

DAFTAR PUSTAKA

- Adekitan, A. I., & Noma-Osaghae, E. (2019). Data mining approach to predicting the performance of first year student in a university using the admission requirements. *Education and Information Technologies*, 24(2), 1527–1543. <https://doi.org/10.1007/s10639-018-9839-7>
- Adekitan, A. I., Salau, O., Ng, A. I., & Adekitan,). (2019). *The impact of engineering students' performance in the first three years on their graduation result using educational data mining*. <https://doi.org/10.1016/j.heliyon.2019>
- Al Fanah, M., & Ansari, M. A. (2019). Understanding E-learners' behaviour using data mining techniques. *ACM International Conference Proceeding Series*, 59–65. <https://doi.org/10.1145/3322134.3322145>
- Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37(January), 13–49. <https://doi.org/10.1016/j.tele.2019.01.007>
- Alghazali, N., Zahruddin, M., Diyana Kamarudin, N., Jusoh, R. M., Aisyah, N., Fataf, A., & Hidayat, R. (2023). Case Study: Using Data Mining to Predict Student Performance Based on Demographic Attributes. *JOIV: International Journal of Informatics and Visualization*. www.joiv.org/index.php/joiv
- Alnomay, I., Alfadhly, A., & Alqarni, A. (2024). A Comparative Analysis for GPA Prediction of Undergraduate Students Using Machine and Deep Learning. *International Journal of Information and Education Technology*, 14(2), 287–292. <https://doi.org/10.18178/ijiet.2024.14.2.2050>
- Alshehri, E., Alhakami, H., Baz, A., & Alsubait, T. (2020). A Comparison of EDM Tools and Techniques. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 11, Issue 12). www.ijacsa.thesai.org
- Altalhan, M., Algarni, A., & Turki-Hadj Alouane, M. (2025). Imbalanced Data problem in Machine Learning: A review. *IEEE Access*, 13, 13686–13699. <https://doi.org/10.1109/ACCESS.2024.0429000>

- Aluko, R. O., Daniel, E. I., Shamsideen Oshodi, O., Aigbavboa, C. O., & Abisuga, A. O. (2018). Towards reliable prediction of academic performance of architecture students using data mining techniques. *Journal of Engineering, Design and Technology*, 16(3), 385–397. <https://doi.org/10.1108/JEDT-08-2017-0081>
- Alyahyan, E., & Düşteğör, D. (2020). Predicting academic success in higher education: literature review and best practices. In *International Journal of Educational Technology in Higher Education* (Vol. 17, Issue 1). Springer. <https://doi.org/10.1186/s41239-020-0177-7>
- Araque, F., Roldán, C., & Salguero, A. (2009). Factors influencing university drop out rates. *Computers and Education*, 53(3), 563–574. <https://doi.org/10.1016/j.compedu.2009.03.013>
- Bakhshinategh, B., Zaiane, O. R., ElAtia, S., & Ipperciel, D. (2017). Educational data mining applications and tasks: A survey of the last 10 years. *Education and Information Technologies*, 23(1), 537–553. <https://doi.org/10.1007/s10639-017-9616-z>
- Bisong, E. (2019). Google Colaboratory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform* (pp. 59–64). Apress. https://doi.org/10.1007/978-1-4842-4470-8_7
- Bothma, J. (2023). *Evaluating Classification Models: Understanding the Confusion Matrix and ROC Curves*. <https://statisticallyrelevant.com/confusion-matrix-and-roc-curves/>
- Bujang, S. D. A., Selamat, A., Ibrahim, R., Krejcar, O., Herrera-Viedma, E., Fujita, H., & Ghani, N. A. M. (2021). Multiclass Prediction Model for Student Grade Prediction Using Machine Learning. *IEEE Access*, 9, 95608–95621. <https://doi.org/10.1109/ACCESS.2021.3093563>
- Cheng, X. (2024). A Comprehensive Study of Feature Selection Techniques in Machine Learning Models. *Comput. Signal Syst*, 1(1). <https://soapubs.com/index.php/ICSS>
- Christian, T. M., & Ayub, M. (2014). Exploration of classification using NBTree for predicting students' performance. *Proceedings of 2014 International Conference on Data and Software Engineering, ICODSE 2014*, 0–5. <https://doi.org/10.1109/ICODSE.2014.7062654>
- Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for

- early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*, 73, 247–256. <https://doi.org/10.1016/J.CHB.2017.01.047>
- Cote, P.-O., Nikanjam, A., Ahmed, N., Humeniuk, D., & Khomh, F. (2024). Data Cleaning and Machine Learning: A Systematic Literature Review. *Autom Softw Eng*. <http://arxiv.org/abs/2310.01765>
- Cunningham, P., Cord, M., & Delany, S. J. (2008). Supervised learning. *Cognitive Technologies*, 21–49. https://doi.org/10.1007/978-3-540-75171-7_2
- David Andrés. (2023, November 30). *Introduction to Decision Trees*. <https://mlpills.dev/machine-learning/introduction-to-decision-trees/>
- De Santos, K. J. O., Menezes, A. G., De Carvalho, A. B., & Montesco, C. A. E. (2019). Supervised learning in the context of educational data mining to avoid university students dropout. *Proceedings - IEEE 19th International Conference on Advanced Learning Technologies, ICALT 2019, 2161-377X*, 207–208. <https://doi.org/10.1109/ICALT.2019.00068>
- Ependi, U., Rochim, A. F., & Wibowo, A. (2023). A Hybrid Sampling Approach for Improving the Classification of Imbalanced Data Using ROS and NCL Methods. *International Journal of Intelligent Engineering and Systems*, 16(3), 345–361. <https://doi.org/10.22266/ijies2023.0630.28>
- Fernandes, E., Holanda, M., Victorino, M., Borges, V., Carvalho, R., & Erven, G. Van. (2019). Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil. *Journal of Business Research*, 94, 335–343. <https://doi.org/10.1016/j.jbusres.2018.02.012>
- Frank, K. (2017). *Hands-On Data Science and Python Machine Learning*. https://www.packtpub.com/mapt/book/big_data_and_business_intelligence/9781787280748
- Genuer, R., Poggi, J. M., Tuleau-Malot, C., & Villa-Vialaneix, N. (2017). Random Forests for Big Data. *Big Data Research*, 9, 28–46. <https://doi.org/10.1016/j.bdr.2017.07.003>
- Gray, G., McGuinness, C., & Owende, P. (2014). An application of classification models to predict learner progression in tertiary education. *Souvenir of the 2014 IEEE International Advance Computing*

Conference, IACC 2014, 549–554.
<https://doi.org/10.1109/IAdCC.2014.6779384>

- Gufroni, A. I., Purwanto, P., Farikhin, F., Wibowo, A., & Warsito, B. (2021). Exploratory Data Analysis To Identify The Most Important Feature Of University Admission Test Criteria Using Random Forest And Neural Network Algorithm. *2021 5th International Conference on Informatics and Computational Sciences (ICICoS)*, 1–5. <https://doi.org/10.1109/ICICoS53627.2021.9651757>
- Guo, Y., Feng, J., Jiao, B., Cui, N., Yang, S., & Yu, Z. (2022). A dual evolutionary bagging for class imbalance learning. *Expert Systems with Applications*, 206. <https://doi.org/10.1016/j.eswa.2022.117843>
- Hakkal, S., & Lahcen, A. A. (2024). XGBoost To Enhance Learner Performance Prediction. *Computers and Education: Artificial Intelligence*, 7, 100254. <https://doi.org/10.1016/j.caeai.2024.100254>
- Han, J., Kamber, M., & Pei, J. (2012). Data Mining Concepts and Techniques. In *The Morgan Kaufmann Series in Data Management Systems*.
- Harefa, E. (2024). Predicting Science Learning Outcomes Of Elementary Students In Rural Area Using Machine Learning Algorithms. *Journal Of Social Science Research*, 4, 3780–3792.
- Hassan, S. M., & Al-Razgan, M. S. (2016). Pre-University Exams Effect on Students GPA: A Case Study in IT Department. *Procedia Computer Science*, 82, 127–131. <https://doi.org/10.1016/j.procs.2016.04.018>
- Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). *Supervised Machine Learning: A Brief Primer*. www.elsevier.com/locate/bt
- Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining*, 15(4), 531–538. <https://doi.org/10.1002/sam.11583>
- Josse, J., Chen, J. M., Prost, N., Varoquaux, G., & Scornet, E. (2024). On the consistency of supervised learning with missing values. *Statistical Papers*. <https://doi.org/10.1007/s00362-024-01550-4>
- Kanetaki, Z., Stergiou, C., Bekas, G., Troussas, C., & Sgouropoulou, C. (2022). A Hybrid Machine Learning Model for Grade Prediction in Online Engineering Education. *International Journal of Engineering Pedagogy*, 12(3), 4–23. <https://doi.org/10.3991/IJEP.V12I3.23873>
- Khan, A., & Ghosh, S. K. (2021). Student performance analysis and prediction in classroom learning: A review of educational data mining studies.

Education and Information Technologies, 26(1), 205–240.
<https://doi.org/10.1007/s10639-020-10230-3>

Khan, K., Ahmad, W., Amin, M. N., Ahmad, A., Nazar, S., & Alabdullah, A. A. (2022). Compressive Strength Estimation of Steel-Fiber-Reinforced Concrete and Raw Material Interactions Using Advanced Algorithms. *Polymers*, 14(15). <https://doi.org/10.3390/polym14153065>

Larose, D. T. (2006). Data Mining Methods And Models. In *Contemporary Psychology: A Journal of Reviews* (Vol. 21, Issue 1). John Wiley & Sons. Inc.

Lemay, D. J., Baek, C., & Doleck, T. (2021). Comparison of learning analytics and educational data mining: A topic modeling approach. *Computers and Education: Artificial Intelligence*, 2, 100016. <https://doi.org/10.1016/J.CAEAI.2021.100016>

Luque, A., Carrasco, A., Martín, A., & de las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, 91, 216–231. <https://doi.org/10.1016/j.patcog.2019.02.023>

Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers and Education*, 103, 1–15. <https://doi.org/10.1016/j.compedu.2016.09.005>

Mayilvaganan, M., & Kalpanadevi, D. (2015). Comparison of classification techniques for predicting the cognitive skill of students in education environment. *2014 IEEE International Conference on Computational Intelligence and Computing Research, IEEE ICCIC 2014*, 113–118. <https://doi.org/10.1109/ICCIC.2014.7238346>

Mengash, H. A. (2020). Using data mining techniques to predict student performance to support decision making in university admission systems. *IEEE Access*, 8, 55462–55470. <https://doi.org/10.1109/ACCESS.2020.2981905>

Miguéis, V. L., Freitas, A., Garcia, P. J. V., & Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach. *Decision Support Systems*, 115, 36–51. <https://doi.org/10.1016/j.dss.2018.09.001>

- Nachouki, M., Mohamed, E. A., Mehdi, R., & Abou Naaj, M. (2023). Student course grade prediction using the random forest algorithm: Analysis of predictors' importance. *Trends in Neuroscience and Education*, 33. <https://doi.org/10.1016/j.tine.2023.100214>
- Nachouki, M., & Naaj, M. A. (2022). Predicting Student Performance to Improve Academic Advising Using the Random Forest Algorithm. *International Journal of Distance Education Technologies*, 20(1). <https://doi.org/10.4018/IJDET.296702>
- Natek, S., & Zwilling, M. (2014). Student data mining solution-knowledge management system related to higher education institutions. *Expert Systems with Applications*, 41(14), 6400–6407. <https://doi.org/10.1016/j.eswa.2014.04.024>
- Nowozin, S., Rother, C., Bagon, S., Sharp, T., Yao, B., & Kohli, P. (2011). Decision tree fields. *2011 International Conference on Computer Vision*, 1668–1675. <https://doi.org/10.1109/ICCV.2011.6126429>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830. <http://jmlr.org/papers/v12/pedregosa11a.html>
- Rezvani, S., Pourpanah, F., Lim, C. P., & Wu, Q. M. J. (2024). *Methods for Class-Imbalanced Learning with Support Vector Machines: A Review and an Empirical Evaluation*. <https://doi.org/10.1007/s00500-024-09931-5>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3). <https://doi.org/10.1002/widm.1355>
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3). <https://doi.org/10.1007/s42979-021-00592-x>
- Sarker, S., Paul, M. K., Thasin, S. T. H., & Hasan, M. A. M. (2024). Analyzing students' academic performance using educational data mining.

Computers and Education: Artificial Intelligence, 7.
<https://doi.org/10.1016/j.caeai.2024.100263>

Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science*, 72, 414–422.
<https://doi.org/10.1016/j.procs.2015.12.157>

Skittou, M., Merrouchi, M., & Gadi, T. (2024). Development of an Early Warning System to Support Educational Planning Process by Identifying At-Risk Students. *IEEE Access*, 12, 2260–2271.
<https://doi.org/10.1109/ACCESS.2023.3348091>

Sravani, B., & Bala, M. M. (2020). *Prediction of Student Performance Using Linear Regression*.

Ul Hassan, C. A., Sufyan Khan, M., & Ali Shah, M. (2018). *Comparison of Machine Learning Algorithms in Data classification*.

Velthoen, J., Dombry, C., Cai, J.-J., & Engelke, S. (2023). Gradient boosting for extreme quantile regression. *Extremes*, 26(4), 639–667.
<https://doi.org/10.1007/s10687-023-00473-x>

Yaacob, W. F. W., Nasir, S. A. M., Yaacob, W. F. W., & Sobri, N. M. (2019). Supervised data mining approach for predicting student performance. *Indonesian Journal of Electrical Engineering and Computer Science*, 16(3), 1584–1592. <https://doi.org/10.11591/ijeecs.v16.i3.pp1584-1592>

Yang, F., & Li, F. W. B. (2018). Study on student performance estimation, student progress analysis, and student potential prediction based on data mining. *Computers and Education*, 123, 97–108.
<https://doi.org/10.1016/j.compedu.2018.04.006>

Zhang, H. (2024). Student Academic Early Warning Prediction based on LSTM networks. *2nd International Conference on Integrated Circuits and Communication Systems, ICICACS 2024*.
<https://doi.org/10.1109/ICICACS60521.2024.10498545>

Peraturan Menteri Riset, Teknologi, dan Pendidikan Tinggi (Permenristekdikti) No. 60 Tahun 2018 tentang Penerimaan Mahasiswa Baru Program Sarjana di PTN.

Peraturan Menteri Pendidikan dan Kebudayaan (Permendikbud) Nomor 6 Tahun 2020 tentang Penerimaan Mahasiswa Baru Program Sarjana pada Perguruan Tinggi Negeri.

Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi (Permendikbudristek) Nomor 48 Tahun 2022 tentang Penerimaan Mahasiswa Baru Program Diploma dan Program Sarjana pada Perguruan Tinggi Negeri.



SEKOLAH PASCASARJANA