significance level of differences indicate significant differences in the parameters of Augmented R-Value for Multi-Class, Class Average Accuracy, and Hamming Loss.

In general, the imbalance ratio and the number of attributes greatly affect the Augmented R-Value for Multi-Class, Class Average Accuracy, and Class Balance Accuracy parameters. As for Hamming Loss, the number of attributes that affect the most is then followed by the imbalance ratio. This study's results indicate a direct relationship between the effect of overlapping and multi-class imbalance on the accuracy of the classification results. Overlapping is more often ignored when compared to class imbalance. However, it can be seen from the results of the study that the higher the overlap (which means that the overlap is more serious), the lower the accuracy of the classification results obtained.

The interesting thing is that for Class Balance Accuracy, although the results show that the Hybrid Approach with Distance Feature gives better results than MultiRandBal, the differences are insignificant. This can be understood because the Class Balance Accuracy is a balance between the accuracy of each existing class. The determination of the sample from the safe region tends to accommodate the handling of multi-class imbalance in the minority class.

IV. Conclusion

Based on the findings in Tables III, IV, and V, it is possible to conclude that both approaches have produced positive outcomes for handling multi-class imbalances. However, the results obtained by the Hybrid Approach with Distance Feature on several parameters are better. There are significant differences in the parameters of Augmented R-Value for Multi-Class, Class Average Accuracy, and Hamming Loss. As for the Class Balance Accuracy parameter, the difference obtained is not significant.

Implementing the Distance Feature in the Hybrid Approach for determining samples in safe regions has proven effective. In addition to dealing with multi-class imbalance problems, it can also handle overlapping. Thus, this study also shows a new approach to the oversampling process in SMOTE. It is hoped that this research can develop methods that can provide better accuracy results on datasets with a large number of attributes.

Acknowledgment

The Ministry of Education, Culture, Research, and Technology of Indonesia supported this study.

References

[1] S. García, Z.-L. Zhang, A. Altalhi, S. Alshomrani, and F. Herrera, "Dynamic ensemble selection for multi-class imbalanced datasets," *Information Sciences*, vol. 445–446, pp. 22–37, Jun. 2018, doi: 10.1016/j.ins.2018.03.002.

[2] M. Temraz and M. T. Keane, "Solving the class imbalance problem using a counterfactual method for data augmentation," *Machine Learning with Applications*, vol. 9, p. 100375, Sep. 2022, doi: 10.1016/j.mlwa.2022.100375.

[3] Y. Zhang, T. Sun, and C. Jiang, "Biomacromolecules as carriers in drug delivery and tissue engineering," *Acta Pharmaceutica Sinica B*, vol. 8, no. 1, pp. 34–50, Jan. 2018, doi: 10.1016/j.apsb.2017.11.005.

[4] X. Chao, G. Kou, Y. Peng, and A. Fernández, "An efficiency curve for evaluating imbalanced classifiers considering intrinsic data characteristics: Experimental analysis," *Information Sciences*, vol. 608, pp. 1131–1156, Aug. 2022, doi: 10.1016/j.ins.2022.06.045.

[5] P. Sadhukhan and S. Palit, "Adaptive learning of minority class prior to minority oversampling," *Pattern Recognition Letters*, vol. 136, pp. 16–24, Aug. 2020, doi: 10.1016/j.patrec.2020.05.020.

[6] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, and G. Bing, "Learning from Class-Imbalanced Data: Review of Methods and Applications," *Expert Systems With Applications*, vol. 73, pp. 220–239, May 2017.

[7] A. Zhang, H. Yu, Z. Huan, X. Yang, S. Zheng, and S. Gao, "SMOTE-RkNN: A hybrid re-sampling method based on SMOTE and reverse k-nearest neighbors," *Information Sciences*, vol. 595, pp. 70–88, May 2022, doi: 10.1016/j.ins.2022.02.038.

[8] M. Koziarski, "Potential Anchoring for imbalanced data classification," *Pattern Recognition*, vol. 120, p. 108114, Dec. 2021, doi: 10.1016/j.patcog.2021.108114.

[9] Z. Chen, J. Duan, L. Kang, and G. Qiu, "A hybrid data-level ensemble to enable learning from highly imbalanced dataset," *Information Sciences*, vol. 554, pp. 157–176, Apr. 2021, doi: 10.1016/j.ins.2020.12.023.

[10] A. S. Desuky and S. Hussain, "An Improved Hybrid Approach for Handling Class Imbalance Problem," *Arab J Sci Eng*, vol. 46, no. 4, pp. 3853–3864, Apr. 2021, doi: 10.1007/s13369-021-05347-7.

[11] T. Pan, J. Zhao, W. Wu, and J. Yang, "Learning imbalanced datasets based on SMOTE and Gaussian distribution," *Information Sciences*, vol. 512, pp. 1214–1233, Feb. 2020, doi: 10.1016/j.ins.2019.10.048.

[12] Q. Li, Y. Song, J. Zhang, and V. S. Sheng, "Multi-class imbalanced learning with one-versus-one decomposition and spectral clustering," *Expert Systems with Applications*, vol. 147, p. 113152, Jun. 2020, doi: 10.1016/j.eswa.2019.113152.

[13] T. R. Hoens, Q. Qian, N. V. Chawla, and Z.-H. Zhou, "Building Decision Trees for the Multi-class Imbalance Problem," in *Advances in Knowledge Discovery and Data Mining*, 2012, pp. 122–134.

[14] J. A. Sáez, B. Krawczyk, and M. Woźniak, "Analyzing the oversampling of different classes and types of examples in multi-class imbalanced datasets," *Pattern Recognition*, vol. 57, pp. 164–178, Sep. 2016, doi: 10.1016/j.patcog.2016.03.012.

[15] D. Elreedy and A. F. Atiya, "A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance," *Information Sciences*, vol. 505, pp. 32–64, Dec. 2019, doi: 10.1016/j.ins.2019.07.070.

[16] A. Fernandez, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary," *1*, vol. 61, pp. 863–905, Apr. 2018.

[17] J. Bi and C. Zhang, "An empirical comparison on state-of-the-art multi-class imbalance learning algorithms and a new diversified ensemble learning scheme," *Knowledge-Based Systems*, vol. 158, pp. 81–93, Oct. 2018, doi: 10.1016/j.knosys.2018.05.037.

[18] M. S. Santos, P. H. Abreu, N. Japkowicz, A. Fernández, and J. Santos, "A unifying view of class overlap and imbalance: Key concepts, multi-view panorama, and open avenues for research," *Information Fusion*, vol. 89, pp. 228–253, Jan. 2023, doi: 10.1016/j.inffus.2022.08.017.

[19] H. K. Lee and S. B. Kim, "An overlap-sensitive margin classifier for imbalanced and overlapping data," *Expert Systems with Applications*, vol. 98, pp. 72–83, May 2018, doi: 10.1016/j.eswa.2018.01.008.

[20] X. Gao *et al.*, "A multi-class classification using one-versus-all approach with the differential partition sampling ensemble," *Engineering Applications of Artificial Intelligence*, vol. 97, p. 104034, Jan. 2021, doi: 10.1016/j.engappai.2020.104034.

[21] B. Chen, S. Xia, Z. Chen, B. Wang, and G. Wang, "RSMOTE: A self-adaptive robust SMOTE for imbalanced problems with label noise," *Information Sciences*, vol. 553, pp. 397–428, Apr. 2021, doi: 10.1016/j.ins.2020.10.013.

[22] V. P. K. Turlapati and M. R. Prusty, "Outlier-SMOTE: A refined oversampling technique for improved detection of COVID-19," *Intelligence-Based Medicine*, vol. 3–4, p. 100023, Dec. 2020, doi: 10.1016/j.ibmed.2020.100023.

[23] K. De Angeli *et al.*, "Class imbalance in out-of-distribution datasets: Improving the robustness of the TextCNN for the classification of rare cancer types," *Journal of Biomedical Informatics*, vol. 125, p. 103957, Jan. 2022, doi: 10.1016/j.jbi.2021.103957.

[24] E. R. Q. Fernandes and A. C. P. L. F. de Carvalho, "Evolutionary inversion of class distribution in overlapping areas for multi-class imbalanced learning," *Information Sciences*, vol. 494, pp. 141–154, Aug. 2019, doi: 10.1016/j.ins.2019.04.052.

[25] N. K. Mishra and P. K. Singh, "Feature construction and smote-based imbalance handling for multi-label learning," *Information Sciences*, vol. 563, pp. 342–357, Jul. 2021, doi: 10.1016/j.ins.2021.03.001.

[26] P. Soltanzadeh and M. Hashemzadeh, "RCSMOTE: Range-Controlled synthetic minority over-sampling technique for handling the class imbalance problem," *Information Sciences*, vol. 542, pp. 92–111, Jan. 2021, doi: 10.1016/j.ins.2020.07.014.

[27] X. Tao *et al.*, "SVDD-based weighted oversampling technique for imbalanced and overlapped dataset learning," *Information Sciences*, vol. 588, pp. 13–51, Apr. 2022, doi: 10.1016/j.ins.2021.12.066.

[28] M. Koziarski, M. Woźniak, and B. Krawczyk, "Combined Cleaning and Re-sampling algorithm for multi-class imbalanced data with label noise," *Knowledge-Based Systems*, vol. 204, p. 106223, Sep. 2020, doi: 10.1016/j.knosys.2020.106223.

[29] N. Nnamoko and I. Korkontzelos, "Efficient treatment of outliers and class imbalance for diabetes prediction," *Artificial Intelligence in Medicine*, vol. 104, p. 101815, Apr. 2020, doi: 10.1016/j.artmed.2020.101815.

[30] Y. Liu, Y. Liu, B. X. B. Yu, S. Zhong, and Z. Hu, "Noise-robust oversampling for imbalanced data classification," *Pattern Recognition*, vol. 133, p. 109008, Jan. 2023, doi: 10.1016/j.patcog.2022.109008.

[31] J. J. Rodríguez, J.-F. Díez-Pastor, Á. Arnaiz-González, and L. I. Kuncheva, "Random Balance ensembles for multi-class imbalance learning," *Knowledge-Based Systems*, vol. 193, p. 105434, Apr. 2020, doi: 10.1016/j.knosys.2019.105434.

[32] P. Vuttipittayamongkol and E. Elyan, "Neighbourhood-based undersampling approach for handling imbalanced and overlapped data," *Information Sciences*, vol. 509, pp. 47–70, Jan. 2020, doi: 10.1016/j.ins.2019.08.062.

[33] Q. Chen, Z.-L. Zhang, W.-P. Huang, J. Wu, and X.-G. Luo, "PF-SMOTE: A novel parameter-free SMOTE for imbalanced datasets," *Neurocomputing*, vol. 498, pp. 75–88, Aug. 2022, doi: 10.1016/j.neucom.2022.05.017.

[34] T. G.s., Y. Hariprasad, S. S. Iyengar, N. R. Sunitha, P. Badrinath, and S. Chennupati, "An extension of Synthetic Minority Oversampling Technique based on Kalman filter for imbalanced datasets," *Machine Learning with Applications*, vol. 8, p. 100267, Jun. 2022, doi: 10.1016/j.mlwa.2022.100267.

[35] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 4, pp. 463–484, Jul. 2012, doi: 10.1109/TSMCC.2011.2161285.

[36] A. Arafa, N. El-Fishawy, M. Badawy, and M. Radad, "RN-SMOTE: Reduced Noise SMOTE based on DBSCAN for enhancing imbalanced data classification," *Journal of King Saud University - Computer and Information Sciences*, Jun. 2022, doi: 10.1016/j.jksuci.2022.06.005.

[37] F. Charte, A. Rivera, M. J. del Jesus, and F. Herrera, "A First Approach to Deal with Imbalance in Multi-label Datasets," in *Hybrid Artificial Intelligent Systems*, Berlin, Heidelberg, 2013, pp. 150–160. doi: 10.1007/978-3-642-40846-5_16.

[38] S. Ruuska, W. Hämäläinen, S. Kajava, M. Mughal, P. Matilainen, and J. Mononen, "Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle," *Behavioural Processes*, vol. 148, pp. 56–62, Mar. 2018, doi: 10.1016/j.beproc.2018.01.004.

[39] P. Branco, L. Torgo, and R. P. Ribeiro, "Relevance-Based Evaluation Metrics for Multi-class Imbalanced Domains," in *Advances in Knowledge Discovery and Data Mining*, Cham, 2017, pp. 698–710. doi: 10.1007/978-3-319-57454-7_54.

[40] L. Mosley, "A balanced approach to the multi-class imbalance problem," *Graduate Theses and Dissertations*, Jan. 2013, doi: https://doi.org/10.31274/etd-180810-3375.

[41] N. K. Mishra and P. K. Singh, "FS-MLC: Feature selection for multi-label classification using clustering in feature space," *Information Processing & Management*, vol. 57, no. 4, p. 102240, Jul. 2020, doi: 10.1016/j.ipm.2020.102240.

[42] A. Frank and A. Asuncion, "UCI Machine Learning Repository." University of California, School of Information and Computer Science, 2010. [Online]. Available: https://archive.ics.uci.edu/ml/index.php.

[43] F. Wilcoxon, "Individual Comparisons by Ranking Methods on JSTOR," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.