Table 5S. Performance of neural network (NN)-based prediction models using the best feature set from Table 4S. Highlighted with bold face is the final NN architecture selected for MutaCYP.

NN architecture [a]	NN learning algorithm [b]	MCC(5f-VS) [c]	MCC(5f-TS) [d]	MCC±SD [e]
5-[10-5]-2	**Rprop**	0.53 0.61 0.71 0.49 **0.65**	0.30 0.40 0.46 0.55 **0.58**	0.46±0.10
5-[10-5]-2	StdBP	0.53 0.67 0.56 0.51 0.61	0.42 0.46 0.36 0.64 0.58	0.49±0.10
5-[5-3]-2	Rprop	0.46 0.61 0.56 0.55 0.61	0.26 0.34 0.39 0.55 0.58	0.42±0.12
5-[5-3]-2	StdBP	0.50 0.67 0.56 0.55 0.61	0.36 0.40 0.36 0.66 0.58	0.47±0.12
5-[10]-2	Rprop	0.41 0.61 0.61 0.44 0.58	0.39 0.34 0.48 0.75 0.54	0.50±0.14
5-[10]-2	StdBP	0.46 0.67 0.74 0.51 0.61	0.43 0.40 0.43 0.64 0.58	0.50±0.10
5-[5]-2	Rprop	0.53 0.61 0.61 0.49 0.61	0.29 0.52 0.40 0.72 0.58	0.50±0.15
5-[5]-2	StdBP	0.49 0.67 0.61 0.44 0.65	0.35 0.58 0.39 0.70 0.45	0.49±0.13
5-[3]-2	Rprop	0.49 0.67	0.31 0.46	0.49±0.11

		0.56	0.46	
		0.47	0.64	
		0.58	0.58	
5-[3]-2	StdBP	0.49	0.36	0.51±0.12
		0.61	0.40	
		0.56	0.50	
		0.49	0.70	
		0.61	0.58	
5-[2]-2	Rprop	0.41	0.39	0.46±0.10
		0.61	0.34	
		0.45	0.40	
		0.43	0.57	
		0.61	0.58	
5-[2]-2	StdBP	0.46	0.36	0.49±0.13
		0.67	0.40	
		0.56	0.39	
		0.51	0.70	
		0.61	0.58	
5-2	Rprop	0.49	0.36	0.53±0.11
		0.67	0.52	
		0.25	0.48	
		0.47	0.67	
		0.42	0.63	
5-2	StdBP	0.41	0.30	0.52±0.12
		0.67	0.58	
		0.38	0.50	
		0.49	0.66	
		0.61	0.58	

[a] Numbers represent the number of nodes in a given layer. The first number is an input layer, the last number is the output layer, and the numbers in square brackets are nodes in the hidden layer(s).
[b] Rprop – resilient backpropagation; StdBP – standard backpropagation learning algorithms.
[c] Based on a validation subset for each of 5 folds (see section Methods for details).
[d] Based on a test subset for each of 5 folds (see section Methods for details).
[e] Based on 5-fold cross-validation (values from column [d]).