

Table 4. Performance of LR and SVM models: test and training sets

Variable group	Data set	Accuracy (%)		Sensitivity (%)		Specificity (%)	
		LR	SVM	LR	SVM	LR	SVM
A. Avg.Lys.Cys + h-CLAT + KeratinoSens + Toolbox + Log P	Training	94	94	94	94	95	95
	Test	92	92	87	93	100	89
B. h-CLAT + Toolbox + Log P	Training	86	88	84	84	91	95
	Test	75	83	73	80	78	89
C. KeratinoSens + Toolbox + Log P	Training	82	85	84	86	76	81
	Test	75	83	73	87	78	78
D. Avg.Lys.Cys + Toolbox + Log P	Training	86	89	88	84	81	100
	Test	75	83	73	73	78	100
E. h-CLAT + Log P	Training	82	82	88	88	67	67
	Test	75	75	80	80	67	67
F. KeratinoSens + Log P	Training	78	78	82	82	67	67
	Test	63	63	60	60	67	67
G. Avg.Lys.Cys + Log P	Training	86	85	82	80	95	95
	Test	75	71	67	67	89	78
H. Log P + Log S + Log VP + MP + BP + MW	Training	81	82	84	80	71	86
	Test	54	58	67	67	33	44
I. Avg.Lys.Cys + h-CLAT + KeratinoSens + Toolbox	Training	93	93	92	92	95	95

Variable group	Data set	Accuracy (%)		Sensitivity (%)		Specificity (%)	
		LR	SVM	LR	SVM	LR	SVM
			Test	92	92	93	93
J. h-CLAT + KeratinoSens + Toolbox	Training	85	90	90	90	71	91
	Test	75	79	87	80	56	78
K. Avg.Lys.Cys + h-CLAT + Toolbox	Training	93	93	92	92	95	95
	Test	92	92	93	93	89	89
L. Avg.Lys.Cys + KeratinoSens + Toolbox	Training	85	90	86	88	81	95
	Test	75	88	73	87	78	89

Avg.Lys.Cys, average depletion for lysine and cysteine peptides from the direct peptide reactivity assay; BP, boiling point; h-CLAT, human cell line activation test; log P, log octanol/water partition coefficient; log S, log water solubility; log VP, log vapor pressure; LR, logistic regression; MP, melting point; MW, molecular weight; SVM, support vector machine; Toolbox, read-across using QSAR Toolbox.

The only variable group that contains all six physicochemical properties is variable group H. The only physicochemical property contained in the other variable groups is log P. The three variable groups yielding the highest accuracy are bolded. The training set of 72 substances contains 51 human sensitizers and 21 non-sensitizers. The test set of 24 substances contains 15 human sensitizers and nine non-sensitizers.