

Table S1. Grid search parameters in each machine learning classifier.

	BLR	SVM	XGB	ANN
C	[0.001, 0.01, 0.1, 1, 10, 100]	[0.001, 0.01, 0.1, 1, 10, 100]		
gamma		[0.001, 0.01, 0.1, 1, 10, 100]		
kernel		rbf, linear		
n_estimator			range(5,30)	
max_depth			range(6,10)	
learning rate			[0.4, 0.45, 0.5, 0.55, 0.6]	
colsample_bytree			[0.6, 0.7, 0.8, 0.9, 1.0]	
optimizer				adam
loss				binary_crossentropy
batch_size				py
Epoch				50
				250

BLR, binary logistic regression; SVM, support vector machine; XGB, extreme gradient boosting; ANN, artificial neural network.

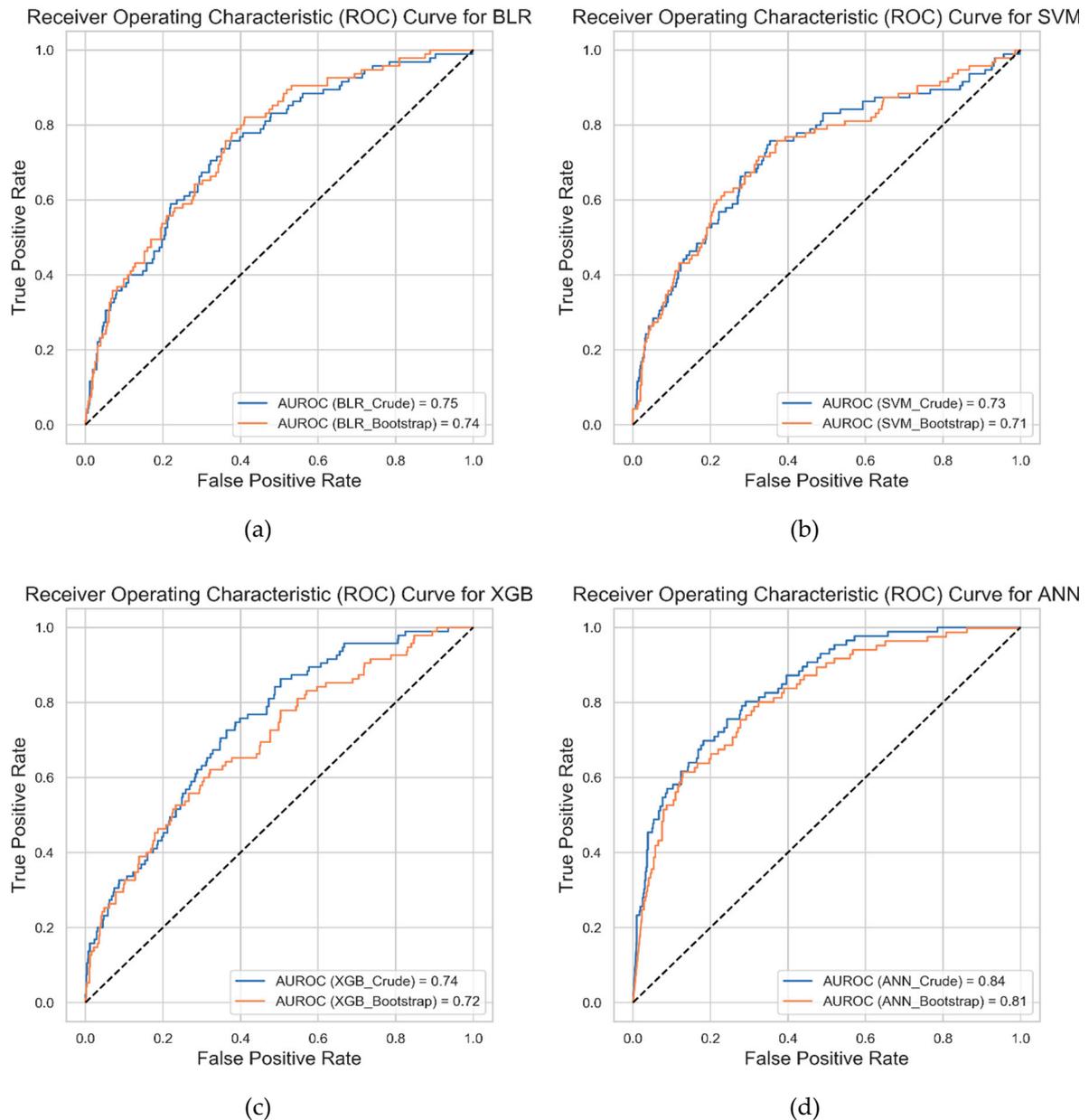


Figure S1. Result of the receiver operating characteristic curve of binary logistic regression (a) and support vector machine (b), extreme gradient boosting (c), and artificial neural network algorithm (d) before and after resampling strategy.

Each bootstrap resampled training set (120 different training set) was used to train each ML model. BLR, binary logistic regression; XGB, extreme gradient boosting; SVM, support vector machine; ANN, artificial neural network; AUROC, area under the receiver operating characteristic curve. 'Crude' refers to the case where the input variables were entered into the raw value, and the rest of them (normalizer, min-max, standard, and robust) were the different type of scaling methods for continuous variables.

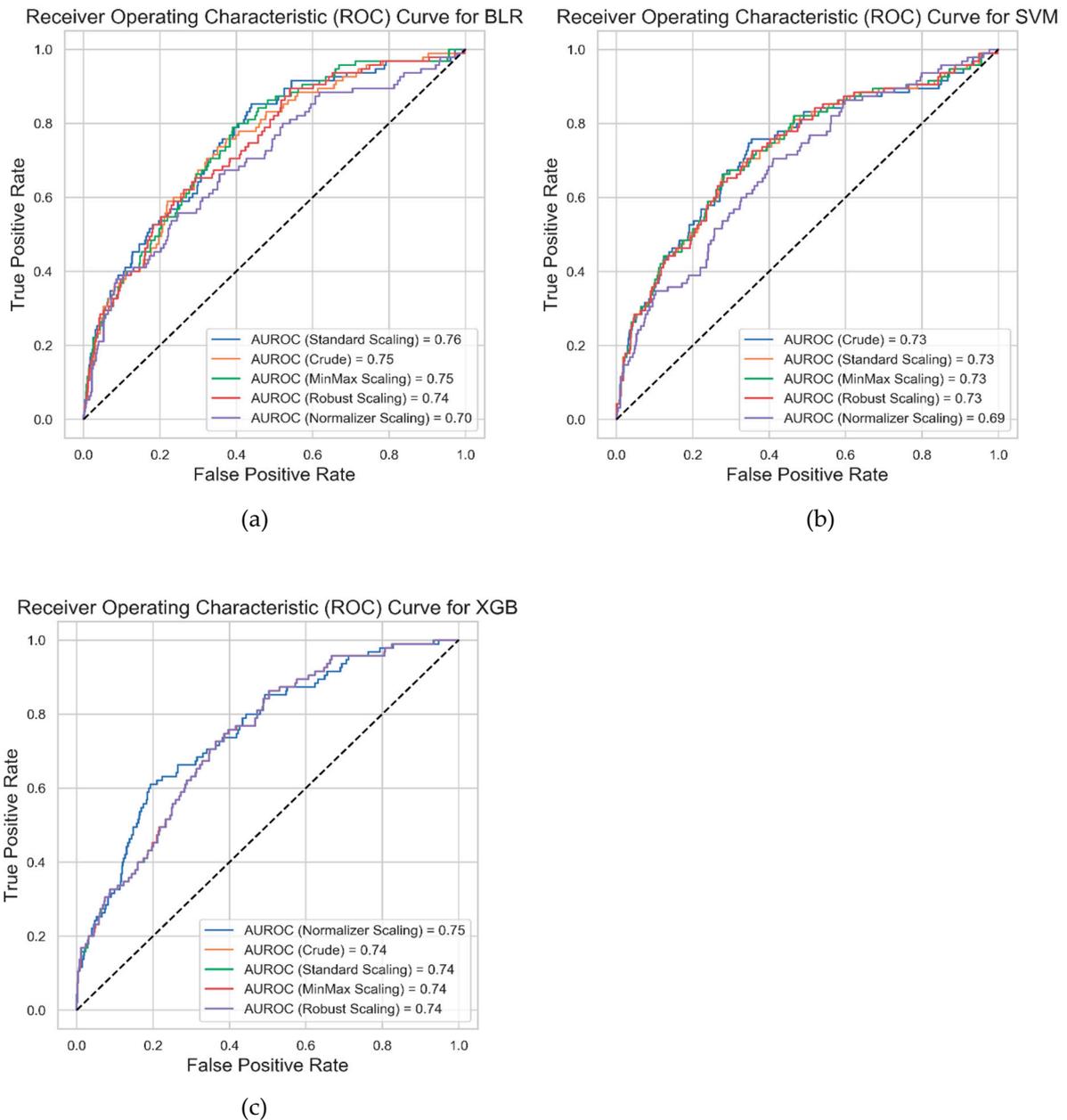


Figure S2. Result of the receiver operating characteristic curve of binary logistic regression (a), support vector machine (b), and extreme gradient boosting (c) before and after different scaling methods.

“Crude” refers to the case where the input variables were entered into the raw value, and the rest of the models were the results when continuous variables were applied to the model using various scaling methods (normalizer, min-max, standard, and robust).