

Supplementary Table 4: Metrics and Hyperparameters of ML models

Data Integration Models

Early Integration Models

Model	AUC-ROC	AUC-PR	F1 Score	Precision	Recall	Hyperparameters
Random Forest	0.923	0.9705	0.9048	0.9383	0.8736	class_weight={0: 0.986, 1: 0.013}, criterion='entropy', max_depth=19, max_features=23, min_samples_leaf=2, min_samples_split=5, n_estimators=250, warm_start=True
MLP	0.9289	0.9713	0.8944	0.973	0.8276	activation='tanh', alpha=0.000139, batch_size=64, early_stopping=True, hidden_layer_sizes=(87, 87, 87), learning_rate='invscaling', momentum=0.837, solver='adam'
Voting Classifier	0.9418	0.9784	0.8875	0.9726	0.8161	voting='soft', combination of RF, GNB, SVC, LR, LightGBM, and MLP
SVC	0.872	0.9592	0.8805	0.9722	0.8046	C=40, class_weight={0: 0.727, 1: 0.273}, decision_function_shape='ovo', gamma='auto', kernel='sigmoid', probability=True
LightGBM	0.7071	0.8427	0.8404	0.7822	0.908	class_weight={0: 0.711, 1: 0.289}, learning_rate=0.28, max_depth=4, num_leaves=40, reg_alpha=1.624, reg_lambda=0.078, subsample=0.833, colsample_bytree=0.722
Gaussian NB	0.868	0.9485	0.8375	0.9178	0.7701	var_smoothing=1.0
Logistic Regression	0.9401	0.9786	0.7832	1.0	0.6437	C=10, class_weight={0: 0.951, 1: 0.049}, l1_ratio=0.3, multi_class='ovr', penalty='l1', solver='liblinear', warm_start=True

Intermediate Integration Models

Model	AUC-ROC	AUC-PR	F1 Score	Precision	Recall
Autoencoder without SMOTE	0.837	0.934	0.852	0.833	0.872
Autoencoder with SMOTE and CV	0.843	0.933	0.828	0.915	0.756

Late Integration Models

Model	AUC-ROC	AUC-PR	F1 Score	Precision	Recall	Best Parameters
XGBoost	0.8919	0.9605	0.875	0.8556	0.8953	subsample=0.5, n_estimators=300, min_child_weight=6, max_depth=12, learning_rate=0.01, gamma=0, colsample_bytree=0.75
Random Forest	0.8906	0.959	0.8736	0.8636	0.8837	warm_start=False, n_estimators=400, min_samples_split=4, min_samples_leaf=9, max_features='log2', max_depth=50, criterion='gini', class_weight=None, bootstrap=True
Stacking Classifier	0.8724	0.9553	0.8621	0.8523	0.8721	final_estimator__solver='saga', final_estimator__penalty='l2', final_estimator__max_iter=200, final_estimator__C=0.359
Voting Classifier	0.8663	0.9528	0.8428	0.9178	0.7791	weights=[1, 1, 1], voting='soft', flatten_transform=False

Individual Modality Models

Biospecimen Modality

Model	AUC-ROC	AUC-PR	F1 Score	Precision	Recall	Hyperparameters
Gaussian NB	0.6343	0.8261	0.8134	0.7391	0.9043	var_smoothing=0.1207
SVC	0.5362	0.8554	0.8039	0.7455	0.8723	C=100, class_weight={0: 0.7273, 1: 0.2727}, gamma='auto', kernel='poly', probability=True, shrinking=False

MLP	0.6337	0.8473	0.7847	0.713	0.8723	activation='tanh', alpha=5.1795, batch_size=32, early_stopping=True, hidden_layer_sizes=(187, 187), learning_rate='adaptive', learning_rate_init=0.0001, max_iter=500, momentum=0.5, solver='sgd'
Logistic Regression	0.6508	0.8471	0.7582	0.7841	0.734	C=0.1, class_weight={0: 0.7000, 1: 0.3000}, l1_ratio=0.5, multi_class='ovr', penalty='l2', solver='liblinear', warm_start=True
MLP with SMOTE	0.6489	0.8422	0.75	0.8049	0.7021	SMOTE(random_state=42, sampling_strategy=1.0), MLPClassifier(alpha=0.0268, batch_size=128, early_stopping=True, hidden_layer_sizes=(60, 60, 60), learning_rate='adaptive', max_iter=500, momentum=0.5)
Voting Classifier	0.6489	0.8422	0.75	0.8049	0.7021	voting='soft', combination of RF, GNB, SVC, MLP with SMOTE
LightGBM	0.6029	0.8218	0.7447	0.7447	0.7447	class_weight={0: 0.6239, 1: 0.3761}, colsample_bytree=0.6667, learning_rate=0.07, max_depth=4, min_child_samples=40, num_leaves=20, reg_alpha=0.1438, reg_lambda=0.0001, subsample=0.6111
Random Forest	0.6467	0.8017	0.716	0.8529	0.617	bootstrap=False, class_weight={0: 0.7353, 1: 0.2647}, criterion='entropy', max_depth=1, max_features=4, n_estimators=180

Clinical Modality

Model	AUC-ROC	AUC-PR	F1 Score	Precision	Recall	Hyperparameters
Voting Classifier	0.9392	0.974	0.9215	0.9362	0.9072	voting='soft', combination of RF, GNB, SVC, MLP

LightGBM	0.9296	0.9689	0.9101	0.9348	0.8866	class_weight={0: 0.767, 1: 0.233}, learning_rate=0.12, max_depth=10, num_leaves=120, reg_alpha=0.0428, reg_lambda=0.00034, subsample=0.7222
Logistic Regression	0.9418	0.976	0.9043	0.9341	0.8763	C=1, class_weight={0: 0.650, 1: 0.350}, l1_ratio=0.5, penalty='l1', solver='saga', multi_class='ovr', warm_start=True
SVC	0.8506	0.9439	0.8995	0.9239	0.8763	C=50, class_weight={0: 0.660, 1: 0.340}, gamma='auto', kernel='sigmoid', probability=True
Random Forest	0.9227	0.9647	0.8973	0.9432	0.8557	bootstrap=False, class_weight={0: 0.800, 1: 0.200}, criterion='entropy', max_depth=6, max_features=7, min_samples_leaf=2, n_estimators=350
MLP	0.9289	0.9689	0.8962	0.8261	0.9794	activation='tanh', alpha=0.00193, batch_size=32, hidden_layer_sizes=(188, 188, 188), early_stopping=True, momentum=0.95, n_iter_no_change=15
Gaussian NB	0.9015	0.9592	0.807	0.9324	0.7113	var_smoothing=0.0339

Genetic Modality

Model	AUC-ROC	AUC-PR	F1 Score	Precision	Recall	Hyperparameters
MLP	0.5332	0.7854	0.7593	0.7736	0.7455	alpha=0.3728, batch_size=128, early_stopping=True, hidden_layer_sizes=(51,), learning_rate_init=0.0001, max_iter=500, momentum=0.8375, solver='sgd'

LightGBM	0.579 1	0.821 9	0.7143	0.75	0.681 8	class_weight={0: 0.608, 1: 0.392}, colsample_bytree=0.5, learning_rate=0.16, max_depth=3, min_child_samples=120, n_estimators=150, num_leaves=60, reg_alpha=0.00034, reg_lambda=0.0001, subsample=0.8889
MLP-SMOTE	0.599 5	0.844 2	0.6935	0.7753	0.627 3	SMOTE pipeline with MLP(activation='tanh', alpha=0.3728, batch_size=128, early_stopping=True, hidden_layer_sizes=(90, 90, 90), learning_rate_init=0.0001, max_iter=500, momentum=0.8375)
Voting Classifier	0.599	0.828 2	0.6632	0.7875	0.572 7	voting='soft', combination of RF, GNB, SVC, MLP with SMOTE
Random Forest	0.623	0.843	0.6559	0.8026	0.554 5	bootstrap=False, class_weight={0: 0.678, 1: 0.322}, criterion='entropy', max_depth=3, max_features=7, min_samples_leaf=4, min_samples_split=10, warm_start=True
SVC	0.556 4	0.835 2	0.6264	0.7917	0.518 2	C=0.1, class_weight={0: 0.698, 1: 0.302}, kernel='linear', probability=True
Gaussian NB	0.600 1	0.824 3	0.5909	0.7879	0.472 7	var_smoothing=1.5999e-07
Logistic Regression	0.595 6	0.814 4	0.5778	0.7429	0.472 7	C=0.1, class_weight={0: 0.703,