Advances in Economics, Business and Management Research, volume 156
2nd International Scientific and Practical Conference on Digital Economy (ISCDE 2020)

The Students Group Detection Based on the Learning Styles and Clustering Algorithms

Dyulicheva Yu.Yu.*, Kosova Ye.A.

V. I. Vernadsky Crimean Federal University, Simferopol, Russia
**Corresponding author E-mail: dyulicheva@gmail.com*

ABSTRACT
The approach to automatically student groups detection based on Honey and Mumford's questionnaire and index of learning styles questionnaire with the help of clustering methods are proposed in the paper. The methodology of our research consists of the following stages: 1) the evaluation optimal number of clusters and clustering students data based on Honey and Mumford's questionnaire and Ward.D2 method realised in R-library NbClust; 2) the evaluation optimal number of clusters and clustering students data based on index of learning style questionnaire and k-Means method; 3) forming the clusters of students that are in one cluster based on both learning styles and description of the student clusters taking into account similarities on learning style preferences and similar interests in the social network.
Keywords: *learning styles, clustering algorithms, student groups detection, R-library NbClust*

1. INTRODUCTION

The priority problems for all universities are the studying of student community in order to create a favourable atmosphere in the team, the increasing of academic performance, motivation, and involvement of students in the educational process, establishing the correspondence between teaching styles and student learning preferences, the practical skills development and the professional competencies acquiring. The learning styles are often used in students community research. Here just a few studies on the student learning styles preferences in education. For example, at the stage of students preparing for project activities in order to develop research skills based on VARK model, the index of learning styles (ILS), Kolb's learning styles (Pershukova, 2010); in the process of the adaptive learning systems development and e-learning environment creation taking into account the individual characteristics and student needs and learning strategies (Popescu, Badica & Moraret L, 2010; Ali, Eassa & Hamed, 2019); in the research process of dominant styles of student activities and their self-organization level based on situation analysis, planning, self-control, correction and volitional efforts (Epachinceva, Kozlovskaya & Kozlovskij, 2019); in the research process of the influence the learning styles on students' critical thinking (Myers & Dyer, 2006). Thus, the learning styles usage for student community studying allows to develop adaptive systems and teaching methods taking into account the personal preferences of each student and helps students to better understand their advantages and disadvantages in the learning process as well as in acquiring students' collaboration skills.

In addition to a tool based on learning styles, researchers also use methods of cluster analysis, for example, to detect the student groups with high academic performance or to identify the groups of students with low motivation and personal interests in learning (Nagesh & Satyamurty, 2018). Unfortunately, the principles of the automatic formation group of students (teams) for successful cooperation remain little studied. Meanwhile, the teachers often face with a problem of forming groups of students, for example, for foreign language learning (Nhan H & Nhan Th, 2019) or for participation in scientific research, competitions. The question is how to automate the search and formation of student groups for participation in various activities, in which students can best demonstrate their capacities.

The learning styles usage together with clustering methods, as shown in our study, creates an analytical tool for the teacher, which allows us to automatically detect the student groups with common interests and perception of information. The aim of the study is to identify regularities that can be used for the automatic student groups formation participated in various activities. Understanding such regularities allows us to create a favourable atmosphere for the development of the students' creative abilities, cognitive, communication, and professional skills. We used Honey and Mumford's questionnaire, index of learning style questionnaire, and data from open profiles in social network VKontakte (VK) to understand student groups that can collaborate in the learning environment.

So, let's enumerate the research questions.

1. How to detect student groups for effective collaboration?
2. What regularities can be drawn from the learning styles, clustering methods and data from the social network to better understand student groups for collaboration?

2. RELATED WORKS

The usage of clustering methods identifies the regularities for student groups description, the individual learning trajectories creation, the teaching and learning strategies development, focused on homogeneous student groups. For example, the researches intending to clusters detection for describing the satisfaction with presence in the student community (Rosa, Fida & Avallone, 2011), and the clusters for identification of career-minded students (Vashisht & Bhawna, 2019).

Numerous studies demonstrate the relationship or its absence between learning styles and academic performance, gender, demographic and other factors as well as preferred learning styles for teaching methods development. For example, no differences are found between index of learning styles such as sensing, visual, reflective, global and gender (Naik & Girish, 2012). At the same time, it is shown that the preferred learning styles are different for men and women, but if the teacher combines the different types of learning content, the academic performance and learning efficiency will be high for any gender (Suryaman, Chaerul & Rahma, 2019).

The relationship between students' stress tolerance and Felder-Silverman's preferred learning styles are investigated by Dangmei (Dangmei, 2019). At the stage of empirical study, a multiple regression model is constructed to detect that the students with strong visual and active learning styles have the greatest stress resistance. Such researches help to better understand the psychological and cognitive characteristics of the students especially at the student groups forming and creation a favorable atmosphere in the student team and aimed at the personal growth of each student. The students' self-motivation and Kolb's learning styles are studied by (Mobeva, Arnado & Buot, 2020). In particular, no significance between learning styles and students' self-motivation has been found but it is noted that students' skills adapted to learning styles can improve their academic performance. The differences between the preferred learning styles of students from different countries and cultures are discussed in (Kamisli, 2019) and the relationship with academic performance are studied (Nja et al., 2019).

3. METHODS

Our research is based on the results of student clustering with the help of the index of learning styles questionnaire, k-Means, and hierarchical clustering methods (Pasina et al., 2019; Liu et al, 2017; Kyprianidou et al., 2012). We propose to extract student groups that are similar by Honey and Mumford's styles and ILS at the same time. Additionally, we use data from the social network VK and describe clusters with recommendations for student activity in harmony with similar interests. So, the methodology of our research consists of the following stages: 1) the evaluation of the optimal number of clusters and clustering student data based on Honey and Mumford's questionnaire and Ward.D2 method, realized in R; 2) the evaluation of the optimal number of clusters and clustering of student data based on ILS questionnaire and k-Means method, realized in R; 3) forming the clusters of students that are in one cluster based on both learning styles and description of the student clusters taking into account similarities on learning style preferences and similar interests in the social network.

4. EXPERIMENTS AND RESULTS

61 students from the Mathematics and Computer Science Department of V. I. Vernadsky Crimean Federal University were participated in our experiment and filled at the same time two questionnaires based on Honey and Mumford's learning styles and ILS. Descriptive statistics of student data are shown in table 1. As can be seen in Table 1, the students have almost identical attitude to Reflector, Theorist, Pragmatist learning styles and Active, Sensing, Visual, Sequential learning styles by Honey and Mumford and ILS respectively. Let's consider the stages of our research.

1st stage: the optimal number of clusters detection based on Honey-Mumford's questionnaire. For identifying the optimal number of clusters we use NbClust function, realized in R, based on varying of the different combinations of number of clusters, distance measure, and clustering algorithms [19]. So, the usage of function with configuration distance = "euclidean" and method = "ward.D2" in NbClust allows us to detect the optimal number of clusters that is equal 4. The results of optimal parameters detection are shown in Figure 1 (left) and the results of student data clustering and their description are shown in Table 2.

Table 1 Descriptive statistics of student data (N=61)

Learning styles	mean	std	min	25%	50%	75%	max
(a) Honey and Mumford's learning styles							
Activist	9.442623	3.662536	1	7	9	12	17
Reflector	14.786885	3.246099	5	12	16	17	20
Theorist	12.918033	3.536925	5	10	14	15	20
Pragmatist	13.704918	3.051471	8	11	14	16	19
(b) Index of learning styles							
Active	2.688525	2.784128	0	0	3	5	11
Reflective	0.967213	1.974565	0	0	0	1	9
Sensing	3.672131	3.096887	0	1	3	5	11
Intuitive	0.639344	1.30426	0	0	0	0	5
Visual	3.245902	3.535911	0	0	3	5	11
Verbal	0.967213	1.949079	0	0	0	1	9
Sequential	2.934426	3.086902	0	0	3	5	11
Global	0.918033	1.715567	0	0	0	1	7

Table 2 Results of students clustering based on Honey and Mumford's questionnaire (N=61)

	Cluster1	Cluster2	Cluster3	Cluster4
student ids	id69; id11; id63; id6; id94; id76; id60; id74; id23; id29; id42; id53	id85; id58; id91; id75; id21; id54; id47; id48; id71; id10; id73; id33; id52; id15; id87; id62	id37; id34; id8; id4; id7; id70; id12; id72; id88; id26; id79; id43; id78; id9; id20	id3; id90; id36; id93; id82; id31; id56; id68; id45; id55; id92; id24; id13; id51; id81; id77; id22; id40
number of students	12	16	15	18
centres of clusters in format [Activist; Reflector; Theorist; Pragmatist]	[11.0833; 10.5; 9; 11]	[6.8125; 16.0625; 11.75; 11.375]	[6.66667; 17.2; 16; 16.3333333]	[13; 14.5; 14; 15.3888889]
clusters description based on Honey and Mumford's learning styles	no strong preferences of one learning style	more preference to Reflector style	less tendency towards Activist style	no strong preferences of one learning style

2nd stage: the detection of the optimal number of clusters based on index of learning styles questionnaire. For the identifying the optimal number of clusters we use NbClust with options: distance = "euclidean" and method = "kmeans". The results of NbClust function are shown in Figure 2 (right) and results of student data clustering and their description are shown in Table 3. The optimal number of clusters is 7.

3rd stage: forming and description of student clusters based on both learning styles taking into account similar interests in the social network VK. We considered sociability index (SI) (SI=1, if the number of friends is 0-100; SI=2, if the number of friends is 101-200 and SI=3, if the number of friends is more than 200); popularity index (PI) (PI=1, if the number of subscribers is 0-100; PI=2, if the number of subscribers is 101- 200 and PI=3, if the number of subscribers is more then 200); index of interests (II) (II=1, if the number of followers is 0-100; II=2, if the number of followers is 100-200, II=3, if the number of followers is more than 200.

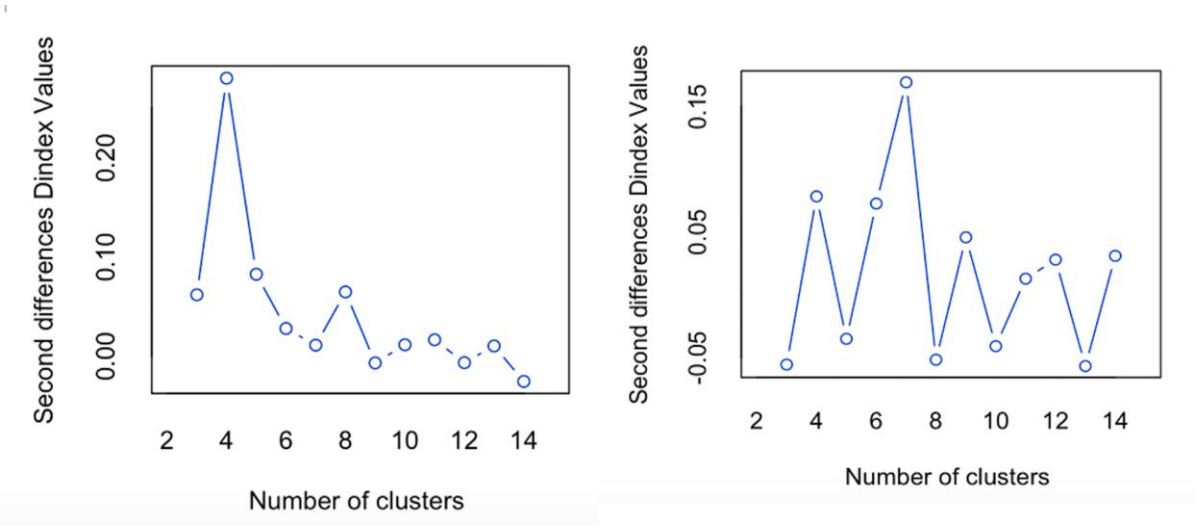

Figure 1 The detection of optimal number of clusters based on Honey and Mumford's questionnaire (left) and index of learning styles questionnaire (right)

Table 3 Results of students clustering based on index of learning styles questionnaire (N=61)

	Cluster1	**Cluster2**	**Cluster3**	**Cluster4**	**Cluster5**	**Cluster6**	**Cluster7**
student ids	id8; id55	id45; id4; id6; id7; id9; id12; id13; id15; id56; id60; id72; id75; id31; id78; id87; id91; id21; id22	id58; id82; id85; id77; id11; id33; id71	id26; id42; id3; id47; id52; id62; id74; id92; id81; id34; id51; id88	id29; id53; id24; id90; id68; id23	id36; id37; id10; id48; id54; id63; id94; id79; id93; id40; id69; id73	id9; id79; id70; id94
number of students	2	18	7	12	6	12	4
centres of clusters in format [Active; Reflective; Sensing; Intuitive; Visual; Verbal; Sequential; Global]	[10; 0; 10; 0; 11; 0; 10; 0]	[2.4444; 0.33333333; 5.666667; 0; 1.5; 0.3889; 1.33333; 1.1111]	[3.7143; 0.1428571; 1; 1.42857; 1.4286; 0.7143; 6.42857143;0]	[2.5; 0.1666667; 4.5; 0; 7.8333; 0; 4.16666667; 0.3333]	[6.3333; 0; 0.333333; 1.33333; 5.6667; 0; 1.33333333; 1.3333]	[0.5; 2.3333333; 0.916667; 1.75; 0.9167; 2.75; 0.66666667; 2]	[0; 5.5; 7; 0; 0; 3.5; 6; 0]
clusters description based on index of learning styles	strong preferences of Active, Sensing, Visual, Sequential styles	moderate preference of Sensing style	moderate preference of Sequential style	strong preferences of Visual style	moderate preference of Active and Visual styles	balanced preference of Reflective, Intuitive, Verbal learning styles	moderate preference of Reflective, Sensing and Sequential style

We chose clusters with the number of students from 3 to 5 and formed clusters for 35 students. Let's describe the results of student clusters creation and description (Honey & Mumford, 1986; Felder, 1996).

Cluster 1 (the students with id45, id13, id56, id31, id22) has the center of cluster [12.8; 12.2; 14; 16] by Honey and Mumford's learning styles and center of cluster [2.6; 0; 5.4; 0; 2; 0.2; 0.8; 2.2] by ILS. The cluster description based on both learning styles: students are oriented on experimentation, problem solving and discussion in groups, too concern with details and prefer concrete and practical information. The cluster description based on open profiles data from VK: students are professionally oriented (mathematics, computer science, programming) (SI=2, PI=3, II=2).

Cluster 2 (the students with id15, id75, id87, id91, id21) has the center of cluster [7.25; 17.75; 13.5; 12.75] by Honey and Mumford's learning styles and center of cluster [1.8; 1.2; 5.8; 0; 1.4; 0.8; 1; 0.8] by ILS. The cluster description based on both learning styles: the students

prefer to use proven methods, well-thought-out strategies, and then act. The cluster description based on open profiles data from VK: the students are inclined to organizational activity (SI=2, PI=2, II=3).
Cluster 3 (the students with id58, id85, id33, id71) has the center of cluster [8.75; 17.5; 8.75; 10.25] by Honey and Mumford's learning styles and center of cluster [4; 0; 0.75; 1.75; 1.5; 1; 6.5; 0] by ILS. The cluster description based on both learning styles: the students are oriented on data preprocessing, implementation of logically sequential steps and revalidation of data. The cluster description based on open profiles data from VK: the students are inclined to creative activity (graphics, drawing, modeling) (SI=2, PI=2, II=1).
Cluster 4 (the students with id26, id34, id88) has the center of cluster [6.(6); 18.(6); 17.(6); 16.(3)] by Honey and Mumford's learning styles and center of cluster [2.(3); 0; 5.(6); 0; 9; 0; 2.(6); 0.(3)] by ILS. The cluster description based on both learning styles: the students are oriented on the experiments, visual presentation of data, search for practical implementation of results, and see the whole thing. The cluster description based on open profiles data from VK is missing because open profiles data is uninformative (SI=1, PI=1, II=1).
Cluster 5 (the students with id3, id92, id81, id51) has the center of cluster [12.5; 15.5; 15.75; 14.5] by Honey and Mumford's learning styles and center of cluster [3.25; 0.25; 4.5; 0; 8.5; 0; 4.25; 0.75] by ILS. The cluster description based on both learning styles: the students are oriented on data preprocessing and structuring, making plots and search of regularities. The cluster description based on open profiles data from VK is missing because open profiles data is uninformative (SI=2, PI=3, II=1).
Cluster 6 (the students with id47, id52, id62) has the center of cluster [6.(3); 17; 11.(6); 12] by Honey and Mumford's learning styles and center of cluster [2; 0.(3); 3.(6); 0; 7; 0; 5; 0] by ILS. The cluster description based on both learning styles: the students are oriented on consistent actions and visual data presentation. The cluster description based on open profiles data from VK is missing because open profiles data is uninformative (SI=2, PI=3, II=2).
Cluster 7 (the students with id24, id90, id68, id93) has the center of cluster [14; 16; 12.75; 16.5] by Honey and Mumford's learning styles and center of cluster [5; 0; 0.25; 1.75; 6.5; 0; 1.25; 1.75] by ILS. The cluster description based on both learning styles: the students are oriented on new ideas search, visual results presentation and look at the issue in the broader perspective. The cluster description based on open profiles data from VK: the students prefer social and organizational activity (SI=3, PI=2, II=2).
Cluster 8 (the students with id10, id48, id54, id73) has the center of cluster [4.75; 14.25; 13.25; 11] by Honey and Mumford's learning styles and center of cluster [0; 3.5; 0.25; 2.75; 0; 3; 0.75; 2.75] by ILS. The cluster description based on both learning styles: the students are oriented on creativity, acquiring of theoretical results, the reports writing and protection of the results. The cluster description based on open profiles data from VK is missing because open profiles data is uninformative (SI=1, PI=2, II=1).
Cluster 9 (the students with id9, id79, id70) has the center of cluster [5; 16.(3); 14.(7); 16.3] by Honey and Mumford's learning styles and center of cluster [0; 7; 5; 0; 0; 1.(6); 2; 1] by ILS. The cluster description based on both learning styles: the students are oriented on learning of detailed data. The cluster description based on open profiles data from VK is missing because open profiles data is uninformative (SI=1, PI=2, II=1).

5. CONCLUSION AND FUTURE WORK

We propose a methodology aimed at automatically student groups detection with the help of learning styles and clustering methods to extract student groups with similar preferred learning styles and interests. Based on our approach we could automatically form student groups only for 35 students from 61 students and we construct a description of clusters that can be used for creating of recommendation for the participation of students in various activities. The results of our research do not mean that there are no regularities for other students. It is necessary to use other principles and approaches to detect groups for such students. Besides, it is important to discuss with teachers and approve with students automatically forming groups.
In further research, we will plan to understand psychological compatibility based on learning styles in automatically detected student groups and investigate students' attitudes to collaborate in the proposed team.

REFERENCES

[1] N. A. Ali, F. Eassa, E. Hamed, Personalized learning style for adaptive e-learning system. International Journal of Advanced Trends in Computer Science and Engineering, 8(1.1) (2019) 223-230.

[2] J. Dangmei, Exploring the link between Felder-Silverman learning style and stress coping ability: An Empirical Study, Int. J. of Management Studies, VI(2(3)) (2019) 44-50.

[3] G.A. Epachinceva, T.N. Kozlovskaya, D.A. Kozlovskij, Features of activities styles and self-organisation among university students, Bulletin of the Orenburg State University, 1 (2019) 80-87.

[4] R.M. Felder, Matters of style ASEE Prism, 6(4) (1996) 18-23.

[5] P. Honey, A. Mumford, Using your learning styles. Peter Honey, Maidenhead, 1986.

[6] H. Kamisli, Learning styles and culture. Elementary Education Online, 18(2) (2019) 451-460.

[7] M. Kyprianidou, S. Demetriadis, T. Tsiatsos, A. Pombortsis, Group formation based on learning styles: can it improve students' teamwork, Education Tech Research Dev, 60 (2012) 83-110.

[8] Q. Liu, Sh. Ba, J. Huang, L. Wu, Ch. Lao, A study on grouping strategy of collaborative learning based on clustering algorithm, Lecture Notes in Computer Science, pp. 284-294, 2017.

[9] B.E. Myers, J.E. Dyer, The influence of student learning style on critical thinking skill, J. of Agricultura, 47(1) (2006) 43-52.

[10] J.C. Mobeva, J.S. Arnado, I.N. Buot, Students' learning styles and self-motivation, Int. J. of Social Sci. Res., 8(2) (2020) 16-29.

[11] A.S. Nagesh, V.S. Satyamurty, Application of clustering algorithm for analysis of student academic performance, Int. J. of Computer Science and Engineering, 6(1) (2018) 381-384.

[12] B. Naik, V.G. Girish, An investigation of the learning styles of South Korean business students, Asia-Pacific J. of Business, 3(1) (2012) 1-9.

[13] NbClust Package For Determining The Best Number Of Clusters. https://www.rdocumentation.org/packages/NbClust/versions/3.0/topics/NbClust

[14] H. Nhan, Th. Nhan Different grouping strategies for cooperative learning in english majored seniors and juniors at Can Tho University. Vietnam Education Sciences, 9(59) (2019) 16.

[15] C.O. Nja, C.-U. B. Umali, E.E. Asuquo, R.E. Orim, The influence of learning styles on academic performance among science education undergraduates at the University of Calabar. Educational Research and Reviews, 14(17) (2019) 618-624.

[16] I. Pasina, G. Bayram, W. Labib, A. Abdelhadi, M. Nurunnabi, Clustering students into groups according to their learning style, MethodsX, 6 (2019) 2189-2197

[17] V.E. Pershukova, Students preparing for project activities by identifying and taking into account individual learning styles, The world of science, culture, education, 5(24) (2010) 90-93.

[18] E. Popescu, C. Badica, L. Moraret, Accommodating Learning Styles in an Adaptive Educational System, Informatica, 34 (2010) 451-462.

[19] V. Rosa, R. Fida, F. Avallone, A cluster analysis of high school students' styles of "Living-Together" in the classroom. Procedia – Social and Behavioral Sciences, 29 (2011) 380-389.

[20] M. Suryaman, A. Chaerul, F.N. Rahma, The learning style preferences based on gender differences. The International Conference on Education, Management, Humanity, Social, Politics, Economics, Science and Technology, Law and Health, 1 (2019). 225-229.

[21] M.G. Vashisht, A.K. Bhawna, Students expectations using opinion mining clustering, Int. J. of Innovative Technology and Exploring Engineering, 8(9S) (2019) 716-720.