

Machine Learning Approach	F1			accuracy			precision			recall			AUC		
	no SMOTE	SMOTE 1:2	SMOTE 1:1	no SMOTE	SMOTE 1:2	SMOTE 1:1	no SMOTE	SMOTE 1:2	SMOTE 1:1	no SMOTE	SMOTE 1:2	SMOTE 1:1	no SMOTE	SMOTE 1:2	SMOTE 1:1
Logistic Regression	0.473	0.531	0.542	0.806	0.802	0.749	0.643	0.591	0.473	0.379	0.489	0.644	0.794	0.785	0.766
Balanced Random Forest Classifier	0.747	0.833	0.847	0.863	0.921	0.933	0.656	0.817	0.891	0.874	0.854	0.813	0.936	0.945	0.959
Balanced Bagging Classifier	0.803	0.848	0.841	0.901	0.929	0.931	0.752	0.844	0.887	0.869	0.857	0.806	0.942	0.950	0.959
Random Forest Classifier	0.797	0.820	0.847	0.919	0.924	0.932	0.934	0.895	0.884	0.701	0.763	0.818	0.957	0.959	0.959
Multilayer Perceptron Classifier	0.399	0.473	0.505	0.802	0.801	0.712	0.690	0.649	0.440	0.301	0.401	0.638	0.777	0.772	0.759

Supplementary Table 2. Performance metrics on each machine learning approach with versus without using Synthetic Minority Oversampling Technique. We compare two different ratios for SMOTE: 1) 1:2 = one positive case to every 2 negative class; and 2) 1:1 = one positive case to every 1 negative class

AUC, area under the receiver operating characteristics curve

SMOTE, Synthetic Minority Oversampling Technique