Journal of Artificial Intelligence and Emerging Technologies
Volume 3, Issue 9, pp 1-9, September-2026
ISSN (online): 3049-1967
DOI: https://doi.org/10.47001/JAIET/2026.309001
Available Online at https://www.jaiet.com

Adaptive Ensemble Weighting with Local Competence for Multi-Class Imbalanced Classification

[1]S. Obe, *[2]D. Matthias, *[3]E. O. Bennett

[1,2,3]Department of Computer Science, Rivers State University, Port-Harcourt, Nigeria
*E-mail: [1]sampson.obe@rsu.edu.ng, [2]daniel.matthias@ust.edu.ng, [3]bennett.okoni@ust.edu.ng

Abstract: Multi-class imbalanced classification remains difficult because minority classes can be poorly recognised even when aggregate performance appears acceptable. Many imbalance-handling methods still rely on a fixed technique, a single selected technique, or one level of adaptation, despite the fact that technique suitability changes with both dataset characteristics and local decision regions. This article presents Hybrid Adaptive Ensemble Weighting with Dynamic Ensemble Selection for multi-class imbalanced classification. The model combines two sources of evidence: dataset-level suitability estimated from meta-features, and instance-level competence estimated around each test sample. Adaptive Ensemble Weighting uses kernel similarity over 13 meta-features to estimate the suitability of 10 imbalance-handling technique pipelines, while Dynamic Ensemble Selection measures local classifier competence from validation-neighbourhood behaviour. Both weight vectors are then combined during prediction. Evaluation was carried out on 30 OpenML benchmark datasets with 3 to 20 classes, 160 to 10,000 instances, 4 to 1,300 features, and imbalance ratios from 1.9:1 to 567.2:1. Under Leave-One-Dataset-Out validation, using macro F1-score as the primary metric, Hybrid AEW-DES achieved the highest average macro F1-score of 0.7308 and the best average rank of 3.83 among the 10 evaluated methods. AEW ranked second with an average macro F1-score of 0.7245 and exceeded the Oracle single-technique reference score of 0.7188. The Friedman test confirmed significant differences among methods ($p = 0.0020$). Wilcoxon signed-rank tests showed significant AEW improvements over Baseline, SingleSelect, KNORA-E, and Stacking, while Hybrid AEW-DES significantly outperformed DES-LA and KNORA-E. These results indicate that combining dataset-level meta-feature weighting with instance-level local competence can improve adaptive technique combination within the evaluated benchmark setting.

Keywords: Multi-class imbalanced classification; adaptive ensemble weighting; dynamic ensemble selection; meta-learning; kernel similarity; meta-features; macro F1-score.

I. INTRODUCTION

Multi-class classification assigns each instance to one of three or more possible classes. It is common in medical diagnosis, fault detection, intrusion detection, document classification, remote sensing, and other applied domains where decisions rarely fall into only two categories. The task becomes more difficult when the classes are unevenly distributed. In such cases, a classifier may learn the majority classes well while giving weak recognition to minority classes that may be important for decision-making [1], [2].

Multi-class imbalance is more complex than binary imbalance because one dataset can contain several minority classes, multiple majority-minority relationships, and overlapping decision regions [3]. This also makes evaluation harder. Overall accuracy can look strong even when rare classes are being missed, so macro F1-score is more informative because each class contributes equally to the final score.

Imbalanced learning has produced several families of techniques. Data-level methods such as SMOTE, ADASYN, Borderline-SMOTE, SMOTE-Tomek, SMOTE-ENN, and Random Undersampling alter the training distribution before classification [4], [5], [6], [7]. Algorithm-level methods such as cost-sensitive learning change the training objective so that minority-class errors receive greater penalty [8]. Ensemble methods such as EasyEnsemble, BalancedBagging, and BalancedRandomForest reduce majority-class dominance by building multiple models or training on balanced subsets [9], [10].

The practical difficulty is choosing which technique to use. No single imbalance-handling method is consistently best across heterogeneous datasets. A method that performs well on a small dataset with mild imbalance may fail on a high-dimensional dataset with severe imbalance, many classes, or strong overlap. Trial-and-error selection is therefore inefficient and often unreliable. Meta-learning offers a more systematic route by

Journal of Artificial Intelligence and Emerging Technologies
Volume 3, Issue 9, pp 1-9, September-2026
ISSN (online): 3049-1967
DOI: https://doi.org/10.47001/JAIET/2026.309001
Available Online at https://www.jaiet.com

relating dataset characteristics to algorithm behaviour [11], [12]. Dynamic Ensemble Selection addresses the problem from another angle by estimating which classifiers are competent around each test instance [13], [14].

Hybrid Adaptive Ensemble Weighting with Dynamic Ensemble Selection, abbreviated as Hybrid AEW-DES, brings these two forms of adaptation together. It uses meta-features to compute dataset-level weights for several imbalance-handling technique pipelines, then refines prediction with instance-level local competence. Its main contribution is a two-level adaptive weighting framework that combines global dataset suitability and local classifier competence for multi-class imbalanced classification.

II. RELATED WORK

Imbalanced learning methods are commonly grouped into data-level, algorithm-level, and ensemble-based approaches. SMOTE creates synthetic minority samples by interpolating between existing minority instances [4]. Borderline-SMOTE shifts generation toward samples near decision boundaries [6], while ADASYN places more synthetic samples in regions that appear difficult to learn [7]. Hybrid methods such as SMOTE-Tomek and SMOTE-ENN combine oversampling with data cleaning to reduce ambiguous boundary samples [5].

Algorithm-level methods handle imbalance inside the learner itself. Cost-sensitive learning assigns higher penalties to errors involving minority classes, encouraging the classifier to pay more attention to under-represented classes [8]. Support vector machine variants, weighted loss functions, and class-weighted tree models follow this general principle [15].

Ensemble methods take a different route by combining learners trained under different sampling or weighting conditions. EasyEnsemble trains several learners on balanced majority-class subsets [9]. BalancedBagging and BalancedRandomForest integrate class balancing into bagging and random forest construction, respectively [10]. Recent surveys show that ensemble and data-augmentation strategies are widely used, but their effectiveness still depends on dataset structure and evaluation setting [2], [16], [17].

Meta-learning treats algorithm choice as a learning problem. Rice [18] formalised the algorithm selection problem, while later work showed that datasets can be represented through measurable descriptors such as sample size, feature count, class entropy, imbalance ratio, class overlap, and landmarking performance [11], [12]. In imbalanced classification, these descriptors can help estimate which technique is likely to work for a new dataset. Kernel-based similarity provides one way to transfer performance evidence from historical datasets to a new target dataset [19]. Dynamic Ensemble Selection adapts at prediction time. Instead of applying one fixed ensemble rule to every sample, DES estimates classifier competence in the neighbourhood of the current test instance. Local accuracy estimation was introduced as a basis for classifier combination by Woods et al. [14], and DESlib later provided practical implementations of methods such as KNORA-E, KNORA-U, and DES-LA [13]. Recent work has also applied DES to multi-class imbalanced settings, supporting its relevance for local adaptation [20].

What remains less developed is a single weighting model that uses both dataset-level technique suitability and instance-level local competence. Meta-learning methods usually recommend or weight methods at the dataset level, while DES methods adapt locally without first using global imbalance characteristics. Hybrid AEW-DES links these two sources of evidence in one prediction framework.

III. METHOD

Hybrid AEW-DES evaluates a pool of imbalance-handling technique pipelines and combines their predictions through adaptive weights. Figure 1 gives the architecture. First, the current dataset is described with meta-features and compared with historical datasets in a meta-knowledge base. AEW uses the resulting similarity scores and historical technique performance to compute dataset-level weights. After the technique pipelines are trained, DES estimates local competence for each test instance, and the hybrid combiner produces the final weighted prediction.

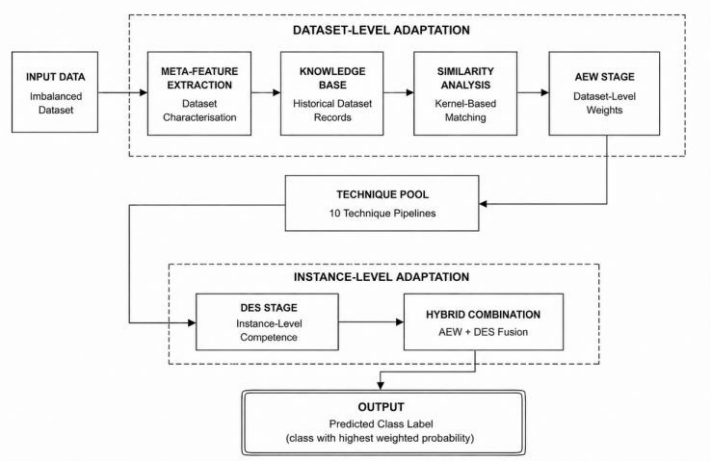

Figure 1: Hybrid AEW-DES Architecture

Journal of Artificial Intelligence and Emerging Technologies
Volume 3, Issue 9, pp 1-9, September-2026
ISSN (online): 3049-1967
DOI: https://doi.org/10.47001/JAIET/2026.309001
Available Online at https://www.jaiet.com

Dataset and Technique Pool

The evaluation used 30 benchmark datasets obtained from OpenML [22]. They covered heterogeneous tabular classification problems with 3 to 20 classes, 160 to 10,000 instances, 4 to 1,300 features, and imbalance ratios from 1.9:1 to 567.2:1. Table 1 summarises the dataset coverage.

Table 1: Benchmark Dataset Coverage

Characteristic	Observed Range or Count
Number of datasets	30
Dataset source	OpenML
Instances	160 to 10,000
Features	4 to 1,300
Classes	3 to 20
Imbalance ratio	1.9:1 to 567.2:1

The technique pool contained 10 imbalance-handling pipelines: SMOTE, SMOTE-Tomek, SMOTE-ENN, ADASYN, Borderline-SMOTE, Random Undersampling, Cost-Sensitive Random Forest, EasyEnsemble, BalancedBagging, and BalancedRandomForest. Where a separate classifier was required, Random Forest was used with 100 estimators and random seed 42.

Meta-Feature Extraction

Each dataset was represented with 13 meta-features: imbalance ratio, number of classes, sample size, number of features, feature-to-data ratio, class entropy, Gini imbalance, class overlap score, mean correlation, mean skewness, mean kurtosis, 1-nearest-neighbour landmarking score, and decision-stump landmarking score. Together, these descriptors capture imbalance severity, class structure, distributional behaviour, data complexity, and simple landmarking performance.

Adaptive Ensemble Weighting

In each Leave-One-Dataset-Out run, the left-out target dataset was denoted as D_t, with meta-feature vector $m(D_t)$. Each historical dataset in the meta-knowledge base was denoted as D_j, with meta-feature vector $m(D_j)$. Similarity between the target dataset and each historical dataset was computed with the RBF kernel in Equation (1):

$$S_j = \exp\left(-\frac{\| m(D_t) - m(D_j) \|^2}{2\sigma^2}\right) \quad (1)$$

where S_j is the similarity score for historical dataset j, and σ is the kernel width estimated using the median distance heuristic.

AEW then estimated technique suitability from historical performance on similar datasets, as shown in Equation (2):

$$w_k^{AEW} = \frac{\sum_{j=1}^{M} S_j \, P_{j,k}}{\sum_{j=1}^{M} S_j} \quad (2)$$

where w_k^{AEW} is the dataset-level weight for technique k, $P_{j,k}$ is the recorded macro F1-score of technique k on historical dataset j, and M is the number of historical datasets in the current meta-knowledge base. The weights were normalised before prediction.

Hybrid AEW-DES Prediction

DES estimated instance-level competence using validation samples near each test instance. For a test instance x_i, the DES component produced a local competence vector w_i^{DES}. Dataset-level and local weights were combined using Equation (3):

$$w_i^H = \alpha w^{AEW} + (1 - \alpha) w_i^{DES} \quad (3)$$

where w_i^H is the final hybrid weight vector for test instance x_i, w^{AEW} is the dataset-level AEW weight vector, w_i^{DES} is the DES local competence vector, and $\alpha = 0.5$.

The 10 technique pipelines then supplied class-probability outputs, which were aggregated according to Equation (4):

$$p(c|x_i) = \sum_{k=1}^{K} w_{i,k}^H \, p_k(c|x_i) \quad (4)$$

where $p_k(c|x_i)$ is the probability assigned to class c by technique pipeline k, and $K = 10$. The class with the highest weighted probability became the final prediction.

Evaluation Protocol

Leave-One-Dataset-Out validation tested generalisation across datasets. In each iteration, one dataset was held out as the target dataset and the remaining 29 datasets formed the meta-knowledge base. The held-out dataset was split into training, validation, and test partitions. Training data fitted the technique pipelines, validation data supported DES competence estimation, and test data produced the final performance score.

Macro F1-score was used as the primary metric and was computed using Equation (5):

Journal of Artificial Intelligence and Emerging Technologies
Volume 3, Issue 9, pp 1-9, September-2026
ISSN (online): 3049-1967
DOI: https://doi.org/10.47001/JAIET/2026.309001
Available Online at https://www.jaiet.com

$$\text{Macro F1} = \frac{1}{C}\sum_{c=1}^{C} F1_c \quad (5)$$

where C is the number of classes and $F1_c$ is the F1-score for class c. The comparison included Baseline, SingleSelect, MajorityVote, KNORA-E, KNORA-U, DES-LA, Stacking, AEW, Hybrid AEW-DES, and Oracle. Oracle was treated as a best single-technique reference, not as an upper bound for weighted ensembles. Friedman and Wilcoxon signed-rank tests were used for statistical comparison across datasets [21].

IV. RESULTS AND DISCUSSION

Individual Technique Performance

Table 2 reports the average macro F1-score of each individual imbalance-handling technique across the 30 datasets. Borderline-SMOTE had the highest average score, 0.7327, but led on only 9 datasets. SMOTE was best on 7 datasets, while CostSensitive was best on 4. Several other techniques led only on small subsets of the benchmark suite. This spread, also shown in Figure 2, makes fixed technique choice unreliable for heterogeneous multi-class imbalanced datasets.

Table 2: Average Macro F1-Score of Individual Techniques

Technique	Category	Avg F1	Std Dev	Best on N Datasets
Borderline-SMOTE	Data-level	0.7327	0.218	9
SMOTE	Data-level	0.7313	0.222	7
SMOTE-Tomek	Data-level	0.7300	0.219	2
CostSensitive	Algorithm-level	0.7221	0.225	4
BalancedRandomForest	Ensemble	0.6977	0.233	2
BalancedBagging	Ensemble	0.6760	0.244	2
RandomUndersampling	Data-level	0.6592	0.255	0
SMOTE-ENN	Data-level	0.6489	0.251	2
EasyEnsemble	Ensemble	0.5268	0.243	0
ADASYN	Data-level	0.3713	0.378	2

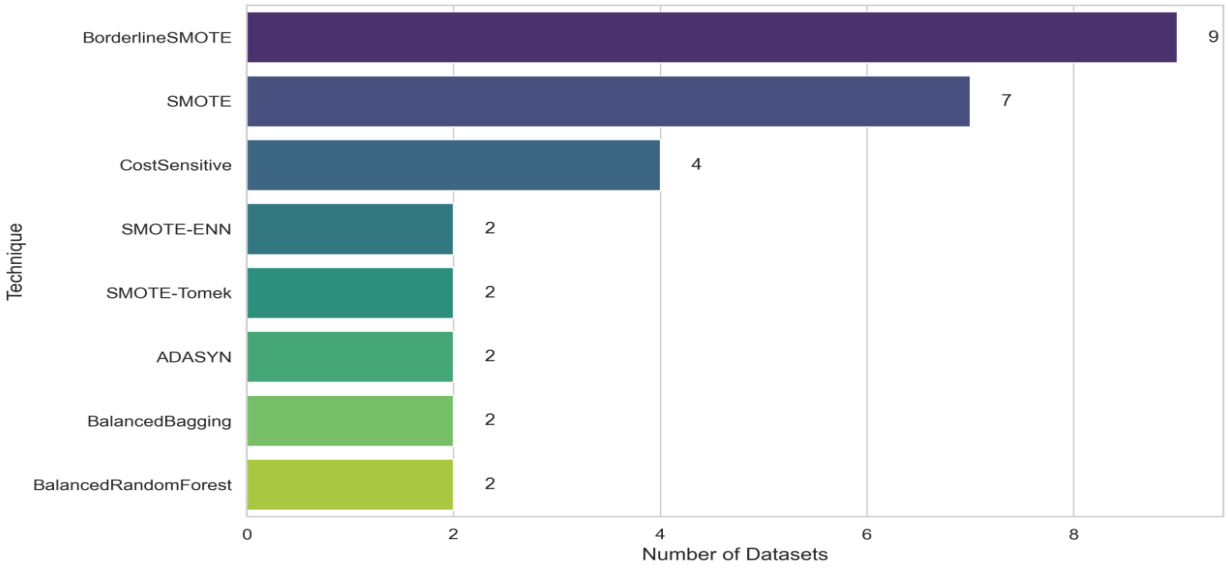

Figure 2: Distribution of Best Techniques Across Datasets

ADASYN recorded the weakest average performance, despite being designed to generate samples in difficult minority regions. A likely explanation is that difficult regions may contain noise or strong class overlap, so extra synthetic samples can increase ambiguity rather than improve representation. This supports weighting techniques according to dataset evidence rather than selecting one method by default.

Journal of Artificial Intelligence and Emerging Technologies
Volume 3, Issue 9, pp 1-9, September-2026
ISSN (online): 3049-1967
DOI: https://doi.org/10.47001/JAIET/2026.309001
Available Online at https://www.jaiet.com

Method Comparison

Table 3 presents the Leave-One-Dataset-Out comparison. Hybrid AEW-DES achieved the highest average macro F1-score, 0.7308. AEW ranked second with 0.7245, followed closely by KNORA-U and DES-LA. The two adaptive weighting methods therefore occupied the top two positions by average score.

Table 3: Leave-One-Dataset-Out Evaluation Results

Method	Avg F1	Std Dev	Median	Min	Max
Hybrid AEW-DES	0.7308	0.217	0.7866	0.224	1.000
AEW	0.7245	0.218	0.7672	0.222	1.000
KNORA-U	0.7237	0.213	0.7733	0.222	1.000
DES-LA	0.7219	0.214	0.7722	0.217	1.000
KNORA-E	0.7199	0.215	0.7677	0.219	1.000
MajorityVote	0.7188	0.214	0.7809	0.218	1.000
Oracle	0.7188	0.225	0.7726	0.212	1.000
Baseline	0.7114	0.223	0.7512	0.212	1.000
Stacking	0.7010	0.232	0.7479	0.161	1.000
SingleSelect	0.6703	0.260	0.7062	0.000	1.000

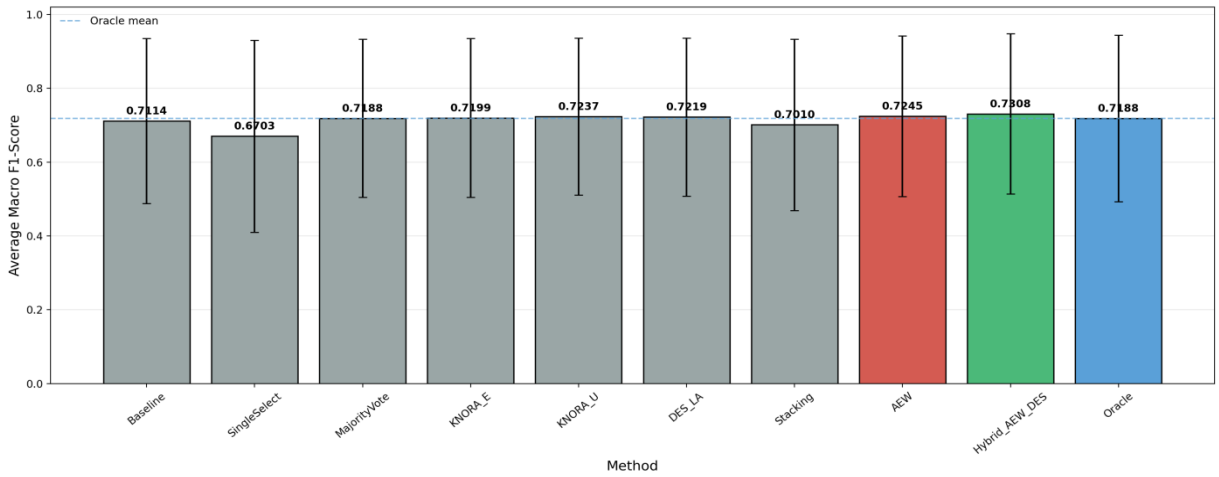

Figure 3: LODO Method Performance Comparison

Two points stand out from the comparison. First, AEW exceeded the Oracle single-technique reference, with 0.7245 compared with 0.7188. Since Oracle represents the best single technique under the evaluation setting, the result shows that weighted combination can outperform selecting one technique. Second, Hybrid AEW-DES improved over AEW by 0.0063, indicating that local competence added useful instance-level information to dataset-level technique weighting.

Statistical Analysis

The Friedman test showed significant differences among the 10 methods, with $\chi^2 = 28.47$, 9 degrees of freedom, and $p = 0.0020$. This rejects the null hypothesis of equivalent performance across the 30 datasets.

Wilcoxon signed-rank tests showed that AEW significantly outperformed Baseline ($p = 0.0013$), SingleSelect ($p = 0.0059$), KNORA-E ($p = 0.0390$), and Stacking ($p = 0.0100$). Its differences from MajorityVote, KNORA-U, and DES-LA were not significant.

Hybrid AEW-DES significantly outperformed DES-LA ($p = 0.0410$) and KNORA-E ($p = 0.0285$), although its difference from KNORA-U was not significant. Average ranks followed the same pattern: Hybrid AEW-DES ranked first at 3.83, followed by AEW at 4.17 and KNORA-U at 4.43. Figure 4 summarises this ranking pattern with the critical difference diagram.

Journal of Artificial Intelligence and Emerging Technologies
Volume 3, Issue 9, pp 1-9, September-2026
ISSN (online): 3049-1967
DOI: https://doi.org/10.47001/JAIET/2026.309001
Available Online at https://www.jaiet.com

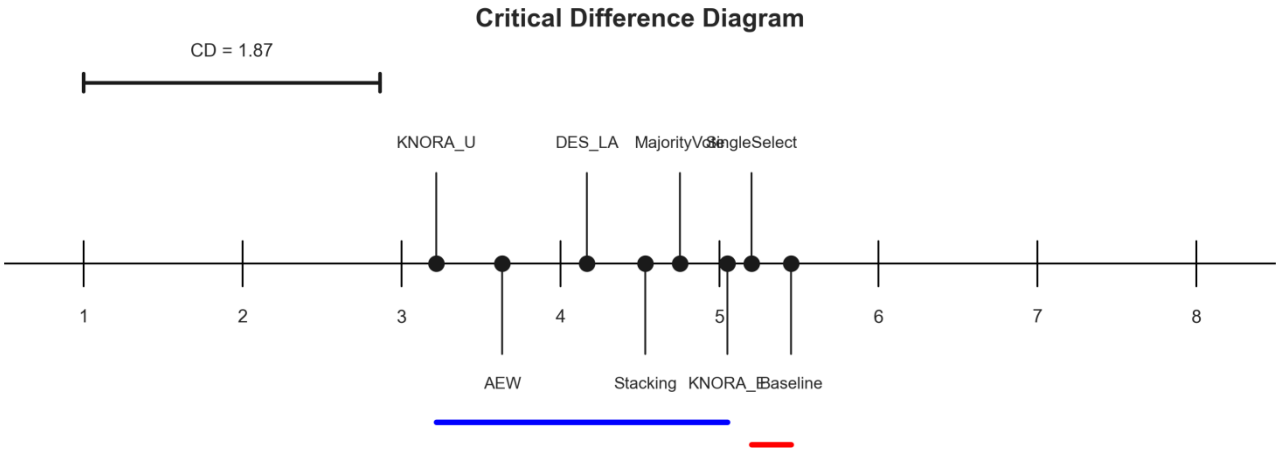

Figure 4: Critical Difference Diagram

Kernel Sensitivity

Table 4 reports the kernel-function ablation for AEW. RBF and Laplacian both achieved an average macro F1-score of 0.7359, while Cosine achieved 0.7358. With a maximum difference of only 0.0001, the meta-feature representation appears to provide stable similarity information across kernel choices. Figure 5 visualises the near-identical kernel performance.

Table 4: Kernel Function Comparison

Kernel	Avg F1	Std Dev	Best or Tied on N Datasets
RBF	0.7359	0.214	28
Laplacian	0.7359	0.214	27
Cosine	0.7358	0.214	25

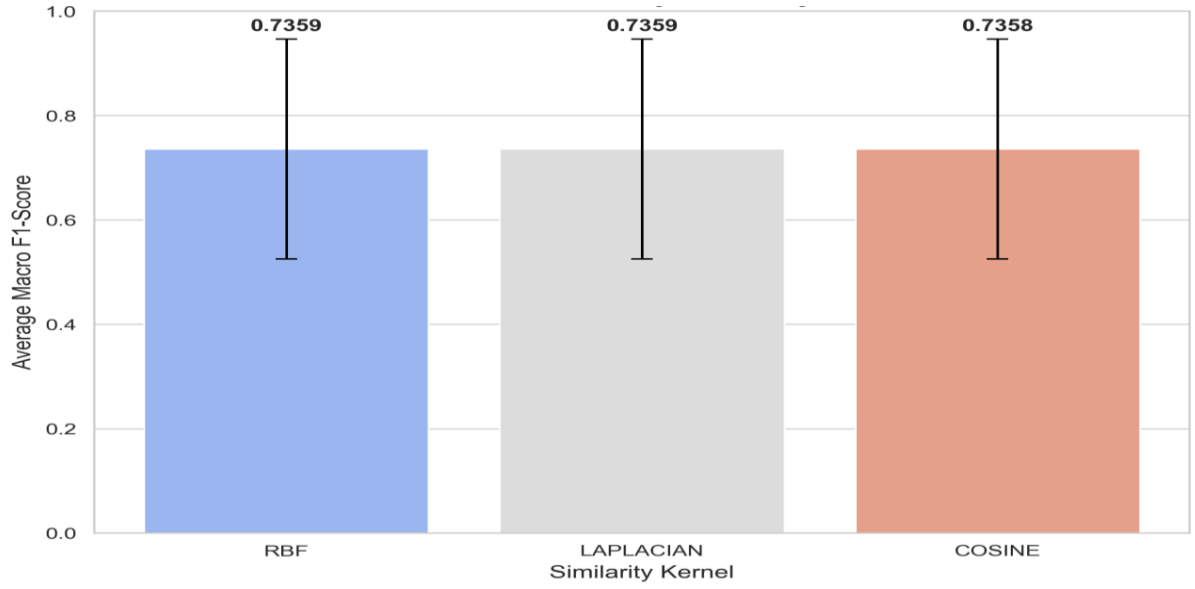

Figure 5: Kernel Function Performance Comparison

Journal of Artificial Intelligence and Emerging Technologies
Volume 3, Issue 9, pp 1-9, September-2026
ISSN (online): 3049-1967
DOI: https://doi.org/10.47001/JAIET/2026.309001
Available Online at https://www.jaiet.com

Per-Dataset Behaviour

Selected per-dataset results show strong Hybrid AEW-DES performance on several easier or moderately difficult datasets. It achieved macro F1-scores of 1.000 on cardiotocograph, 0.997 on analcatdata_aut, 0.992 on wall-robot-navi, 0.950 on dna, 0.933 on collins, and 0.895 on satimage. Some datasets remained difficult: yeast recorded 0.604, cmc recorded 0.530, and RandomRBF_50_1E_159 recorded 0.224. Figure 6 presents the wider per-dataset result pattern across methods.

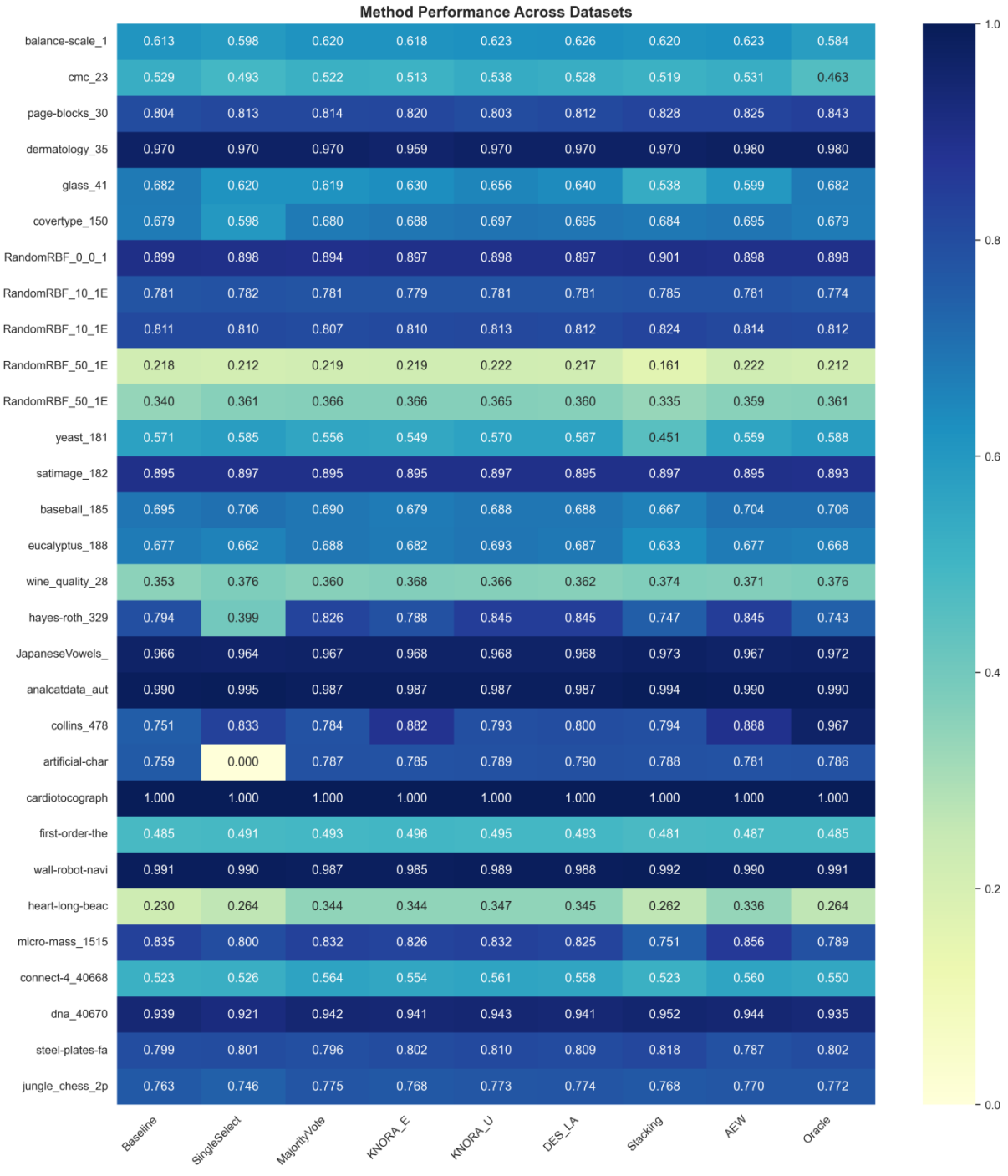

Figure 6: Per-Dataset Macro F1-Score Heatmap

Journal of Artificial Intelligence and Emerging Technologies
Volume 3, Issue 9, pp 1-9, September-2026
ISSN (online): 3049-1967
DOI: https://doi.org/10.47001/JAIET/2026.309001
Available Online at https://www.jaiet.com

These lower scores were also visible for competing methods, which indicates that the difficult datasets were challenging across the full comparison, not only for Hybrid AEW-DES. The pattern also marks a realistic limit: adaptive weighting can improve method selection and combination, but it does not remove the effect of severe overlap, sparse minority classes, or weak local separability.

Overall, the results support the value of combining dataset-level and instance-level adaptation. AEW identifies techniques that suit the dataset as a whole, while DES adjusts technique influence around each test instance. The gains are modest in absolute value, but they occur in a competitive setting where the strongest baselines already perform closely.

V. CONCLUSION

This article presented Hybrid AEW-DES, an adaptive weighting model for multi-class imbalanced classification. The model uses 13 dataset meta-features to compute dataset-level technique weights through kernel similarity, then combines those weights with instance-level DES competence during prediction. By joining these two views, the final prediction is guided by both the general behaviour of similar datasets and the local competence of classifiers around each test sample.

Experiments on 30 OpenML benchmark datasets showed that Hybrid AEW-DES achieved the highest average macro F1-score of 0.7308 and the best average rank of 3.83 among the 10 evaluated methods. AEW ranked second with 0.7245 and exceeded the Oracle single-technique reference score of 0.7188. Statistical testing confirmed significant overall method differences, while pairwise tests showed significant gains for AEW over several competing methods and for Hybrid AEW-DES over selected DES baselines. These findings suggest that meta-feature-guided weighting is useful for heterogeneous imbalanced datasets, and that local competence can refine this weighting at prediction time. Future work should evaluate larger and domain-specific dataset collections, adaptive tuning of the hybrid parameter, additional base classifiers, expanded meta-feature sets, and online or streaming imbalanced classification settings.

REFERENCES

[1] W. Chen, K. Yang, Z. Yu, Y. Shi, and C. L. P. Chen, "A survey on imbalanced learning: Latest research, applications and future directions," *Artificial Intelligence Review*, vol. 57, Article 162, 2024. https://doi.org/10.1007/s10462-024-10759-6

[2] M. Altalhan, A. Algarni, and M. Turki-Hadj Alouane, "Imbalanced data problem in machine learning: A review," *IEEE Access*, vol. 13, pp. 13686-13699, 2025. https://doi.org/10.1109/ACCESS.2025.3531662

[3] Y. Wang, M. M. Rosli, N. Musa, and F. Li, "Multi-class imbalanced data classification: A systematic mapping study," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14183-14190, 2024. https://doi.org/10.48084/etasr.7206

[4] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, 2002. https://doi.org/10.1613/jair.953

[5] G. E. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 20-29, 2004. https://doi.org/10.1145/1007730.1007735

[6] H. Han, W. Y. Wang, and B. H. Mao, "Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning," *in Advances in Intelligent Computing. ICIC 2005. Lecture Notes in Computer Science*, vol. 3644, pp. 878-887, Springer, 2005. https://doi.org/10.1007/11538059_91

[7] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," *in 2008 IEEE International Joint Conference on Neural Networks*, pp. 1322-1328, IEEE, 2008. https://doi.org/10.1109/IJCNN.2008.4633969

[8] I.Araf, A. Idri, and I. Chairi, "Cost-sensitive learning for imbalanced medical data: A review," *Artificial Intelligence Review*, vol. 57, Article 80, 2024. https://doi.org/10.1007/s10462-023-10652-8

[9] X. Y. Liu, J. Wu, and Z. H. Zhou, "Exploratory undersampling for class-imbalance learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 2, pp. 539-550, 2009. https://doi.org/10.1109/TSMCB.2008.2007853

[10] G. Lemaitre, F. Nogueira, and C. K. Aridas, "Imbalanced-learn: A Python toolbox to tackle the curse of imbalanced datasets in machine learning," *Journal of Machine Learning Research*, vol. 18, no. 17, pp. 1-5, 2017. https://jmlr.org/papers/v18/16-365.html

[11] J. Vanschoren, "Meta-learning: A survey," *arXiv preprint arXiv*:1810.03548, 2018. https://arxiv.org/abs/1810.03548

[12] A.Rivolli, L. P. F. Garcia, C. Soares, J. Vanschoren, and A. C. P. L. F. de Carvalho, "Meta-features for meta-

Journal of Artificial Intelligence and Emerging Technologies
Volume 3, Issue 9, pp 1-9, September-2026
ISSN (online): 3049-1967
DOI: https://doi.org/10.47001/JAIET/2026.309001
Available Online at https://www.jaiet.com

learning," *Knowledge-Based Systems*, vol. 240, Article 108101, 2022. https://doi.org/10.1016/j.knosys.2021.108101

[13] R. M. O. Cruz, L. G. Hafemann, R. Sabourin, and G. D. C. Cavalcanti, "DESlib: A dynamic ensemble selection library in Python," *Journal of Machine Learning Research*, vol. 21, no. 8, pp. 1-5, 2020. https://jmlr.org/papers/v21/18-144.html

[14] K. Woods, W. P. Kegelmeyer, and K. Bowyer, "Combination of multiple classifiers using local accuracy estimates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 4, pp. 405-410, 1997. https://doi.org/10.1109/34.588027

[15] S. Rezvani, F. Pourpanah, C. P. Lim, and Q. M. J. Wu, "Methods for class-imbalanced learning with support vector machines: A review and an empirical evaluation," *Soft Computing*, vol. 28, pp. 11873-11894, 2024. https://doi.org/10.1007/s00500-024-09931-5

[16] A.A. Khan, O. Chaudhari, and R. Chandra, "A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation," *Expert Systems with Applications*, vol. 244, Article 122778, 2024. https://doi.org/10.1016/j.eswa.2023.122778

[17] Y. Yang, H. A. Khorshidi, and U. Aickelin, "A review on over-sampling techniques in classification of multi-class imbalanced datasets: Insights for medical problems," *Frontiers in Digital Health*, vol. 6, Article 1430245, 2024. https://doi.org/10.3389/fdgth.2024.1430245

[18] J. R. Rice, "The algorithm selection problem," *Advances in Computers*, vol. 15, pp. 65-118, 1976. https://doi.org/10.1016/S0065-2458(08)60520-3

[19] L. Deng and M. Xiao, "Hyperparameter recommendation via automated meta-feature selection embedded with kernel group Lasso learning," *Knowledge-Based Systems*, vol. 306, Article 112706, 2024. https://doi.org/10.1016/j.knosys.2024.112706

[20] K. Aziz, F. Chen, M. Ahmad, M. S. Khan, M. M. Sabri Sabri, and H. Almujibah, "An interpretable dynamic ensemble selection multiclass imbalance approach with ensemble imbalance learning for predicting road crash injury severity," *Scientific Reports*, vol. 15, Article 24666, 2025. https://doi.org/10.1038/s41598-025-08935-x

[21] J. Demsar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine Learning Research*, vol. 7, pp. 1-30, 2006. https://jmlr.org/papers/v7/demsar06a.html

[22] J. Vanschoren, J. N. van Rijn, B. Bischl, and L. Torgo, "OpenML: Networked science in machine learning," *ACM SIGKDD Explorations Newsletter*, vol. 15, no. 2, pp. 49-60, 2014. https://doi.org/10.1145/2641190.2641198.

Citation of this Article:

S. Obe, D. Matthias, & E. O. Bennett. (2026). Adaptive Ensemble Weighting with Local Competence for Multi-Class Imbalanced Classification. *Journal of Artificial Intelligence and Emerging Technologies (JAIET).* 3(9), 1-9. Article DOI: https://doi.org/10.47001/JAIET/2026.309001

***** End of the Article *****