Supplementary Document of "Phage Host Prediction Using Deep Neural Network with Multi-source Protein Language Models and Squeeze-and-Excitation Attention Mechanism"

Peng Gao, Long Xu, Yuan Bai, Qiuzhen Lin, *Member, IEEE*, Junkai Ji, and Lijia Ma, *Member, IEEE*

This is the supplementary document to the paper entitled "Phage Host Prediction Using Deep Neural Network with Multi-source Protein Language Models and Squeeze-and-Excitation Attention Mechanism" and submitted to IEEE Journal of Biomedical and Health Informatics. The supplementary document includes performance comparisons, statistical analyses, impact of parameter settings, and additional results that further validate the robustness and effectiveness of the proposed PHPRBP model.

I. Comparison of Prediction Performance and Class Distributions

In this part, we first present the performance comparison of PHPRBP and baseline methods on 5-fold cross-validation (see Table S1). Next, the class distributions for Datasets 2 and 3 are shown (see Fig. S1). We then compare the performance of PHPRBP and baseline methods at different taxonomic level across three datasets (see Table S2). Following this, we introduce various pre-trained PLMs and present their performance along with PHPRBP on 5-fold cross-validation (see Table S3 and Table S4). We also examine the performance of PHPRBP with different class imbalance mitigation strategies (see Table S5), followed by a comparison of the model's performance before and after ADASYN data augmentation for specific minority classes (see Table S6). Finally, we present an ablation study of PHPRBP and its variants based on 5-fold cross-validation (see Table S7).

II. Statistical Analysis of Prediction Performance Comparisons

We conduct paired sample t-tests of PHPRBP with the baseline methods on five evaluation metrics across the three datasets, and the results are shown in Tables S8, S9, and S10. All statistical analyses are based on the results of 5-fold cross-validation with 10 replications, and the validity of the t-tests is ensured by verifying that the data meets the normality assumption using the Shapiro-Wilk test. These test results reveal the statistical significance of the performance improvement and confirm the robustness of the PHPRBP model. In all comparisons, PHPPRBP shows significant performance improvement on most metrics, demonstrating the superiority of its model. For example, PHPRBP achieves statistically significant advantages over all the baseline methods on all metrics in the experiments on Dataset 1 (p-values are all much less than 0.05). However, in some cases, such as the MSRF and MSXGB models in Dataset 2, PHPRBP's performance on ACC, Precision, and Sensitivity fails to reach statistical significance. This may be due to the lower diversity in Dataset 2, which affects the model's adaptability and stability. Notably, PHPRBP remains statistically significance on two comprehensive metrics, MCC and F1.

III. Impact of Hyperparameter Settings on Model Performance

In this part, we analyze the impacts of key parameters on the host prediction performance of PHPRBP, focusing on hyperparameters such as batch size n_b and reduction factor r. We first test various batch sizes n_b (16, 32, 64, and 128). After identifying the optimal n_b, we then explore different reduction factors r (4, 8, 16, and 32).

The batch size n_b affects the stability of the PHPRBP model during training. As shown in Table S11, PHPRBP performs best in the host prediction task with $n_b = 32$. This optimal batch size likely strikes a balance between stability and learning flexibility, thereby enhancing the effectiveness of batch normalization. Additionally, the results further show that higher batch sizes (i.e., $n_b = 32$, 64, and 128) generally yield better performance than the lower batch size (i.e., $n_b = 16$). This can be attributed to the instability caused by smaller batch sizes, which negatively impacts batch normalization and leads to poorer prediction performance. In this case, a batch size setting of 32 is deemed appropriate for PHPRBP.

In the SE attention mechanism, the reduction factor r is crucial for learning the attention coefficients. Table S12 illustrates the variation in the prediction performance of PHPRBP with different values of r. The results indicate that the prediction performance initially increases and then decreases as r increases, reaching optimal performance at $r = 16$. This may be because smaller r values (i.e., $r = 2$, 4, and 8) do not sufficiently suppress unimportant features, limiting the effectiveness of the attention mechanism. Conversely, a larger r value (i.e., $r = 32$) may place too much emphasis on important features while neglecting the contribution of other potential features, leading to a decline in performance. Thus, setting r to 16 in PHPRBP is an appropriate choice.

References

[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 5998–6008.

[2] B. E. Suzek, H. Huang, P. McGarvey, R. Mazumder, and C. H. Wu, "Uniref: comprehensive and non-redundant uniprot reference clusters," *Bioinformatics*, vol. 23, no. 10, pp. 1282–1288, 2007.

[3] M. Steinegger, M. Mirdita, and J. Söding, "Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold," *Nature methods*, vol. 16, no. 7, pp. 603–606, 2019.

[4] M. Steinegger and J. Söding, "Clustering huge protein sequence sets in linear time," *Nature communications*, vol. 9, no. 1, 2018, Art. no. 2542.

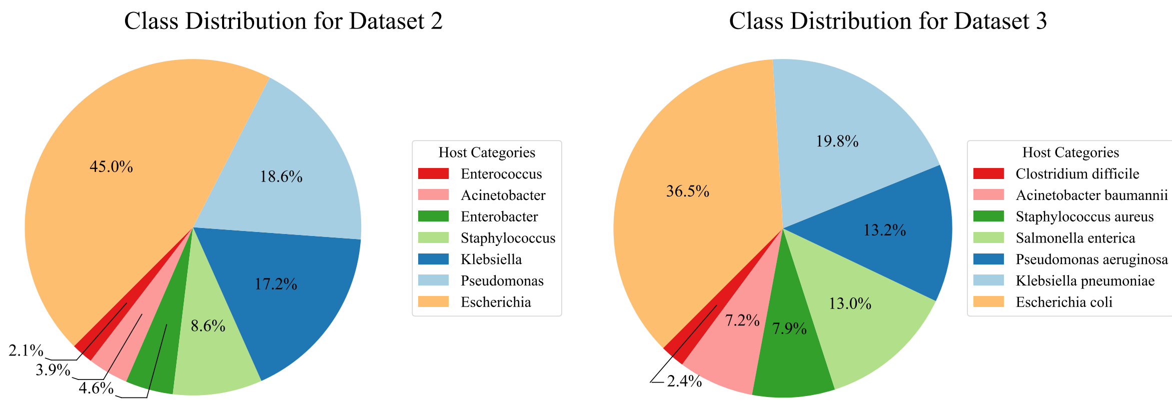
[5] M. E. Peters, M. Neumann, L. Zettlemoyer, and W. Yih, "Dissecting contextual word embeddings: Architecture and representation," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018, pp. 1499–1509.

TABLE S1: The performance comparison of PHPRBP and baseline methods on Datasets 1, 2, and 3.The best performance are highlighted in boldface.

Dataset	Model	ACC	F1	MCC	Precision	Recall
Dataset 1	**PHPRBP**	**0.8804±0.0027**	**0.8764±0.0025**	**0.8693±0.0030**	**0.8788±0.0035**	**0.8804±0.0027**
	HFRF	0.8085±0.0077	0.8008±0.0075	0.7895±0.00817	0.8178±0.0067	0.8085±0.0077
	ESMMLP	0.8132±0.0060	0.8044±0.0065	0.7924±0.0070	0.8070±0.006	0.8132±0.0060
	PHIEmbed	0.8097±0.0045	0.8047±0.0041	0.7910±0.0047	0.8159±0.0045	0.8097±0.0045
	DeepHost	0.7618±0.0123	0.7559±0.0116	0.7409±0.0125	0.7683±0.0091	0.7618±0.0123
	MSKNN	0.8158±0.0018	0.8117±0.0022	0.8031±0.0024	0.8172±0.0029	0.8158±0.0018
	MSRF	0.8634±0.0039	0.8614±0.0036	0.8525±0.0042	0.8648±0.0039	0.8634±0.0039
	MSXGB	0.8646±0.0025	0.8622±0.0027	0.8540±0.0027	0.8667±0.0020	0.8646±0.0025
Dataset 2	**PHPRBP**	**0.9294±0.0187**	**0.9293±0.0184**	**0.9135±0.0226**	**0.9325±0.0162**	**0.9294±0.0187**
	HFRF	0.8850±0.0152	0.8841±0.0151	0.8480±0.0211	0.8893±0.0162	0.8850±0.0152
	ESMMLP	0.9064±0.0141	0.9063±0.0143	0.8795±0.0179	0.9094±0.0134	0.9064±0.0141
	PHIEmbed	0.8862±0.0171	0.8846±0.0171	0.8537±0.0207	0.8928±0.0181	0.8862±0.0171
	DeepHost	0.8738±0.0334	0.8732±0.0323	0.8385±0.0427	0.8806±0.0307	0.8738±0.0334
	MSKNN	0.8741±0.0240	0.8732±0.0234	0.8465±0.0292	0.8784±0.0237	0.8741±0.0240
	MSRF	0.9203±0.0039	0.9195±0.0041	0.8933±0.0054	0.9218±0.0037	0.9203±0.0039
	MSXGB	0.9218±0.0155	0.9219±0.0154	0.9037±0.0179	0.9249±0.0145	0.9218±0.0155
Dataset 3	**PHPRBP**	**0.9370±0.0111**	**0.9368±0.0112**	**0.9221±0.0150**	**0.9391±0.0112**	**0.9370±0.0111**
	HFRF	0.8935±0.0057	0.8876±0.0058	0.8453±0.0082	0.8948±0.0059	0.8935±0.0057
	ESMMLP	0.9096±0.0050	0.9038±0.0046	0.8710±0.0062	0.9035±0.0051	0.9096±0.0050
	PHIEmbed	0.8886±0.0031	0.8857±0.0029	0.8389±0.0040	0.8900±0.0031	0.8886±0.0031
	DeepHost	0.8759±0.0051	0.8715±0.0053	0.8219±0.0070	0.8742±0.0047	0.8759±0.0051
	MSKNN	0.8833±0.0042	0.8849±0.0035	0.8508±0.0042	0.8893±0.0025	0.8833±0.0042
	MSRF	0.9275±0.0214	0.9273±0.0211	0.9098±0.0262	0.9303±0.0203	0.9275±0.0214
	MSXGB	0.9279±0.0037	0.9270±0.0040	0.9037±0.0051	0.9280±0.0042	0.9279±0.0037

TABLE S2: The performance comparison of PHPRBP and baseline methods at different taxonomic level on Datasets 1, 2, and 3. The best performance are highlighted in boldface.

Level	Dataset	Model	ACC	F1	MCC	Precision	Recall
Family	Dataset 1	**PHPRBP**	**0.9447±0.0035**	**0.9446±0.0036**	**0.9391±0.0039**	**0.9449±0.0036**	**0.9447±0.0035**
		HFRF	0.8977±0.0036	0.8960±0.0034	0.8864±0.0040	0.9011±0.0039	0.8977±0.0036
		ESMMLP	0.9066±0.0066	0.9056±0.0066	0.8963±0.0074	0.9058±0.0068	0.9066±0.0066
		PHIEmbed	0.9035±0.0085	0.9035±0.0083	0.8949±0.0091	0.9110±0.0074	0.9035±0.0085
		DeepHost	0.8465±0.0056	0.8448±0.0049	0.8291±0.0054	0.8500±0.0047	0.8465±0.0056
		MSKNN	0.8844±0.0038	0.8825±0.0043	0.8751±0.0044	0.8909±0.0048	0.8844±0.0038
		MSRF	0.9293±0.0027	0.9300±0.0025	0.9229±0.0027	0.9323±0.0021	0.9293±0.0027
		MSXGB	0.9315±0.0049	0.9316±0.0048	0.9248±0.0053	0.9325±0.0048	0.9315±0.0049
Family	Dataset 2	**PHPRBP**	**0.9877±0.0007**	**0.9876±0.0007**	**0.9850±0.0006**	**0.9878±0.0007**	**0.9877±0.0007**
		HFRF	0.9650±0.0024	0.9640±0.0025	0.9576±0.0027	0.9656±0.0024	0.9650±0.0024
		ESMMLP	0.9647±0.0039	0.9645±0.0039	0.9578±0.0049	0.9652±0.0038	0.9647±0.0039
		PHIEmbed	0.9549±0.0052	0.9537±0.0057	0.9450±0.0063	0.9573±0.0053	0.9549±0.0052
		DeepHost	0.9445±0.0060	0.9435±0.0065	0.9319±0.0074	0.9463±0.0052	0.9445±0.0060
		MSKNN	0.9594±0.0027	0.9598±0.0026	0.9538±0.0032	0.9629±0.0022	0.9594±0.0027
		MSRF	0.9822±0.0016	0.9822±0.0016	0.9785±0.0019	0.9826±0.0016	0.9822±0.0016
		MSXGB	0.9828±0.0041	0.9828±0.0042	0.9794±0.0049	0.9830±0.0041	0.9828±0.0041
Genus	Dataset 3	**PHPRBP**	**0.9370±0.0111**	**0.9368±0.0112**	**0.9221±0.0150**	**0.9391±0.0112**	**0.9370±0.0111**
		HFRF	0.8935±0.0057	0.8876±0.0058	0.8453±0.0082	0.8948±0.0059	0.8935±0.0057
		ESMMLP	0.9096±0.0050	0.9038±0.0046	0.8710±0.0062	0.9035±0.0051	0.9096±0.0050
		PHIEmbed	0.8886±0.0031	0.8857±0.0029	0.8389±0.0040	0.8900±0.0031	0.8886±0.0031
		DeepHost	0.8759±0.0051	0.8715±0.0053	0.8219±0.0070	0.8742±0.0047	0.8759±0.0051
		MSKNN	0.8833±0.0042	0.8849±0.0035	0.8508±0.0042	0.8893±0.0025	0.8833±0.0042
		MSRF	0.9203±0.0039	0.9195±0.0041	0.8933±0.0054	0.9218±0.0037	0.9203±0.0039
		MSXGB	0.9218±0.0155	0.9219±0.0154	0.9037±0.0179	0.9249±0.0145	0.9218±0.0155
Family	Dataset 3	**PHPRBP**	**0.9801±0.0113**	**0.9802±0.0112**	**0.9652±0.0192**	**0.9811±0.0104**	**0.9801±0.0113**
		HFRF	0.9480±0.0131	0.9403±0.0159	0.8798±0.0331	0.9514±0.0116	0.9480±0.0131
		ESMMLP	0.9492±0.0124	0.9475±0.0123	0.8860±0.0275	0.9495±0.0120	0.9492±0.0124
		PHIEmbed	0.9319±0.0161	0.9228±0.0196	0.8420±0.0371	0.9376±0.0136	0.9319±0.0161
		DeepHost	0.9400±0.0130	0.9347±0.0130	0.8676±0.0304	0.9400±0.0138	0.9400±0.0130
		MSKNN	0.9353±0.0108	0.9344±0.0115	0.8900±0.0129	0.9387±0.0081	0.9353±0.0108
		MSRF	0.9621±0.0060	0.9613±0.0063	0.9300±0.0112	0.9643±0.0053	0.9621±0.0060
		MSXGB	0.9671±0.0081	0.9668±0.0080	0.9415±0.0149	0.9686±0.0077	0.9671±0.0081



(a) (b)

Fig. S1: The class distributions for Dataset 2 (a) and Dataset 3 (b) .

TABLE S3: Pre-trained PLMs for generating phage RBP sequence embeddings. "-" represents that the model does not use a Transformer-based architecture.

Pre-trained PLM	Architecture	Layers of Transformer Encoder	Protein Database	Amino acid embedding size
ESM1	Transformer [1]	34	UniRef50 [2]	1280
ESM1b	Transformer [1]	33	UniRef50 [2]	1280
ESM2	Transformer [1]	33	UniRef50 [2]	1280
ProtBert	Transformer [1]	30	UniRef100 [2], BFD100 [3], [4]	1024
ProtXLNet	Transformer [1]	30	UniRef100 [2]	1024
ProtAlbert	Transformer [1]	12	UniRef100 [2]	4096
ProtT5	Transformer [1]	24	UniRef50 [2], BFD100 [3], [4]	1024
SeqVec	ELMo [5]	-	UniRef50 [2]	1024

TABLE S4: The performance comparison of PHPRBP and various pre-trained PLMs on Datasets 1, 2, and 3. The best performance are highlighted in boldface.

Dataset	Model	ACC	F1	MCC	Precision	Recall
Dataset 1	**PHPRBP**	**0.8804±0.0027**	**0.8764±0.0025**	**0.8693±0.0030**	**0.8788±0.0035**	**0.8804±0.0027**
	ESM1	0.8658±0.0028	0.8617±0.0022	0.8542±0.0028	0.8670±0.0036	0.8658±0.0028
	ESM1b	0.8736±0.0015	0.8681±0.0019	0.8612±0.0020	0.8721±0.0021	0.8736±0.0015
	ESM2	0.8765±0.0017	0.8714±0.0023	0.8648±0.0022	0.8764±0.0016	0.8765±0.0017
	ProtAlbert	0.8698±0.0029	0.8658±0.0023	0.8583±0.0027	0.8694±0.0031	0.8698±0.0029
	ProtBert	0.8550±0.0019	0.8490±0.0020	0.8420±0.0022	0.8566±0.0019	0.8550±0.0019
	ProtT5	0.8754±0.0024	0.8706±0.0030	0.8636±0.0032	0.8745±0.0036	0.8754±0.0024
	ProtXLNet	0.8230±0.0019	0.8160±0.0021	0.8083±0.0017	0.8274±0.0033	0.8230±0.0019
	SeqVec	0.8406±0.0053	0.8360±0.0057	0.8275±0.0057	0.8411±0.0055	0.8406±0.0053
	ESM1b_ProtT5	0.8784±0.0034	0.8746±0.0030	0.8672±0.0034	0.8763±0.0031	0.8784±0.0034
Dataset 2	**PHPRBP**	**0.9294±0.0187**	**0.9293±0.0184**	**0.9135±0.0226**	**0.9325±0.0162**	**0.9294±0.0187**
	ESM1	0.9080±0.0044	0.9050±0.0044	0.8763±0.0053	0.9070±0.0046	0.9080±0.0044
	ESM1b	0.9180±0.0052	0.9138±0.0058	0.8890±0.0072	0.9159±0.0058	0.9180±0.0052
	ESM2	0.9101±0.0032	0.9052±0.0036	0.8779±0.0047	0.9074±0.0039	0.9101±0.0032
	ProtAlbert	0.9146±0.0044	0.9123±0.0040	0.8852±0.0057	0.9137±0.0046	0.9146±0.0044
	ProtBert	0.8940±0.0058	0.8872±0.0081	0.8563±0.0086	0.8928±0.0043	0.8940±0.0058
	ProtT5	0.9181±0.0044	0.9153±0.0049	0.8894±0.0061	0.9168±0.0047	0.9181±0.0044
	ProtXLNet	0.8945±0.0034	0.8908±0.0035	0.8587±0.0045	0.8931±0.0036	0.8945±0.0034
	SeqVec	0.9012±0.0034	0.8985±0.0031	0.8676±0.0040	0.9012±0.0026	0.9012±0.0034
	ESM1b_ProtT5	0.9234±0.0031	0.9225±0.0022	0.8980±0.0035	0.9231±0.0021	0.9234±0.0031
Dataset 3	**PHPRBP**	**0.9370±0.0111**	**0.9368±0.0112**	**0.9221±0.0150**	**0.9391±0.0112**	**0.9370±0.0111**
	ESM1	0.9329±0.0078	0.9330±0.0076	0.9168±0.0087	0.9366±0.0079	0.9329±0.0078
	ESM1b	0.9282±0.0216	0.9279±0.0222	0.9116±0.0278	0.9307±0.0222	0.9282±0.0216
	ESM2	0.9336±0.0116	0.9334±0.0116	0.9170±0.0147	0.9355±0.0111	0.9336±0.0116
	ProtAlbert	0.9237±0.0112	0.9231±0.0109	0.9047±0.0140	0.9250±0.0108	0.9237±0.0112
	ProtBert	0.9212±0.0192	0.9213±0.0194	0.9026±0.0234	0.9262±0.0174	0.9212±0.0192
	ProtT5	0.9308±0.0181	0.9308±0.0181	0.9142±0.0224	0.9324±0.0175	0.9308±0.0181
	ProtXLNet	0.9157±0.0236	0.9155±0.0240	0.8957±0.0291	0.9174±0.0238	0.9157±0.0236
	SeqVec	0.9039±0.0215	0.9028±0.0223	0.8793±0.0268	0.9064±0.0202	0.9039±0.0215
	ESM1b_ProtT5	0.9325±0.0177	0.9326±0.0175	0.9166±0.0216	0.9342±0.0162	0.9325±0.0177

TABLE S5: The performance comparison of PHPRBP and various class imbalance mitigation strategies on Datasets 1, 2, and 3. The best performances are highlighted in boldface.

Dataset	Model	ACC	F1	MCC	Precision	Recall
Dataset 1	**PHPRBP**	**0.8804±0.0027**	**0.8764±0.0025**	**0.8693±0.0030**	**0.8788±0.0035**	**0.8804±0.0027**
	RUS	0.7714±0.0062	0.7677±0.0060	0.7634±0.0063	0.7721±0.0064	0.7714±0.0062
	GAN	0.6743±0.0029	0.6650±0.0034	0.6552±0.0034	0.6642±0.0034	0.6743±0.0029
	SMOTE	0.8703±0.0049	0.8671±0.0046	0.8594±0.0050	0.8692±0.0046	0.8703±0.0049
	SVMSMOTE	0.8641±0.0032	0.8611±0.0031	0.8530±0.0033	0.8626±0.0037	0.8641±0.0032
	BorderlineSMOTE	0.8683±0.0033	0.8643±0.0031	0.8568±0.0034	0.8664±0.0032	0.8683±0.0033
Dataset 2	**PHPRBP**	**0.9294±0.0187**	**0.9293±0.0184**	**0.9135±0.0226**	**0.9325±0.0162**	**0.9294±0.0187**
	RUS	0.8450±0.0219	0.8382±0.0265	0.8101±0.0279	0.8393±0.0238	0.8450±0.0219
	GAN	0.8166±0.0039	0.8074±0.0035	0.7746±0.0040	0.8036±0.0034	0.8166±0.0039
	SMOTE	0.9216±0.0028	0.9199±0.0027	0.8956±0.0035	0.9204±0.0026	0.9216±0.0028
	SVMSMOTE	0.9242±0.0038	0.9234±0.0033	0.8995±0.0041	0.9236±0.0034	0.9242±0.0038
	BorderlineSMOTE	0.9193±0.0027	0.9174±0.0026	0.8922±0.0036	0.9177±0.0032	0.9193±0.0027
Dataset 3	**PHPRBP**	**0.9370±0.0111**	**0.9368±0.0112**	**0.9221±0.0150**	**0.9391±0.0112**	**0.9370±0.0111**
	RUS	0.8908±0.0394	0.8907±0.0385	0.8707±0.0459	0.8960±0.0361	0.8908±0.0394
	GAN	0.8090±0.0130	0.8082±0.0132	0.7764±0.0156	0.8128±0.0131	0.8090±0.0130
	SMOTE	0.9251±0.0179	0.9254±0.0178	0.9089±0.0213	0.9303±0.0146	0.9251±0.0179
	SVMSMOTE	0.9136±0.0184	0.9140±0.0182	0.8944±0.0231	0.9182±0.0180	0.9136±0.0184
	BorderlineSMOTE	0.9203±0.0090	0.9201±0.0092	0.9023±0.0108	0.9237±0.0079	0.9203±0.0090

TABLE S6: The performance comparison of PHPRBP before and after ADASYN-based data enhancement for specific minority classes. The best performances are highlighted in boldface.

Dataset	Host Category	ACC		MCC	
		AF	BE	AF	BE
Dataset 1	*Helicobacter*	**1.0000±0.0000**	0.9800±0.0400	**1.0000±0.0000**	0.9897±0.0206
	Providencia	**0.8795±0.0976**	0.5200±0.1406	**0.9213±0.0519**	0.6198±0.1108
	Pantoea	**0.7484±0.0193**	0.4694±0.0839	**0.7215±0.0260**	0.5518±0.0769
	Pseudoalteromonas	**0.8385±0.0654**	0.3182±0.1076	**0.8316±0.0661**	0.4743±0.1183
	Achromobacter	**0.9656±0.0324**	0.6894±0.0857	**0.9573±0.0239**	0.6916±0.0723
	Prevotella	**1.0000±0.0000**	0.9667±0.0408	**1.0000±0.0000**	0.9678±0.0163
	Lactobacillus	**0.9451±0.0561**	0.7792±0.0900	**0.9613±0.0288**	0.8077±0.0361
	Clostridium	**0.9000±0.0789**	0.7100±0.0599	**0.9106±0.0547**	0.7491±0.0528
	Caulobacter	**0.9879±0.0242**	0.8000±0.0471	**0.9852±0.0133**	0.8049±0.0489
	Arthrobacter	**0.9943±0.0114**	0.7438±0.1291	**0.9834±0.0132**	0.7751±0.0781
Dataset 2	*Enterococcus*	**0.9895±0.0001**	0.9191±0.0030	**0.8610±0.0006**	0.8582±0.0064
	Acinetobacter	**0.9749±0.0000**	0.8374±0.0009	**0.9667±0.0070**	0.8732±0.0276
	Enterobacter	**0.6223±0.0010**	0.3464±0.0026	**0.6694±0.0047**	0.4628±0.0048
Dataset 3	*Clostridium difficile*	**0.9926±0.0002**	0.8897±0.0064	**0.9445±0.0176**	0.8993±0.0494

"AF" represents after data augmentation, while "BE" indicates before data augmentation.

TABLE S7: The performance comparison of PHPRBP and the variant models on Datasets 1, 2, and 3. The best performances are highlighted in boldface.

Dataset	Model	Accuracy	F1-score	MCC	Precision	Recall
Dataset 1	**PHPRBP**	**0.8804±0.0027**	**0.8764±0.0025**	**0.8693±0.0030**	**0.8788±0.0035**	**0.8804±0.0027**
	PHPRBP-E	0.8778±0.0025	0.8734±0.0027	0.8665±0.0026	0.8765±0.0022	0.8778±0.0025
	PHPRBP-P	0.8766±0.0039	0.8722±0.0038	0.8650±0.0040	0.8744±0.0040	0.8766±0.0039
	PHPRBP-C	0.8764±0.0034	0.8712±0.0034	0.8648±0.0035	0.8763±0.0032	0.8764±0.0034
	PHPRBP-S	0.8780±0.0023	0.8743±0.0023	0.8668±0.0023	0.8759±0.0022	0.8780±0.0023
	PHPRBP-A	0.8288±0.0057	0.8223±0.0055	0.8107±0.0065	0.8248±0.0068	0.8288±0.0057
Dataset 2	**PHPRBP**	**0.9294±0.0187**	**0.9293±0.0184**	**0.9135±0.0226**	**0.9325±0.0162**	**0.9294±0.0187**
	PHPRBP-E	0.9243±0.0032	0.9224±0.0032	0.8981±0.0044	0.9230±0.0029	0.9243±0.0032
	PHPRBP-P	0.9212±0.0048	0.9200±0.0051	0.8943±0.0067	0.9199±0.0051	0.9212±0.0048
	PHPRBP-C	0.9164±0.0025	0.9130±0.0034	0.8867±0.0032	0.9141±0.0024	0.9164±0.0025
	PHPRBP-S	0.9232±0.0034	0.9214±0.0042	0.8971±0.0044	0.9218±0.0036	0.9232±0.0034
	PHPRBP-A	0.9163±0.0047	0.9116±0.0046	0.8795±0.0064	0.9112±0.0052	0.9163±0.0047
Dataset 3	**PHPRBP**	**0.9370±0.0111**	**0.9368±0.0112**	**0.9221±0.0150**	**0.9391±0.0112**	**0.9370±0.0111**
	PHPRBP-E	0.9336±0.0254	0.9331±0.0255	0.9169±0.0316	0.9349±0.0242	0.9336±0.0254
	PHPRBP-P	0.9213±0.0142	0.9212±0.0133	0.9028±0.0160	0.9242±0.0113	0.9213±0.0142
	PHPRBP-C	0.9141±0.0273	0.9135±0.0270	0.8929±0.0330	0.9173±0.0253	0.9141±0.0273
	PHPRBP-S	0.9284±0.0161	0.9282±0.0161	0.9108±0.0202	0.9301±0.0157	0.9284±0.0161
	PHPRBP-A	0.9098±0.0198	0.9098±0.0198	0.8838±0.0247	0.9126±0.0191	0.9098±0.0198

TABLE S8: Paired t-test results comparing PHPRBP with baseline methods on Dataset 1.

Methods	Metrics	t-Stat	p-value	Significance (Y/N)
PHPRBP & HFRF	ACC	57.4781	5.49e-07	Y
	F1	52.4932	7.88e-07	Y
	MCC	45.7041	1.37e-06	Y
	Precision	32.6802	5.23e-06	Y
	Sensitivity	57.4781	5.49e-07	Y
PHPRBP & ESMMLP	ACC	64.8814	3.38e-07	Y
	F1	68.0543	2.79e-07	Y
	MCC	65.4058	3.27e-07	Y
	Precision	39.4048	2.48e-06	Y
	Sensitivity	64.8814	3.38e-07	Y
PHPRBP & PHIEmbed	ACC	42.5686	1.82e-06	Y
	F1	43.8896	1.61e-06	Y
	MCC	46.7230	1.26e-06	Y
	Precision	39.8922	2.36e-06	Y
	Sensitivity	42.5686	1.82e-06	Y
PHPRBP & DeepHost	ACC	18.3502	5.19e-05	Y
	F1	19.6776	3.93e-05	Y
	MCC	21.8411	2.60e-05	Y
	Precision	21.6950	2.67e-05	Y
	Sensitivity	18.3502	5.19e-05	Y
PHPRBP & MSKNN	ACC	57.4781	5.49e-07	Y
	F1	52.4932	7.88e-07	Y
	MCC	45.7041	1.37e-06	Y
	Precision	32.6802	5.23e-06	Y
	Sensitivity	57.4781	5.49e-07	Y
PHPRBP & MSRF	ACC	11.2793	3.52e-04	Y
	F1	10.0318	5.55e-04	Y
	MCC	10.7750	4.21e-04	Y
	Precision	9.71820	6.28e-04	Y
	Sensitivity	11.2793	3.52e-04	Y
PHPRBP & MSXGB	ACC	11.6893	3.06e-04	Y
	F1	11.5022	3.26e-04	Y
	MCC	11.1043	3.74e-04	Y
	Precision	8.45510	1.07e-03	Y
	Sensitivity	11.6893	3.06e-04	Y

TABLE S9: Paired t-test results comparing PHPRBP with baseline methods on Dataset 2.

Methods	Metrics	t-Stat	p-value	Significance (Y/N)
PHPRBP & HFRF	ACC	8.3362	1.13e-03	Y
	F1	4.1686	1.40e-02	Y
	MCC	12.7294	2.19e-04	Y
	Precision	3.0280	3.89e-02	Y
	Sensitivity	8.3362	1.13e-03	Y
PHPRBP & ESMMLP	ACC	5.0067	7.45e-03	Y
	F1	5.2762	6.19e-03	Y
	MCC	16.0825	8.74e-05	Y
	Precision	5.0440	7.26e-03	Y
	Sensitivity	5.0067	7.45e-03	Y
PHPRBP & PHIEmbed	ACC	6.0899	3.68e-03	Y
	F1	5.5447	5.17e-03	Y
	MCC	40.3353	2.26e-06	Y
	Precision	3.1439	3.47e-02	Y
	Sensitivity	6.0899	3.68e-03	Y
PHPRBP & DeepHost	ACC	6.6947	2.59e-03	Y
	F1	7.3160	1.86e-03	Y
	MCC	26.5424	1.20e-05	Y
	Precision	5.9821	3.93e-03	Y
	Sensitivity	6.6947	2.59e-03	Y
PHPRBP & MSKNN	ACC	8.7542	9.38e-04	Y
	F1	6.7049	2.58e-03	Y
	MCC	29.3393	8.04e-06	Y
	Precision	4.2474	1.32e-02	Y
	Sensitivity	8.7542	9.38e-04	Y
PHPRBP & MSRF	ACC	1.9876	1.18e-01	N
	F1	2.8275	4.75e-02	Y
	MCC	2.8582	4.60e-02	Y
	Precision	2.4054	7.39e-02	N
	Sensitivity	1.9876	1.18e-01	N
PHPRBP & MSXGB	ACC	0.9801	3.83e-01	N
	F1	2.2657	4.53e-02	Y
	MCC	2.8727	3.09e-02	Y
	Precision	1.0619	3.48e-01	N
	Sensitivity	0.9801	3.83e-01	N

TABLE S10: Paired t-test results comparing PHPRBP with baseline methods on Dataset 3.

Methods	Metrics	t-Stat	p-value	Significance (Y/N)
PHPRBP & HFRF	ACC	7.0291	2.16e-03	Y
	F1	7.4822	1.71e-03	Y
	MCC	7.3811	1.80e-03	Y
	Precision	6.9044	2.31e-03	Y
	Sensitivity	7.0291	2.16e-03	Y
PHPRBP & ESMMLP	ACC	5.3446	5.91e-03	Y
	F1	2.8234	4.77e-02	Y
	MCC	4.5377	1.05e-02	Y
	Precision	7.4693	1.72e-03	Y
	Sensitivity	5.3446	5.91e-03	Y
PHPRBP & PHIEmbed	ACC	6.0824	3.69e-03	Y
	F1	6.1415	3.56e-03	Y
	MCC	6.7734	2.48e-03	Y
	Precision	4.9896	7.55e-03	Y
	Sensitivity	6.0824	3.69e-03	Y
PHPRBP & DeepHost	ACC	6.4535	2.97e-03	Y
	F1	6.4208	3.02e-03	Y
	MCC	6.3295	3.19e-03	Y
	Precision	6.2492	3.34e-03	Y
	Sensitivity	6.4535	2.97e-03	Y
PHPRBP & MSKNN	ACC	8.1659	1.22e-03	Y
	F1	8.5876	1.01e-03	Y
	MCC	8.1570	1.23e-03	Y
	Precision	8.0200	1.31e-03	Y
	Sensitivity	8.1659	1.22e-03	Y
PHPRBP & MSRF	ACC	2.9827	4.06e-02	Y
	F1	3.2082	3.26e-02	Y
	MCC	3.2463	3.15e-02	Y
	Precision	2.5924	6.05e-02	N
	Sensitivity	2.9827	4.06e-02	Y
PHPRBP & MSXGB	ACC	1.1557	3.12e-01	N
	F1	2.7983	4.89e-02	Y
	MCC	2.9672	4.13e-02	Y
	Precision	2.2445	8.82e-02	N
	Sensitivity	1.1557	3.12e-01	N

TABLE S11: The performance comparison of PHPRBP with different batch sizes n_b on Datasets 1, 2, and 3. The best performances are highlighted in boldface.

Dataset	Parameter	ACC	F1	MCC	Precision	Recall
Dataset 1	$n_b = 16$	0.8756±0.0033	0.8713±0.0027	0.8640±0.0032	0.8738±0.0032	0.8756±0.0033
	$\boldsymbol{n_b = 32}$	**0.8804±0.0027**	0.8764±0.0025	0.8693±0.0030	**0.8788±0.0035**	**0.8804±0.0027**
	$n_b = 64$	0.8799±0.0032	0.8766±0.0034	0.8692±0.0037	0.8780±0.0036	0.8799±0.0032
	$n_b = 128$	0.8802±0.0027	**0.8771±0.0024**	**0.8694±0.0028**	0.8776±0.0028	0.8802±0.0027
Dataset 2	$n_b = 16$	0.9247±0.0018	0.9225±0.0020	0.8987±0.0022	0.9234±0.0020	0.9247±0.0018
	$\boldsymbol{n_b = 32}$	**0.9294±0.0187**	**0.9293±0.0184**	**0.9135±0.0226**	**0.9325±0.0162**	**0.9294±0.0187**
	$n_b = 64$	0.9242±0.0015	0.9225±0.0021	0.8984±0.0024	0.9228±0.0020	0.9242±0.0015
	$n_b = 128$	0.9243±0.0030	0.9234±0.0032	0.8991±0.0037	0.9238±0.0030	0.9243±0.0030
Dataset 3	$n_b = 16$	0.9275±0.0134	0.9270±0.0138	0.9095±0.0176	0.9283±0.0133	0.9275±0.0134
	$\boldsymbol{n_b = 32}$	**0.9370±0.0111**	**0.9368±0.0112**	**0.9221±0.0150**	**0.9391±0.0112**	**0.9370±0.0111**
	$n_b = 64$	0.9303±0.0173	0.9305±0.0170	0.9152±0.0211	0.9343±0.0157	0.9303±0.0173
	$n_b = 128$	0.9351±0.0153	0.9349±0.0154	0.9201±0.0194	0.9384±0.013	0.9351±0.0153

TABLE S12: The performance comparison of PHPRBP with different reduction factors r on Datasets 1, 2, and 3. The best performances are highlighted in boldface.

Dataset	Parameter	ACC	F1	MCC	Precision	Recall
Dataset 1	$r = 4$	0.8779±0.0046	0.8744±0.0048	0.8669±0.0052	0.8763±0.0052	0.8779±0.0046
	$r = 8$	0.8782±0.0038	0.8742±0.0035	0.8670±0.0038	0.8760±0.0035	0.8782±0.0038
	$\boldsymbol{r = 16}$	**0.8804±0.0027**	**0.8764±0.0025**	**0.8693±0.0030**	**0.8788±0.0035**	**0.8804±0.0027**
	$r = 32$	0.8785±0.0033	0.8744±0.0030	0.8671±0.0034	0.8763±0.0035	0.8785±0.0033
Dataset 2	$r = 4$	0.9235±0.0027	0.9219±0.0021	0.8972±0.0028	0.9224±0.0025	0.9235±0.0027
	$r = 8$	0.9238±0.0016	0.9222±0.0015	0.8977±0.0026	0.9224±0.0016	0.9238±0.0016
	$\boldsymbol{r = 16}$	**0.9294±0.0187**	**0.9293±0.0184**	**0.9135±0.0226**	**0.9325±0.0162**	**0.9294±0.0187**
	$r = 32$	0.9223±0.0014	0.9201±0.0016	0.8953±0.0018	0.9206±0.0018	0.9223±0.0014
Dataset 3	$r = 4$	0.9265±0.0140	0.9261±0.0138	0.9085±0.0170	0.9289±0.0141	0.9265±0.0140
	$r = 8$	0.9361±0.0185	0.9356±0.0189	0.9197±0.0239	0.9380±0.0186	0.9361±0.0185
	$\boldsymbol{r = 16}$	**0.9370±0.0111**	**0.9368±0.0112**	**0.9221±0.0150**	**0.9391±0.0112**	**0.9370±0.0111**
	$r = 32$	0.9322±0.0166	0.9323±0.0164	0.9162±0.0197	0.9350±0.0142	0.9322±0.0166