

SUPPLEMENTARY MATERIALS S2

Table S1. Model performance metrics of the different machine learning methods on the *International* dataset (Adamson).

Model	Classification	Precision		Recall		F1-score		Area under ROC
Logistic Regression	Low stress	0.85	0.94	0.71	0.71	0.75	0.81	0.7843
	High stress		0.29		0.72		0.41	
Decision Tree	Low stress	0.80	0.89	0.77	0.83	0.78	0.86	0.6056
	High stress		0.26		0.38		0.31	
Gaussian Naïve Bayes	Low stress	0.84	0.92	0.81	0.86	0.82	0.88	0.7879
	High stress		0.36		0.51		0.42	
Support Vector Machine	Low stress	0.82	0.91	0.69	0.71	0.73	0.80	0.7284
	High stress		0.24		0.57		0.34	

Table S2. Model performance metrics of the ensemble models versus the previous Gaussian Naïve Bayes model on the *International* dataset (Adamson).

Model	Classification	Precision		Recall		F1-score		Area under ROC
Gaussian Naïve Bayes	Low stress	0.84	0.92	0.81	0.86	0.82	0.88	0.7879
	High stress		0.36		0.51		0.42	
Random Forest	Low stress	0.82	0.89	0.83	0.91	0.82	0.90	0.6814
	High stress		0.36		0.30		0.33	
Stacking Model	Low stress	0.86	0.95	0.73	0.73	0.77	0.82	0.7895
	High stress		0.31		0.74		0.44	

Table S3. Model performance metrics of the deep learning approach on the International dataset (Adamson).

Model	Classification	Precision		Recall		F1-score		Area under ROC
Deep Learning	Low stress	0.82	0.91	0.75	0.79	0.78	0.84	0.7348
	High stress		0.29		0.53		0.37	

Table S4. Model performance metrics of the logistic regression model with added interaction features on the *International* dataset (Adamson).

Model	Classification	Precision		Recall		F1-score		Area under ROC
Logistic Regression with Interaction	Low stress	0.84	0.93	0.72	0.72	0.76	0.81	0.7805
	High stress		0.28		0.68		0.40	

