

will lead to better analysis and more effective predictions. Overall, future work should prioritize resolving the issues raised by imbalanced datasets and developing techniques to improve the accuracy and usefulness of prediction models in educational settings.

ACKNOWLEDGMENT

This work was supported by the Ministry of Higher Education, Malaysia (MOHE) under the Fundamental Research Grant Scheme (FRGS), FRGS/1/2023/ICT02/UTM/02/8.

REFERENCES

- [1] S. D. A. Bujang *et al.*, "Multiclass Prediction Model for Student Grade Prediction Using Machine Learning," *IEEE Access*, vol. 9, pp. 95608–95621, 2021, doi: 10.1109/ACCESS.2021.3093563.
- [2] D. Solomon, "Predicting Performance and Potential Difficulties of University Student using Classification: Survey Paper," *Int. J. Pure Appl. Math.*, vol. 118, no. 18, pp. 2703–2707, 2018.
- [3] E. B. Costa, B. Fonseca, M. A. Santana, F. F. de Araújo, and J. Rego, "Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses," *Comput. Human Behav.*, vol. 73, pp. 247–256, 2017, doi: 10.1016/j.chb.2017.01.047.
- [4] Y. Zhang, Y. Yun, H. Dai, J. Cui, and X. Shang, "Graphs regularized robust matrix factorization and its application on student grade prediction," *Appl. Sci.*, vol. 10, no. 5, pp. 1–19, 2020, doi:10.3390/app10051755.
- [5] A. Hellas *et al.*, "Predicting academic performance: a systematic literature review," in *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, in ITiCSE 2018 Companion. New York, NY, USA: Association for Computing Machinery, 2018, pp. 175–199. doi:10.1145/3293881.3295783.
- [6] S. T. Jishan, R. I. Rashu, N. Haque, and R. M. Rahman, "Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique," *Decis. Anal.*, vol. 2, no. 1, 2015, doi: 10.1186/s40165-014-0010-2.
- [7] I. Khan, A. Al Sideiri, A. Ahmad, and N. Jabeur, "Tracking Student Performance in Introductory Programming by Means of Machine Learning," Feb. 2019, pp. 1–6. doi:10.1109/ICBDSC.2019.8645608.
- [8] M. A. Al-Barrak and M. Al-Razgan, "Predicting Students Final GPA Using Decision Trees: A Case Study," *Int. J. Inf. Educ. Technol.*, vol. 6, no. 7, pp. 528–533, 2016, doi: 10.7763/IJIEET.2016.V6.745.
- [9] M. Agaoglu, "Predicting Instructor Performance Using Data Mining Techniques in Higher Education," *IEEE Access*, vol. 4, pp. 2379–2387, 2016, doi: 10.1109/ACCESS.2016.2568756.
- [10] L. Ismail, H. Materwala, and A. Hennebelle, "Comparative Analysis of Machine Learning Models for Students' Performance Prediction," in *Advances in Intelligent Systems and Computing*, Springer Science and Business Media Deutschland GmbH, 2021, pp. 149–160. doi:10.1007/978-3-030-71782-7_14.
- [11] B. Flanagan, R. Majumdar, and H. Ogata, "Early-warning prediction of student performance and engagement in open book assessment by reading behavior analysis," *Int. J. Educ. Technol. High. Educ.*, vol. 19, no. 1, Dec. 2022, doi: 10.1186/s41239-022-00348-4.
- [12] A. Polyzou and G. Karypis, "Grade prediction with models specific to students and courses," *Int. J. Data Sci. Anal.*, vol. 2, no. 3–4, pp. 159–171, Dec. 2016, doi: 10.1007/s41060-016-0024-z.
- [13] F. Ahmad, N. H. Ismail, and A. A. Aziz, "The prediction of students' academic performance using classification data mining techniques," *Appl. Math. Sci.*, vol. 9, no. 129, pp. 6415–6426, 2015, doi:10.12988/ams.2015.53289.
- [14] T. Anderson, "Applications of Machine Learning to Student Grade Prediction in Quantitative Business Courses," 2017.
- [15] E. C. Abana, "A decision tree approach for predicting student grades in Research Project using Weka," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 7, pp. 285–289, 2019, doi:10.14569/ijacsa.2019.0100739.
- [16] I. Khan, A. Al Sadiri, A. R. Ahmad, and N. Jabeur, "Tracking Student Performance in Introductory Programming by Means of Machine Learning," in *2019 4th MEC International Conference on Big Data and Smart City (ICBDSC)*, 2019, pp. 1–6. doi:10.1109/ICBDSC.2019.8645608.
- [17] E. Wakelam, A. Jefferies, N. Davey, and Y. Sun, "The potential for student performance prediction in small cohorts with minimal available attributes," *Br. J. Educ. Technol.*, vol. 51, no. 2, pp. 347–370, Mar. 2020, doi: 10.1111/bjet.12836.
- [18] Y. Prityanto, N. A. Setiawan, and I. Ardiyanto, "Hybrid resampling to handle imbalanced class on classification of student performance in classroom," Feb. 2017, pp. 207–212. doi:10.1109/ICICOS.2017.8276363.
- [19] X. Zhang, R. Xue, B. Liu, W. Lu, and Y. Zhang, "Grade Prediction of Student Academic Performance with Multiple Classification Models," Feb. 2018, pp. 1086–1090. doi:10.1109/FSKD.2018.8687286.
- [20] A. Saifudin, Ekawati, Yulianti, and T. Desyani, "Forward Selection Technique to Choose the Best Features in Prediction of Student Academic Performance Based on Naïve Bayes," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, 2020. doi:10.1088/1742-6596/1477/3/032007.
- [21] C. Chen, A. Liaw, and L. Breiman, "Using Random Forest to Learn Imbalanced Data," *Discovery*, no. 1999, pp. 1–12.
- [22] R. Couronné, P. Probst, and A. L. Boulesteix, "Random forest versus logistic regression: A large-scale benchmark experiment," *BMC Bioinformatics*, vol. 19, no. 1, pp. 1–15, 2018, doi: 10.1186/s12859-018-2264-5.
- [23] C. Y. J. Peng, K. L. Lee, and G. M. Ingersoll, "An introduction to logistic regression analysis and reporting," *J. Educ. Res.*, vol. 96, no. 1, pp. 3–14, 2002, doi: 10.1080/00220670209598786.
- [24] P. Brous and M. Janssen, "Trusted decision-making: Data governance for creating trust in data science decision outcomes," *Adm. Sci.*, vol. 10, no. 4, 2020, doi: 10.3390/admsci10040081.
- [25] M. Tsiakmaki, G. Kostopoulos, S. Kotsiantis, and O. Ragos, "Implementing autoML in educational data mining for prediction tasks," *Appl. Sci.*, vol. 10, no. 1, pp. 1–27, 2020, doi:10.3390/app10010090.