

**Table S4: Performance of ML-models on training and validation dataset developed using four types of Binary profile-based features of N, C, NC-terminal of gram-positive ABPs**

Feature type	Terminal	Training set						Validation set					
		Sn	Sp	Acc	AUC	AUPRC	MCC	Sn	Sp	Acc	AUC	AUPRC	MCC
<b>AAB (RF)</b>	<b>N</b>	81.3	80.9	81.1	0.90	0.91	0.62	78.0	88.2	83.1	0.92	0.92	0.67
	<b>C</b>	79.2	79.6	79.4	0.87	0.89	0.59	75.3	79.6	77.4	0.86	0.86	0.55
	<b>NC</b>	84.1	84.1	84.1	0.92	0.92	0.68	80.7	90.3	85.5	<b>0.93</b>	0.92	0.71
<b>DPB (SVC)</b>	<b>N</b>	79.8	79.2	79.5	0.88	0.89	0.59	77.4	85.5	81.5	0.89	0.90	0.63
	<b>C</b>	77.0	76.9	77.0	0.85	0.88	0.54	68.8	83.3	76.1	0.85	0.85	0.53
	<b>NC</b>	82.3	81.9	82.1	0.90	0.91	0.64	79.0	88.2	83.6	<b>0.91</b>	0.90	0.68
<b>AIB (ET)</b>	<b>N</b>	79.8	79.0	79.4	0.89	0.90	0.59	78.0	86.6	82.3	0.91	0.91	0.65
	<b>C</b>	78.2	78.1	78.2	0.88	0.89	0.56	79.0	77.4	78.2	0.88	0.87	0.57
	<b>NC</b>	82.8	82.1	82.5	0.91	0.91	0.65	80.1	88.2	84.1	<b>0.93</b>	0.91	0.69
<b>PCB (RF)</b>	<b>N</b>	82.5	81.2	81.9	0.90	0.91	0.64	77.4	85.5	81.5	0.92	0.92	0.63
	<b>C</b>	79.8	80.9	80.4	0.88	0.89	0.61	75.8	80.7	78.2	0.86	0.85	0.57
	<b>NC</b>	82.9	81.5	82.2	0.91	0.91	0.64	82.3	87.6	85.0	<b>0.93</b>	0.92	0.70

# **Sn**: Sensitivity, **Sp**: Specificity, **Acc**: Accuracy, **MCC**: Matthews Correlation Coefficient, **AUC**: Area Under the Receiver Operating Characteristic curve, **AUPRC**: Area Under the Precision-Recall Curve, **AAB**: Amino acid-based binary profile, **DPB**: Dipeptide-based binary profile, **PCB**: Physico-chemical properties based binary profile, **AIB**: Amino-acid indices based binary profile, **RF**: Random Forest classifier, **ET**: Extra-tree classifier, **SVC**: Support vector classifier