

An Analysis of Data Mining Approaches For Learning Pattern

Narinder Kumar¹, Anshu sharma²

¹M.tech Student, Dept. of CSE, CT Institute of Technology & Research, Jalandhar, India

²Assistant Professor, Dept. of CSE, CT Institute of Technology & Research, Jalandhar, India

Abstract - There are so many techniques with data mining, the major step to creating a good model is to determine what type of technique to use. The system provides a rich set of powerful machine learning algorithms for data mining task, few of task are not found in popular data mining systems. These include basic statistics and visualization tools, at same time tools for pre-processing, classification, and clustering, all available through an easy to use graphical user interface. It uses the GNU General Public License (GPL). Clustering is the unsupervised classification of patterns (data item, attribute vectors, or observations) into groups (clusters). K-Means clustering generates a specific number of disjoint, flat (non-hierarchical) clusters. It is well suited to generating global cluster.

Keywords: Learning Style; classification; data mining; integrated model

1. INTRODUCTION

Data mining is a process by which we can able to get only essential details out of large volume of data. Data mining is done by following steps: selection, pre-processing, transformation, data mining. In this paper we have discuss various learning techniques by which apply one can able to faster and understand things in a better manner. There are number of learning styles which are as follows: chalk and board, audio, listener, charts, maps, presentations, tutorial, education set. Along with various learning techniques there are number factors which affect the learning of the student. The term of learning style refers to aspirate approach to learning based on strength, weakness, and preference. If there is proper interaction between student and teacher

then a teacher must proper understand that how she will make things clear to students. The broad use of this term is generally understood to mean an individual's preferred method to learning and gaining knowledge. Moreover, it can include personalized emotionality, physical needs and sociological needs. Everyone has a different learning style and prefers different kind of learning content. Understanding their learning style can help students to develop their learning ability as well. Therefore, many researchers have been conducted to discern useful approach for understanding and describing different learning styles, or model, are to classify them into different groups based on what the model measures or describes.

Education play vital role in every aspect of the life. Every human being has a different style of the learning. It is up to the student which method he/she has been learnt, In this paper, we work on the learning style so that there will be improvement on the learning method. Here we have used two concepts i.e. clustering and classification. Clustering is process of dividing things into some specified groups without any advanced knowledge and Classification is basically a supervised data which is imposed on clustering.

2. RELATED WORKS

a. Learning Domain

There are three main domain of learning and all teachers should know about them and use them to build lessons. These domains are thinking, emotion, feeling, and physical / kinesthetic. Each other domain on this page has on similar with it. Taxon is simple a word for a classification. All of the taxonomies down are organized so that they proceed from

the simplest level to more complex levels. It is interesting to note that even though the taxon correlated with cognition is commonly referred to as Bloom's Taxonomy. It is based on the 1956 work, the Handbook I-Cognitive Domain; behavioral objectives were separated into subsets. These subsets were arranged into taxonomy and listed according to the cognitive difficulty – simpler to more complex forms. The two highest forms of cognitive have been reversed. In the older version the listing from simple to more complex functions was ordered as knowledge, realization, application, analysis, combination, and education. In the newer version the steps changes to verb and are arranged as knowing, understanding, applying, analyzing, evaluating, and last, creating. Secondly, the affective domain, According to Krathwoh the affective domain can also be divided into a ranking. This area is concerned with feelings or emotions. There are some categories which are receiving the phenomena, responding to phenomena, valuing, organization and characterization. Finally, the psychomotor domain is those specific to discrete physical functions, reflex actions and interpretive movements. There are some categories, which consist of perception, set, guided response, mechanism, complex over response, adaptation and origination.

b. Learning Style

The term learning style refers to aspirate approach to learning based on strength, weakness, and preference. Many people admit that every person prefers different learning styles and methods. Learning style group common ways that people learn. Some people may find that they have a dominant style of learning, so far less use of the other styles. Others can be finding that they use different styles in different circumstances. There is no right mix, neither are your styles fixed. You can develop skill in less dominant styles, as well as further develop styles that you earlier use well. The broad use of this term is understood to mean an individual's preferred method to learning knowledge and gaining knowledge. Moreover, it can include personalized

emotional, physical needs and sociological needs. Everyone has a different learning style and prefers different kind of learning content. Understanding their learning style can help students to develop their learning ability as well. Therefore, many researchers have been conducted to identify learning styles, the ways of gathering information and how to enhance learning performance. There are several learning style theories such as the Felder- Silverman Index of Learning Styles [4].

Bhattacharyya and Shariff [5] reported that learners prefer a variety of learning styles in order to promote their creativity

and critical thinking. VAK learning style was applied to identify students' preference. The results indicated that males

and females accepted different kinds of leaning styles. Thus, learning action should be designed and supported among both male and female students in order to enhance their creativity, knowledge and innovation. However, this study found that individual learning styles can change with different types of activities. According to Scott et al. [6].with the learner centered approach student, skills and preferences are important factors in personalizing the teaching. Therefore, agile methodology is applied to record and analyze student's activities in software engineering course. A data mining technique such as association rule is used to extract student's knowledge. Also, this methodology can discover the relationship between student's performance and their learning styles according to the Felder-Silverman learning styles model. It was found that applying a learning styles model to understand students' preferences should be beneficial as the first step of learning activities. If the learning style of a student is accurately investigated; this can help the instructor to prepare suitable

course materials which appropriately match each student's learning style.

c. Learning Style Classification

There are two types of learning style: Questionnaire and computer model prediction. The questionnaire approach needs the learner to fill in the information regarding the available model. For example, the Felder and Silverman [4] and Klačnjak-Milićević [7] learning models provide the questionnaire to the user and calculate an appropriate score based on the model. Then, the learning style for each individual learner is derived. It can be used at further stages such as providing personalized course materials. The other approach uses computational techniques. The learning style prediction model is created based on available information. For instance, Yang and Wu [8] focused on an attribute-based learning system using ant colony optimization. Some research has implemented combined classification techniques such as k nearest neighbor and genetic algorithm [9].

According to David Kolb's, a professor at Case Western Reserve University, developed a theory of learning and suggested four styles of learning are based on the basic thinking (reasoning) and acting. Learning are being categorized depending on which phase is influential during learning process like sensing and observation, observation and thinking, action and thinking, action and perception. Adaptive learning is a methodology that adapts learning based on students' learning styles by integration of concepts and theories of education and information technology such as monitoring student activities, interpreting results and understanding students' requirements. Consequently, students can learn more productively and have improved understanding of the knowledge domain [10], [11]. In addition, there is a research focusing on learning style using the Felder-Silverman learning style model. The students'

behavior was captured during their online course participation and used for learning style identification [12].

d. Data Mining and Clustering and Classification Techniques

According to Ed Colet, two common data mining techniques for searching the hidden patterns in the data that are clustering and classification analyses. Whereas the classification and clustering are often mentioned in the same wind, they are different analytical approaches. In this column, In this work to explain similarities and differences between these related, but distinct approaches. The problem context can be interpreted data by mining techniques. The wide range of techniques such as classification, prediction, association and detection can be used. Wu et al. [13] have shown that some commonly used algorithms in data mining that are k-mean, SVM, Apriori, and Page Rank including the Naïve Bayes, that is the Classification techniques. Also have been incorporated into learning style prediction. In this paper, simple K-mean clustering algorithm and Make Density Base clustering algorithm used and the Naive Bayes is classification techniques that are apply on the clustering algorithm and to show the better result. In this paper, the Make Density based clustering algorithm is used. It is play a vital role in finding nonlinear shapes of structure. It can even find a cluster completely surrounded by (but not connected to) a different cluster.

Steps of k-mean algorithm: [12]

1. Choose randomly k data objects from given dataset which works as centroids for k clusters initially.
2. Compute distance of each data object from k centroids and then allocate each data object to the closest cluster with minimum centroid distance.
3. Compute new centroid for each cluster by taking mean of the all data objects belonging to particular

cluster. Calculate the total mean-square quantization error function. If error function reduces from previous one than these centroids will work as new centroids.

4. Repeat step 2 and 3 until error function get constant.

3. PROPOSED METHODOLOGY

In this study, we firstly collect the suggestion of the student by taking the survey with the help of college administrator. By analysis all of this, we built up 15 categories on the basis of their pattern. After collection of dataset, preprocessing and data cluster was performed in which clustering algorithms used. There are following steps shows the pseudo codes for the proposed framework.

1. The very first step is preprocessing of data. In this step we choose the attributes which we used for applying the algorithm. In this we cleansed the data into proper format.
2. Second step of proposed technique is to apply k-mean algorithm on the unsupervised data.
3. Now apply the density base algorithm on the unsupervised data. .
4. Now we have to impose the naïve bayes net classification on the results of visualization logs of distance formulas.

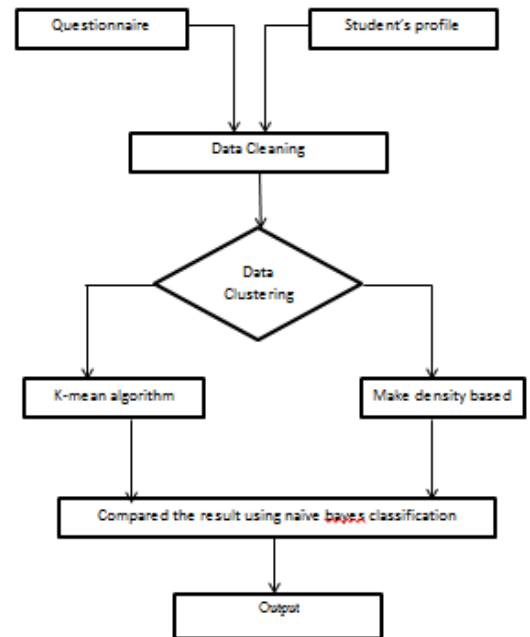


Fig.1: Proposed Methodology

4. RESULT AND DISCUSSION

From the Table 1, it can be inferred that while comparing the k-mean algorithm and make density based algorithm, Paireekreng and Prexawanprasut (2015) used the k-mean algorithm, however this is very sensitive towards centroid selection. So, make density based was used to data classification. Correctly classification instances were higher in make density based (98.11%) than k-mean (97.72%). Similarly the average magnitude of error measure by mean absolute error (MAE) was higher in k-mean (0.0402) as compared to farthest first. The accuracy of farthest first measure was again calculated while using the formula of root mean square error (RMSE), while inferred that RMSE of (.1414) as compared to k-mean (.1449)

From the above discussion it can be concluded that make density based method is accurate and measure in data classification. Next step is to pre-process data against null values followed by normalization and prioritize dataset attribute according to expert knowledge. [17]

Table 1: Comparison of results by k mean and make density based method.

Particulars	K-mean algorithm	Make density based
Correctly classified instances	97.72%	98.11%
Incorrectly classified Instances	2.279%	1.883%
Mean absolute error	0.0402	0.0397
Root mean squared error	0.1449	0.1414
Relative absolute error	11.89%	11.48%
Root relative squared error	35.27%	34.01%

From the above discussion, the below two figures depicts the efficiency of the k-mean and make density based algorithm and the result depicted in the table 1.

```

Correctly Classified Instances      986      97.7205 %
Incorrectly Classified Instances    23      2.2795 %
Kappa statistic                    0.9319
Mean absolute error                0.0402
Root mean squared error            0.1449
Relative absolute error            11.8902 %
Root relative squared error        35.2763 %
Total Number of Instances          1009

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area
      0.935   0.011   0.958     0.935   0.946     0.988
      0.989   0.065   0.982     0.989   0.986     0.988
Weighted Avg.   0.977   0.053   0.977     0.977   0.977     0.988

=== Confusion Matrix ===

  a  b  <-- classified as
203 14 | a = cluster0
  9 783 | b = cluster1
    
```

Fig. 2: Results of k-mean clustering algorithm

```

Correctly Classified Instances      990      98.1169 %
Incorrectly Classified Instances    19      1.8831 %
Kappa statistic                    0.9447
Mean absolute error                0.0397
Root mean squared error            0.1414
Relative absolute error            11.4814 %
Root relative squared error        34.015 %
Total Number of Instances          1009
    
```

```

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC A
      0.938   0.006   0.977     0.938   0.957     0.9
      0.994   0.063   0.982     0.994   0.988     0.9
Weighted Avg.   0.981   0.05   0.981     0.981   0.981     0.9

=== Confusion Matrix ===

  a  b  <-- classified as
210 14 | a = cluster0
  5 780 | b = cluster1
    
```

Fig. 3: Results of make density based clustering algorithm

5. CONCLUSION AND FUTURE SCOPE

Determining the most suitable learning style for learners is important for any educational system. Data mining techniques which helps the improving the learning style of the students. Different learners should be served different course materials and content in order to encourage the most efficient learning for them. Some approaches may need the user to complete a questionnaire, while others may need to classify the learning style using learning information from the user. Using the K-Mean clustering to classify the learning style of the students. The experiment result shows that with the help of naïve bayes classification on make density clustering gives better result than k-mean clustering.

References

[1] Hawk, Thomas F., and Amit J. Shah., "Using learning style instruments to enhance student learning." Decision Sciences Journal of Innovative Education vol. 5, pp. 1-19, 2007.

[2] Bozhilov D., Stefanov K., and Stoyanov S., "Effect of adaptive learning style scenarios on learning achievements." International Journal of Continuing Engineering Education and Life Long Learning (IJCELL), vol. 19, pp.381-398, 2009.

[3] Klačnja-Miličević A., Vesin b., and Budimac Z., "E-Learning personalization

based on hybrid recommendation strategy and learning style identification." *Computers & Education*, vol. 56, pp. 885-899, 2011.

[4] Bhattacharyya E., and Shariff A.B.M.S., "Learning Style and its Impact in Higher Education and Human Capital Needs." *Procedia-Social and Behavioral Sciences*, vol.123, pp. 485-494, 2014.

[5] Paireekreng, W. and Takorn P., "An integrated model for learning style classification in university students using data mining techniques." *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2015 12th International Conference on.IEEE*, 2015.

[6] Wu, X. et al., "Top 10 Algorithms in Data Mining," *Knowledge and information System*, pp.1-37, 2008.

[7] Pairekreg W.and Wong K., "Mobile Content Personalization using Intelligent User Approach," in *Proceeding of Third international Conference on Knowledge Discovery and Data Mining*, pp 241-244,2010.

[8] Grigoriadou M. and Papanikolaou K., "Authoring Personalised Interactive Content," presented at the *First International Workshop on Semantic Media Adaptation and Personalization (SMAP '06)*,2006.

[9] Graf S.and Kinshuk., "Identifying Learning Styles in Learning Management System by using Indications from Students'behaviour," in *IEEE International Conference on Advanced Learning Technologies (ICALT'08)*,2008.

[10] Cufoglu A., et al., "A Comparative Study of Selected Classifiers with Classification Accuracy in User Profiling," presented at the *2009 World Congress on Computer Science and Information Engineering Los Angeles, USA 2009*.