

- imbalance problem," *Expert Systems with Applications*, vol. 168, no. 114301, p. 114301, 2021. <https://doi.org/10.1016/j.eswa.2020.114301>.
- [7] A. D. Pozzolo, O. Caelen, and G. Bontempi, "When is undersampling effective in unbalanced classification tasks?," in *Machine Learning and Knowledge Discovery in Databases*, Cham: Springer International Publishing, 2015, pp. 200–215. https://doi.org/10.1007/978-3-319-23528-8_13
- [8] C.-F. Tsai, W.-C. Lin, Y.-H. Hu, and G.-T. Yao, "Under-sampling class imbalanced datasets by combining clustering analysis and instance selection," *Information Sciences*, vol. 477, pp. 47–54, 2019. <https://doi.org/10.1016/j.ins.2018.10.029>
- [9] L. Gong, S. Jiang, L. Bo, L. Jiang, and J. Qian, "A novel class-imbalance learning approach for both within-project and cross-project defect prediction," *IEEE Transactions on Reliability*, vol. 69, no. 1, pp. 40–54, 2020. <https://doi.org/10.1109/TR.2019.2895462>
- [10] D. Devi, S. K. Biswas, B. Purkayastha, "A review on solution to class imbalance problem: Undersampling approaches," In: *International Conference on Computational Performance Evaluation (ComPE)*, IEEE, 2020.
- [11] W. Zheng and M. Jin, "The effects of class imbalance and training data size on classifier learning: An empirical study," *SN Computer Science*, vol. 1, no. 2, 2020. <https://doi.org/10.1007/s42979-020-0074-0>
- [12] H. Guo, J. Zhou, and C. A. Wu, "Ensemble learning via constraint projection and undersampling technique for class-imbalance problem," *Soft Computing* vol. 24, no. 7, pp. 4711–27, 2020.
- [13] B. Liu, and G. Tsoumakas, "Dealing with class imbalance in classifier chains via random undersampling," *Knowledge- Based Systems* vol. 192, no. 105292, p. 105292, 2020.
- [14] T. Pan, J. Chen, J. Xie, Z. Zhou, and S. He "Deep feature generating network: A new method for intelligent fault detection of mechanical systems under class imbalance," *IEEE Transactions on Industrial Informatics* vol. 17, no. 9, pp. 6282–93, 2021.
- [15] J. Guan, X. Jiang, and B. Mao, "A method for class-Imbalance Learning in Android malware detection," *Electronics (Basel)*, vol. 10, no. 24, p. 3124, 2021. <https://doi.org/10.3390/electronics10243124>
- [16] H. Hartono, E. Ongko, and D. Abdullah, "Hybrid approach redefinition with cluster-based instance selection in handling class imbalance problem," *International Journal of Advances in Intelligent Informatics*, vol. 7, no. 3, p. 345, 2021. <https://doi.org/10.26555/ijain.v7i3.515>
- [17] B. Mirzaei, F. Rahmati, and H. Nezamabadi-pour, "A score-based preprocessing technique for class imbalance problems," *Pattern Analysis and Applications*, vol. 25, no. 4, pp. 913–931, 2022. <https://doi.org/10.1007/s10044-022-01084-1>
- [18] R. Walsh and M. Tardy, "A comparison of techniques for class imbalance in deep learning classification of breast cancer," *Diagnostics (Basel)*, vol. 13, no. 1, p. 67, 2022. <https://doi.org/10.3390/diagnostics13010067>
- [19] S. Goyal, "Handling class-imbalance with KNN (neighbourhood) under-sampling for software defect prediction," *Artificial Intelligence Review*, vol. 55, no. 3, pp. 2023–2064, 2022. <https://doi.org/10.1007/s10462-021-10044-w>
- [20] F. Hooshmand, and S. A. MirHassani, "A novel two-phase clustering-based under-sampling method for imbalanced classification problems," *Expert Systems with Applications*, vol. 213, no. 119003, p. 119003, 2023. <https://doi.org/10.1016/j.eswa.2022.119003>
- [21] T. Kohonen, "Essentials of the self-organizing map," *Neural Networks*, vol. 37, pp. 52–65, 2013. <http://dx.doi.org/10.1016/j.neunet.2012.09.018>
- [22] J. S. Shirabad and T. J. Menzies, *The PROMISE repository of software engineering databases*. School of Information Technology and Engineering, University of Ottawa, Canada, 2005. [Online]. Available: <http://promise.site.uottawa.ca/SERepository>
- [23] K. Kaur and A. Kumar, "MCDM-EFS: A novel ensemble feature selection method for software defect prediction using multi-criteria decision making," *Intelligent Decision Technologies*, vol. 17, no. 4, pp. 1283–1296, 2023. <https://doi.org/10.3233/IDT-230251>
- [24] N. A. Bhat and S. U. Farooq, "An improved method for training data selection for cross-project defect prediction," *Arabian Journal for Science & Engineering*, vol. 47, no. 2, pp. 1939–1954, 2022.
- [25] M. Nevendra and P. Singh, "A survey of software defect prediction based on deep learning," *Archives Computational Methods in Engineering*, vol. 29, no. 7, pp. 5723–5748, 2022.
- [26] S. Loov, "Comparison of undersampling methods for prediction of casting defects based on process parameters," Master's thesis, University of Skovde, 2021.
- [27] Kluyver, T. et al., 2016. Jupyter Notebooks – a publishing format for reproducible computational workflows. In F. Loizides & B. Schmidt, eds. *Positioning and Power in Academic Publishing: Players, Agents and Agendas*. pp. 87–90.
- [28] L. Gong, S. Jiang, and L. Jiang, "Tackling class imbalance problem in software defect prediction through cluster-based over-sampling with filtering," *IEEE Access*, vol. 7, pp. 145725–145737, 2019. <https://doi.org/10.1109/ACCESS.2019.2945858>
- [29] M. Alamri and M. Ykhlef, "Survey of Credit Card Anomaly and Fraud Detection Using Sampling Techniques," *Electronics*, vol. 11, no. 23, pp. 4003, 2022.
- [30] P. Kumar, R. Bhatnagar, K. Gaur, and A. Bhatnagar, "Classification of imbalanced data: Review of methods and applications," *IOP Conference Series: Materials Science & Engineering*, vol. 1099, no. 1, p. 012077, 2021.
- [31] Benavoli, G. Corani, J. Demšar, and M. Zaffalon, "Time for a change: a tutorial for comparing multiple classifiers through Bayesian analysis", *The Journal of Machine Learning Research*, vol. 18 no. 1, pp.2653-2688, 2017. <http://jmlr.org/papers/v18/16-305.html>



Ajay Kumar is assistant professor with Department of Information Technology, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India. He completed his Ph.D. in Computer Science and Engineering (CSE) from the University School of Information, Communication & Technology, Guru Gobind Singh Indraprastha University (GGSIPU), New Delhi, India. He received his master degree in CSE from National Institute of Technical Teachers Training and Research, Chandigarh. He has done his Bachelor of Engineering degree in CSE from Dr. B. R. A University Agra, Uttar Pradesh, India. His research fields include software engineering, multi-criteria decision-making, soft computing, artificial intelligence, and machine learning.