

Additional file 14: Table S6: Model hyperparameters

Model	Hyperparameters	Tuning range
Logistic regression with L_1 penalty	$\alpha = 0.01$	$[1e-6, 1e-4, 1e-2, 1, 1e2, 1e4]$
Logistic regression with L_2 penalty	$\alpha = 1$	$[1e-6, 1e-4, 1e-2, 1, 1e2, 1e4]$
Random forest classifier	n_estimators = 1000 max_depth = 1000	$[1000, 100]$ $[None, 1000, 100, 3]$
NN _{logreg}	weightdecay = 0.1 dropout = 0.0 epochs = 20	$[0, 1e-5, 1e-3, 1e-1, 1]$ $[0, 0.1, 0.4, 0.75]$ $[10 - 20 - 30 - 50 - 75 - 100]$
NN _{dense}	weightdecay = 0.001 dropout = 0.4 epochs = 30	$[0, 1e-5, 1e-3, 1e-1, 1]$ $[0, 0.1, 0.4, 0.75]$ $[10 - 20 - 30 - 50 - 75 - 100]$
NN _{biosparse}	weightdecay = 0.001 dropout = 0.4 epochs = 30	$[0, 1e-5, 1e-3, 1e-1, 1]$ $[0, 0.1, 0.4, 0.75]$ $[10 - 20 - 30 - 50 - 75 - 100]$