines
	outline_stay	Total time spent on outlines
Exercises	exercise_visit	Number of visits to exercises
	exercise_stay	Total time spent on exercises
Course overview	course_ovview_ visit	Number of visits to course overview page
	course_ovview_ stay	Total time spent on course overview page
Self-assessment tests	selfasses_visit	Performed self-assessment questions
	selfassess_stay	Total time spent on self-assessment tests
	selfassess_revision	Number of times a learner revised self-assessment test
	quest_graphics	Number of graphical questions which are correctly answered
	quest_text	Number of textual questions which are correctly answered
Examples	example_visit	Number of visits to examples
	example_stay	Total time spent on examples
Forum for assignment related queries	forum_assignment_visit	Total number of visits to forums for assignments
	forum_assignment_post	Total number of posts in the forums for assignment related queries
	forum_assignment_post_reply	Total number of replies to the posts in the forums for assignment related queries
Forum related to the content objects	forum_content_ visit	Total number of visits to forums for content objects
	forum_content_post	Total number of posts in the forums related to content objects
	forum_content_post_repl	Total number of replies to the posts in the forums related to content objects
Assignments	assignment_ revision	Total number of times a student revised his/her assignment after the initial submission
	assignment_stay	Total time spent on assignments
Navigation	navigation_skip	Total number of content objects skipped by the learner

The thresholds regarding visiting and performing exercises were used with 25% and 75% of the total

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

available exercises. Regarding the time stay at exercises, thresholds of 50% and 75% were used, with respect to the expected learning time to spend on this type of object.

Regarding course overview visits, thresholds of 10% and 20% were used in terms of total learning objects visits. Similarly, with regard to course overview stay i.e. total time students were expected to spend, thresholds of 50% and 75% of predefined time estimate were used. The thresholds regarding the visits of the self-assessment tests were used with 50% and 75% of the number of self-assessment questions. For stay at self-assessment tests, thresholds of 50% and 75% were used, in relation to the expected learning time estimates of students with high interest in that kind of object. The thresholds regarding revision of the self-assessment tests were used with 25% and 50% of the total number of self- assessment tests. With respect to the performance on specific types of questions, thresholds of 50% and 75% were used, based on the number of questions that were answered correctly. The thresholds regarding visits of the examples were used, with 50% and 100% of the total number of examples that were available. Regarding the stay at examples, thresholds of 50% and 75% were used in relation to the expected learning time estimate of students with high interest in that kind of object. *a)*

For discussion forums, thresholds regarding the visits of the assignment related forum were used with 2 and 4 for each individual assignment. Regarding the queries posted on the assignment related forum, thresholds of 1 and 2 were used for each individual assignment. Similarly, regarding replies to queries posted on the forum related to assignments, thresholds of 1 and 2 were used for each individual assignment. For the discussion forum related to content objects, thresholds regarding visits of the forum were used with 2 and 4 per week. Moreover, regarding the posts, thresholds of 1 and 2 were considered per week. Similarly, regarding replies/commenting to posts, thresholds of 1 and 2 per week were used. For the assignments, thresholds regarding the number of revisions were set with 1 to 2 for each individual assignment. For stay at assignments, thresholds of 50% and 75% were used with respect to the expected submission time of students, with greater attention on that kind of learning object. Regarding skipping learning objects, 1% and 2% thresholds were used with respect to the total number of visited learning objects.

3) RELEVANT PATTERNS FOR IDENTIFYING LEARNING STYLE DIMENSIONS

The patterns for identification of each learning style dimension and information regarding whether a low or high existence of respective patterns data is relevant, are presented, in this section. The research literature concerning the FSLSM [26] is used as a basis for the relevant patterns and the information about their occurrence. The patterns regarding active or reflective, sensing or intuitive, visual or verbal, and sequential or global dimensions are introduced in the respective subsections.

According to FSLSM, each learning style dimension comprises of two opposite poles thus, when a high existence of respective pattern data indicates one pole of the dimension; a low existence of the same pattern data indicates opposite pole of the same dimension. Therefore, for the reflective, global, verbal, and intuitive dimension the relevant occurrences of patterns are simply opposite. The respective patterns indicate that the identification of each learning style dimension comprises of comparatively greater number of patterns. Considering a great number of patterns provides more elaborated information and is essential particularly for formulating an approach which has the capability of identifying the learning styles in the LMSs in general instead of in the one particular system, owing to the possibility that information related to some patterns are unavailable. Table 3 illustrates the patterns for each learning dimensions.

a: ACTIVE/REFLECTIVE DIMENSION

Learners who are more active in processing information are categorized as active learners. In order to process the information, active learners usually discuss it, apply it, and explain it to the other. On the other hand, reflective learners have a great preference towards thinking about material prior to acting and they preferably work alone. Regarding discussing, participation in discussions through the discussion forums such as "discussion/peer rating forum" and "assignment forum" are used to get indications regarding the preference of students for active or reflective learning. For the purpose of discussing something, active learners are supposed to post a great number of times while reflective learners are expected to participate in a passive manner by reading the "assignments related forum" and the "discussion and peer rating forum" posts instead of posting actively in these forums. Therefore, the number of posts regarding the "discussion/peer rating forum" can be used as a pattern for the identification of an active or reflective learning style.

In terms of applying (testing and trying things out), active learners are supposed to attempt a great number of exercises and self-assessment tests. Moreover, overall, active learners are expected to stay at self-assessment tests for less time, however, they are expected to stay at exercises for more time. Similarly, regarding preference of explaining, active learners are supposed to reply to the queries posted in assignment related forum. In addition, they are expected to comment on newly posted queries of other students in the discussion and peer rating forum regarding the content objects.

Reflective learners, on the other hand, have a great tendency towards thinking about the learning material and reflecting about it, therefore, they are supposed to have a great number of visits to the learning material such as content objects. Furthermore, they are expected to stay at outlines and content objects for a long time. Moreover, for the purpose of producing good results, they are expected to stay at exercises and self-assessment tests for a long time.

b: SENSING/INTUITIVE DIMENSION

Learners gravitating towards concrete material, for instance, facts and data are characterized as sensing learners. Using well-established methodologies, they (sensing learners) tend to solve their problems. Moreover, such learners are slow at working and have patience but cautiousness while dealing with details, and more often perform well when they use repetition as a learning tool. Regarding learning from concrete material and also from existing approaches, sensing learners prefer examples. Thus, the total visits as well as the stay at examples can be used as a pattern for the identification of sensing or intuitive learning style. Regarding the preference to work cautiously but slowly, sensing learners are assumed to be the ones who take more time for the submission of exercises and self-assessment tests. With regard to the use of repetition as a learning tool, sensing learners are supposed to repeat the self-assessment test and are also expected to get a satisfactory score in the final attempt.

According to FSLSM, intuitive learners, on the other hand, welcome challenges in learning and do not like details. Furthermore, they are innovative and do not like repetitions. Additionally, unlike sensing learners, intuitive learners work at a high pace. Regarding learner's behavior when faced with challenges, intuitive learners are expected to spend less time on assignments and have a low number of assignment revisions. Having a high number of content objects' visits, spending more time; and having fewer examples' visits, spending less time, indicates the behavior of students in terms of using examples exclusively as a secondary material and also that they are disinterested in niceties. No revision of self-assessment tests on getting a moderate or satisfactory grade in the initial attempt indicates the students' behavior in terms of disliking repetitions.

c: VISUAL/VERBAL DIMENSION

Learners that learn best using graphical material such as images, graphics and flow charts are categorized as visual learners. Therefore, correctly performing a great number of graphical questions gives us an indication about the visual dimension of a learner. On the other hand, verbal learners are those learners which prefer input in the form of words, irrespective of the input being written or oral.

Therefore, a great number of visits to and more time spent on content objects gives us an indication about the students' verbal dimension. In addition, verbal learner have a great tendency to discuss and communicate with others. Therefore, having a high number of posting in the discussion forums as well as frequent replies to forum posts indicates students' behavior in terms of verbalizing.

d: *SEQUENTIAL/GLOBAL DIMENSION*

Sequential learners feel comfort while dealing with the details and like to be logical in problem-solving. Since such learners navigate through the course in a linear fashion, therefore, the information related to the sequential dimension of students can be obtained using navigation behavior of students. According to FSLSM, global learners, on the other hand, tend to get an idea/overview of the contents and do not go into much details. They obtain big picture regarding the course contents and have a high number of navigation skips. Therefore, frequent visits to course overview and chapter outlines indicate the behavior of students in terms of globalizing.

4) RELEVANT PATTERNS FOR IDENTIFYING AFFECTIVE STATES

This section introduces relevant patterns for the identification of students' affective states such as confidence, independence, effort, and confusion. The patterns and the information about their existence is based on the relevant research literature concerning the confidence [27], independence [29], effort [30]; [16]; [35], and confusion [16]; [36]. The patterns regarding confidence, independence, effort, and confusion are introduced in the following sub sections. Since each affective state consists of two opposite levels therefore, for one level of the affective state, indications are obtained using high occurrence of a specific pattern data, while for opposite level of the same affective state, indications are obtained using a low occurrence of the same pattern data.

Table 4 describes a high occurrence of the relevant patterns data for high level of each affective state and vice versa. Moreover, the identification of each affective state comprises relatively a great number of patterns. Considering such a great number of patterns provides more elaborated information and is quite significant for the development of an approach with the capability of identifying affective states in the LMSs in general instead of in one particular system, owing to the unavailability of information with regard to some of the patterns in some

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

cases. A detailed description regarding relevant patterns for each affective state is given in the following subsections.

TABLE 2
THRESHOLD VALUES FOR EACH PATTERN

Patterns	Description of thresholds for each pattern	Lower Threshold	Higher Threshold
content_visit	content objects visits (based on the total available content objects)	75%	100%
content_stay	stay at content objects (based on a predefined expected time value)	10%	20%
outline_visit	outlines visits (based on the total available outlines)	75%	150%
outline_stay	stay at outlines (based on a predefined expected time value)	50%	75%
exercise_visit	exercises visits (based on the total available exercises)	25%	75%
exercise_stay	stay at exercises (based on a predefined expected time value)	50%	75%
course_ovview_visit	course overview visits (based on the number of visited learning objects)	10%	20%
course_ovview_stay	stay at course overview (based on a predefined expected time value)	50%	75%
selfasses_visit	performed self-assessment questions(based on the total number of available questions)	50%	75%
selfassess_stay	stay at self-assessment tests (based on a predefined expected time value)	50%	75%
selfassess_revision	number of times a learner revised self-assessment test (based on the total number of self-assessment tests)	25%	50%
quest_graphics	Correctly answered questions about graphics (based on the total number of available questions about graphics)	50%	75%
quest_text	Correctly answered questions about text (based on the total number available questions about content objects presented in textual form)	50%	75%
example_visit	examples visits (based on the total available examples)	50%	100%
example_stay	stay at examples (based on a predefined expected time value)	50%	75%
forum_assignment_visit	number of visits in a forum (based on the number of assignments offered during the course)	2	4
forum_assignment_post	number of postings in the forum (based on the number of assignments offered during the whole course)	1	2
forum_assignment_post_repl	number of post replies in the forum (based on the number of queries, posted related to each assignment during the whole course)	1	2
forum_content_ visit	number of visits in a forum during each week	2	4
forum_content_post	number of posts in the forum during each week	1	2
forum_content_post_repl	number of posts replies in the forum (based on the number of posts related to content objects during the course)	1	2
assignment_revision	number of times a student revised his/her assignment	1	2

IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

	after the initial submission		
assignment_stay	stay at assignments (% of assignments submitted well before the deadline)	50%	75%
navigation_skip	number of times learning object is skipped (based on the number of visited learning objects)	1%	2%

TABLE 3
PATTERNS FOR EACH LEARNING STYLE DIMENSION

Active / Reflective	Sensing/ Intuitive	Visual/ Verbal	Sequential/ Global
content_visit(-)	content_visit (-)	quest_graphics(+)	outline_visit(-)
content_stay(-)	content_stay(-)	ques_text(-)	outline_stay(-)
outline_stay(-)	example_visit(+)	content_visit(-)	course_ovview_visit(-)
forum_content_post(+)	example_stay(+)	content_stay(-)	course_ovview_stay(-)
forum_content_post_reply (+)	selfasses_visit(+)	forum_content_post(-)	navigation_skip(-)
forum_assignment_post_reply (+)	selfasses_stay(+)	forum_content_post_reply (-)	
selfassess_visit(+)	exercise_visit(+)		
selfassess_stay(-)	exercise_stay(+)		
exercise_visit(+)	selfassess_revision (+)		
exercise_stay(+)	assignment_revision (+)		
	assignment_stay (+)		

a: *CONFIDENCE*

Stankov et al., [37] highlighted the importance of academic confidence by describing it to be the best predictor of students' achievement. Similarly, Shaukat & Bashir [28] conducted a study based on Paul Sander & Sanders [38] in order to measure the students' academic confidence. Their findings indicate the different levels of academic confidence amongst students. In the same way, for the measurement of students' academic confidence, P. Sander & Sanders [27] conducted a study. This research study produced six factors namely *studying, verbalizing, understanding, clarifying, grades,* and *attendance.* Among the six factors mentioned by Sander and Sanders, we considered five factors for the identification of academic confidence. Since *grades* is a factor which does not correlate with the learning behavior of students. Therefore, this factor is exempted.

Regarding studying, visits to the content objects, examples, and outlines can demonstrate the students' behavior with respect to *studying*. Similarly, the total number of exercises' and self-assessment tests' visits, and as a consequence successfully attempting a great number of questions, indicate students' behavior with regard to *understanding*. Regarding students behavior with respect to *verbalizing*, posting frequently over the discussion and peer rating forums regarding the content objects as well as comments on a high number of posts in such forums indicate students' behavior in terms of *verbalizing*. The total number of visits to the queries posted in assignment related forums and also visits to discussion and peer rating forum posts regarding content objects indicate students' behavior with respect to *clarifying.* Counting all discussion and peer rating forum posts regarding content-objects, comments and peer rating of posts, and replies to forum posts posted in the forums relating to queries about assignments indicate students' behavior in terms of attendance.

b: *INDEPENDENCE*

According to Kesten [39], independence or independent learning is a type of learning in which learners, with the help of other relevant learners, can make decisions essential for their own learning requirements. Many recent researchers, for instance, Cukurova [40], Laurillard [41] have accepted and used this definition of independence. Therefore, the definition of Kensten about independence is accepted for this research study. In addition, in order to look into students' abilities to work independently during web-based learning, Singh & EMBI [29] described the significance of five factors namely *planning, organizing, monitoring, evaluating*, and *computer abilities*. Among five factors mentioned by Singh and Embi, we considered four factors for the identification of autonomous abilities.

Computer abilities is a factor which has been exempted on the basis of an assumption that there is similarity in students' abilities in terms of accessing the course materials as well as the related links for the accomplishment of different learning tasks using the LMSs. *Planning and organizing* refers to the student's ability regarding the formulation of techniques and materials, learning objectives, and a timetable for the accomplishment of different learning. Therefore, visits to outlines, content objects, examples, and forum postings as well as the visits to discussion and peer rating forum posts regarding content objects indicate the behavior of students with regard to *planning*. *Monitoring* relates to the student's ability for checking, verifying, and correcting themselves during different learning tasks. Therefore, peer rating of discussion and peer rating forum posts regarding content objects, assignment submissions, even in various attempts indicate the students' behavior in terms of *monitoring*. *Evaluating* refers to the ability of a student

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

regarding judgment, evaluation, and decision making with respect to performance for the achievement of different learning tasks. Therefore, the number of attempts at the exercises as well as the self-assessment tests indicate the behavior of students with regard to *evaluating*.

c: EFFORT

According to Attribution Theory, effort is not a stable factor, although students have a high level of control on it [42]. For instance, by trying hard, students can control their effort or students who are failing repeatedly in a course that is relatively tough, can succeed by registering to an easy one.

TABLE 4
PATTERNS OF BEHAVIOR FOR EACH AFFECTIVE STATE

Affective State	Factors		Patterns for each factor
Confidence	i.	Studying	content_visit outline_visit example_visit
	ii.	Understanding	exercise_visit selfassess_visit
	iii.	Verbalizing	forum_content_post forum_content_post_reply
	iv.	Clarifying	forum_assignment_visit forum_content_visit
	v.	Attendance	forum_content_post forum_content_post_reply forum_assignment_post_repl
Effort			selfassess_visit selfassess_stay exercise_visit exercise_stay forum_content_visit forum_content_post_repl assignment_revision
Independence	i.	Planning & Organizing	content_visit outline_visit example_visit forum_content_visit forum_content_post
	ii.	Monitoring	forum_content_post_repl assignment_revision
	iii.	Evaluating	selfassess_visit selfassess_stay selfassess_revision exercise_visit exercise_stay
Confusion			selfassess_visit exercise_visit example_visit example_stay forum_assignment_post assignment_revision content_stay forum_content_visit forum_content_post_repl

Students' attribution of failure to the factors which are unstable, for example, luck and effort, is a preservation of expectations to the success in future tasks and expedites performance [43]. For instance, attributing failure to the low ability by a student, results in the expectation of failures in future tasks as well. Because, there is no immediate remedy to enhance one's ability. Conversely, if students attribute their failure towards low effort, they can work harder in the upcoming tasks, and can gain success. Jiga-Boy et al., [44] also highlighted the investment of effort with respect to the potential success in a particular task. Similarly, Shanabrook et al., [45] studied the students' level of effort from the log data generated through touch interactions with the tablet.

Following the motivational theory concept Bernard Weiner [42], Pintrich & DeGroot [30], Qu et al., [16], and De Vicente [35] verified rules regarding effort, information regarding student's effort in LMSs is obtained with the help of following behavior. A great number of exercise and self-assessment attempts indicate the students' behavior in terms of exerting high effort. Moreover, a great number of visits to the forums as well as a high number of forum post replies in the discussion and peer rating indicates a high level of effort by student. Submitting the assignments well before deadline and assignment revisions in case of low grades in the initial submission indicates the behavior of student in terms of exerting high effort.

d: CONFUSION

Arguel et al., [46] argued that confusion is an affective state which arises while dealing with complex learning material. This emotional state on the one hand, can be helpful since it can promote further engagement, thus resulting into deep understanding of the topic. On the other hand, if learners are not able to overcome the state of confusion in time, it can badly affect their learning. Such experiences are more concerning within web-based learning environments, where instructor is not physically present to look into the learners' engagement. Due to this lacking, it becomes difficult to offer the different teaching strategies accordingly. Similarly, B. Lehman & Graesser [47] argued that the confusion can have a positive or negative impact. It depends on whether a learner is able to identify the source of confusion and also has the tendency to mitigate it accordingly. The resolution of confusion is highly significant because it can take the learner to the state of frustration Lehman et al., [48].

Following the Arguel et al., [46], B. Lehman & Graesser [47], B Lehman et al., [48], Qu et al., [16] and Baker et al., [36], in the state of confusion, students can be divided into two categories 1) stuck and 2) gamer. Therefore, the information related to the students' confusion in LMSs can be obtained using the following behavior. Solving a low number of exercises and self-assessment tests indicates that the students are stuck. Furthermore, stuck students are those students who do not attempt most of the questions in the self-assessment tests as well as the exercises. A great number of visits to the examples as well as a great spending of time on each example gives an indication regarding stuck students. In terms of assignment submissions, students who post messages repeatedly over the discussion forums for assignment related queries, are considered to be the stuck students. Furthermore, students having a great number of assignment revisions are assumed to the stuck students. They spend much time on content objects and are supposed to have a great number of visits to forums for content objects. Moreover, students forwarding a low number of peer ratings regarding the content objects in discussion and peer rating forum are supposed to be the stuck students. On the other hand, misusing the available system is an attribute of gamer students. They tend to involve in gaming activities during self-assessment tests, for example, gamer students input answers to question quickly and repeatedly until they get a positive feedback. Information regarding students' confusion level can be obtained by using these patterns of stuck and gamer students.

B. FROM BEHAVIOR TO LEARNING STYLES AND AFFECTIVE STATES

The previous section describes relevant patterns for identifying learning style dimensions and affective states, and the classification of data related to these patterns for the differentiation of a high, low, or moderate existence of the respective patterns. The high, low or moderate existence of these patterns indicates a particular learning style preference as well as a particular affective state level. On the basis of this available information, data related to students' preferences and behavior can be used for calculating the hints for particular learning style dimension and also the particular affective state level. This section gives a description about the calculation of learning styles and affective states, beginning from the unrefined data related to students' learning preferences and behavior in the LMS's database.

In the first step, ordered data regarding each pattern is calculated and arranged in a form that it can be used as input for both of the suggested approaches, each for the identification of learning styles and affective states. In the next subsection, an approach related to this is explained. Afterwards, both of the suggested approaches including the literature-based approach with the use of rule-base mechanism and the data-driven approach with the use of Bayesian networks, each for the identification of learning styles and affective states are introduced.

1) METHOD FOR DATA PRE-PROCESSING

Data pre-processing involves data collection, data cleaning, conversion of continuous data into a suitable shape, attribute selection, data integration, etc. The purpose of data pre-processing is to pre-process usage information of students which is gathered in the duration of their course usage and stored in the LMS's database as well as the log files.

In the first step, data about students' learning preferences and behavior, as mentioned in Section III part 1 (Features and Pattern Selection), is required to be extracted from the LMSs' database. Afterwards, on the basis of available information, mapping of data about students' learning preferences and behavior is performed on a 4-item scale. To be more specific, let O be the ordered data matrix representing rows the students and columns the available patterns. For the classification of each student's preferences and behavior regarding each pattern, the value in a range of 0 to 3 is assigned, where 1, 2, and 3 represent a low, moderate and high occurrence respectively. The unavailability of information related to the respective pattern is specified by a zero (0). This data mapping is based on the threshold values described in Section III part 2.

Afterwards, the matrix LS_{dim} was built regarding each learning style dimension and the matrix AS_{fac} was built regarding each affective state. The LS_{dim} and AS_{fac} include the ordered data from matrix O. Each LS_{dim} matrix representing students in rows and the relevant patterns related to the respective learning style dimension *dim* in columns, as mentioned in Section III part 3. Similarly, each AS_{fac} matrix representing students in rows and the relevant patterns for each affective state *fac* in columns, as mentioned in Section III part 4. In order to calculate the respective learning style dimension and also the respective affective state, the four matrices, each for the learning styles and the affective states, are used as input data.

2) MINING STUDENTS' LEARNING STYLES AND AFFECTIVE STATES USING BAYESIAN NETWORKS (DATA DRIVEN APPROACH)

Data mining is a process that uses one or more machine learning techniques such as Bayesian Networks, Neural Networks etc., for automatic analysis and knowledge extraction from the data stored in a database. Data mining techniques explore fundamental trends, patterns, and relationships concealed in the available data. (Chang & Chen [49] described that unexpected relationship in a dataset can be explored using data mining techniques, and declared it to be one its advantages. Moreover, Lee et al., [50] described that the data mining techniques can be classified into the two main groups namely the supervised and the unsupervised learning. Supervised learning deals with the assignment of objects to predefined classes or

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

categories. On the other hand, the unsupervised learning refers to dividing the data into groups having similar objects. Song et al., [51] described that for finding applicability of a supervised or an unsupervised learning technique, consideration of the known classes or categories is necessary. This indicates that if classes or categories are unknown, unsupervised learning approach should be used and vice versa.

Since, this study aims at investigating the relationships between the learning styles and learning preferences of students as well as the affective states and learning behavior of students, so, there are known categories i.e. learning styles and affective states. Therefore, for this study, a supervised learning approach was adopted.

The most extensively used supervised learning technique for extracting useful information from a dataset is classification [52]. Regarding the classification, one of the most widely used model is the Bayesian Networks (BNs), which is a graphical model for probability relationships among a set of variable (features). Kotsiantis [53] described that BNs' most fascinating feature as opposed to other machine learning techniques such as neural networks and decision trees etc., is most probably the feasibility of considering previous information about the problem given, with respect to the structural connections amongst its features. Therefore, due to its effectives, several researchers have exploited BNs, particularly to classify students' learning preferences and behavior.

Garcia et al., [32], for instance, used BNs for identifying learning styles on the basis of learning styles model defined by (Felder & Silverman [26]. In this regard, a study was conducted comprising 27 computer science students taking a web-based course on artificial intelligence. Afterwards, the results of their approach using BNs were compared with ILS Questionnaire, demonstrating the effectiveness of BNs in the prediction of learning styles of students with higher precision. In a similar manner, Graf [54] adopted a data-driven approach for the identification of learning styles using BNs and also a literature based approach for finding out the effectiveness of suggested concept with the use of one or the other approach for the identification of learning styles. Furthermore, Danine [55] designed and developed an intelligent tutoring system that exploits BNs. The BNs were employed for the evaluation of students' knowledge, for the perception of their strategies, for predicting their actions while solving a problem and for making a highly reliable diagnosis.

BNs have been used extensively also to develop detectors of students' affective states. Qu et al., [16] presented a model for pedagogical agents to use the attention of learners for the detection of learners' motivation factors in the interactive learning environments. In this regard, BN was employed that combine evidence from the learner's interface actions and eye gaze, and consequently infers the focus of attention of a learner. After inferring the focus of attention, it was combined with the information's inferred by the plan recognizer, about the learner's activities. The purpose of this approach was to identify learners' confidence level, effort, and confusion. Effectiveness of the model was confirmed by conducting an experimental study with 24 students at the University of California (UC). The accuracy of their model were reported to be 82% for confidence, 76.3% for effort, and 76.8% for confusion. Furthermore, for the prediction of students' affective states such as thinking, recalling, tiredness, and satisfaction during a classroom lecture, Abbasi et al., [56] explored the possibility of employing unintended hand gestures of students. They employed a BN model for the prediction of one, out of the four affective states observed in video recordings of a classroom lecture using five hand gestures. They reported that the generalization accuracy of model is 100% over the cases in which student reported an affective state and 79.4% over the cases in which student reported no affective state.

The mentioned studies indicate that BNs is a suitable approach and has the potential for inferring learning styles and affective states. In this section, therefore BNs was employed to analyze the results of an empirical study for the investigation of relationships between the learning styles and learning preferences of students as well as between students' affective states and their behavior in LMSs. An introduction of BNs is presented very briefly in the next subsection. Afterwards, BNs application for this study is presented in a subsection.

3) BAYESIAN NETWORKS AN INTRODUCTION

Bayesian networks belong to the family of probabilistic graphical models, and describe the relationships between causes and effects. They consist of nodes and arcs. Each node in the graph represents a random variable, and arcs between the nodes represent probabilistic dependencies among the corresponding random variables.

Moreover, BNs correspond to another graphical model structure known as directed acyclic graph (DAG). DAG consists of two sets. The two sets include the set of nodes and the set of directed edges. The nodes describe random variables and are represented as circles labelled by the variable names. The edges describe direct dependence among the variables and are represented by arrows between nodes. For example, an edge from node A to node B represents a statistical dependence between the corresponding variables and the arrow indicates that a value taken by the variable B depends on the value taken by the variable A or in other words the variable A influences the variable B. Node A is then referred to as parent of B and, similarly, B is referred to as the child of A. A simple BN regarding learning styles as well as affective states is shown in figure 1 and 2 respectively.

BN shown in figure 1 demonstrates the active/reflective dimension of the FSLSM. The patterns used for identifying the active/reflective learning style preference referred to as parent nodes and the preference for an active/reflective learning style referred to as the child node, influenced by all parent nodes.

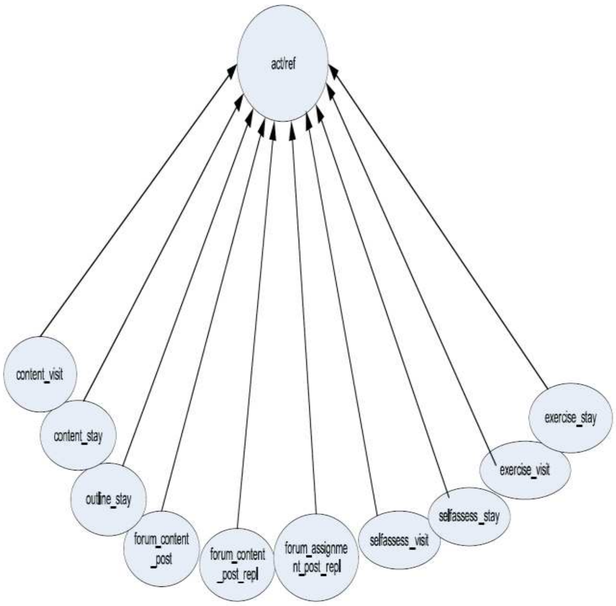

Figure 1. **BN for Active/Reflective dimension**

Similarly, BN shown in figure 2 demonstrates the affective state such as confusion. The patterns used for identifying the affective state such as confusion, act as parent nodes and the behavior for affective state such as confusion is the child node, influenced by all parent nodes. The BN shown in figure 1 and 2 consists only of converging connections; however BNs can also include diverging, serial or a mix of these types of connections.

Each node in a BN is associated with a conditional probability table (CPT), which quantifies the relationship between each respective node and its parents in the network (parameter learning.) For example, the CPT for variable B specifies the conditional probability distribution of B given its parents. The CPT is described mathematically by the notation (B/parent (B)) and specifies the probability of each possible state of the node given each possible combination of states of its parents. If a node does not have any parent then probabilities are not conditioned and the table consist of prior probabilities.

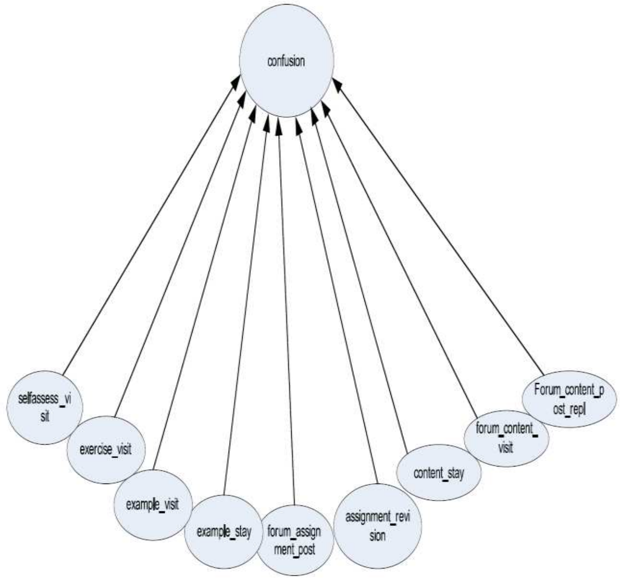

Figure 2. **BN for affective states such as confusion**

The task of learning a BN consists of two sub tasks such as the learning of the DAG structure of the network, and then the determination of its parameters. Probabilistic parameters in the form of local conditional distribution of a variable given its parents are encoded into a set of tables, one for each variable. Given the structure of the network and probabilistic parameters, several BN inference algorithms such as likelihood sampling, backward sampling etc, can be used for drawing inferences.

4) BAYESIAN NETWORKS APPLICATION FOR THE CALCULATION OF LEARNING STYLES AND AFFECTIVE STATES

Darwiche [57] described three main methods for generating BN. The first method deals with reflecting one's own knowledge or the knowledge of others (typically, perceptions about causal influences) and then capturing them into a BN. On the other hand, the next method is based on automatically synthesising the BN from some other type of formal knowledge. The last method deals with learning BN from the data. This method is usually applied for generating BNs, since acquiring data is cheap and also amount of available information growing rapidly. On the other hand, knowledge acquisition is an expensive process and often we don't have an expert. Friedman [58] described that learning BNs from data is a field of research which is growing rapidly and currently, a significant deal of activity can be seen in this field. Therefore, for this study the method that deals with learning BN from the data was adopted. The structure of

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

the BN regarding learning styles and affective states was deduced from literature, giving an indication about relevant patterns for the learning style dimensions and also the affective states. Based on the corresponding patterns, BNs were manually constructed for each learning style dimension and also each affective state.

In the BN, parameter learning was employed for determining the conditional probability distribution. For input data regarding parameter learning, the matrix LS_{dim} related to each learning styles dimension was extended by further inclusion of one column representing scale data with regard to learning style preference for each student. The extended matrix is known as D_{dim}. Similarly, AS_{fac}, the corresponding matrix of each affective state was also extended by additional inclusion of one column representing the scale data about learning behavior for each student and the extended matrix is known as D_{fac}. The scale data approach was adopted from the ILS questionnaire regarding learning styles and affective states. For example, the strength of the learning style preferences in ILS questionnaire is indicated by values ranging from -11 to +11 in steps of 2. Since the aim of this student modelling approach is the identification of learning styles and affective states on a three-item scale, for instance, regarding the learning styles it distinguishes between an active, reflective, and balanced learning style. Similarly, regarding the affective states such as confidence; it distinguishes between low, balanced, and high level of confidence. Therefore, the values of ILS were scaled in a range of -1 and +1. Firstly, -1 specifies a value of ILS which is either smaller than -5 or is equal to -5 and therefore, indicates a preference for reflective, global, verbal, and intuitive learning style, depending on the investigated learning style dimension. Secondly, the value 0 specifies a value of ILS ranging from -3 to +3 and consequently, a preference for balanced learning style. Lastly, +1 specifies a value of ILS which is either greater than or equal to 5 and therefore gives an indication towards a preference for an active, sequential, visual, and sensing learning style.

For parameter learning and evaluation of BN, both D_{dim} and D_{fac} were divided at random into four training and testing sets with partitions of 80, 70, 60, and 50 percent respectively. The training datasets were used for parameter learning or in other words, training aims at making predictions for future cases in which only the inputs to the network are known, and testing data set were used for evaluating the BN. This data division is necessary, since the purpose of training the network with respect to parameter learning, is to construct a network that can classify available data in a best possible way with regard to the reference values. For constructing the BN and learning the parameters, Waikato Environment for Knowledge Analysis abbreviated as WEKA [59], was used. Moreover, a simple and very fast learning algorithm, called K2 [60] was used for learning BNs.

5) A LITERATURE-BASED APPROACH FOR INFERRING LEARNING STYLES AND AFFECTIVE STATES USING A SIMPLE RULE-BASED METHOD

An approach presented in the previous subsection, uses data for building a model to infer the learning styles and affective states. This section presents an approach for inferring learning styles and the affective states, which is completely derived from literature. This approach is developed on a concept, where each relevant pattern introduced in Section III part 3 regarding learning style dimensions, provides a hint regarding the learning styles of students. On the basis of this information, and the information regarding whether a low or high occurrence of particular learning preference is supporting the respective learning style, the total number of matching hints can be computed, provided the students' preferences. Similarly, each relevant pattern introduced in Section III part 4 regarding affective states, provides a hint about student's affective state. This information, and the information regarding whether a low or high occurrence of particular learning behavior is indicating a respective affective state, the total number of matching hints can be computed, provided the behavior of students.

The approach for calculation of hints with respect to learning style preferences and also the levels of respective affective state is founded on an approach developed by Graf et al., [12] for calculating learning styles from the behavioral patterns. Therefore, four values are used for denoting hints, that is, 0 to 3, where 3 indicates that the student's preference gives a strong indication towards the respective learning style dimension, and similarly, if 3 is used for affective state level, it gives a strong indication as well for that affective state. The hint value 2 gives an indication that the student's preference and behavior is either average or balanced, and do not give a specific hint, towards the respective learning style dimension and affective state. Similarly, hint value 1 demonstrates that the student's preference and behavior is in dissent with the corresponding learning style dimension and affective state. Finally, the hint value 0 demonstrate that no information are available. In order to classify the preferences and behavior of students for each pattern, into four values, thresholds described in Section III part 2 are used as a basis.

$$Precision = \frac{\sum_{i=0}^{n} x_i}{n} \qquad where\ 0 \le x_i \le 3 \tag{1}$$

IEEE Access Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

A measure for the each learning style dimension and also the affective states are computed individually by summing up all the hint values initially, and then dividing it by the total number of the patterns giving available information. An equivalent mathematical notation is shown in formula (1), where the hint value giving available information for each pattern is denoted by x. The x can have a value ranging from 0 to 3, the total number of patterns giving available information is denoted by n, and a particular pattern number is denoted by i. For both the learning styles and affective states, this measure is normalized on a scale of 0 to 1, where 1 indicates the strong positive preference towards a respective learning style and affective state. On the other hand, 0 indicates a strong negative preference for a particular learning style and affective state. However, no conclusion can be drawn, in case of unavailability of information for all patterns regarding a learning style dimension and affective state.

IV. EVALUATION OF TWO APPROACHES

The two preceding sections described the conceptual framework and approaches for the identification of learning styles and affective states from students' learning preferences and behavior. This section presents the evaluation of suggested concept regarding automatic student modelling, using either a literature based or a data driven approach to infer the learning styles and affective states.

The evaluation of proposed approach regarding automatic student modelling was done by tracking students' learning preferences and behavior in a course about "Human Computer Interaction", which was offered through web based LMS such as Moodle. Moreover, the students were instructed to fill out online the ILS, ACS, and independence survey questionnaire. ILS questionnaire for the identification of students' learning styles, whereas ACS and independence survey questionnaire for the identification of students' affective states such as academic confidence and independence respectively. In this study, eighty one students participated.

For tracking the data regarding students' learning preferences and behavior, most of the patterns based on certain features are either fully or partially supported by the LMSs. For example, the pattern selfassess_stay is fully supported by Moodle. Regarding this Moodle offers the fields starting time and closing time of self-assessment tests. However, to calculate the stay at each self-assessment test some kind of calculations is still needed. These kinds of calculations were implemented in a tool (AsLim) mentioned in the Section V. On the other hand, some patterns such as content_stay are partially supported by the Moodle. Regarding this Moodle offers a field that describe only the time of visit of that particular learning object and does not contain a field that describe the closing time of that object. Therefore, to calculate the stay at such kind of learning objects, some kind of calculation is needed. These kinds of calculations were also implemented in AsLim.

The next subsection introduces the "Human Computer Interaction" course and its structure. The thresholds mentioned in Section III part 2 were maintained for the classification of data related to the occurrence of learning preferences and behavior, due to the characteristics that matches with the "Human Computer Interaction" course mentioned in Section IV part A. Afterwards, the evaluation method as well and the results of automatic student modelling with the use of a literature based and data driven approach are reported and discussed.

A. DESCRIPTION OF THE INVESTIGATED COURSE

For the purpose of getting data regarding the students' learning preferences and behavior, a course of "Human Computer Interaction" is used as a basis for this study, which was delivered in the summer semester 2018, for about 10 weeks at COMSATS University Islamabad, Attock Campus. The course included lectures and assignments. The contents of the lecture were applied in assignments to support practical understanding of the material. The entire course was uploaded and managed through Moodle. The main purpose of using a LMS was to facilitate learning by presenting additional material and opportunities to students for learning. Moreover, the course also included exams. Exams were conducted into two parts, midterm exam and final exam. Midterm exam was scheduled after 4 weeks and final exam after 9 weeks. For midterm exam the course included six chapters and final exam included 8 chapters.

This study was conducted during the first four weeks i.e. before the midterm exam. The study included the course chapters such as usability engineering, usability benchmarking, goal oriented interaction design, prototyping, usability inspection methods, and usability testing methods. Moreover, the study also included the course features such as outlines, examples, and exercises. Overall, the chapters included 147 content objects. For all of the chapters, the outline, examples, exercises, and the self-assessment tests were made available. Overall, there were 100 questions in the self-assessment tests. The exercises included overall 65 questions, and were based on the concept that students can solve online exercises related to each chapter and consequently get feedback regarding their performance. Moreover, discussion/peer rating forum was provided for the course, where students' can post, reply, and view discussions on a specific topic for discussion. For examining the students' knowledge, 3 marked assignments were included. Two of the assignments included questions which were to be answered online and one of the assignment answers were to be uploaded. Finally, the course formatting was done in such a way that learning everything was required by all the students and they were tested regarding each chapter content objects by a self-assessment test, exercise, and discussion/peer rating forum; hence the course had great appropriateness for investigating the individual learning.

B. EVALUATION METHOD

This section describes the evaluation of suggested concept regarding automatic student modelling, against the data-driven approach, in order to determine the suggested concept usefulness. The data gathered from the "Human Computer Interaction" course is used as a basis for this evaluation. For verifying the predicted learning styles and also affective states of the two approaches, the students were instructed to fill out online the ILS questionnaire for the identification of learning styles, and the ACS and independence survey questionnaire for identification of students' affective states such as academic confidence and independence respectively.

The number of students initially willing to participate in the study were 125. For this study, a criteria was defined that contains three requirements, and data meeting those requirements was used as input data. Firstly, data from the students was not considered who either did not submit the concerned questionnaires online or the submitted questionnaires were incomplete. Secondly, data from only those students was considered, who secured out of the assignments total allotted points, at least 50% points, , which was required for a positive outcome. This was opted to exclude the data of drop out students. Thirdly, data from only those students was included who registered as well as appeared in the midterm exam, which was also a requirement for positive outcome. This requirement has great significance as it assures that students' preparation for the terminal exam is incorporated in the data. Finally, for this study, data related to the 81 students was used.

Subsequently, the two subsections describe in detail the evaluation of concept for automatic student modelling based on either of the two approaches namely the data-driven and literature-based approach.

C. EVALUATION METHOD BASED ON THE DATA DRIVEN APPROACH

The use of data in certain proportion for training a model is the key idea of a data-driven approach. For this purpose the available learning preferences and behavior data was thus divided into the training data and test data respectively. For training the BN, training data-set was used whereas for evaluating the effectiveness of resulting BN for the identification of learning styles and affective states based on students' learning preferences and behavior respectively, testing data-set was used. Then a comparison was made between the learning styles predicted by the BN and the information regarding the learning styles identified by the ILS questionnaire. Similarly, the affective states such as confidence and independence, predicted by BN were then compared with the information about the affective states identified by the ACS and independence survey questionnaire respectively. Since the approach was developed for the identification of learning styles and the affective states by differentiating between three values, for example an active, reflective and balanced learning style and also low, balanced, and high affective state level. Therefore, both the learning styles based on either the values of ILS or predicted by the BN are on a three-item scale. Similarly, both the affective states reported either by the ACS and independence survey questionnaire or predicted by the BN are on a 3-item scale.

For measuring the closeness of predicted learning style and learning style based on values of ILS, and also the closeness of predicted affective state and affective state based on either the values of ACS or independence survey questionnaire, the subsequent measure suggested by Garcia et al., [32], and also reported by Graf [54] and Atman et al., [61] regarding the learning styles, was used.

$$Precision = \frac{\sum_{i=1}^{n} Sim(LS_{predicted}, LS_{ILS})}{\text{n}} \,.\, 100 \qquad (2)$$

Where $LS_{predicted}$ indicate the learning style predicted by the BN, LS_{ILS} relates to the learning style identified by the ILS questionnaire, mapped to a three-item scale, and n denotes the total number of students. The comparison between the two parameters i.e., the learning style predicted by the BN ($LS_{predicted}$) and the learning style identified by the ILS questionnaire (LS_{ILS}) is made using the function *Sim*. If both the parameters are equal, the function returns 1, whereas if both parameters are opposite, the function returns 0. Furthermore, if the function value is 0.5, this demonstrate that a balanced learning style has been reported by one parameter, and a preference towards one of two poles of the learning style dimension is represented other parameter.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

Similarly, for measuring the closeness of predicted affective state and the affective state based on either ACS or independence survey questionnaire values, the measure used for learning styles was adopted.

$$Precision = \frac{\sum_{i=1}^{n} Sim(AS_{predicted}, AS_{Quest})}{n} \cdot 100 \qquad (3)$$

Where $AS_{predicted}$ indicate the affective state predicted by the BN, AS_{Quest} refers to the affective state identified by the concerned questionnaire, mapped to a three-item scale, and n denotes the total number of students. The comparison between the two parameters i.e., the affective state predicted by the BN $(AS_{predicted})$ and affective state identified by the concerned questionnaire (AS_{Quest}) is made using the function *Sim*. If both the parameters i.e. $AS_{predicted}$ and AS_{Quest} are equal, the function returns 1. If one parameter depicts a balanced affective state and the other parameter a behavior towards one of the two extreme levels of affective state, then the function returns 0.5. If both the parameters are opposite, the function returns 0.

For the purpose of getting more reliable results, 4 runs were performed for each BN regarding the learning styles and affective states. Each run includes the training and the testing set with the partitions of 80 percent, 70 percent, 60 percent, and 50 percent respectively. BN was built for the identification of each learning style dimension mentioned by FSLSM. Similarly, BN was also constructed for the identification of affective states such as confidence, and independence. For the respective BN, the average calculated precision of the 4 runs was used as a result.

D. EVALUATION METHOD BASED ON THE LITERATURE BASED APPROACH

The model for the calculation of learning styles and affective states in a data driven approach is trained with sample data. On the other hand, model which is literature based, is developed without using the sample data. Thus, to verify literature based model, the entire dataset can be used.

In order to make literature based model comparable to data-driven approach regarding each the learning styles and the affective states, the same measure was used. For example, the values of ILS questionnaire were scaled again to the values ranging from -1 to +1. In order to scale the results of literature-based approach between 0 and 1 as described in Section 3.4, threshold values of 0.25 and 0.75 were used. [54] is used as a basis for these thresholds, describing that the use of first and last quarter for the indication of learning style preferences for one or the other extreme of particular dimension and the use of second and third quarter for the demonstration of a balanced learning style produces better as well as reliable results than the division of range into 3 parts. Therefore, regarding learning styles, on the basis of scaled results of the literature based approach $(LS_{predicted})$ and scaled ILS values, the formula 2 was used and the results were used as a measure for literature-based approach. Similarly, regarding the affective states, on the basis of scaled results of literature-based approach $(AS_{predicted})$ and scaled AS_{Quest} values, the formula 3 was applied. AS_{Quest} describe either the ACS or independence survey questionnaire.

E. EVALUATION RESULTS

Table 5 and 6 reports the results of data driven approach using BNs for inferring learning styles and affective states respectively. The table 5 presents the four training and the testing set with various partitions, and their corresponding results. Furthermore, it describe the average results regarding each learning style dimension. The average results give an indication towards a moderate precision, ranging from values between 63.56 and 71.21.

TABLE 5
RESULTS OF THE DATA-DRIVEN APPROACH REGARDING LEARNING STYLES

Training Data (in %)	**Testing Data (in %)**	**Act/Ref (in %)**	**Sen/Int (in %)**	**Vis/Ver (in %)**	**Seq/Glo (in %)**
80	20	72.76	85.17	77.21	82.00
70	30	65.24	73.06	71.52	83.10
60	40	65.85	58.22	62.34	67.31
50	50	58.01	43.38	43.17	52.42
Average:		65.47	64.96	63.56	71.21

Similarly, table 6 also presents the four training and the testing set with various partitions, and the corresponding results of the affective states such as confidence and independence as well as the average results for the corresponding affective states. The average results indicate a moderate precision, ranging from values between 67.86 and 73.92.

TABLE 6
RESULTS OF THE DATA-DRIVEN APPROACH REGARDING AFFECTIVE STATES

Training Data (in %)	Testing Data (in %)	Confidence (in %)	Independence (in %)
80	20	79.20	83.11
70	30	71.41	77.32
60	40	67.07	75.27
50	50	53.78	60.01
Average:		67.86	73.92

Table 6 presents results regarding the affective states such as confidence and independence, while the results regarding the affective states such as confusion and effort

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

cannot be reported according to the formula 3 due to the unavailability of the concerned questionnaires. However, from the reported results tendency regarding the learning styles and affective states, it can be concluded that the average results regarding the affective states such as confusion and effort may also demonstrate a moderate precision.

TABLE 7
COMPARISON OF THE RESULTS REGARDING LEARNING STYLES OBTAINED BY USING THE TWO APPROACHES

	Act/Ref (in %)	Sen/Int (in %)	Vis/Ver (in %)	Seq/Glo (in %)
Data driven approach	65.47	64.96	63.56	71.21
Literature based approach	81.25	78.17	76.31	74.87

literature-based and data-driven approach regarding the learning styles is presented in Table 7. This indicates clearly that literature-based approach yields better results than data-driven approach regarding each learning style dimension. The results reported by the literature based approach indicate high precision, ranging from values between 74.87 and 81.25, and therefore can be declared as good results. A comparison of average results, reported by the data-driven approach and the literature-based approach regarding the affective states is presented in Table 8. The comparison indicates clearly that as compared to data-driven approach, literature-based approach produces good results regarding affective states such as confidence and independence. The results reported by the literature based approach indicate high precision, ranging from values between 79.05 and 82.31, and therefore can be declared as good results.

TABLE 8
COMPARISON OF THE RESULTS REGARDING AFFECTIVE STATES OBTAINED BY USING THE TWO APPROACHES

	Confidence (in %)	Independence (in %)
Data driven approach	67.86	73.92
Literature based approach	82.31	79.05

Table 8 presents results regarding the affective states such as confidence and independence, here again the results of the literature based approach regarding the affective states such as confusion and effort cannot be reported due to the unavailability of the concerned questionnaires. However, from the reported results tendency of literature based approach as compared to the data driven approach, it can be concluded that the literature based approach regarding the affective states such as confusion and effort may also demonstrate a high precision.

V. ASLIM – A TOOL FOR AFFECTIVE STATES AND LEARNING STYLES IDENTIFICATION AND MEASUREMENT

AsLim stands for "Affective States and Learning Styles Identification and Measurement" and is developed on the basis of an approach for automatic identification of affective states and learning styles. The AsLim extends this approach by allowing teachers to specify the required information regarding learning objects more easily, and thus enabling them to automatically detect the affective states and the learning styles of their students. It is a stand-alone tool, and based on features commonly available in most of the LMSs. For the identification of affective states and learning styles, it uses an implicit modelling mechanism by analyzing the interaction of students with web-based LMSs in the form of learning behavior and preferences. On the one hand, these identified affective states behavior and learning styles preferences provide teachers with more information about their students, and on the other hand, can act as a basis for enhancing adaptivity in terms of affective states and learning styles. AsLim implements literature-based approach for the purpose of inferring affective states and learning styles, since this approach produces better results than the data-driven approach. The next section describes the architecture of AsLim tool in more detail. Afterwards, an introduction to the tool is presented by displaying the necessary and the possible interaction of user for the ease of teachers' use.

A. ARCHITECTURE OF ASLIM

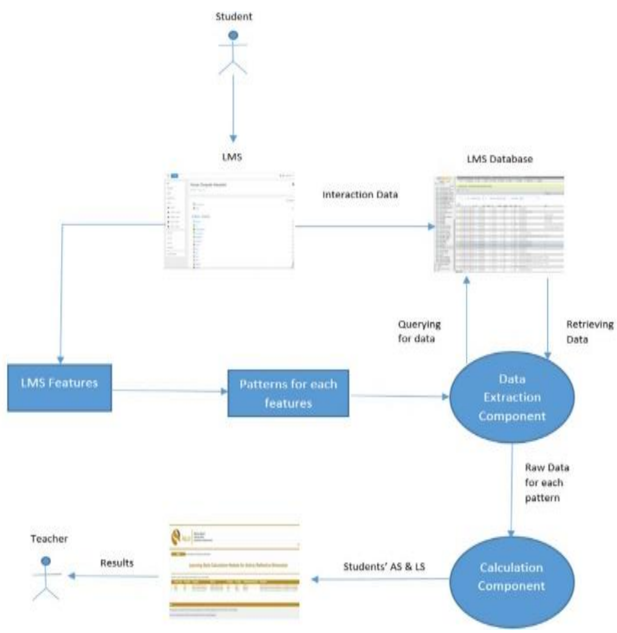

Figure 3. Architecture of AsLim tool

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

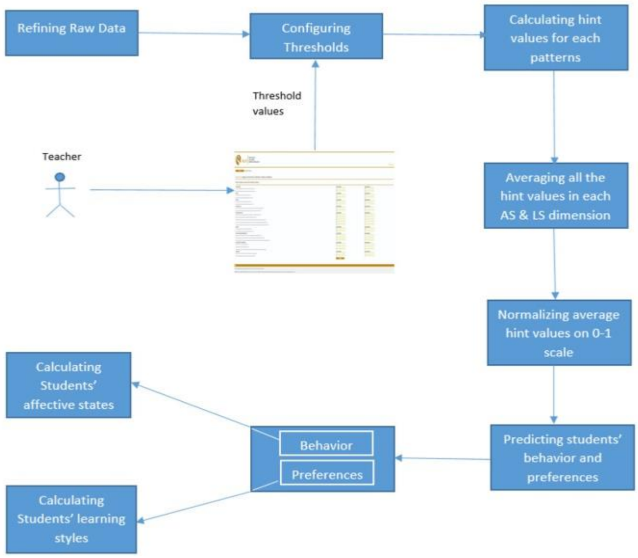

Figure 4. Calculation Process

Figure 3 illustrates the tool architecture. It comprises of two components namely the data extraction component and the calculation component. The data extraction component based on the information provided by a teacher or course designer extracts the patterns raw data regarding the learning preferences behavior and behavior from the database. The data extracted by the data extraction component is then passed on to the calculation component for further processing.

The calculation component, initially refines the raw data associated with each pattern and then calculates the Affective states (AS) and Learning styles (LS) of each student. The results obtained are then visible to a teacher. The figure 4 gives a detailed overview of the calculation process (already described in section III part D) followed by the AsLim.

B. ASLIM AND USER INTERACTION

This section presents a brief description of AsLim's features in relation to user interaction. After logging-in, a course selection page is displayed to the user, a page where teacher can select a single or multiple courses in which he/she wants to identify students' learning styles and affective states. The next step deals with configuration of thresholds for the expected stay at learning objects.

In this step, the teacher determines how much time a student should spent on content objects, outlines, course overviews and examples as shown in figure 6. Since the expected stay for self-assessment tests and exercises is required at the time of its creation of an exercise and self-assessment test and this time is stored in the Moodle database, therefore, AsLim directly retrieves this time estimate from Moodle database rather than getting the time as input from the teacher.

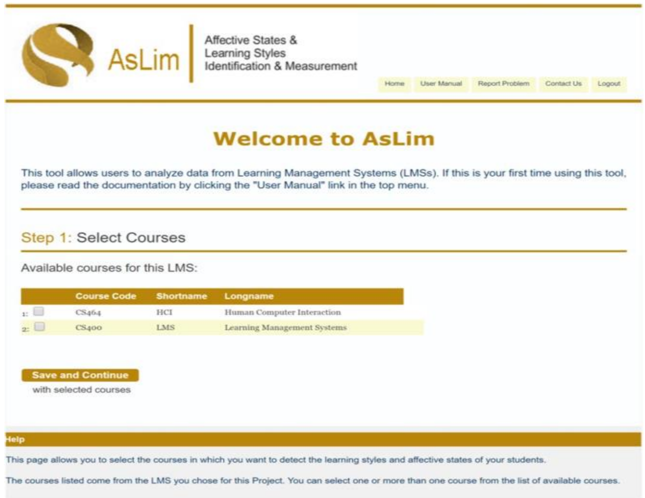

Figure 5. Course Selection Page

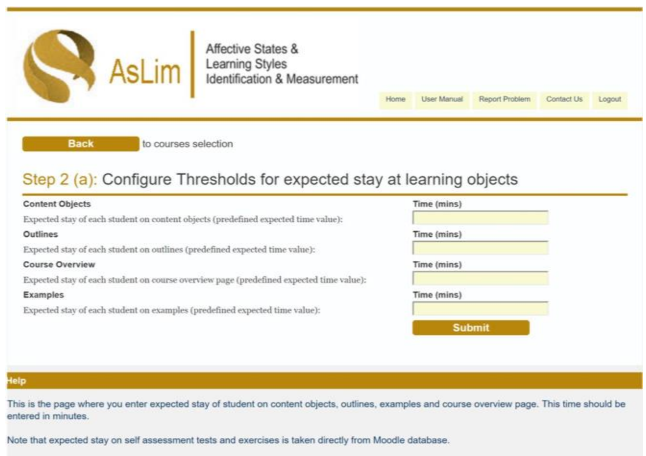

Figure 6. Configuration of thresholds for expected stay at learning objects

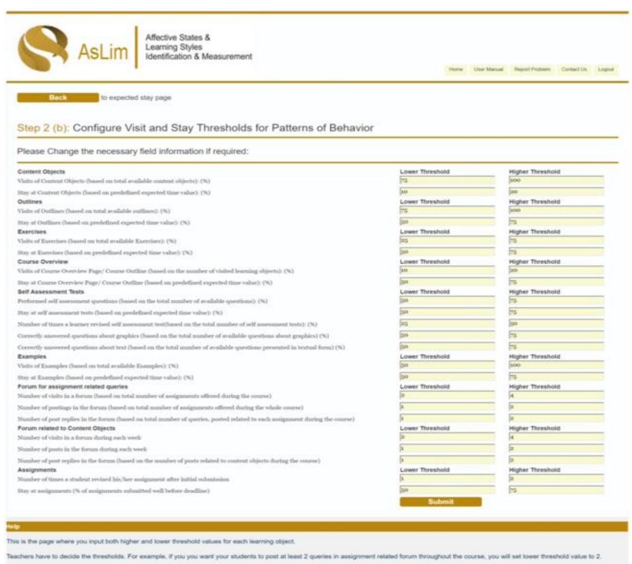

Figure 7. Configuration of visit and stay thresholds

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

Following this step, the teacher is required to configure threshold values for visits and stay (mentioned in section III part 2) for different learning objects as illustrated in figure 7. After this, the teacher selects one of the two actions which are "calculate learning styles" and "calculate affective states". In the final step, results are displayed to the teacher about the learning styles and affective states of his/her students.

VI. DISCUSSION & CONCLUSION

In this paper, we presented automatic approach of student modelling, for the identification of learning styles based on FSLSM, and the affective states based on the literature reported by (P. Sander & Sanders [27], Singh & EMBI [29], Pintrich & DeGroot [30], and Qu et al., [16]. In this regard, a simple rule based mechanism is applied, assuming that preferences and behavioral patterns can provide relevant hints for the identification of students' learning styles and affective states. The basic idea of this approach is very similar to that of the questionnaire approaches, except that, this approach exploits preferences and behavioral data implicitly, instead of asking students about their preferences and behavior explicitly. Furthermore, our approach is applicable to most of the LMSs due to its generic nature.

The proposed approach regarding learning styles was evaluated by comparing the results with the data driven approach using BN, and also with the two recent studies reported by [6] and [62]. They reported consecutively, the results of 70 % and 65% for the active/reflective dimension. Similarly, for the sensing/intuitive dimension results with an accuracy of 66% and 75% were reported. Regarding the visual/verbal dimension, marginally higher results (75% and 76.25%) were reported. For sequential/global dimension results of 80% and 77.5% were achieved. Comparison of results reported by the two studies with those of our results, demonstrate that our approach yields high precise results for the active/reflective, sensing/intuitive, and visual/verbal dimensions, whereas, for the sequential/global dimension our results are slightly lower.

With regards to the results of affective states' identification through behavioral patterns, there is no such study reported in literature (to the best of our knowledge) which demonstrates the results. However, by comparing the results of two approaches i.e. data-driven and literature-based that we have exploited, it can be seen that the literature-based approach produces good results for affective states such as confidence and independence. Regarding confidence and independence, literature-based approach achieved the results of 82.31% and 79.05% respectively, whereas the results of data-driven approach were significantly lower which accounted for 67.86% and 73.92% for confidence and independence respectively. Owing to the unavailability of concerned questionnaires, the results of literature-based approach regarding confusion and effort cannot be reported. However, from the reported results tendency of literature based approach as compared to the data driven approach, it can be concluded that the literature based approach regarding the affective states namely the confusion and effort may also demonstrate a high precision.

The evaluation of the suggested approach showed promising results, demonstrating that the approach is appropriate for the identification of learning styles and affective states. Therefore, the proposed approach is implemented by developing a learning styles and affective states identification and measurement tool (AsLim), thus making the teachers capable to know about their students preferences and behavior.

There are some limitations to generalizability of the results of our studies. Firstly, participants of the studies were limited to only COMSATS University Islamabad, Attock Campus, studying a course "Human Computer Interaction (HCI)". Therefore it might be interesting to confirm our results with university students from other countries and cultures. Secondly, for the convenience of collecting data, results of the studies reported are based on HCI course, so it might also be interesting to confirm our results in a course other than HCI. Thirdly, all the participants of our studies were from the university due to the fact that universities are major target groups of using LMSs, so it would be interesting to confirm our results with non-university students.

In the following paragraphs, possible future researches that can be carried out based on the presented research, are discussed in more detail.

Regarding the identification of learning styles and affective states, this research proposed a static student modelling approach. This approach first gathers students' preferences and behavioral data over a period of time and then used that data to calculate learning styles and affective states. On the other hand, dynamic student modelling approach processes the students' preferences and behavioral data immediately, consequently updating the student model. This frequent update identifies transition in learning styles as well as in the affective states, if exist. Specifically, the affective states are not static but it changes over time as a result of the changing learning environment and the particular interpretation of the situation by each student. Therefore, research reported can be seen as the basis for the development of an accurate real-time assessment of student learning styles and affective states. AsLim, a tool for the

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

identification of learning styles and affective states can be extended by introducing a dynamic student modelling approach as well.

Future studies can focus, more specifically with incorporating the information about the relationship between learning styles and affective states into the identification process of learning styles and affective states. Moreover, future studies can also focus on the identification of the learning style dimensions and the affective states that contribute more towards students' learning gains.

Another direction of the future research is to provide adaptivity in LMSs, based on the identified learning styles and affective states. This will involve different learning style and affective state strategies, and the tailor instructional design in order to improve students' learning performance. Appropriate and feasible, learning styles and affective states based strategies need to be developed for different target groups. Moreover, facilitating teachers to define which types of learning objects they want to include in the adaptation process as well as defining respective adaptation features. This will enable teachers to adjust the adaptation mechanism to their courses rather than the other way around.
Moreover, to get aware and knowing students about their own learning styles and affective states, and helping them to understand their weaknesses and strengths in a better way during the process of learning, the future work is also aimed at developing an approach of open learner modeling.

ACKNOWLEDGMENT

The authors would like to thank COMSATS University Islamabad, Attock Campus for supporting this research.

REFERENCES

[1] H. M. Truong, "Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities," *Computers in Human Behavior,* vol. 55, pp. 1185-1193, 2016.

[2] J. Feldman, A. Monteserin, and A. Amandi, "Automatic detection of learning styles: state of the art," *Artificial Intelligence Review,* 2014.

[3] J. Kay and B. Kummerfeld, "Lifelong learner modeling," in *Adaptive Technologies for Training and Education*: Cambridge University Press, 2012, pp. 140-167.

[4] M. P. P. Liyanage, K. S. L. Gunawardena, and M. Hirakawa, "Detecting Learning Styles in Learning Management Systems Using Data Mining," *Journal of Information Processing,* vol. 24, pp. 740–749, 2016.

[5] V. Shute and D. Zapata-Rivera, *Adaptive Educational Systems* (Adaptive Technologies for Training and Education). 2012.

[6] I. Karagiannis and M. Satratzemi, "An adaptive mechanism for Moodle based on automatic detection of learning styles," *Education and Information Technologies,* vol. EAIT 23(3), pp. 1331-1357, 2018.

[7] E. Popescu, Badica, C., and Moraret, L., "Welsa: An intelligent and adaptive web-based educational system," *In Intelligent Distributed Computing (IDC),* pp. 175-185, 2009.

[8] D. Kelly and B. Tangney, "Adapting to intelligence profile in an adaptive educational system," *Interacting with Computers,* vol. 18, pp. 385-409, 2006.

[9] H. J. Cha, Y. S. Kim, S. H. Park, T. B. Yoon, Y. M. Jung, and J.-H. Lee, "Learning Styles Diagnosis Based on User Interface Behaviors for the Customization of Learning Interfaces in an Intelligent Tutoring System," in *Intelligent Tutoring Systems'06*, 2006, pp. 513-524.

[10] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE,* vol. 77, pp. 257-286, 1989.

[11] M. H. Dunham, *Data Mining: Introductory and Advanced Topics*. Prentice Hall, 2002.

[12] S. Graf, Kinshuk, and T.-C. Liu, "Supporting Teachers in Identifying Students' Learning Styles in Learning Management Systems: An Automatic Student Modelling Approach," *Educational Technology & Society,* vol. 12, pp. 3-14, 2009.

[13] R. Rajendran, "Automatic Identification of Affective States Using Student Log Data in ITS," in *Artificial Intelligence in Education*, vol. LNAI 6738: Springer, 2011, pp. 612-615.

[14] R. S. J. d. Bakera, S. K. D'Mello, M. M. T. Rodrigo, and A. C. Graesser, "Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments," *International Journal of Human-Computer Studies,* vol. Volume 68, no. Issue 4, pp. 223-241, 2010.

[15] S. Salmeron-Majadas, O. C. Santos, and J. G. Boticario, "An evaluation of mouse and keyboard interaction indicators towards non-intrusive and low cost affective modeling in an educational context," in *18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems - KES2014*, 2014, vol. 35, pp. 691-700: Elsevier B.V.

[16] L. Qu, N. Wang, and W. L. Johnson, "Using Learner Focus of Attention to Detect Learner Motivation Factors," in *User Modeling 2005*, 2005, vol. 3538, pp. 70-73: Springer Berlin / Heidelberg.

[17] M. Dessouky, S. Verma, D. Bailey, and J. Rickel, "A methodology for developing a web-based factory simulator for manufacturing education," *IEEE Transactions,* vol. 33, pp. 167-180, 2001.

[18] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers Inc., 1988.

[19] A. de Vicente and H. Pain, "Informing the Detection of the Students' Motivational State: an Empirical Study," presented at the Proceedings of the Sixth International Conference on Intelligent Tutoring Systems, Heidelberg, Berlin, 2002.

[20] A. Hershkovitz and R. Nachmias, "Developing a Log-based Motivation Measuring Tool," in *Proceedings of the first international conference on Educational Data Mining (EDM)*, 2008, pp. 226-233.

[21] R. Nachimas and A. Hershkovitz, "A Case Study of Using Visualization for Understanding the Behavior of the Online Learner," *Proceedings of the International Workshop on Applying Data Mining in e-Learning,* pp. 43-52, 2007.

[22] M. Cocea and S. Weibelzahl, "Eliciting Motivation Knowledge from Log Files Towards Motivation Diagnosis for Adaptive Systems," in *Proceedings of the 11th international conference on User Modeling*, Corfu, Greece, 2007, pp. 197-206: Springer-Verlag.

[23] A. Bandura, *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1986.

[24] S. R. Viola, "E-Learning Process Characterization using data driven approaches," in *Proceedings of the International Workshop on Applying Data Mining in e-Learning*, 2007, pp. 63-72.

[25] S. Graf and P. Kinshuk, "An Approach for Detecting Learning Styles in Learning Management Systems," in *Sixth IEEE*

International Conference on Advanced Learning Technologies (ICALT'06), Kerkrade, Netherlands, 2006, pp. 161-163: IEEE.

[26] R. M. Felder and L. K. Silverman, "Learning and Teaching Styles in Engineering Education.," *International Journal of Engineering Education,* vol. 78, pp. 674-68, 1988, Art. no. 2846193.

[27] P. Sander and L. Sanders, "Measuring Confidence in academic study: a summary report," *Electronic Journal of Research in Educational Psychology,* vol. 1, pp. 113-130, 2003.

[28] S. Shaukat and M. Bashir, "University Students' Academic Confidence: Comparison between Social Sciences and Natural Science Disciplines," *Journal of Elementary Education,* vol. 25, pp. 113-123, 2016.

[29] R. K. A. P. G. Singh and M. A. EMBI, "Learner autonomy through computer mediated communication," *Jurnal Teknologi,* vol. 46, pp. 99-112, 2007.

[30] P. R. Pintrich and E. V. DeGroot, "Motivational and self-regulated learning components of classroom academic performance," *Journal of Educational Psychology,* vol. 82, pp. 33-40, 1990.

[31] G. Alberer *et al.*, "The Learn@WU Learning Environment," in *Wirtschaftsinformatik (1)'03*, 2003, pp. 593-612.

[32] P. Garcia, A. Amandi, S. Schiaffino, and M. Campo, "Evaluating Bayesian networks' precision for detecting students' learning styles," *Comput. Educ.,* vol. 49, pp. 794-808, 2007.

[33] A. G. Picciano, "Beyond student perceptions: Issues of interaction, presence, and performance in an online course," *Journal of Asynchronous Learning Networks,* vol. 6, pp. 21-40, 2002.

[34] R. Cheng, "User Motivation and Persuasion Strategy for Peer-to-Peer Communities," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 2005, pp. 3--6.

[35] A. De Vicente, "Towards tutoring systems that detect students' motivation an investigation," School of Informatics, University of Edinburgh, UK, 2003.

[36] R. S. Baker, A. T. Corbett, K. R. Koedinger, and A. Z. Wagner, "Off-Task Behavior in the Cognitive Tutor Classroom: When Students Game the System," in *SIGCHI conference on Human factors in computing systems*, 2004.

[37] L. Stankov, J. Lee, W. Luo, and D. J. Hogan, "Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety?," *Learning and Individual Differences,* vol. 22, pp. 747-758, 2012.

[38] P. Sander and L. Sanders, "Understanding Academic Confidence," *The British Psychological Society,* vol. 1 (12), pp. 29-4, 2006.

[39] C. Kesten, *Independent Learning: A Common Essential Learning: A Study Completed for the Saskatchewan Department of Education Core Curriculum Investigation Project*. 1987.

[40] M. Cukurova, "An Investigation of an Independent Learning Approach in University Level Chemistry: The Effects on Students' Knowledge, Understanding and Intellectual Attributes," Doctor of Philosophy, Department of Education, University of York, 2014.

[41] D. Laurillard, *Rethinking university teaching: A conversational framework for the effective use of learning technologies*, Second ed. 2013, p. 284.

[42] B. Weiner, *An Attributional Theory of Motivation and Emotion*. New York: Springer-Verlag 1986.

[43] B. Weiner, H. Heckhausen, and W. Meyer, "Causal ascriptions and achievement behavior: A conceptual analysis of effort and reanalysis of locus of control," *Journal of Personality and Social Psychology,* vol. 21(2), pp. 239-248, 1972.

[44] G. M. Jiga-Boy, C. Toma, and O. Corneille, "Work More, Then Feel More: The Influence of Effort on Affective Predictions," *PLoS ONE,* vol. 9(7): e101512, 2014.

[45] D. H. Shanabrook, I. Arroyo, and B. P. Woolf, "Using Touch as a Predictor of Effort: What the iPad Can Tell Us about User Affective State," in *UMAP'12 Proceedings of the 20th international conference on User Modeling, Adaptation, and Personalization*, Montreal, Canada, 2012, pp. 322-327.

[46] A. Arguel, L. Lockyer, O. V. Lipp, and J. M Lodge, "Inside Out: Detecting Learners Confusion to Improve Interactive Digital Learning Environments," *Journal of Educational Computing Research,* 2017.

[47] B. Lehman and A. Graesser, "To Resolve or not to Resolve? that is the Big Question About Confusion," *Artificial Intelligence in Education,* 2015.

[48] B. Lehman, S. D'Mello, and A. Graesser, "Confusion and complex learning during interactions with computer learning environments," *Internet and Higher Education,* vol. 15(3), pp. 184–194, 2012.

[49] L.-Y. Chang and W.-C. Chen, "Data mining of tree-based models to analyze freeway accident frequency," *Journal of Safety Research,* vol. 36, pp. 365 - 375, 2005.

[50] M. W. Lee, S. Y. Chen, K. Chrysostomou, and X. Liu, "Mining students' behavior in web-based learning programs," *Expert Syst. Appl.,* vol. 36, pp. 3459-3464, 2009.

[51] S. Song, A. J. Akande, R. O. Idem, and N. Mahinpey, "Inter-relationship between preparation methods, nickel loading, characteristics and performance in the reforming of crude ethanol over Ni/Al2O3 catalysts: A neural network approach," *Eng. Appl. Artif. Intell.,* vol. 20, pp. 261-271, 2007.

[52] M. Baglioni, B. Furletti, and F. Turini, "DrC4.5: Improving C4.5 by means of prior knowledge," in *Proceedings of the 2005 ACM symposium on Applied computing*, 2005, pp. 474-481, Santa Fe, New Mexico: ACM.

[53] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," in *Proceeding of the 2007 conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*, 2007, pp. 3-24, Amsterdam, The Netherlands: IOS Press.

[54] S. Graf, "Adaptivity in Learning Management Systems Focusing on Learning Styles," Vienna University of Technology, 2007.

[55] A. a. L. Danine, B. and Mayers, A., "TIDES - Using Bayesian Networks for Student Modeling," in *Proceedings of the Sixth IEEE International Conference on Advanced Learning Technologies*, 2006, pp. 1002 -1007, Washington, DC, USA: IEEE Computer Society.

[56] A. R. Abbasi, M. N. Dailey, N. V. Afzulpurkar, and T. Uno, "Probabilistic Prediction of Student Affect from Hand Gestures," in *Proceedings of the International Conference on Automation, Robotics and Control Systems, ARCS-08*, 2008, pp. 58-63, Orlando, Florida, USA: ISRST.

[57] A. Darwiche, *Modeling and Reasoning with Bayesian Networks*, First Edition ed. Cambridge University Press, 2009.

[58] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian Network Classifiers," *Journal of Machine Learning,* vol. 29, pp. 131-163, 1997.

[59] E. Frank, I. H. Witten, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, Third Edition ed. Morgan Kaufmann, 2011.

[60] G. F. Cooper and E. Herskovits, "A Bayesian Method for the Induction of Probabilistic Networks from Data," *Mach. Learn.,* pp. 309-347, 1992.

[61] N. Atman, M. M. Inceoğlu, and B. G. Aslan, "Learning Styles Diagnosis Based on Learner Behaviors in Web Based Learning," presented at the International Conference on Computational Science and Its Applications, 2009.

[62] M. P. P. Liyanage, K. S. L. Gunawardena, and M. Hirakawa, "Using Learning Styles to Enhance Learning Management Systems," *International Journal on Advances in ICT for Emerging Regions (ICTer),* vol. 7(2), pp. 1–10, 2014.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2937178, IEEE Access

Farman Ali Khan is an Assistant Professor at COMSATS University Islamabad, Attock Campus. He received his PhD degree in Computer Science from Vienna University of Technology, Austria. His research interests include technology enhanced learning, adaptivity and personalization, learning analytics, student modeling, image analysis, and collaborative learning technologies. He has numerous research publications including journal articles and conference papers.

Awais Akbar received his bachelor degree in Computer Science from COMSATS University Islamabad, Attock Campus. His research interest includes Technology Enhanced Learning. He has worked on various projects in the field of Educational Technology.

Muhammad Altaf completed his PhD in the field of Multimedia and Networking from university of Essex, United Kingdom. His research interests include video and image compression, video streaming over wireless and wired line networks, Signal Processing and Artificial Intelligence. He has published numerous technical papers and two book chapters on topics related to multimedia and networking and signal processing. He has worked in London South University as visiting research fellow from Jan 2015 till June 2016 and worked on projects related to signal and image processing. Currently he is working as Associate Professor in the Department of Electrical Engineering COMSATS University, Islamabad, Wah Campus, Pakistan

Shujaat Ali Khan Tanoli is currently working as Assistant Professor in Electrical Engineering Department, COMSATS University Islamabad, Attock campus, Pakistan. Mr. Tanoli received his MS leading to PhD degrees in Telecommunications Engineering from the Asian Institute of Technology, Thailand, in 2011. Mr. Tanoli was affliated with Telecoms SANS Frontieres (TSF), Regional Office for Asia and The Pacific, Bangkok, Thailand in 2008 for the establishment of Emergency deployments to operate a remote Satellite Telecom Centre in areas affected by earthquake, flood and cyclone. During his PhD studies, he earned special research exchange doctoral student scholarship award for pursuing research at the Wireless Signal Processing and Networking Lab, Tohoku University, Sendai Japan in 2010. Currently, Mr. Tanoli is involved in funded projects of HEC, Pakistan including Integration of Wireless Sensor Network (WSN) with Radio Frequency Identification (RFID) for remote pipeline monitoring and extending the vision of Internet of Thing (IoT) to include passive identification and sensing

Ayaz Ahmad is currently serving as Assistant Professor in the Department of Electrical Engineering, COMSATS University Islamabad, Wah Cantt., Pakistan. He obtained his MS and Ph.D. degrees both in Telecommunication from Ecole Superieure d'Electricite (Supelec), Gif-sur-Yvette, France, in 2008 and 2011, respectively. Ayaz has several years of research experience and has authored/coauthored several scientific publications in various refereed international journals and conferences. He has also published several book chapters and is the leading co-editor of the book entitled, "Smart Grid as a Solution for Renewable and Efficient Energy" published in 2016. He is regularly serving as TPC member for several international conferences including IEEE GLOBECOM, IEEE ICC, and IEEE PIMRC, and as reviewer for several renowned international journals.