Supplement Table 3. Packages, platforms and hyperparameters for machine learning models

	Platforms and packages			Hyperparameters
	Stata	R	python	
logistic regression with LASSO or Ridge	LASSOPACK, PDSLASSO	glmnet	scikit-learn	λ
support vector machine	svmachines	e1071	scikit-learn	kernel, cost (if applicable), gamma (if applicable)
linear discriminant analysis	discrim	MASS	scikit-learn	NA
k-nearest neighbor	discrim	class	scikit-learn	k
random forest	not recommend	randomforest	scikit-learn	number of variables randomly sampled at each split, number of trees to grow
gradient boosting model	not recommend	gbm, xgboost	scikit-learn, Xgboost	minimum number of samples which are required in a node to be considered for splitting, minimum samples required in a terminal leaf, maximum depth of a tree, maximum number of terminal eaves in a tree, gamma, learning rate, number of trees to grow
fully connected neural network	not recommend	keras	keras, tensorflow	number of hidden layers, number of neurons in each hidden layer, learning rate, optimizer, rate of drop out
convolutional neural network	not recommend	keras	keras, tensorflow	structure, pre-set weights (transfer learning), learning rate, optimizer, rate of drop out, batches

NA: not applicable; LASSO: least absolute shrinkage and selection operator