International Journal of Engineering & Technology, 7 (3.15) (2018) 63-67

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



Evaluation of Clustering Methods for Student Learning Style Based Neuro Linguistic Programming

Marina Yusoff[*,1] , Muhammad Najib Bin Fathi[2]

1Advanced Analytic Engineering Center (AAEC), Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia.
2Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia.
**Corresponding author, e-mail: marinay@tmsk.uitm.edu.my*

Abstract

Students' performance is a key point to get a better first impression during a job interview with an employer. However, there are several factors, which affect students' performances during their study. One of them is their learning style, which is under Neurolinguistic Programming (NLP) approach. Learning style is divided into a few behavioral categories, Visual, Auditory and Kinesthetics (VAK). This paper addresses the evaluation of clustering methods for the identification of learning style based on system preferences. It starts with the distribution of questionnaires to acquire the information on the VAK for each student. About 167 respondents in the Faculty of Computer and Mathematical Science are collected. It is then pre- processed to prepare the data for clustering method evaluations. Three clustering methods; Simple K-Mean, Hierarchical and Density-Based Spatial Clustering of Applications with Noise are evaluated. The findings show that Simple K-Mean offers the most accurate prediction. Upon completion, by using the dataset, Simple K-Means technique estimated four clusters that yield the highest accuracy of 74.85 % compared to Hierarchical Clustering, which estimated four clusters and Density- Based Spatial Clustering of Applications with Noise which estimated three clusters with 52.69% and 61.68 % respectively. The clustering method demonstrates the capability of categorizing the learning style of students based on three categories; visual, auditory and kinesthetic. This outcome would be beneficial to lecturers or teachers in university and school with an automatically clustering the students' learning style and would assist them in teaching and learning, respectively.

Keywords: *Clustering; Hierarchical Clustering; K-Means; Learning Style; Neuro Linguistic Programming*

1. Introduction

This The performance of a student is important to result in a good student in terms of academics or in other activity involvement [1]. Their performance can be affected by a few variables, external or internal depending on the student's capabilities of completing their tasks, learning style, intentions, environment, communication skills and societies. Everyone varies from the way they learn in various kinds of environment. Some of them prefer to listen to the lectures, some of them prefer to write all the notes that have been given in the slides, while some of them even prefers a lot of images during their presentation. This learning style has long been researched by other professionals, as the studies started from a psychological study on individual differences in the 1960s and the 1970s range [2]. Various tools and theory had been developed by the researchers to understand how an individual learns.

As the technology is advancing and growing from time to time, students are more exposed to different method of learning. A good performance depends on how the students managed their studies, their approach during their studies. Every learner has their own different approach of learning according to some factors such as learning style [3]. Student in university uses different approach most of the time to study [4]. A suitable learning style is important to ensure students can achieve a better result. In universities, there is still a lack of measurements to see the effectiveness of the system preference in learning style. It is very hard to identify which system preferences that influence the student performances. Learning styles is the way in which individuals perceive and process information in learning situations [5]. To the best of our knowledge, there is a need to develop a system to assist the identification of student system preferences related to learning style.

Every person uses multiple senses as such visual, taste, sight, hearing and touch from time to time. For instance, some people tend to remember things easily by remembering the voice of the person, some can recognize their face but unable to recall their names and some people are able to remember the person by the things they had done together. These senses are very useful during learning. In teaching and learning, listening to the lectures is easier for some students to memorize the points, some of them needs to do something while lecture and some of them system preferences, people may exhibit a certain behavior or unique characteristic for themselves. This kind of system preference often explains in Neuro Linguistic Programming (NLP) communication model. Many people use this idea to rapport, improve the communication skills, understanding other people behavior and learning style identifications. VAK is an acronym for Visual, Auditory and Kinesthetic. Learning styles are methods for system preferences established in NLP basic communicator because different people learn differently [6]. Therefore, research on learning style based NLP is a prom ising idea to assist lecturers or teachers to better understanding student behavior and match to a suitable learning style. This paper addresses the clustering methods, evaluation to categorize students learning style based on VAK system preferences.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2. Neuro Linguistic Programming System Preferences

As for visual, the person prefers the depiction of diagrams, graphs, or charts during their study. Symbols also helped them to understand better rather than audio or physical activities. This preference is unique as they prefer to use all the images and symbols, which could have been depicted, by using words or sentences. In a visual sense, people who prefer this type of learning style learn best by reading and watching, as they must see it for them to understand it [7]. They prefer to visual things in every kind of things they try to understand as if they have a movie camera playing in their mind while they think. By using this "movie camera", they always recall things easily from what they had captured.

The people with this type of learning style prefer information, which is heard or spoken rather than seeing and visualizing. They learn best from verbal discussions during classes instead of watching the slides. They must hear first to learn and understand what was being explained [7]. A person with this style of learning prefers to talk loud and even talking to oneself to recall things accurately. Unlike visual style of learning, they prefer to speak first rather than organizing and sorting their ideas beforehand. They are also a good listener, but this may become a disadvantage to them as they are easily distracted by sounds. Kinesthetic people are the one who prefers physical activities involved. They are poor listeners as they prefer to learn by doing practically and have an outgoing personality. These types of people tend to learn best by doing hands-on activities [7]. By doing so, they can remember most of the things they had done, unlike visual and auditory person. However, one of the disadvantages of these types of learners are they are easily distracted or having a hard time paying attention as they prefer to do things while learning. They connect to the reality, as they need experiences, practices or simulations. These also include videos of how to do things, or a demonstration by an expert or even case studies.

Multimodality is a term used when a person prefers a few learning styles in any of the combinations such as visual and auditory, or kinesthetic and auditory, writing and writing and kinesthetic or such. The person who has multimodality easily adapts with surrounding very well rather than someone with a singular modality [8]. However, the person with multimodality is uncommon as they need to be able to multitask while doing something to be able to learn faster. This type of modalities is divided into two types. First is someone who can do things with all their learning styles simultaneously, or second, the one who can change their learning style one after another.

3. Clustering Implementation

3.1 Data Acquisition and Pre-processing

A survey of VAK learning style is distributed to students in the Faculty of Computer and Mathematical Sciences (FSKM), University Teknologi MARA. About 167 questionnaires were answered. It is then pre-processed to prepare the data for clustering method evaluations. VAK scoring is calculated for each respondent. VAK score is transformed, normalized and organized in CSV format. A normalization technique as such z - score scaling formula is used to reduce the range of the data and the difference of the data will not be too large. For categorical data such as gender, a common technique, which converts the data into a binomial form, is also used. The example of pre-processed data is shown in Figure 1.

Gender	Semester	CGPA	V	A	K	Subject	Course	Style	Total
1	6	3.65	4	1	7	ITT575	CS245	K	12
1	6	2.68	2	6	5	ITT575	CS245	A	13
1	6	2.64	6	1	5	ITT575	CS245	V	12
1	6	3.31	7	1	4	ITT575	CS245	V	12
0	6	2.56	7	1	4	ITT575	CS245	V	12
0	6	2.97	3	7	2	ITT575	CS245	A	12
1	6	2.79	5	1	6	ITT575	CS245	K	12
1	6	3.75	7	1	4	ITT575	CS245	V	12
0	6	2.71	5	4	3	ITT575	CS245	V	12
1	6	2.77	4	5	3	ITT575	CS245	A	12
1	6	3.22	6	1	5	ITT575	CS245	V	12
0	6	3.35	9	3	0	ITT575	CS245	V	12

Fig. 1: Sample of pre-processed data

3.2 Clustering Methods

Clustering methods have been highlighted in many research and applied in many domains [9-13]. In clustering the idea is not to predict the target class as like classification, it is more ever trying to group the similar kind of things by considering the most satisfied conditions all the items in the same group should be similar and no two different group items should not be similar [14]. To group the similar kind of items in clustering, different similarity measures should be considered. This paper highlights the evaluation of the three most common techniques in clustering; Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Hierarchical Clustering and Simple K-Means. These clustering methods were developed in Phyton and the plotting of the graph were using Matplotlib library. Figure 2 is a flowchart to demonstrate the employment of clustering technique steps. Firstly, the dataset is loaded into the system. Then, a clustering method is chosen, whether DBSCAN, Simple K-Mean, or Hierarchical clustering technique is used to cluster the data. With the techniques chosen, the number of clusters is then determined manually, or by using Elbow Method [15]. The distance of each instance coordinates is calculated with the centroids. This step determines the cluster of each centroid. Finally, the clusters are plotted into a graph for visualization.

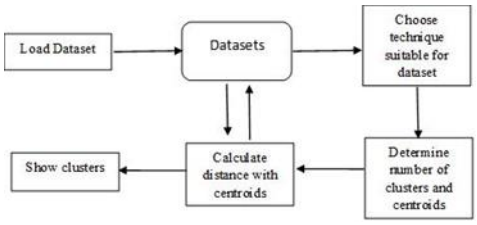


Fig. 2: Flowchart of Clustering Employment Steps

4. Analysis of Results

The experiments were performed measure the accuracy of the three methods; DBSCAN, Simple K-Mean and Hierarchical. Table 1 shows the number of clusters identified for the three techniques. For Simple K-Mean clustering and Hierarchical clustering, the number of clusters predicted were both 4 clusters; V, A, K and Multimodalities (M). M is a mix of more than two of learning style meanwhile DBSCAN predicted it to have 3 clusters; V, A, K instead. This is due to the method of calculating the number of clusters are different for each type of clustering methods.

Table 1: Number of clusters for each technique

Clustering Method	Number of Clusters	Learning Style
DBSCAN	3	V, A, K
Simple K-Means	4	V, A, K, M
Hierarchical	4	V, A, K. M

4.1 Results of Simple K- Means

Figure 3 shows the data cluster for the technique Simple K- Means which consists of 4 clusters. The data are well represented in Table 3 shows the exact value of each clustered data. Figure 3 is the clustered data in the Visual against Cumulative Grade Point Average (CGPA), Auditory against Cumulative Grade Point Average, Kinesthetic against Cumulative Grade Point Average and the last one is showing the clustered data and identification of Visual for cluster 1, Auditory for cluster 2, Kinesthetic for cluster 3 and Multimodal for cluster 4 for the students in FSKM.

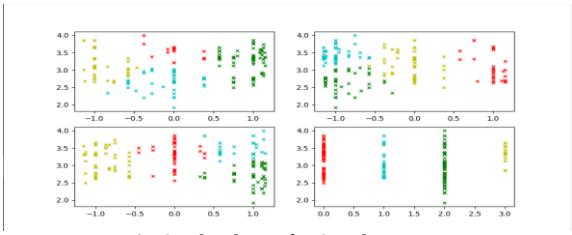


Fig. 3: The cluster for SimpleKMeans

In Table 3, the confusion matrix shows the correctly clustered data with a total of 167 data. The data are accurately clustered which the cluster 0, Cluster 1, Cluster 2 and Cluster 3 represent the learning style of M, A, V and K respectively. As demonstrated in Table 2, there are 46 instances which are incorrectly clustered. 34 of them were misclustered under Cluster 0, and 12 of them were misclustered under Cluster 1.

Table 2: Confusion matrix for SimpleKMeans

System		Cluster No		
Preferences	0	1	2	3
K	26	0	0	36
A	0	29	0	0
V	8	0	49	0
M	11	12	0	0

4.3 Result of Hierarchical Clustering

Figure 4 shows the data cluster for the technique Hierarchical Clustering. The data is well represented in Table 4 which shows the exact value of each clustered data. Figure 4 is the clustered data for the Visual against Cumulative Grade Point Average, Auditory against Cumulative Grade Point Average, Kinesthetic against Cumulative Grade Point Average and the last one is showing the clustered data and identification of Visual for cluster 0, Auditory for cluster 1, Kinesthetic for cluster 2 and Multimodal for cluster 3 for the students in FSKM by using the Hierarchical Clustering technique.

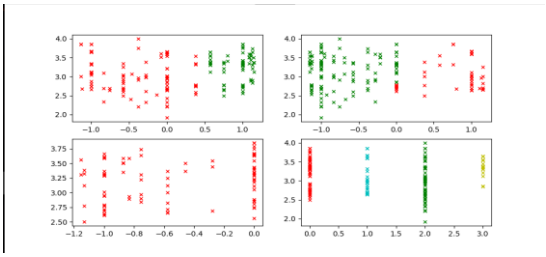


Fig. 4: Data cluster of Hierarchical Clustering

From the Table 3, it shows there are 79 instances, which is incorrectly clustered. There are 75 misclustered data in Cluster 0 and 4 misclustered data in Cluster 2. The performance of Hierarchical clustering is not as good as Simple K-Means technique.

Table 3: Confusion matrix for Hierarchical cluster

System		Cluster No		
Preferences	0	1	2	3
K	62	0	0	0
A	13	20	14	0
V	57	0	0	0
M	13	0	0	6

In Figure 4, the data is clustered into 3 clusters. This data is well represented in the Table 4. Many noisy data are represented with the black spots.

4.3 Result of DBSCAN

Figure 5 shows the data cluster for the technique DBSCAN which consists of 4 clusters. The data is well represented in Table 5 which shows the exact value of each clustered data. Figure 5 is the clustered data for the Visual against Cumulative Grade Point Average, Auditory against Cumulative Grade Point Average, Kinesthetic against Cumulative Grade Point Average and the last one is showing the clustered data and identification of Visual for cluster 1, Auditory for cluster 2, Kinesthetic for cluster 3 and Multimodal for cluster 4 for the students in FSKM by using DBSCAN technique.

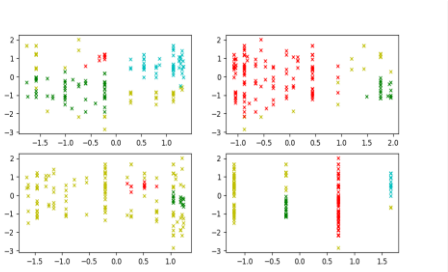


Fig. 5: Data cluster of DBSCAN

Table 4 shows there are 70 instances which is incorrectly clustered by using DBSCAN technique. There are 38 misclustered data in Cluster 0, 2 misclustered data in Cluster 1 and 24 misclustered data in Cluster 2. The performance of DBSCAN is almost likely to hierarchical clustering.

Table 4:.Confusion matrix for DBSCAN

System	Cluster No		
Preferences	0	1	2
K	44	0	18
A	0	0	29
V	27	30	0
M	14	5	6

Table 5 demonstrates the performance of the three clustering methods in terms of accuracy of correctness in clustering. Simple K- Means has the highest value which is higher than the accuracy for DBSCAN and Hierarchical clustering. Simple K- Means, which has 53.29% accuracy is significantly higher than DBSCAN and Hierarchical clustering with 61.68% and 52.69% respectively. This means that the SimpleKMeans technique has good capability to cluster the dataset according to its correctness in clustering the learning styles of the students.

Table 6: Performances of Simple K- Means, DBSCAN and Hierarchical clustering

Clustering Method	Accuracy
DBSCAN	61.68 %
SimpleKMeans	74.85 %
Hierarchical	52.69 %

4.4 Further Results of Simple K-Means based CGPA

From the results obtained from the 3 techniques of clustering; Hierarchical Clustering, DBSCAN and SimpleKMeans, Simple K-Means provides the highest accuracy of 74.85%. Figure 6 demonstrates the clustered data shows a significant clustered of data within the range of -1.0 until -0.5. This clustered data means the students of FSKM's learning style is not preferably an Auditory style of learning since the clustered data shows students with higher CGPA has little value in Auditory learning style.

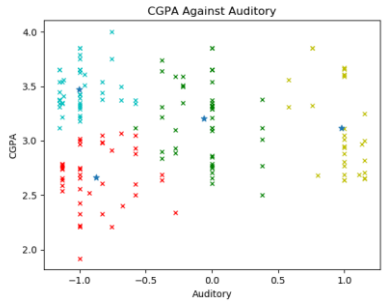


Fig. 6: CGPA against Auditory graph

Figure 7 shows that the clustered data displays a very dense clustered of data within the range of 0.5 until 1.0. This clustered data shows the students of FSKM's learning style has a high value of visual learning style. From the clustered data, we can predict the learning style of students are more to visual learning style. This is a very common for university students since they are more likely to be exposed to many visual presentations during lectures. It is one of the effective methods for students to adapt to the learning environment of university's life.

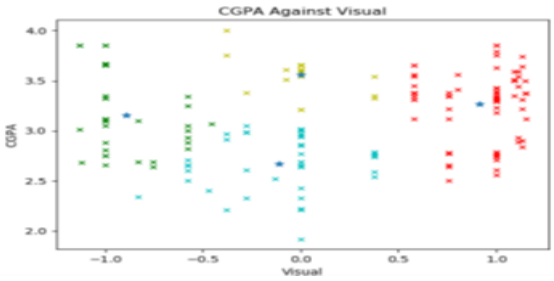


Fig. 7: CGPA against Visual graph

Figure 8 shows the clustered data shows a significant value of clustered data with high density of cluster from the range of 0.5 to 1.0. This shows the students has a mixed of Kinesthetic learning style with the other 2 learning style. With the mixed of 2 or more learning style, the students are well adapted to their study environment and can utilize all the learning style to achieve a higher CGPA in their study.

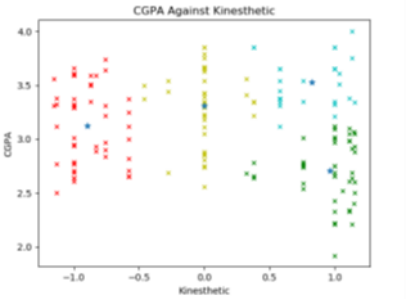


Fig. 8: CGPA against Kinesthetic graph

5. Conclusion

This paper addresses the evaluation of clustering methods for the identification of learning style based on system preferences. This project starts with the distribution of questionnaires to acquire the information on the VAK for each student. Three clustering methods were compared namely; DBSCAN, Hierarchical Clustering and Simple K-Means. The use of clustering methods indicates the capability of clustering learning style can assist educators to determine the learning style of their students in the early semester. Hence, it will help in improving the performance of students. It is evident from the details analysis of CGPA with the association of learning styles, the findings can be seen clearly as an aid to educators in adapting an appropriate learning style in school or university. In addition, the developed system engine based on Simple K-Means can be a suitable clustering method that can be automatically predicted the students' learning styles.

Acknowledgement

The authors express a deep appreciation to the Institute of Research and Innovation, Universiti Teknologi MARA a for the grant of 600-IRMI/DANA 5/3/ARAS (0020/2016) and the Information System Department, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Malaysia for providing essential support and knowledge for the work.

References

[1] Ali, S., Haider, Z., Munir, F., Khan, H., & Ahmed, A. (2013). Factors Contributing to the Students Academic Performance: A Case Study of Islamia University Sub-Campus. America Journal of Educational Research, 1(8), 283-289.

[2] Curry, L. (1987). Integrating concepts of cognitive or learning style: A review with attention to psychometric standards. Ottawa, ON: Canadian College of Health Service Executives.

[3] Reid, J. M. (1987). The Learning Style Preferences of ESL Students. TESOL Quarterly, 21(1),87

[4] Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1-13.

[5] Brown, H. D. (2000). Principles of language teaching and learning, (4th ed.). White Plains, NY: Longman.

[6] O'Connor, J., & Seymour, J. (2011). *Introducing NLP: Psychological skills for understanding and influencing people*. Conari Press.

[7] Kanar, C. C. (1995). The confident student. Boston: Houghton Mifflin Company.

[8] Krishna, T. S., Babu, A. Y., & Kumar, R. K. (2018). Determination of Optimal Clusters for a Non-hierarchical Clustering Paradigm K-Means Algorithm. In *Proceedings of International Conference on Computational Intelligence and Data Engineering* (pp. 301-316). Springer, Singapore.

[9] Ooi, T. H., Ngah, U. K., Khalid, N. E. A., & Venkatachalam, P. A. (2000). Mammographic Calcification Clusters Using The Region

Growing Technique. In *Proceedings of The New Millennium International Conference on Pattern Recognition, Image Processing and Robot Vision* (pp. 157-163).

[10] Seman, A., Bakar, Z. A., & Isa, M. N. (2012). An efficient clustering algorithm for partitioning Y-short tandem repeats data. *BMC research notes*, *5*(1), 557.

[11] Pindah, W., Nordin, S., Seman, A., & Said, M. S. M. (2015, April). Review of dimensionality reduction techniques using clustering algorithm in reconstruction of gene regulatory networks. In *Computer, Communications, and Control Technology (I4CT), 2015 International Conference on* (pp. 172-176). IEEE.

[12] Seman, A., Bakar, Z. A., Sapawi, A. M., & Othman, I. R. (2013). A Medoid-based Method for Clustering Categorical Data. *Journal of Artificial Intelligence*, *6*(4), 257-265.

[13] Khalid, N. E. A., Yusoff, M., Kamaru-Zaman, E. A., & Kamsani, I. I. (2014). Multidimensional Data Medical Dataset Using Interactive Visualization Star Coordinate Technique. *Procedia Computer Science*, *42*, 247-254.

[14] Raja, R., & Tien, N. (2018). Exploring multi-modality tools of Neuro-Linguistic Programming (NLP) to facilitate better learning among Primary School students.

[15] Bajorski, P., & Bajorski, P. (2011). Clustering – Unsupervised Learning. In Statistics for

Imaging, Optics, and Photonics (pp. 297-327): John Wiley & Sons, Inc.