Abstract

Currently, academic instructors in Ghana have some difficulty in grouping students for projects-based courses because of increasing student numbers. One of the recent challenges educational institutions and instructors are facing is the explosive growth of educational data and how to use this data to improve the quality of teaching. K-means clustering is an unsupervised Data Mining technique for grouping large datasets with insightful similarity patterns to expose hidden trends and behavior in each cluster. The purpose of this research is to apply K-means clustering algorithm to analyze students' clusters for centered project-based learning. This research uses K clusters of 20. The clustering gave a low within cluster Sum of Square Error (SSE) of 3.60889. Clusters 1 and 6 have the highest member set of 32 each whiles clusters 8 and 9 have the lowest member set of 2. The results show that the K-means clustering algorithm is effective in grouping learners based on similar characteristics that indicate their performance. Assessments can also be tailored to suit all categories of learners for efficient results in project-based courses.
Using K-Means to Determine Learner Typologies for Project-based Learning: A Case Study of the University of Education, Winneba

References

1. Baradwaj, B. K., & Pal, S. (2012). Mining educational data to analyze students’ performance. arXiv preprint arXiv:1201.3417.
2. Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40(6), 601-618.
3. Bilgin, I., Karakuyu, Y., & Ay, Y. (2015). The effects of project based learning on undergraduate students’ achievement and self-efficacy beliefs towards science teaching. Eurasia Journal of Mathematics, Science & Technology Education, 11(3), 469-477.
4. Perera, D., Kay, J., Koprinska, I., Yacef, K., & Zaïane, O. R. (2009). Clustering and sequential pattern mining of online collaborative learning data. IEEE Transactions on Knowledge and Data Engineering, 21(6), 759-772.
5. Han, S., Capraro, R., & Capraro, M. M. (2015). How science, technology, engineering, and mathematics (STEM) project-based learning (PBL) affects high, middle, and low achievers differently: The impact of student factors on achievement. International Journal of Science and Mathematics Education, 13(5), 1089-1113.
6. Shovon, M., Islam, H., & Haque, M. (2012). An Approach of Improving Students Academic Performance by using k means clustering algorithm and Decision tree. arXiv preprint arXiv:1211.6340.
7. Bahr, P. R. (2010). The bird’s eye view of community colleges: A behavioral typology of first-time students based on cluster analytic classification. Research in Higher Education, 51(8), 724-749.
8. Paquette, G., Léonard, M., Lundgren-Cayrol, K., Mihaila, S., & Gareau, D. (2006). Learning design based on graphical knowledge-modelling. Educational Technology & Society, 9(1), 97-112.
9. Niu, K., Niu, Z., Zhao, X., Wang, C., Kang, K., & Ye, M. (2016). A Coupled User Clustering Algorithm for Web-based Learning Systems. In EDM (pp. 175-182).
10. Dutt, A., Aghabozrgi, S., Ismail, M. A. B., & Mahroieian, H. (2015). Clustering algorithms applied in educational data mining. International Journal of Information and Electronics Engineering, 5(2), 112.
11. Akbar, S., Gehringer, E. F., & Hu, Z. (2018). Improving formation of student teams: a clustering approach. In Proceedings of the 40th International Conference on Software Engineering: Companion Proceedings (pp. 147-148). ACM.
12. Li, L., Luo, X., & Chen, H. (2015). Clustering Students for Group-Based Learning in Foreign Language Learning. International Journal of Cognitive Informatics and Natural Intelligence (IJCINI), 9(2), 55-72.
13. Sarkar, A., Seth, D., Basu, K., & Acharya, A. (2015). A new approach to collaborative group formation. International journal of computer applications, 128(3).
14. Stage, F. K., Muller, P. A., Kinzie, J., & Simmons, A. (1998). Creating learning centered classrooms: What does learning theory have to say? Washington DC: The George Washington University. Graduate School of Education and Human Development and Association for the Study of Higher Education.
15. Ester, M., Frommelt, A., Kriegel, H. P., & Sander, J. (1998). Algorithms for characterization and trend detection in spatial databases. In Proc. 4th Int. Conf. on Knowledge Discovery and Data Mining (pp. 44-50). New York City, NY.
16. Kaufman, L and Rousseeuw, J. P. (2005). Finding Groups in Data: An Introduction to Cluster Analysis (Wiley Series in Probability and Statistics). Wiley-Interscience

17. MacQueen, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability (pp. 281–297). University of California Press.

18. Jian, B., & Vemuri, B.C. (2005). A robust algorithm for point set registration using mixture of Gaussians. Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, 2, 1246-1251 Vol. 2.

Index Terms

Computer Science

Information Sciences

Keywords

K-means, Clustering, Educational Data Mining, Data Mining, Project-Based Learning.