Applying adaptive learning by integrating semantic and machine learning in proposing student assessment model

Kamilia Hosny, Abeer Elkorany
Faculty of Computers and Artificial Intelligence, Cairo University, Cairo, Egypt

Article Info	ABSTRACT
	Adaptive learning is one of the most widely used data driven approach to teaching and it received an increasing attention over the last decade. It aims to meet the student's characteristics by tailoring learning courses materials and assessment methods. In order to determine the student's characteristics, we need to detect their learning styles according to visual, auditory or kinaesthetic (VAK) learning style. In this research, an integrated model that utilizes both semantic and machine learning clustering methods is developed in order to cluster students to detect their learning styles and recommend suitable assessment method(s) accordingly. In order to measure the effectiveness of the proposed model, a set of experiments were conducted on real dataset (Open University Learning Analytics Dataset). Experiments showed that the proposed model is able to cluster students according to their different learning activities with an accuracy that exceeds 95% and predict their relative assessment method(s) with an average accuracy equals to 93%. **

Corresponding Author:

Abeer Elkorany
Faculty of Computers and Artificial Intelligence, Cairo University
Cairo, Egypt
Email: a.korani@fci-cu.edu.eg

1. INTRODUCTION

Education plays a very important role in our life. It is well-known that different students have different capabilities and skills; and thus, they learn differently. However; conventional educational systems do not always accommodate these differences, as in such systems, the typical education process is one-way; from instructors to students. Adaptive learning has received an increasing attention over the last decade. It basically adapts learning courses to meet the student's characteristics. It also provides flexibility, as students are not constrained to a specific class schedule or a predefined content. Adaptive learning [1] is a wide term that applied to an incredibly broad range of technologies and techniques in different educational fields. The common thread when using this terminology is that; they all involve software that observes some aspect of student performance and adjusts what it presents to each student based on those observations. Most educational systems are designed around a set of values that govern all educational choices in that system. Such choices include curriculum, physical classroom design, student-teacher interactions, assessment methods, class size, educational activities, and more. One of the learning factors is the assessment methods; and since the adaptive learning as a concept is to personalize the learning factors in a way that meet the student learning style. Therefore, it is important to provide a method that predict the suitable assessment for each student based on her/his learning style, as it represents one of the most significant issues for adaptive learning. The main objective of this paper is to develop an integrated model that could recommend suitable assessment method to student based on her/his learning data. In order to achieve this objective, two main phases are applied. The first one is concerned with identifying learning style of each student based on visual,

auditory or kinaesthetic (VAK) learning style [2]. The second phase is to recommend the suitable assessment method based on deduced learning style according to [3]. During the first phase, two main steps are applied each utilize different techniques. The first step is to apply semantic mapping techniques to map learner data into VAK Learning style model (which stands for visual (V), auditory (A), and the kinaesthetic (K)). The second step is to cluster students based on their data into one of the three VAK learning styles (visual, auditory or kinaesthetic). Finally, in the second phase, prediction of the suitable assessment method is applied based on student learning style. A set of experiments using Open University Learning Analytics dataset [4] has been applied to evaluate the proposed model. The paper is organized as follow. Section 2 describes the required backgrounds about student learning style, machine learning as well as semantic technology. Section 3 discusses the related work of different technologies of student modelling and different usage of machine learning in education fields. In section 4, the overall recommendation process and the proposed model are described. Section 5 provides the experimental steps using the Open University Learning Analytics dataset. Section 6 presents the conclusion as well as the future work.

2.   LITERATURE REVIEW

In this section we will present the literature review that covers the important parts of this research which are adaptive learning, student modelling, semantic association and machine learning. The proposed model aims to expose the power of the semantic association that is integrarted with supervised machine learning techniques in the improvement of personalized learning as will be illustrated here.

2.1. Adaptive learning

Distance learning (E-learning) [5] is the use of electronic technology in learning or teaching outside the educational institutions. Learning management systems (LMS) [6] apply the concept of distance learning to offer a complete learning process to students through interactive online learning environment. The huge amount of data generated from theses LMS is the nucleus for the adaptive learning since it has a great impact in enhancing the detection of students' learning style. Different learning style models have been presented. These models determine the way every student understands processes, comprehends, retains and deals with information. Among these learning style models is VAK learning style. VAK model [2] is one of the most common and widely-used type of learning styles that categorize learners based on the sensory modalities which are involved in taking information. It assumes that learning will be effective by providing an appropriate learning process to each student based on her/his sensory modalities. As (visual) prefers to learn by remember, (auditory) prefers learning by hearing and (kinaesthetic) prefers learning with movement and emotion. Students should be able to utilize their distinctive sensory modalities in learning activities [7]. VAK learning style does not involve intelligence or inherent skill but is closely related to how we acquire or understand information or new knowledge [8].

Semantic association [9] is a direct or indirect relation between two entities (in our case entities are words). This relation must be considered as meaningful, which means that relations are useful in a certain context for an application. In other words, Semantic association considered as the semantic connection between two textual units (words, sentences or even documents). Semantic association [10] between entities could be in form of semantic similarity, semantic relatedness, semantic distance or other. Accordingly, semantic association computation is the process of converting the semantic connection between two textual units to a number representing the strength of the semantic connection. This computation is based on different kinds of semantic relations. Simply, Semantic association computation process takes two pieces of text as input then produces a real number representing the strength of the semantic connection between them as an output. Semantic association computation quantifies the association between two concepts by identifying a chain of possible lexical and semantic connections between them. It requires an understanding of the implicit relations of concepts based on deeper level of world knowledge. Computing the semantic associations requires background knowledge source. The performance of the semantic association measure is relying on the choice of background knowledge that supports the relationship of the concepts. There are two types of knowledge sources [10], informal knowledge source and formal knowledge source. Informal knowledge source contains the semantic connections in the form of distributional context and occurrence patterns in unstructured way. Formal knowledge source explicitly organizes the semantic relation in a structured way such as Wordent. Wordnet [11] is a formal linguistic knowledge source where the advantages of the huge knowledge coverage with explicit semantic encoding in a structured way exist so it is a rich approach in computing semantic association. Wordnet built upon a structured network model representing concepts as structured elements like nodes in a semantic graph.

Machine learning technology is used to teach machines how to improve automatically the handling of the data through experience. It is also an efficiently study and design of computer algorithms as an alternative to the conventional engineering approach [12], [13]. Two main classes of machine learning

techniques [14]: supervised learning and unsupervised learning. In supervised learning, the training set contains pairs of input and output, with a goal to learn how to map input to output spaces. While in unsupervised learning the training set consists of inputs without any assigned desired output. In the proposed model, the semantic association is integrarted with supervised machine learning algorithms to detect the students' learning styles using the real-life dataset that was generated from learning managemt systems. Next, suitable assessment method is recommended based on student learning style for each student as a part of adative learning concept.

2.2. Student modeling

Different research works introduce personalized learning models built upon the idea of student modelling such as the mathematical model introduced by [15] for profiling learners to determine their learning style based on myers and brigs type indicator questionnaire. Then the learner profile is used to adapt the teaching strategies by providing different learning paths to each learner. 50 students studying information technology were used for test cases in this research. Test cases show that 78% of students passed and the remaining 22% passed when the strategy was revised. Aeiad and Meziane [16] a generic architecture for the development of personalized and adaptable E-learning systems that recommend learning material is introduced. This architecture also relying on a learner model that built upon VAK questionnaire [8] to develop programs' contents from freely available resources on the WWW using knowledge engineering approaches. However, the questionnaire is not considered as a reliable method to detect student learning style since learners may tend to choose answers arbitrarily. While [17] suggests architecture of a personalized learning environment to recommend learning content based on learner modelling. This modelling built upon learner monitoring unit that tracks the learner behaviour using dynamic bayesian network. Then a learner model is built in order to map learner's attribute to learning resources' metadata although this architecture was not implemented yet. Unlike the proposed model; most of the researches depend on the questionnaire to identify student's learning style which is not a reliable method since learners may tend to choose answers arbitrarily.

2.3. Machine learning

Other research works use machine learning in the detection of student learning style such as [18] that extracts features for fleder and silver's man learning style model (FSLSM) [19] dimensions based on other relevant research works. Then, apply k-means clustering to create group of learners according to their degree of preference to specific FSLSM dimension. The clusters' labels are identified based on an average of the features related to each learning dimension. Paper conducted an experiment on a dataset extracted from the edX platform, from two sessions of the ''statistical learning stat'' course delivered by the University of Stanford. The experiment was to assess the quality of clustering according to calinski-harabasz (CH) and silhouette (SI) indexes. CH and SI indices were compared among many clustering algorithms, and the result shows that k-means was the best performing algorithm. Then the clustering was followed by an aggregation process to specify each cluster with a label. Finally, a decision tree classification is conducted on the resultant labelled data and achieved an accuracy of 98%.

Hmedna et al. [20], they extracted features describing FSLSM learning style in order to construct a vector of characteristics for each learner. Then apply k-means clustering to partition learners into different groups according to their learning style preferences. However, features are tailored with FSLSM learning style characteristics manually. The proposed approach is tested and validated on 5,482 learners enrolled in "statistical learning" course administered via Stanford's Logunita platform. The quality of the clustering was assessed by Calinski-Harabasz.

Reyes-González et al. [21] a model that uses conceptual clustering algorithms was introduced to group the students' models. Students' models are grouped in clusters according to their degree of similarity and then distinctive features that characterize each cluster are determined. A described method that depends on human experts was conducted to validate the model feasibility and result of clustering through intelligent tutoring system, and seven experts. The validation method stated that the proposed model provides a feasible and effective method. It also ensured that there are no students in other groups with the same characteristics. Moreover, the validation method showed effectiveness in the correctness of student modelling classification to 100% of the selected students' models. However, there were a miss mapping to any specific learning style standard for each cluster.

Sya'iyah et al. [22] authors employ K-Mean's algorithm to determine the characteristics of students who have excellent performance, standard performance, and underperformance levels. This was achieved by clustering of 724 students over three groups (which was a small number of records). The main goal of their research was to identify the correlation between student performance per each cluster, and variable of grade point average, length of study duration, length of thesis duration, and English proficiency score. The study

concludes that student data are grouped successfully using the K-Means algorithm however they did not provide any clustering validation method.

Others extend the concept of machine learning to make predictions for student performance such as [23] that used the hierarchical clustering methods to determine the relations between student variables and in identifying the main indicators that could help in predicting courses' grades in small datasets. This research proved the possibility of predicting student performance for each course in new universities. The support vector machine classifier with radial kernel was the one which proved its efficiency (among the rest of classifiers) in predicting students' performance in all courses' grades. Balanced accuracy performance measure was conducted for the evaluation and has an accuracy result of 73%.

Aljohani *et al.* [24] use time-series sequential classification to identify students at risk of failure on the basis of their behaviour in the virtual learning environment using accessible Open University Learning Analytics dataset. The research made an early prediction for students at risk of failure in a sequential manner. This research calculated the efficiency of the deployed model, in terms of the difference between the actual and predicted values which results in 95.23% learning accuracy. However, in this research they consider only the frequency of students' interaction ignoring the type of activity and the assessment grades during the year work. In summary, most of the developed research work that utilized machine learning in student performance prediction, did not consider any student standard learning models as a benchmark for student's behaviour data extraction and mapping. Also, the research works that utilize learning standard models to detect the students' learnin style depend on a manual mapping between the learning style characteristics and students' behavior; which may contain a lot of drawbacks. Unlike the above, the proposed model utilizes VAK standard to detect students' learning style using semantic mapping method between the used standard' characteristics and students' behavior.

2.4. Semantic association

Su *et al.* [25] authors introduce a data processing technique for enhancing the text classification named tree-structured multi-linear principal component analysis (TMPCA). TMPCA facilitate the machine-learning task that follows by reducing the dimension of the entire sentence data while preserving the sequential order of words' composition. The authors analyzed the mathematical complexity of TMPCA to demonstrate its computational efficiency over the traditional principal component analysis (PCA). Finally, authors conduct experiments using TMPCA method on four different datasets and the result shows that TMPCA achieves the lowest error rates among benchmarking methods.

Su *et al.* [26], authors extend the TMPCA mathematically in order to enhance the sequence-based text classification tasks by preserving strong mutual information between its input and output. Su and Kuo [27] and Su *et al.* [28] authors conduct mathematical analysis on three RNN cells; the simple recurrent neural network, the long short-term memory and the gated recurrent unit. RNN cells defined as a function that maps an element in a sequence to output, for natural language processing purposes. Based on the analysis, they propose the extended-long short-term memory, to extend the memory length of a cell to address the issue of low memory to establish complicated language tasks such as dependency parsing. Compared with the proposed work, that utilise the semantic association computation to indicate the relation as it considers implicit relations of concepts based on deeper level of world knowledge; and thus enhance the mapping of learning activities to VAK learning style categories.

Sheeba and Krishnan [29], the authors analysed learner's weblog in order to create a learner profile using Wordnet. The detection pass through the following steps; initially, weblog files was processed to retrieve the learner's frequently visited documents then extracts the important keywords. Next step is to compute the term frequency (TF) to calculate how frequently a term occurs in a document and calculate inverse document frequency (IDF) which determines the importance of a term. Final step was to assign weight for each term by multiplying TF*IDF. Second stage is to represent the documents selected by the learner in ontology using WordNet in order to identify the semantic concepts related to the document keywords. In this step, a list of three words with the highest weigh is initialized for each document. For this list; Wordnet is invoked to retrieve each term synset in order to identify the concept for a term. The concepts of the selected documents reflect the true learner interest which in turn is used to retrieve the relevant learning contents. The experiment shows a successful completion of extraction and updating of learner interest to the learner profile. However, the proposed model did not use any learning style standard to adapt the learning content.

While in [30] authors focused on the automatically categorization of exam's questions into the learning levels of Bloom's taxonomy [31]. Researchers introduced a rule-based exam question classification model that depends on natural language processing (NLP) pre-processing techniques, the cosine similarity algorithm and WordNet similarity algorithm. WordNet similarity algorithm was mainly applied on the verbs that were extracted from the questions. This classification process enables evaluators to redesign their exam papers. The proposed model was examined on sample of exam questions obtained from the Faculty of

Applied Sciences, Wayamba University, Sri Lanka. The training set consists of 53 questions and the test dataset consists of 35 questions. Identified question categories were evaluated by a domain expert. Although that the result indicated more than 70% categorization accuracy, the used evaluation criteria were not a reliable method.

Unlike the above work, the proposed model recommends the suitable assessment method type for each student based on her/his auto detected learning style following a specific standard with a reliable evaluation experiment. Assessment methods recommendation is a part of tailored learning environment since it considers the differences between students in the way of input processing. Therefore, the way each student should be assessed is differ based on his/her learning style. Thus, the main objective of the proposed model is the identification of student learning style by tracking his/her learning activities. Unlike all the above-mentioned research work, the proposed model integrates machine learning techniques and semantic techniques to recommend suitable assessment method. First, it utilized suitable clustering techniques to determine students' preferred learning activities. Next, detection of each student learning style is constructed based on a semantic mapping between students' learning activities and VAK model. Learning style detection is followed by a suitable assessment method recommendation according to VAK assessment methods introduced in [3]. Finally, the proposed work was evaluated through realiable evaluation criteris using expermintes on a real dataset.

3. PROPOSED MODEL

This section aims to provide details about the proposed model for recommending a suitable assessment method to student based on her\his learning style. The model is decomposed from three main components as shown in Figure 1. The first one is responsible for applying different mapping techniques to map learner data into VAK model. The second component aims to cluster students based on their data into one of the three VAK learning styles (visual, auditory or kinaesthetic). Finally, the third component targets prediction of suitable assessment method based on student learning style.
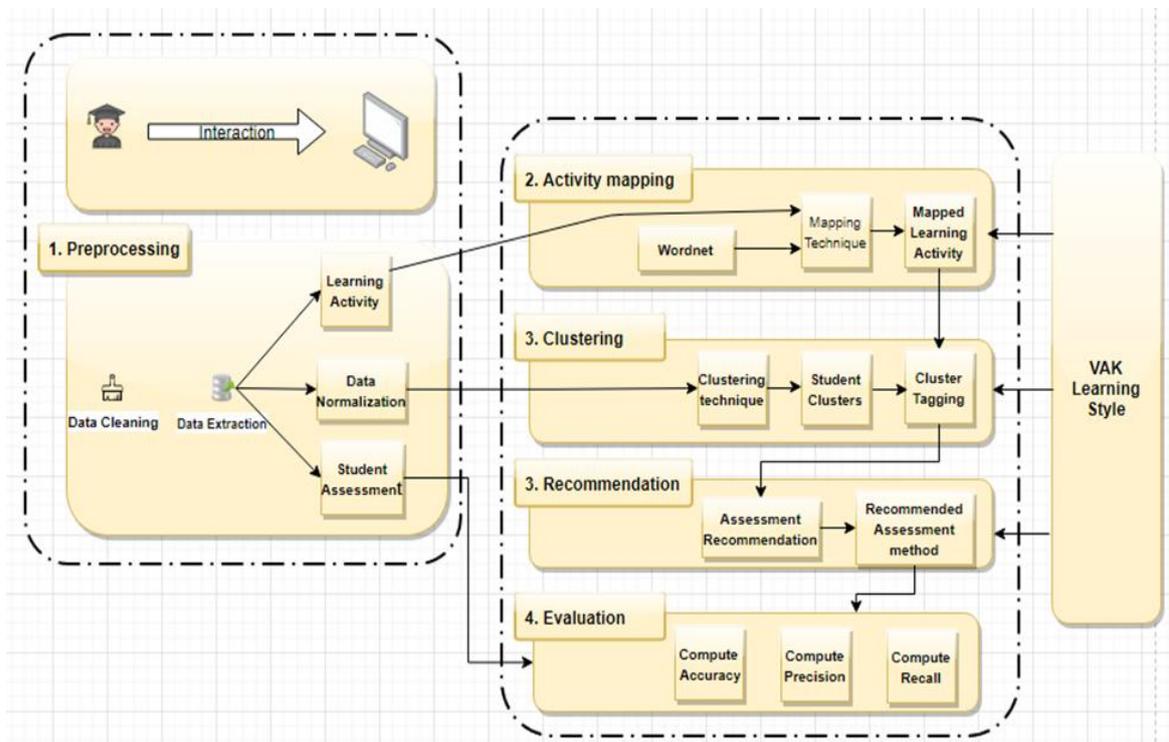
Figure 1. Proposed model for assessment recommendation based on student modelling

3.1. Mapping student learning activities to VAK learning style

In this phase, mapping is applied between learner data and VAK model. Mapping is applied using semantic technology. In the proposed model, the computation of semantic association between learner data

and VAK categories utilize Wordnet [11] as a formal linguistic knowledge source integrated with the advantage of semantic similarity techniques [10], [32]. Therefore, the process of mapping is divided into the following two steps: First a corpus of student VAK learning style is built. Next, mapping of student current activities is applied. It is significant to mention that, the mapping process is generally applicable to any students' activities, however, in this paper, a case study from Open University data set is presented.

3.1.1. Building VAK learning style corpus

According to [3], VAK concepts were first developed by psychologists and teaching specialists such as Fernald, Keller, Orton, Gillingham, Stillman and Montessori, starting in the 1920's. Currently VAK model is provided by explaining a person's preference in terms of specific description of text. Therefore, VAK text is processed and the tokens that represent nouns are extracted and any irrelevant words are excluded. The resultant of this process is three vectors each represents one of the three learning styles, such that each vector represents a list of keywords that characterize that learning style. Then we built learning style dictionary. After the tokenization process, WordNet was invoked to build a dictionary for the resultant learning style keywords with synonym and related words. In this step, WordNet get-synonym and get- related words functions are used respectively. Get related words function retrieves related words based on the semantic relatedness relation between concepts in the same context. Therefore, by the end of this step, each learning style category' vector would be a matrix that contains all the synonyms and related words that represents this learning style category.

3.1.2. Mapping process

This process is concerned with the mapping of students' learning activities to VAK learning style categories. Mapping process could be categorized into two main categories: Full matching and semi-matching as what will be discussed in this section. In this step, mapping between learning activities in the dataset and VAK learning style dictionary is constructed through two matching categories mentioned above. First, each learning activity available in the dataset is processed and the tokens that represent nouns are extracted. The result stored in form of vectors for each learning activity to be compared with the dictionary of the learning style categories.

a.   Full matching category

In this category, a complete matching based on Wordnet synonym relation [32], which group the index words into a synonym set, was applied. In this mapping type, the learning activity description was included in the learning style characteristics synonyms. This relation represents a strong relation between the activity and the resultant learning style (100% mapping) which resulted in a complete matching using Dictionary based Mapping. During this process, each token in the learning activity vector is compared with VAK learning style dictionary. During applying this process, there were two main issues. The first one, when the majority of learning activity tokens are localized in one learning style category. In this case, this activity is directly mapped to that learning style category. While, the second case when learning activity tokens are included in more than one category with the same number of tokens, then this activity is mapped to both categories.

b.   Semi-matching category

This stage depends on semantic similarity between concepts [32]. Semantic similarity [11], [33] represents the semantic connection between two concepts that have similar nature or attributes. Thus, this category would include semi-similar concepts which are related to original concept using specific threshold. According to Tversky's cognitive psychology theory [33], the semantic similarity between concepts can be measured and quantified by the topological parameters edge and information content (IC) in the taxonomy. Such as hierarchy constructed by "is-a" relations, the features based on the has-part relation that provides the component evidence of concepts, and the has-member/member-of relations with the taxonomic significance. There are many equations [32], [34] that compute the semantic similarity and relatedness between concepts; Information content-based measure IC, Path and depth-based measure and hybrid measure. In the proposed model, we depend on the depth of the two senses in the taxonomy and that of their Least Common Subsume (most specific ancestor node). In case that, there is no complete matching between any of the learning activity and one of the learning style categories. Thus, similarity mapping would be applied using Wordnet similarity function which calculates how far two concepts are similar to each other within the same IC. In this case the mapping between the learning activity and the learning style category will be based on the semantic similarity measure using specific threshold. In the proposed model, Wu and palmer similarity measure [35] are used for calculating the Semantic Sentence Similarity. Wu-Palmer Similarity [34] returns a score denoting how similar two words are, based on the depth of the two senses in the taxonomy and that of their (LCS) Least Common Subsumer (most specific ancestor node). Note that the LCS does not necessarily feature in the shortest path connecting the two senses, as it is by definition the common ancestor deepest in the taxonomy, not closest to the two senses. Typically, in case, multiple candidates for the LCS exist, that

whose shortest path to the root node is the longest will be selected. Where the LCS has multiple paths to the root, the longer path is used for the purposes of the calculation. Wu and Palmer calculated the semantic similarity measure between concepts C1 and C2 as per (1),

$$Similarity\ (C1, C2) = \frac{(2 \times N3)}{N1 + N2 + (2 \times N3)} \tag{1}$$

where N1 is the length given as number of nodes in the path from C1 to C3 which is the minimum collective super concept of C1 and C2. N2 is the length given in number of nodes on a path from C2 to C3. N3 signifies the global depth of the hierarchy and it serves as the scaling factor.

3.2. Detection of student' learning style using clustering techniques

Clustering plays a significant role in the detection of student behaviour as it classifies the students in well-defined groups [36]. Clustering builds upon the idea that; instances within the same cluster be similar as much as possible; while Instances in different clusters be different as much as possible. Also, the measurement for similarity and distance must be clear and have a practical meaning since they are the basis for clustering construction. In this research Open University dataset was used as a cased study for student behaviour, where student features are recognized to identify the relationship among data because of quantitative data feature [37]. Several searches divide clusters to many categories, most known and common are partition (centroid) based clustering, density-based clustering, hierarchal based clustering, model-based clustering and distribution-based clustering [37], [38]. Centroid models are iterative clustering algorithms in which the similarity is derived by the closeness distance of data points to the centers of the clusters. These models run iteratively to find the local optimal cluster centroid. K-means clustering algorithm is the most popular algorithm of this category. These models require the number of clusters to be identified in advance. Density based clustering basically finds the places that have dense data points and calls those clusters. These algorithms failed to deal with data of varying densities and high dimensions. Model-based clustering algorithm selects a particular model for each cluster then search for the best fitting points for that model. The disadvantages of these types of algorithms are the relatively high time complexity in general, the clustering result sensitive to the selected model parameters and the premise not completely correct. Distribution models [38], [39] are built based on the probability that all data points in the cluster are belong to the same distribution. These models often suffer from over fitting. Since most of the researches agreed that the detection of the clustering method is highly depends on the experiments. In the proposed model, we will depend on the upcoming experiments to determine the best clustering model.

3.3. Assessment methods recommendation

According to [3] each VAK learning style category related to specific assessment methods where student can perform better. In this phase, Assessment method type recommendation is identified based on the student detected learning style. VAK learning style assessment method types were used according to [3] where auditory learners prefer subjective assessment methods. Visual learners prefer subjective assessment methods and kinesthetic learners prefers objective assessment methods. Subjective assessment includes assessment methods that did not have definite one right answer which could be considered as tutor marked assessment (TMA) type such as case studies and long essay questions. While objective assessment includes assessment methods that required one correct answer which could be considered as computer marked assessment (CMA) type such as multiple-choice questions, true or false questions.

4.   EXPERMINTEL EVALUATION

In order to measure the effectiveness of the proposed approach three experiments were conducted. The first and second experiments target the clustering of students in order to identify the significant feature of each student as well as detecting students' learning style. While the last one targets the prediction of student assessment method based on her/his learning style according to [3].

4.1. Data collection

In order to evaluate the effectiveness of the proposed model, Open University Learning Analytics Dataset (OU, 2017) [4] was used in all experiments. What makes the dataset unique is aggregated stream data of students' interactions in the virtual learning environment (VLE) using Moodle platform learning activities. This enables the analysis of student behavior, represented by their activities. The interaction with the VLE was further categorized into 20 different activity types with each activity referring to a specific action, such as downloading or viewing lectures, course content, or quizzes. The dataset contains the information about 7

courses, 32,593 students, their assessment results, and logs of their interactions with the VLE represented by daily summaries of student clicks (10,655,280 entries). Furthermore, the dataset contains their assessments, quizzes, and course performances.

4.2. Data preprocessing

In order to obtain more meaningful and accurate results, we filter out the students who made course withdrawal. The resultant Dataset includes only 22,593 students, their assessment results, and logs of their interactions with the VLE represented by daily summaries of student clicks (9,921,259 entries). Using SQL queries the dataset is divided into 2 parts. First part used to construct the proposed model as it contains the student's interaction with the VLE. Second part used in the proposed model evaluation as it contains students' assessments methods and grades. Data were obtained by joining 7 different tables using the unique identifiers. The resultant table contains row for each student with student ID, registered course, total number of student activities on the registered course(s) and the student interaction frequency per each learning activity.

4.3. Applying different clustering techniques

In this experiment, it is required to detect the students' VAK learning style by applying different clustering mechanisms.

4.3.1. Setup

A comparison between different types of clustering techniques was done using Weka [40] as an open-source tool. As per the Table 1 and Figure 2, k-means cluster algorithm provided clear limits between clusters with relatively low time complexity and high validation result with less mean square error. On the other hand, model-based clustering suffers from over fitting with relatively high time complexity. Finally, the result of density-based clustering provided unclear limits between clusters as the dataset is very dense.

Table 1. Comparison between different clustering models accuracy

	Centroid Based Clustering (K-means)	Density Based Clustering (DBScan)	Model Based Clustering (EM)
Cross Validation	95.9023%	91.7105%	94.1541%
RootMean-Square Error	0.1616	0.2271	0.1794
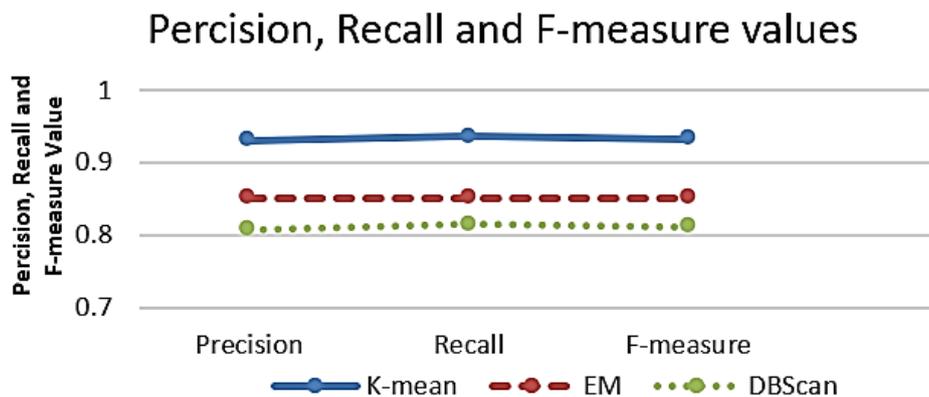
Figure 2. Comparison between different clustering models precision, recall and f-measure

4.3.2. Results and discussion

Based on the above results obtained, the following remarks are highlighted:
− Density-based approaches can find the unknown number of clusters of similar density but it only considers the small neighbourhood of nearby points and so fails to consider the full dataset.
− Model based approach clustering resufiglt is sensitive to the selected model parameters so the premise not completely correct.

However, all those approaches fail at accurate clustering across the full dataset except centroid models. In a sense, partitioning (centroid) clustering considers every point in the dataset and uses that information to construct clusters with the highest evaluation result and less mean square error. Also, the

resultant clusters have a clear limits and distinct dominant features that help in the mapping to learning style categories. Therefore, in the next experiment, different partitioning (centroid) clustering will be applied.

4.4. Applying partition (centroid) based clustering algorithms

The main objective of this experiment is to identify the most significant features of each cluster in order to detect validity of the proposed clustering techniques. Based on experiment1, centroid based clustering algorithms provided the most accurate results, therefore, k-means++ and k-means [41], [42] from centroid based clustering algorithms are applied and a comparison between them is showed in Figure 3.
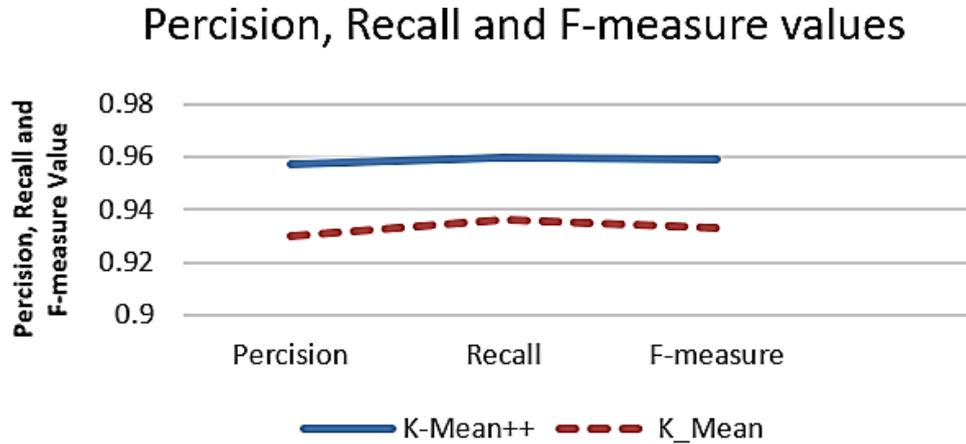
Figure 3. Comparison between k-mean and K-means++ algorithms

4.4.1. Set up

Partition based clustering algorithms are applied through four steps:

− Data normalization step: in centroid clustering, calculation of the similarity between two individuals is accomplished through combining all the feature data for those individuals into a numeric value. Combining feature data requires that the data have the same scale.
− Applying the centroid clustering: using WEKA tool, students per each registered course are divided into three clusters.
− Clustering validation: using WEKA tool, cross validation has been applied over 5 folds
− Detection for dominant activities per each cluster: activities per each cluster were selected based on the mean as an indication for the average use and the variance. Ideally, we obtain significantly different means for most, if not all features, used in the analysis. The magnitude of the Mean values performed on each feature is an indication of how the respective activity discriminates between clusters.

After applying K-Mean and K-Mean++, K-Mean++ clustering algorithm provided the highest cross validation result as per Tables 2-3 and Figure 3 respectively.

Table 2. Partition based clustering algorithms (k-means) result of students registered on one course

	Cluster 1	Cluster 2	Cluster 3
Clusters size	1,343 (25%)	1,638 (31%)	2,342 (44%)
Dominant learning activities of (OU) students	DataPlus	Forumng	OuContent
	DualPane	Ouwiki	Resources
	Subpage	Html Activity	URL
	Glossary	Ouelluminate	OUcollaborate
	Questionnaire		
VAK category	Kinesthetic	Auditory	Visual
Cross Validation		95.5075 %	
	Correctly Classified students: 5,081		
	Incorrectly Classified students: 239		

4.4.2. Result discussion

The difference between K-Means ++ and K-Means [41], [42] is based on the centre of the cluster initialization phase as in the K-Means cluster centre initialization is done randomly while, K-Means++

cluster centres have been proportional to the distance to the centre of the cluster using squared Euclidean distance. Therefore, K-Means++ aims to avoid a poor outcome resulting grouping of central election early. Thus, we will depend on K-Mean++ clustering algorithm as it has the highest cross validation result as showed in Tables 2-3 and Figure 3. K-Means++ clustering algorithm is applied for students in each of the seven courses available in the Open University Dataset with fifteen different learning activities types.

Table 3. Partition based clustering algorithm (k-means++) result of students registered on one course

	Cluster 1	Cluster 2	Cluster 3
Clusters size	1,340 (25%)	1,638 (31%)	2,342 (44%)
Dominant learning activities of (OU) students	DataPlus	Forumng	OuContent
	DualPane	Html Activity	Resources
	Ouwiki	Ouelluminate	URL
	Subpage		OUcollaborates
	Glossary		
	Questionnaire		
VAK category	kinesthetic	Auditory	Visual
Cross Validation		95.5263 %	
	Correctly Classified students: 5,082		
	Incorrectly Classified students: 238		

4.5. Assessment recommendation

According to [43], the most common and powerful metrics to measure the quality of a recommendation system are accuracy, precision, recall, and F1-score. Accuracy [44] is the ratio between right made predictions to all predictions. Precision and recall [44] are computed based on what is selected and not selected, and what is relevant and irrelevant. So, in order to evaluate our recommendation model using precision and Recall, we need to define relevance.

4.5.1. Set up

In our case precision indicates how many recommended assessment methods are relevant to students. While recall indicates how many relevant assessment methods are recommended where relevance is defined by student assessment grade. Thus, according to Open University dataset, the assessment method has a rating scale of 0-100 and is commonly transformed into a binary scale by converting every rating of medium grades or highest grades to "relevant" and all ratings below 50 to "not relevant." SUCH THAT LOWEST GRADE (below 50), medium grades (from 50-75) and highest grade (76-100). For precision and recall, we also need to separate the item set into the set that was returned to the user (selected/recommended), and the set that was not selected as shown in Table 4.

Table 4. Relevance and irrelevance definition

	Selected	Not Selected	Total
Relevant	Nrs (TP)	Nrn (FN)	Nr
Irrelevant	Nis (FP)	Nin(TN)	Ni
Total	Ns	Nn	N

Where:
- True positives (TP): These are the correctly predicted positive values which means that the predicted assessment method was the same like the true method E.g., Student actual method=relevant CMA and Student predicted method=CMA.
- Student actual method relevant TMA and student predicted method=TMA.
- True negative (TN): These are the correctly predicted negative values which mean that the recommended assessment method was true but not relevant to student.
- False positives (FP): When actual class is not as predicted class which means that the recommended assessment method was false but relevant to student.
- False negatives (FN): This is the worst case when recommended assessment method was false and not relevant to student.

4.5.2. Result and discussion

First, we exclude AAA course with his students (611 students) and EEE course with (2069 students) as they have only TMA assessment method type. Then, we apply the explained evaluation metrics on 18891 students in 5 different courses data and its results are showed in Table 5. According to the table, the model provided good accuracy results as well as high F-score, compared to what has been presented in the litrarture.

Table 5. Evaluation metrics of students' assessment method prediction

Course	Accuracy	Precision	Recall	F1-score
BBB	0.87	0.92	0.97	0.94
CCC	0.95	0.94	0.99	0.96
DDD	0.96	0.98	0.99	0.98
FFF	0.92	0.98	0.99	0.98
GGG	0.95	0.96	0.98	0.97

5.   CONCLUSION AND FUTURE WORK

This paper introduced an adaptive model for assessment recommendation based on the student model. By involving both semantic technology and machine learning technology, the student VAK learning style was detected through the tracing of student learning activities. Experiments were conducted using Open University Learning Analytics Dataset to prove the system ability in student learning style detection and assessment methods recommendation. The quality of the recommendation system is assessed by accuracy, precision, recall and F1-score. The experiments show that the system is able to detect the student learning style with an accuracy exceeding 95% while the assessment recommendation accuracy average of 93% and F1 score average of 96.6%. Our future work is to make on time adaptation for learning parameters such as learning contents, assessment methods and more, considering student's real time learning activities.

REFERENCES

[1]   A. H. Jonsdottir, A. Jakobsdottir, and G. Stefansson, "Development and use of an adaptive learning environment to research online study behaviour," *Educational Technology and Society*, vol. 18, no. 1, pp. 132–144, 2015.
[2]   S. S. K. and T. C. Helena, "Styles of learning based on the research of fernald, keller, orton, gillingham, stillman , montessori and neil d fleming," *International Journal for Innovative Research in Multidisciplinary Field*, vol. 3, no. 4, pp. 17–25, 2017.
[3]   S. Wickramasinghe and S. Hettiarachchi, "Relationship among students' learning styles, assessment methods and students' performances," in *Third International Conference on Education and Distance Learning*, 2017, pp. 1–23.
[4]   J. Kuzilek, M. Hlosta, and Z. Zdrahal, "Open university learning analytics dataset," *Scientific Data*, vol. 4, no. 1, Dec. 2017, doi: 10.1038/sdata.2017.171.
[5]   H. Warni, "The relevance of e-Learning in higher education," *Journal Kajian Pendidikan*, vol. 3, no. 2, pp. 181–194, 2013.
[6]   D. Turnbull, R. Chugh, and J. Luck, "Learning management systems: an overview," in *Encyclopedia of Education and Information Technologies*, 2019, pp. 1–7.
[7]   M. T. A. N. Hardina and P. Suyata, "The effectiveness of VAK (visual, auditory, kinesthetic) model in learning of summary writing," *International Journal of Research and Review*, vol. 5, no. 8, pp. 43–49, 2018.
[8]   N. Othman and M. H. Amiruddin, "Different perspectives of learning styles from VARK model," *Procedia - Social and Behavioral Sciences*, vol. 7, pp. 652–660, 2010, doi: 10.1016/j.sbspro.2010.10.088.
[9]   A. Sheth *et al.*, "Semantic association identification and knowledge discovery for national security applications," *Journal of Database Management*, vol. 16, no. 1, pp. 33–53, Jan. 2005, doi: 10.4018/jdm.2005010103.
[10]   S. Jabeen, X. Gao, and P. Andreae, "Semantic association computation: a comprehensive survey," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 3849–3899, Aug. 2020, doi: 10.1007/s10462-019-09781-w.
[11]   D. Soergel, *WordNet: an electronic lexical database*. Cambridge: The MIT Press, 1998.
[12]   O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 648–664, Dec. 2018, doi: 10.1109/TCCN.2018.2881442.
[13]   A. Dey, "Machine learning algorithms : a review," *International Journal of Computer Science and Information Technologies*, vol. 7, no. 3, pp. 1174–1179, 2016.
[14]   M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, "A systematic review on supervised and unsupervised machine learning algorithms for data science," in *Unsupervised and Semi-Supervised Learning*, 2020, pp. 3–21.
[15]   O. Sunday Adewale, O. Catherine Agbonifo, and O. Lauretta Osajiuba, "Development of a myers-briggs type indicator based personalised e-Learning system," *International Journal of Computer (IJC)*, vol. 35, no. 1, pp. 101–125, 2019.
[16]   E. Aeiad and F. Meziane, "An adaptable and personalised E-learning system applied to computer science programmes design," *Education and Information Technologies*, vol. 24, no. 2, pp. 1485–1509, Mar. 2019, doi: 10.1007/s10639-018-9836-x.
[17]   N. S. Raj and V. G. Renumol, "Architecture of an adaptive personalized learning environment (APLE) for content recommendation," in *Proceedings of the 2nd International Conference on Digital Technology in Education - ICDTE 2018*, 2018, pp. 17–22, doi: 10.1145/3284497.3284503.
[18]   B. Hmedna, A. El Mezoury, and O. Baz, "A predictive model for the identification of learning styles in MOOC environments," *Cluster Computing*, vol. 23, no. 2, pp. 1303–1328, Jun. 2020, doi: 10.1007/s10586-019-02992-4.
[19]   S. Graf, S. R. Viola, T. Leo, and Kinshuk, "In-depth analysis of the felder-silverman learning style dimensions," *Journal of Research on Technology in Education*, vol. 40, no. 1, pp. 79–93, Sep. 2007, doi: 10.1080/15391523.2007.10782498.
[20]   B. Hmedna, A. El Mezoury, and O. Baz, "How does learners' prefer to process information in MOOCs? a data-driven study," *Procedia Computer Science*, vol. 148, pp. 371–379, 2019, doi: 10.1016/j.procs.2019.01.045.
[21]   Y. Reyes-González, N. Martínez-Sánchez, A. Díaz-Sardiñas, and M. de la C. Patterson-Peña, "Conceptual clustering: a new approach to student modeling in intelligent tutoring systems," *Revista Facultad de Ingeniería Universidad de Antioquia*, no. 87, pp. 70–76, Jun. 2018, doi: 10.17533/udea.redin.n87a09.
[22]   K. Sya'iyah, H. Yuliansyah, and I. Arfiani, "Clustering student data based on k-means algorithms," *International Journal of Scientific and Technology Research*, vol. 8, no. 8, pp. 1014–1018, 2019.
[23]   L. M. Abu Zohair, "Prediction of of student's performance by modelling small dataset size," *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, Dec. 2019, doi: 10.1186/s41239-019-0160-3.
[24]   N. R. Aljohani, A. Fayoumi, and S.-U. Hassan, "Predicting at-risk students using clickstream data in the virtual learning

environment," *Sustainability*, vol. 11, no. 24, Dec. 2019, doi: 10.3390/su11247238.

[25] Y. Su, Y. Huang, and C.-C. J. Kuo, "Efficient text classification using tree-structured multi-linear principal component analysis," in *24th International Conference on Pattern Recognition (ICPR)*, Aug. 2018, pp. 585–590, doi: 10.1109/ICPR.2018.8545832.

[26] Y. Su, R. Lin, and C.-C. Jay Kuo, "Tree-structured multi-stage principal component analysis (TMPCA): theory and applications," *Expert Systems with Applications*, vol. 118, pp. 355–364, Mar. 2019, doi: 10.1016/j.eswa.2018.10.020.

[27] Y. Su and C. C. J. Kuo, "On extended long short-term memory and dependent bidirectional recurrent neural network," *Neurocomputing*, vol. 356, pp. 151–161, Sep. 2019, doi: 10.1016/j.neucom.2019.04.044.

[28] Y. Su, Yuzhong Huang, and C.-C. Jay Kuo, "Dependent bidirectional RNN with extended-long short-term memory," *ICLR 2018 Conference Blind Submission*, no. 1990, pp. 1–14, 2018.

[29] T. Sheeba and R. Krishnan, "Extracting and updating of learner interest using WordNET in an e-Learning system," *International Journal of Engineering Research and Technology (IJERT)*, vol. 5, no. 5, 2016.

[30] K. Jayakodi, M. Bandara, I. Perera, and D. Meedeniya, "WordNet and cosine similarity based classifier of exam questions using bloom's taxonomy," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 11, no. 4, Apr. 2016, doi: 10.3991/ijet.v11i04.5654.

[31] B. S. Bloom, *Taxonomy of educational objectives: the classification of educational goals*. D. McKay, 1956.

[32] X. Zhu, X. Yang, Y. Huang, Q. Guo, and B. Zhang, "Measuring similarity and relatedness using multiple semantic relations in WordNet," *Knowledge and Information Systems*, vol. 62, no. 4, pp. 1539–1569, Apr. 2020, doi: 10.1007/s10115-019-01387-6.

[33] A. Tversky, "Features of similarity," *Psychological Review*, vol. 84, no. 4, pp. 327–352, 1977, doi: 10.1037/0033-295X.84.4.327.

[34] M. A. Tayal, M. Raghuwanshi, and L. Malik, "Word net based method for determining semantic sentence similarity through various word senses," in *Proceedings of the 11th International Conference on Natural Language Processing*, 2014, pp. 139–145.

[35] Z. Wu and M. Palmer, "Verbs semantics and lexical selection," in *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, 1994, pp. 133–138, doi: 10.3115/981732.981751.

[36] C. Anuradha, T. Velmurugan, and R. Anandavally, "Clustering algorithms in educational data mining : a review," *International Journal of Power Control and Computation(IJPCSC)*, vol. 7, no. 1, pp. 47–52, 2015.

[37] K. Chitra and D. Maheswari, "A comparative study of various clustering algorithms in data mining," *International Journal of Computer Science and Mobile Computing*, vol. 6, no. 8, pp. 109–115, 2017.

[38] D. Xu and Y. Tian, "A comprehensive survey of clustering algorithms," *Annals of Data Science*, vol. 2, no. 2, pp. 165–193, Jun. 2015, doi: 10.1007/s40745-015-0040-1.

[39] N. Sharma, A. Bajpai, and R. Litoriya, "Comparison the various clustering algorithms of weka tools," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 5, pp. 73–80, 2012.

[40] R. Agrawal and J. Agrawal, "Analysis of clustering algorithm of weka tool on air pollution dataset," *International Journal of Computer Applications*, vol. 168, no. 13, pp. 1–5, Jun. 2017, doi: 10.5120/ijca2017914522.

[41] A. M. Rukmi and I. M. Iqbal, "Using k-means++ algorithm for researchers clustering," in *AIP Conference Proceedings*, 2017, vol. 1867, doi: 10.1063/1.4994455.

[42] B. Aubaidan, M. Mohd, and M. Albared, "Comparative study of k-means and k-means++ clustering algorithms on crime domain," *Journal of Computer Science*, vol. 10, no. 7, pp. 1197–1206, Jul. 2014, doi: 10.3844/jcssp.2014.1197.1206.

[43] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowledge-Based Systems*, vol. 46, pp. 109–132, Jul. 2013, doi: 10.1016/j.knosys.2013.03.012.

[44] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Transactions on Information Systems*, vol. 22, no. 1, pp. 5–53, Jan. 2004, doi: 10.1145/963770.963772.

BIOGRAPHIES OF AUTHORS

Kamilia Hosny ⓘ 🔗 SC Ⓟ PreMaster degree at Faculty of Computer and Information, works as reasearcher and developer in United Europe projects such as "Energy-Aware Smart Building", "CALAMARI" and "Semantic Advetrisment Platform". Have many research publications in semantic technology, RFID, SOA, and semantic reasoning. She can be contacted at email: kHosny@ansr.cu.edu.eg.

Abeer Elkorany ⓘ 🔗 SC Ⓟ received Ph.D. in Electronics and Communications Engineering, February 2002 from Cairo University, Faculty of Engineering. With more than 55 publications international journals and conferences, which qualified her to obtain the scientific publication award many times from the Cairo University. Research interest of Prof. Abeer includes: quality assurance and knowledge-based system measurement, semantic network, ontology development, knowledge management, social network analysis, and recommendation systems. Prof. Abeer had supervised several international and national projects. She can be contacted at email: a.korani@fci-cu.edu.eg.