

LEARNING ANALYTIC FRAMEWORK FOR STUDENTS' ACADEMIC PERFORMANCE AND CRITICAL LEARNING PATHWAYS

(Rangka Kerja Analitik Pembelajaran untuk Prestasi Akademik Pelajar dan Laluan Pembelajaran Kritikal)

JESSICA TAN YEN LYN*, GOH YONG KHENG, LAI AN CHOW & NGEOW YOKE MENG

ABSTRACT

In the domain of higher education, the need to leverage data-driven insights for understanding and enhancing student academic performance is becoming increasingly critical. To address this, a unified learning analytics framework is proposed, aimed at deciphering complex student academic journeys and fostering data-informed decision-making for educational institutions. This framework's methodology involves several key steps, starting with standardized data collection and pre-processing. Subsequently, dimensionality reduction techniques like Principal Component Analysis (PCA) and Non-negative matrix factorization (NMF) are applied to capture the most influential course components and grade information. The resulting reduced dataset is then subjected to various clustering algorithms, including partition-based clustering (K-means), hierarchical clustering, and density-based clustering (DBSCAN). These algorithms group students based on academic performance and course profiles, facilitating the identification of clusters with similar characteristics and academic trajectories. Furthermore, a collective network graph is constructed to analyze course relationships and program pathways, identify critical courses, and reveal influential factors affecting student performance and outcomes. This network analysis enables educators to identify bottleneck courses and areas that may require additional support or improvement, fostering a data-driven approach to curriculum design and enhancement. To showcase the framework's efficacy, a case study was conducted on 3550 undergraduates from an engineering program at a Malaysian private university. The student dataset used in this study spans from 2005 to 2021, covering a wide range of academic years for analysis. The results demonstrate the framework's capability to unveil valuable insights into students' academic journeys, revealing key factors contributing to their success. By providing a holistic perspective of student performance and course interactions, the proposed learning analytics framework holds great promise for educational institutions seeking data-driven strategies to enhance student outcomes and optimize learning experiences.

Keywords: learning analytic framework; Principal Component Analysis (PCA); Non-negative Matrix Factorization (NMF); clustering; network graph

ABSTRAK

Dalam domain pendidikan tinggi, keperluan untuk memanfaatkan cerapan dipacu data untuk memahami dan meningkatkan prestasi akademik pelajar menjadi semakin kritikal. Untuk menangani perkara ini, rangka kerja analisis pembelajaran bersatu dicadangkan, bertujuan untuk mentafsir perjalanan akademik pelajar yang kompleks dan memupuk pembuatan keputusan berdasarkan data untuk institusi pendidikan. Metodologi rangka kerja ini melibatkan beberapa langkah utama, bermula dengan pengumpulan data dan pra-pemprosesan piawai. Selepas itu, teknik pengurangan dimensi seperti Principal Component Analysis (PCA) dan Non-negative Matrix Factorization (NMF) digunakan untuk menangkap komponen kursus dan maklumat gred yang paling berpengaruh. Dataset terkurang yang terhasil kemudiannya tertakluk kepada pelbagai algoritma pengelompokan, termasuk pengelompokan berasaskan partition (K-means), pengelompokan hierarki dan pengelompokan berasaskan ketumpatan (DBSCAN). Algoritma ini mengumpulkan pelajar berdasarkan prestasi akademik dan profil kursus, memudahkan pengenalanpastian kluster dengan ciri dan trajektori akademik yang serupa. Tambahan pula, graf rangkaian kolektif dibina untuk menganalisis hubungan kursus dan laluan program, mengenal

pasti kursus kritikal dan mendedahkan faktor-faktor yang mempengaruhi prestasi dan hasil pelajar. Analisis rangkaian ini membolehkan pendidik mengenal pasti kursus kesesakan dan bidang yang mungkin memerlukan sokongan atau penambahbaikan tambahan, memupuk pendekatan terdorong data untuk reka bentuk dan peningkatan kurikulum. Untuk mempamerkan keberkesanan rangka kerja tersebut, satu kajian kes telah dijalankan ke atas 3550 mahasiswa dari program kejuruteraan di universiti swasta Malaysia. Hasilnya menunjukkan keupayaan rangka kerja untuk mendedahkan pandangan berharga tentang perjalanan akademik pelajar, mendedahkan faktor utama yang menyumbang kepada kejayaan mereka. Dengan menyediakan perspektif holistik prestasi pelajar dan interaksi kursus, rangka kerja analitik pembelajaran yang dicadangkan memegang janji besar untuk institusi pendidikan yang mencari strategi dipacu data untuk meningkatkan hasil pelajar dan mengoptimumkan pengalaman pembelajaran.

Kata kunci: rangka kerja analisis pembelajaran; Principal Component Analysis (PCA); Non-negative Matrix Factorization (NMF); pengelompokan; graf rangkaian

References

- Abu Saa A. 2016. Educational data mining & students' performance prediction. *International Journal of Advanced Computer Science and Applications* 7(5): 212–220.
- Alhaji S., Alhaji A. & Özyer S.T. 2022. Combining multiple clustering and network analysis for discoveries in gene expression data. *ASONAM '21: Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 502–509.
- Almond-Dannenbring T., Easter M., Feng L., Guarcello M., Ham M., Machajewski S., Maness H., Miller A., Mooney S., Moore A. & Kendall E. 2022. *A Framework for Student Success Analytics*. EDUCAUSE Publications.
- Bharara S., Sabitha S. & Bansal A. 2018. Application of learning analytics using clustering data mining for students' disposition analysis. *Education and Information Technologies* 23(2): 957–984.
- Borgatti S.P. & Everett M.G. 2006. A graph-theoretic perspective on centrality. *Social Networks* 28(4): 466–484.
- Caliński T. & Harabasz J. 1974. A dendrite method for cluster analysis. *Communications in Statistics - Theory and Methods* 3(1): 1–27.
- Dol S.M. & Jawandhiya P.M. 2023. Classification technique and its combination with clustering and association rule mining in educational data mining — a survey. *Engineering Applications of Artificial Intelligence* 122: 106071.
- Govindasamy K. & Thambusamy V. 2018. Analysis of student academic performance using clustering techniques. *International Journal of Pure and Applied Mathematics* 119(15): 309–322.
- Greenacre M., Groenen P., Hastie T., Iodice D'Enza A., Markos A. & Tuzhilina E. 2022. Principal component analysis. *Nature Reviews Methods Primers* 2(1): 100.
- Grunspan D.Z., Wiggins B.L. & Goodreau S.M. 2014. Understanding classrooms through social network analysis: A primer for social network analysis in education research. *CBE Life Sciences Education* 13(2): 167–178.
- Jain A.K., Murty M.N. & Flynn P.J. 1999. Data clustering: A review. *ACM computing surveys (CSUR)* 31(3): 264–323.
- Jan S.K., Vlachopoulos P. & Parsell M. 2019. Social network analysis and learning communities in higher education online learning : A systematic literature review. *Online Learning* 23(1).
- Joshi A., Desai P. & Tewari P. 2020. Learning analytics framework for measuring students' performance and teachers' involvement through problem based learning in engineering education. *Procedia Computer Science* 172: 954–959.
- Khalil M., Prinsloo P. & Slade S. 2022. A comparison of learning analytics frameworks: a systematic review. *LAK22: 12th International Learning Analytics and Knowledge Conference*, pp. 152–163.
- Križanić S. 2020. Educational data mining using cluster analysis and decision tree technique: A case study. *International Journal of Engineering Business Management* 12: 1847979020908675.
- Leuschner J., Schmidt M., Fernsel P., Lachmund D., Boskamp T. & Maass P. 2019. Supervised non-negative matrix factorization methods for MALDI imaging applications. *Bioinformatics* 35(11): 1940 – 1947.

- Omar T., Alzahrani A. & Zohdy M. 2020. Clustering approach for analyzing the student's efficiency and performance based on data. *Journal of Data Analysis and Information Processing* **8**(3): 171 – 182.
- Samanta P., Sarkar D., Kole D.K. & Jana P. 2021. Social network analysis in education: A study. *Computational Intelligence in Digital Pedagogy* ((eds.). A. Deyasi, S. Mukherjee, A. Mukherjee, A. K. Bhattacharjee & A. Mondal), pp. 65 – 83. Singapore: Springer.
- Sagr M., Fors U., Tedre M. & Nouri J. 2018. How social network analysis can be used to monitor online collaborative learning and guide an informed intervention. *PLoS One* **13**(3): e0194777.
- Shahapure K.R. & Nicholas C. 2020. Cluster quality analysis using silhouette score. *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 747–748.
- Su Y.S., Lin Y.D. & Liu T.Q. 2023. Applying machine learning technologies to explore students' learning features and performance prediction. *Front Neurosci* **16**(3): 1018005.

Lee Kong Chian Faculty of Engineering & Science

Universiti Tunku Abdul Rahman

43200 UTAR Kajang

Selangor, MALAYSIA

E-mail: tanyenlyn97@gmail.com, gohyk@utar.edu.my, laiac@utar.edu.my, ngeowym@utar.edu.my*

Received: 5 September 2023

Accepted: 18 January 2024

*Corresponding author