

**Best Models Used:**

Model	Hyperparameters
Decision Tree	criterion= 'gini', max_depth= None, min_samples_leaf= 5, min_samples_split= 2
Bagging	DecisionTreeClassifier(), max_samples= 0.5, n_estimators= 100
Random Forest	criterion= 'gini', max_depth= None, max_samples= 0.4, n_estimators= 90
Gradient Boosting	learning_rate= 0.1, max_depth= 7, n_estimators= 200
XGBoost(XGBoost3)	learning_rate=0.02, max_depth=7, n_estimators=800, use_label_encoder=False, eval_metric='logloss'
Logistic Regression	C= 10, penalty= 'l2', solver= 'liblinear'
Neural Network (MLP)	max_iter= 1000, learning_rate_init= 0.01, learning_rate= 'constant', hidden_layer_sizes= (60, 60, 50), batch_size= 32, alpha= 0.001, activation= 'tanh'

**Metrics For All Models:**

Model	Optimal Threshold	Accuracy	Precision	Recall	F1-Score	AUC-ROC
XGBoost3	0.5617	0.8427	0.5779	0.7302	0.6452	0.8608
XGBoost	0.5533	0.8358	0.5630	0.7215	0.6324	0.8573
XGBoost2	0.5528	0.8354	0.5618	0.7255	0.6332	0.8571
Bagging (PCA)	0.5598	0.8321	0.5541	0.7297	0.6299	0.8516
Random Forest	0.5528	0.8257	0.5409	0.7276	0.6205	0.8481
Gradient Boosting	0.5155	0.8195	0.5278	0.7425	0.6170	0.8554
Decision Tree (PCA)	0.6364	0.8048	0.5011	0.6762	0.5756	0.8106
MLP Classifier	0.6077	0.7181	0.3672	0.6075	0.4577	0.7078
Logistic Regression	0.4767	0.5835	0.2829	0.7343	0.4084	0.6315

**Champion Model:**

XGBoost3 is our champion model. XGBoost3 has the highest accuracy, precision, f1-score, and AUC-ROC score. In terms of computation time, this grid search for this model executed within 30 minutes. Additionally, when edits were made to hyperparameters within a model fit, it

executed within a reasonable time of 3-5 minutes. This is an exceptional time given the number of observations present in our dataset. We recognize that this model could be improved upon by more complex grid search, if we were not constrained by time. Also, additional experimentation within the data cleaning process may result in more insightful data for model building. Overall, we find our model XGBoost3 to be the best, yet there are many improvements that could be made to the model.