(Working Title): Gaze-Based Mind Windering Detection using Deep Learning

Derek Harter
Subroto Singha
Texas A&M University - Commerce

September 15, 2021

Abstract

Appendix of all figures, tables and materials generated by the project. Not all figures and tables may be used in final document, so can refer to this document for details of data analysis, visualizations, summary tables and other products of this project.

Appendix

Tables

Table 1: Summary of all pipeline transformer and estimator parameters used in grid search to determine best estimator parameters.

Model	Pipeline	Parameter	Search Values
LogisticRegression	ClassImbalanceTransformer	balancer_type	allknn, smote-enn
	SelectKBest	k	12, 15, 18, 21, 24, 31
		score_func	f_classif
	LogisticRegression	C	10.00, 19.68, 38.75, 76.27, 150.13, 2
		max_iter	50000
		solver	lbfgs, liblinear, newton-cg
	WinsorizationOutlierTransformer	outlier_threshold	0.00, 3.00
	FeatureScalerTransformer	type_of_scaling	standard
	VIFThresholdTransformer	score_threshold	0
kNN	ClassImbalanceTransformer	balancer_type	none, random-undersampler, smote
	SelectKBest	k	15, 31, 46, 62
		score_func	f_classif, f_regression
	KNeighborsClassifier	metric	euclidean, manhattan
		n_neighbors	3, 5, 11, 19
		weights	uniform, distance

Continued on

Table 1: Summary of all pipeline transformer and estimator parameters used in grid search to determine best estimator parameters.

Model	Pipeline	Parameter	Search Values
	WinsorizationOutlierTransformer	outlier_threshold	0.00, 2.00, 3.00
	FeatureScalerTransformer	type_of_scaling	standard
	VIFThresholdTransformer	score_threshold	0
SVM	ClassImbalanceTransformer	balancer_type	allknn, smote-enn
	SelectKBest	k	12, 15, 18, 21, 24, 31
		score_func	f_classif
	SVC	C	0.10, 1, 10, 100
		gamma	0.10, 0.01, 0.00, 0.00
		kernel	rbf, poly
	WinsorizationOutlierTransformer	outlier_threshold	0.00, 3.00
	FeatureScalerTransformer	type_of_scaling	standard
	VIFThresholdTransformer	score_threshold	0
DecisionTree	ClassImbalanceTransformer	balancer_type	allknn, smote-enn
	SelectKBest	k	12, 15, 18, 21, 24, 31
		score_func	f_classif
	DecisionTreeClassifier	criterion	gini, entropy
		max_depth	3, 4, 5, 6, 7, 8, 9
		max_leaf_nodes	3, 4, 5, 6, 7, 8, 9
	WinsorizationOutlierTransformer	outlier_threshold	0.00, 3.00
	FeatureScalerTransformer	type_of_scaling	standard
	VIFThresholdTransformer	score_threshold	0
RandomForest	ClassImbalanceTransformer	balancer_type	allknn, smote-enn
	SelectKBest	k	12, 15, 18, 21, 24, 31
		score_func	f_classif
	RandomForestClassifier	criterion	gini, entropy
		max_depth	4, 5, 6, 7, 8
		max_features	auto, sqrt, log2
		n_estimators	200, 500
	WinsorizationOutlierTransformer	outlier_threshold	0.00, 3.00
	FeatureScalerTransformer	type_of_scaling	standard
	VIFThresholdTransformer	score_threshold	0
NaiveBayes	ClassImbalanceTransformer	balancer_type	allknn, smote-enn
	SelectKBest	k	12, 15, 18, 21, 24, 31
		score_func	f_classif
	GaussianNB	var_smoothing	0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.0
	WinsorizationOutlierTransformer	outlier_threshold	0.00, 3.00
	FeatureScalerTransformer	type_of_scaling	standard
	VIFThresholdTransformer	score_threshold	0

Table 2: Comparison of performance for best standard ML estimator found for each type by parameter grid search.

Model name	k-fold aucroc	final aucroc	accuracy	recall	precision
LogisticRegression	0.6258	0.5660	0.4485	0.8248	0.3090
kNN	0.6337	0.8304	0.7983	0.9012	0.5848
SVM	0.6580	0.6272	0.5999	0.6873	0.3735
DecisionTree	0.6428	0.6157	0.5559	0.7475	0.3524
RandomForest	0.6535	0.6273	0.5773	0.7376	0.3646
NaiveBayes	0.6565	0.6165	0.6349	0.5759	0.3868

Figures

Summary Figures

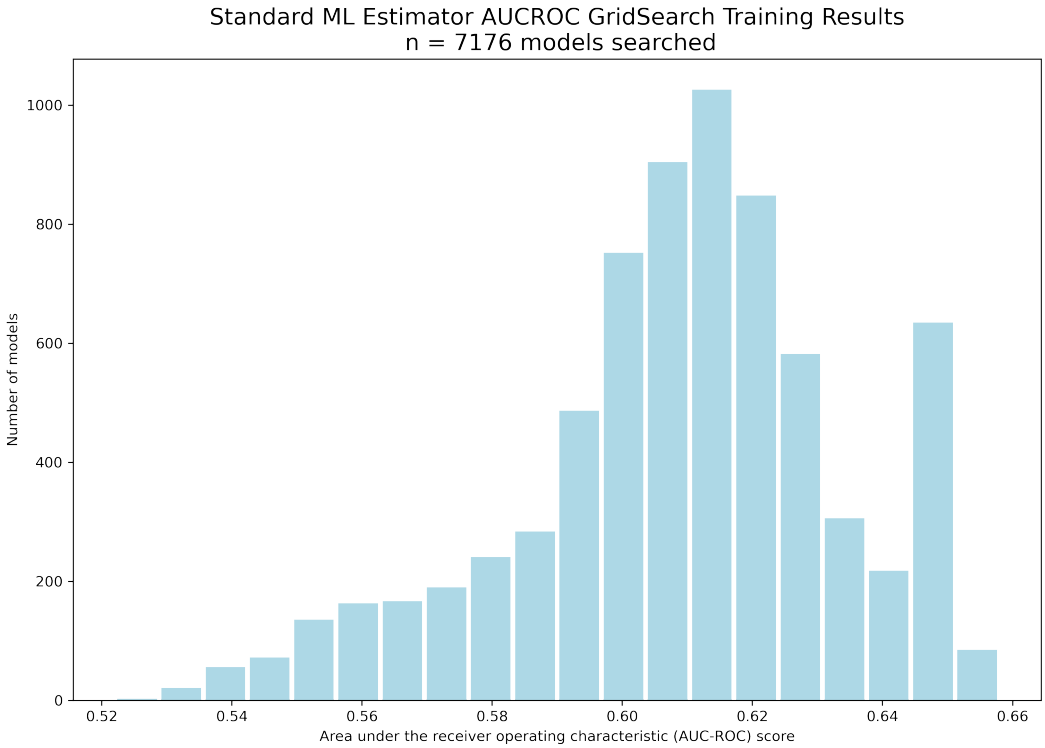

Figure 1: Histogram of AUCROC performance for parameter search of all standard ML trained models

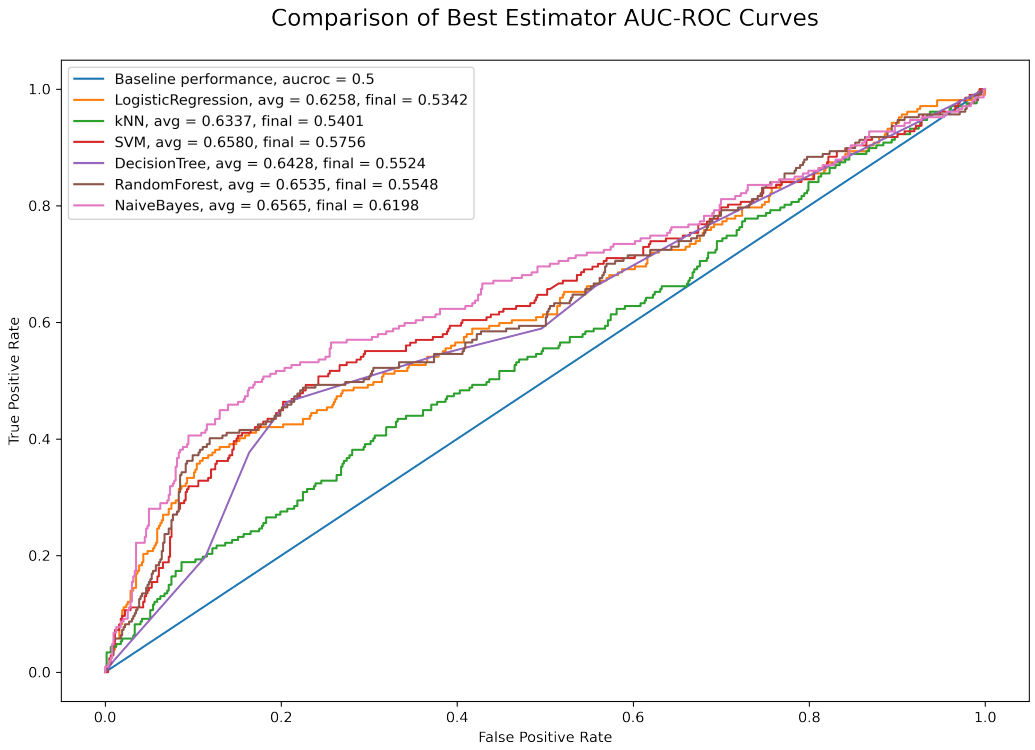

Figure 2: Comparison of AUCROC scores achieved by best standard ML models in each type of estimator that was explored.

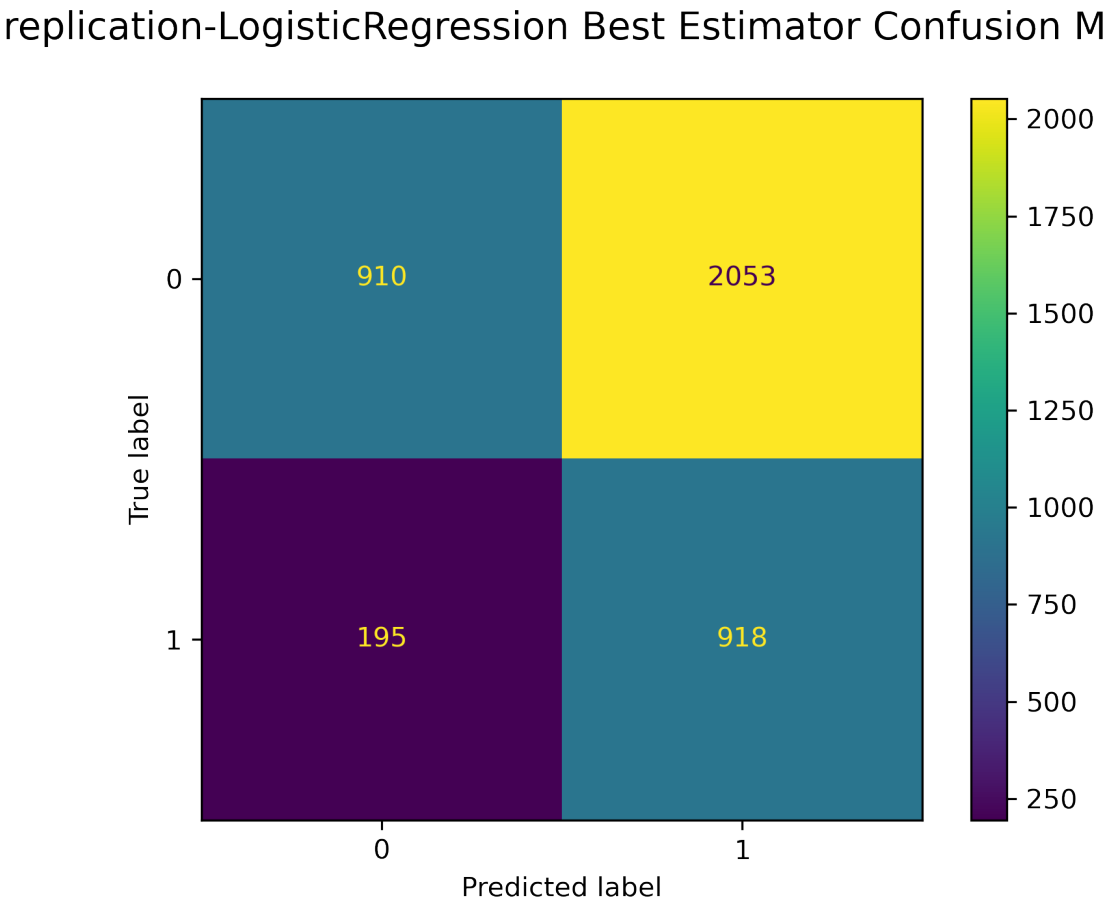

Figure 3: Logistic Regression best estimator final confusion matrix.

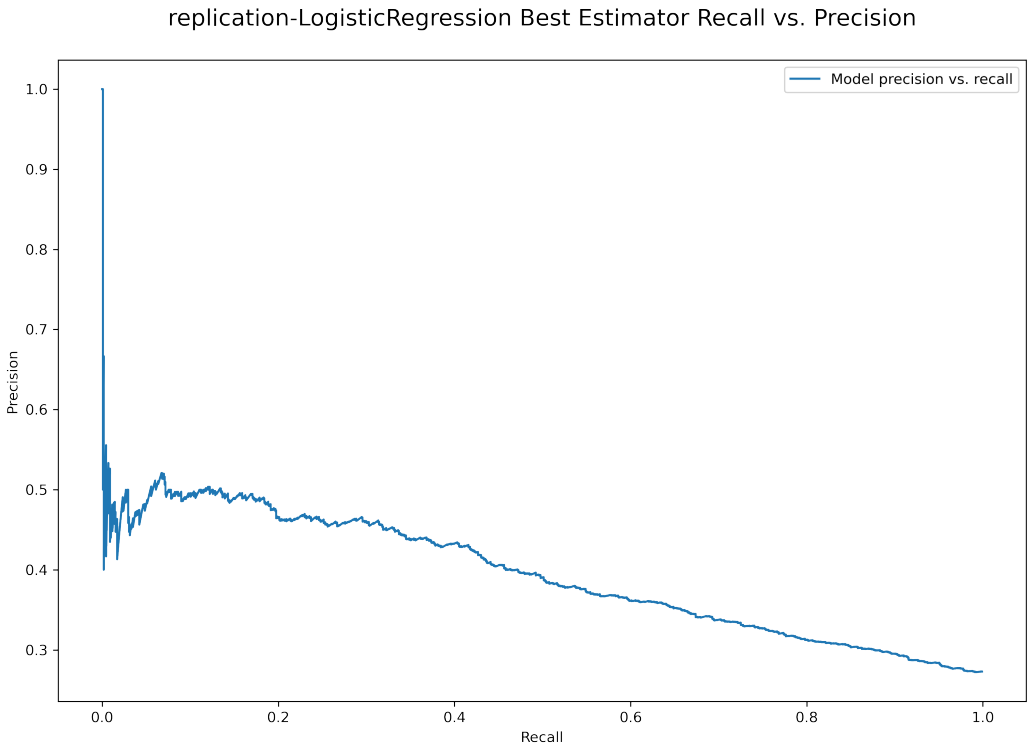

Figure 4: Logistic Regression best estimator recall vs. precision.

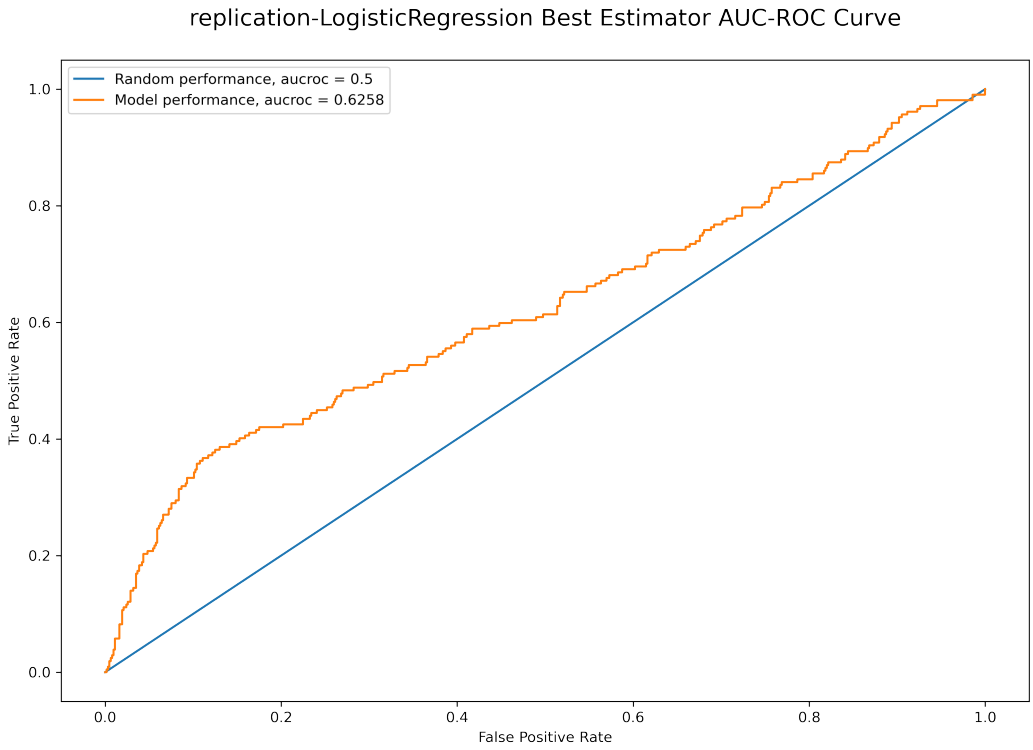

Figure 5: Logistic Regression best estimator auc-roc curve.

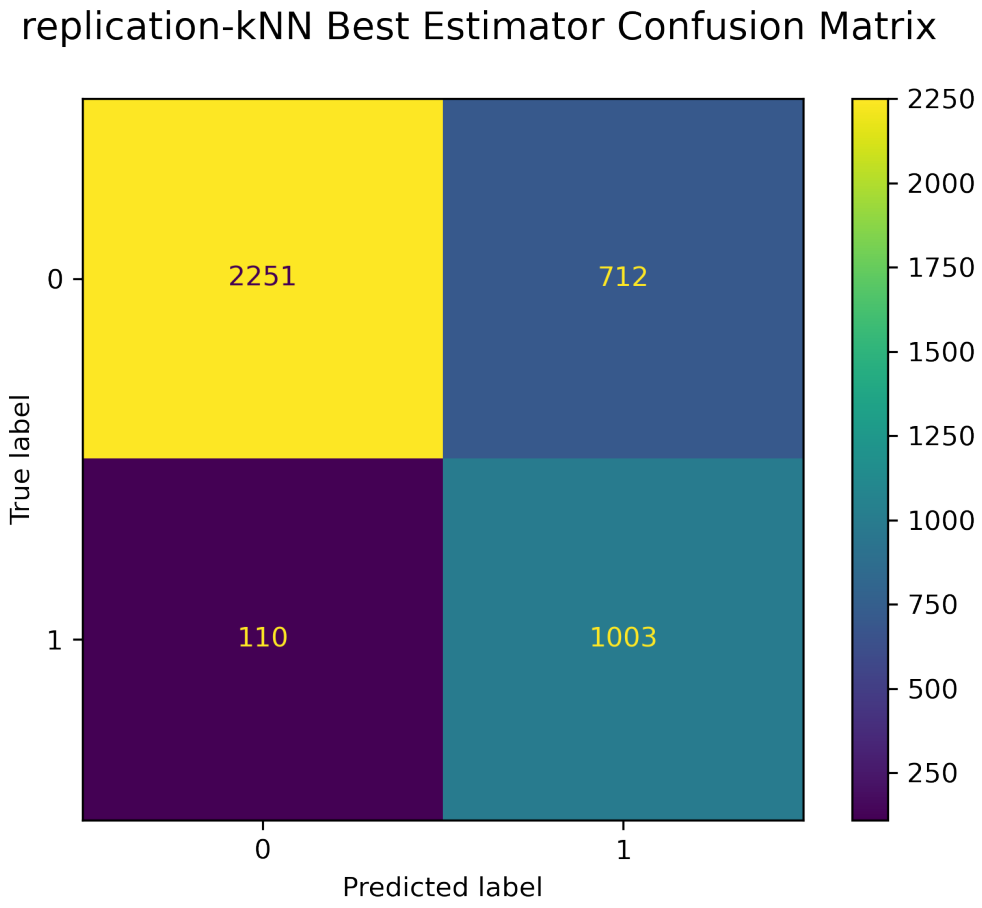

Figure 6: kNN k Nearest Neighbors best estimator final confusion matrix.

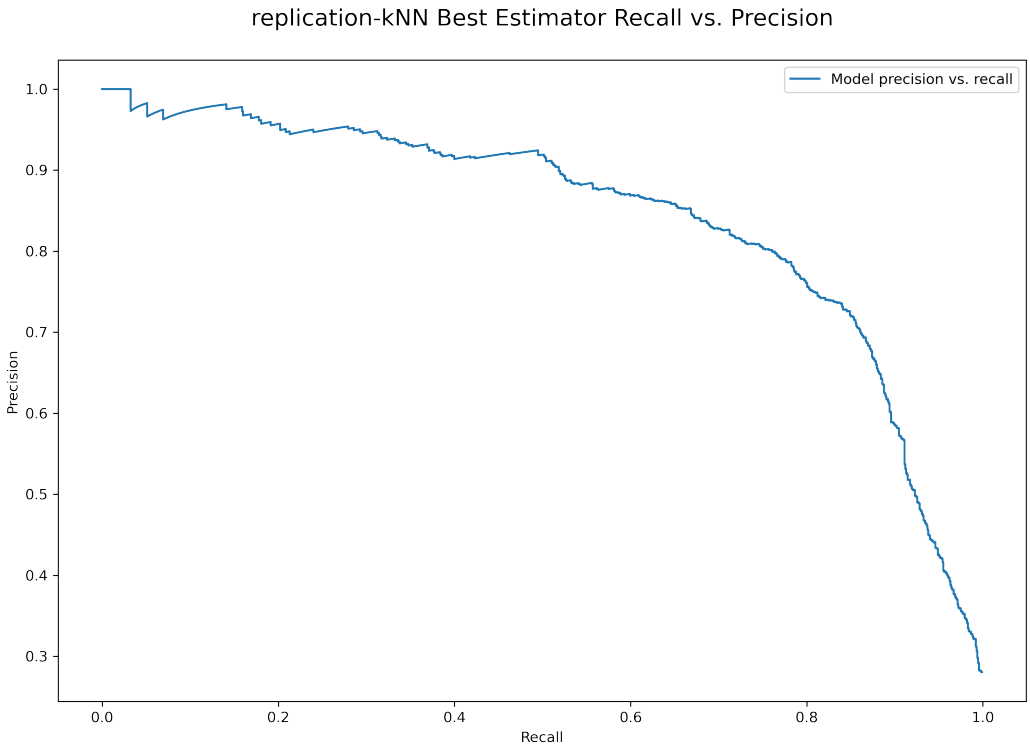

Figure 7: kNN k Nearest Neighbors best estimator recall vs. precision.

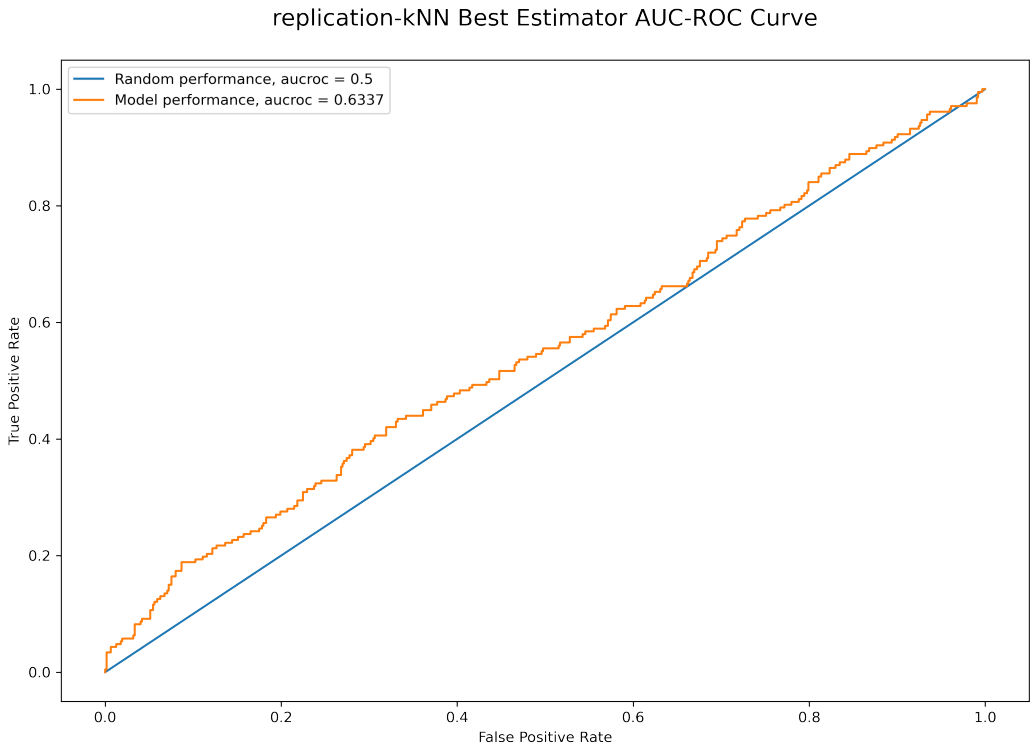

Figure 8: kNN k Nearest Neighbors best estimator auc-roc curve.

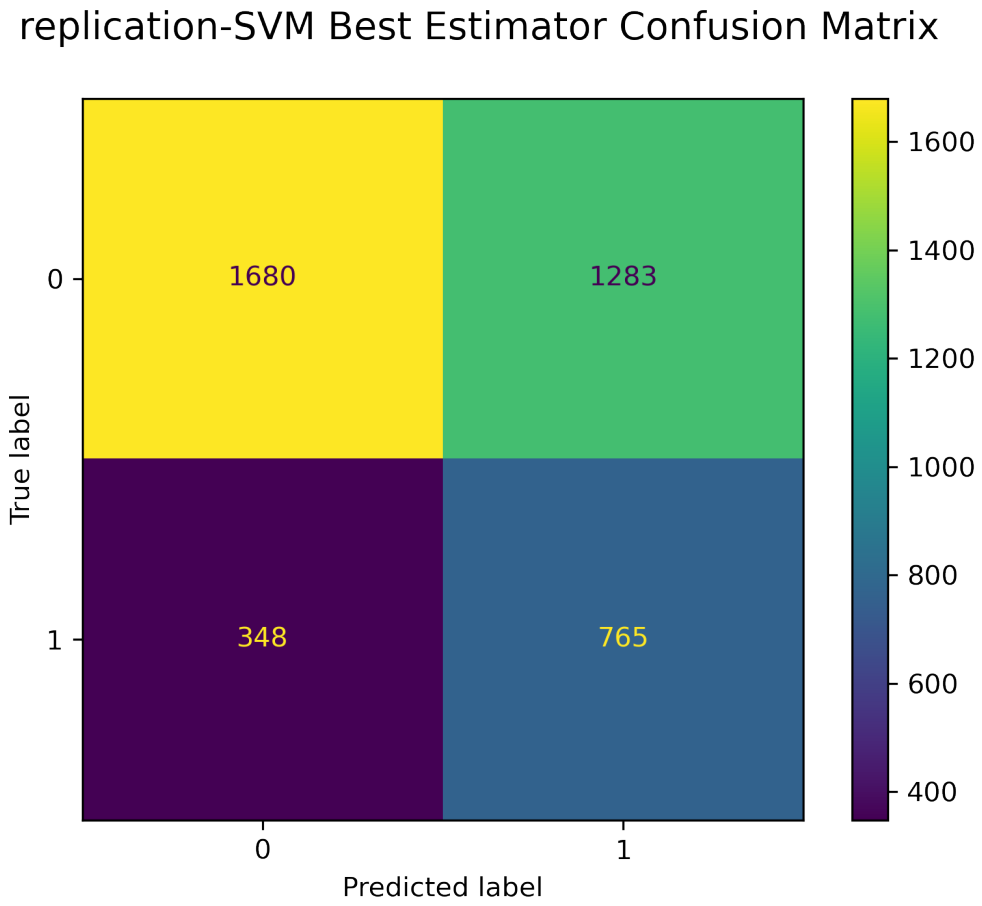

Figure 9: SVM Support Vector Machine best estimator final confusion matrix.

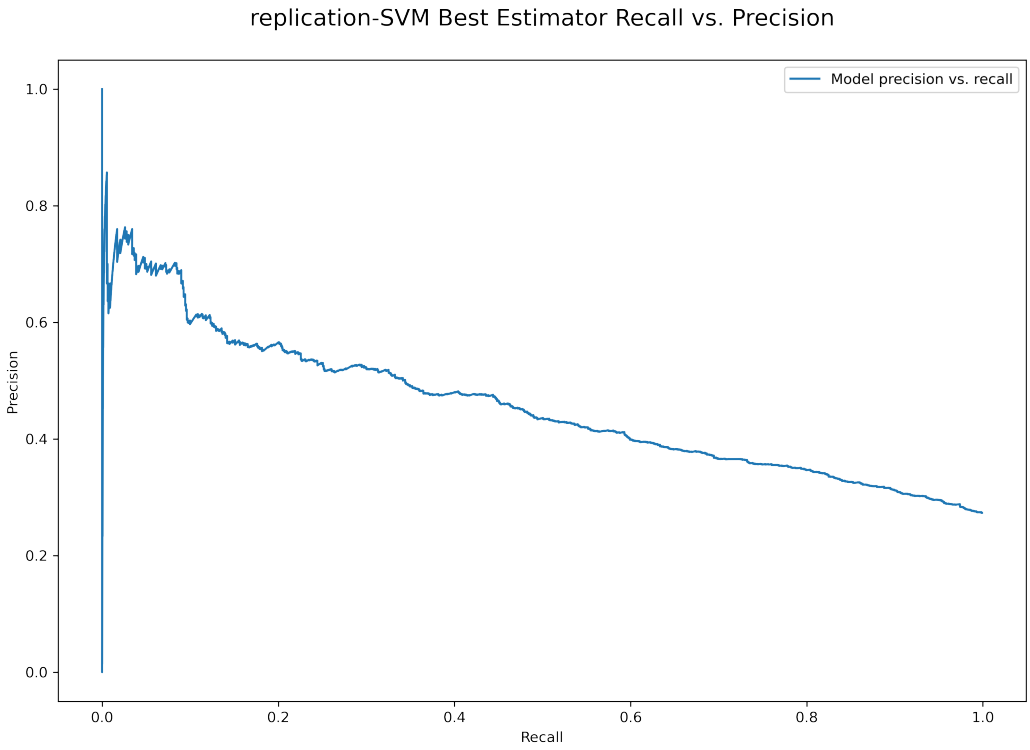

Figure 10: SVM Support Vector Machine best estimator recall vs. precision.

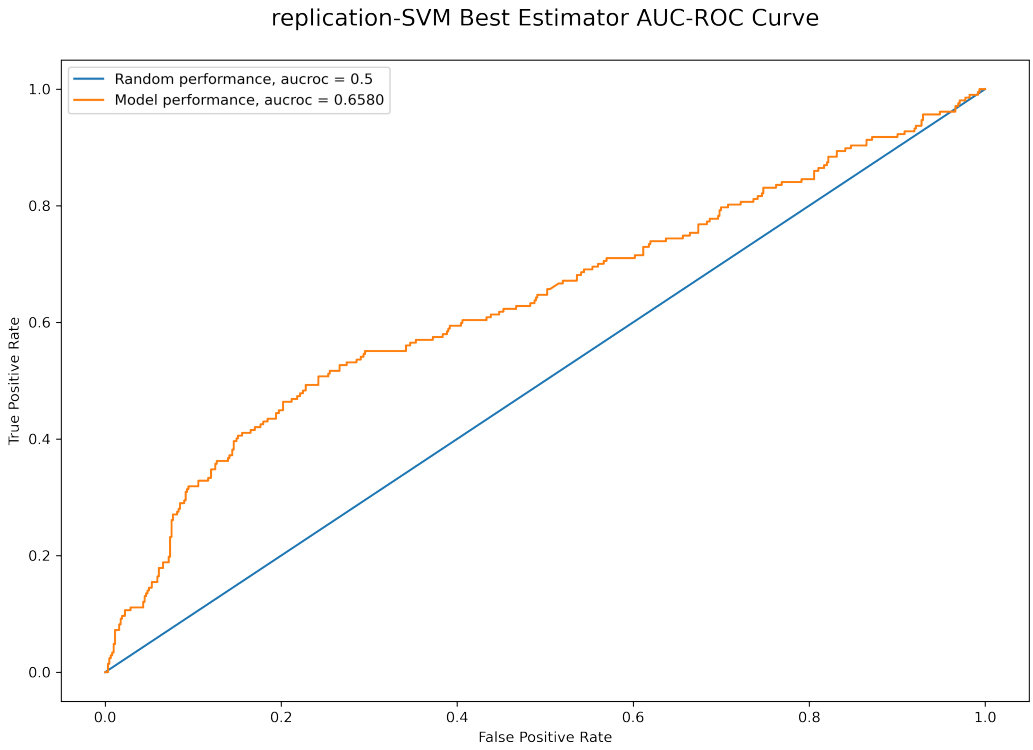

Figure 11: SVM Support Vector Machine best estimator auc-roc curve.

Logistic Regression

kNN k Nearest Neighbors

SVM Support Vector Machine

Decision Tree

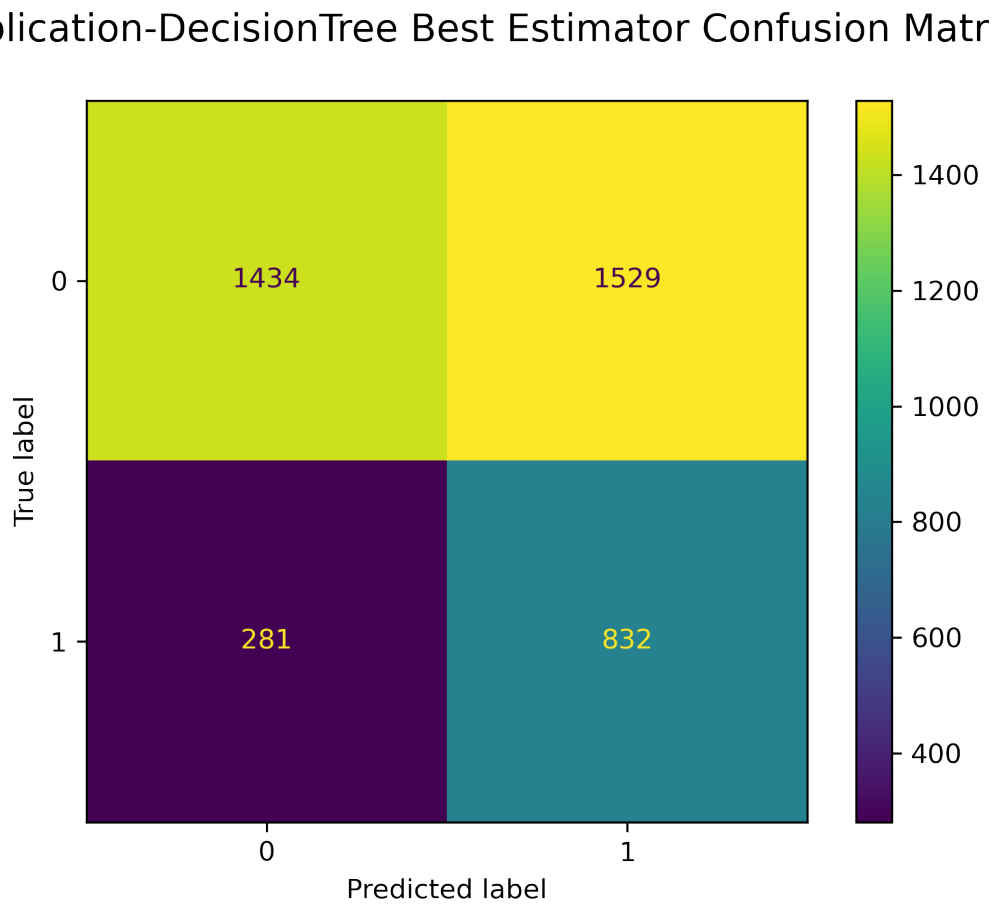

Figure 12: Decision Tree best estimator final confusion matrix.

Random Forest

Naive Bayes

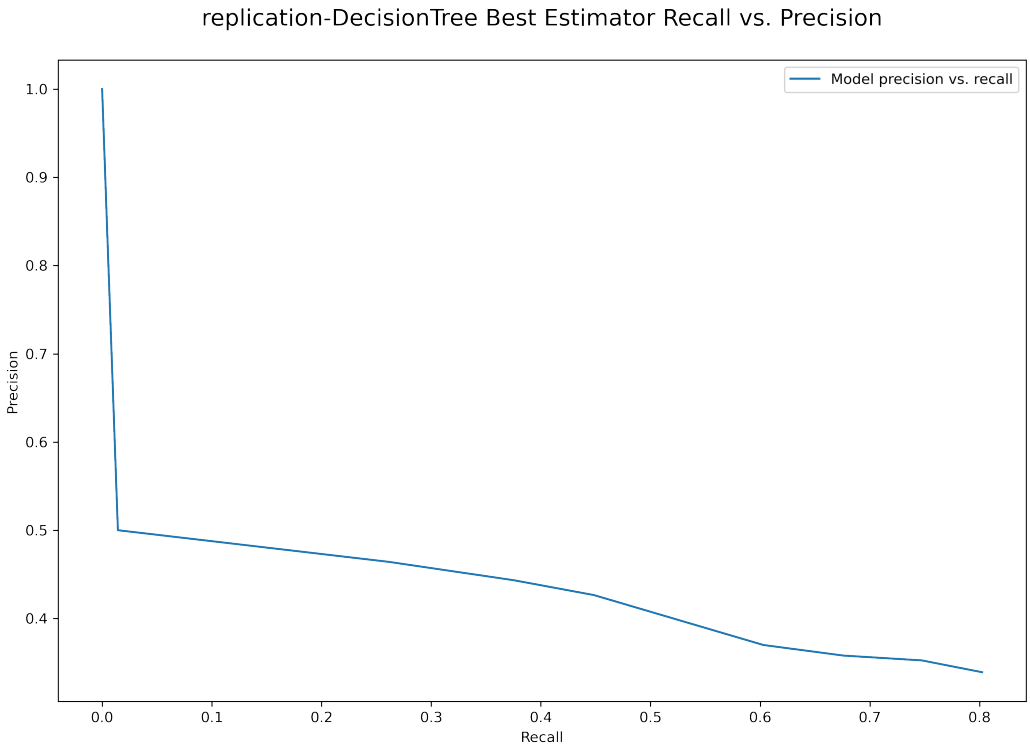

Figure 13: Decision Tree best estimator recall vs. precision.

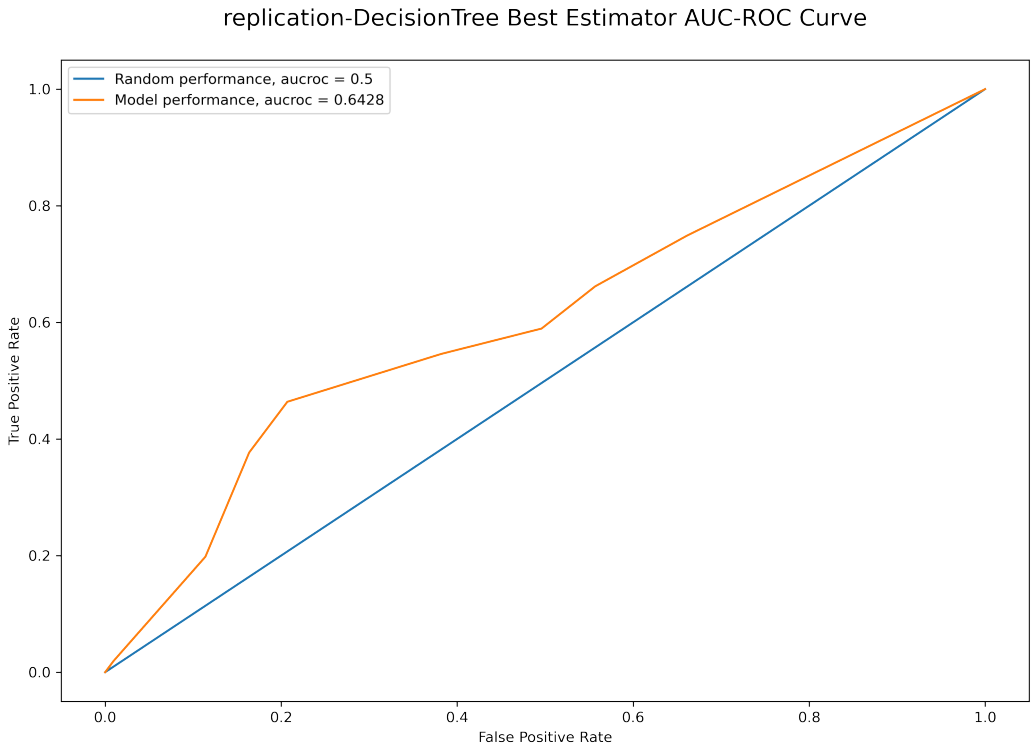

Figure 14: Decision Tree best estimator auc-roc curve.

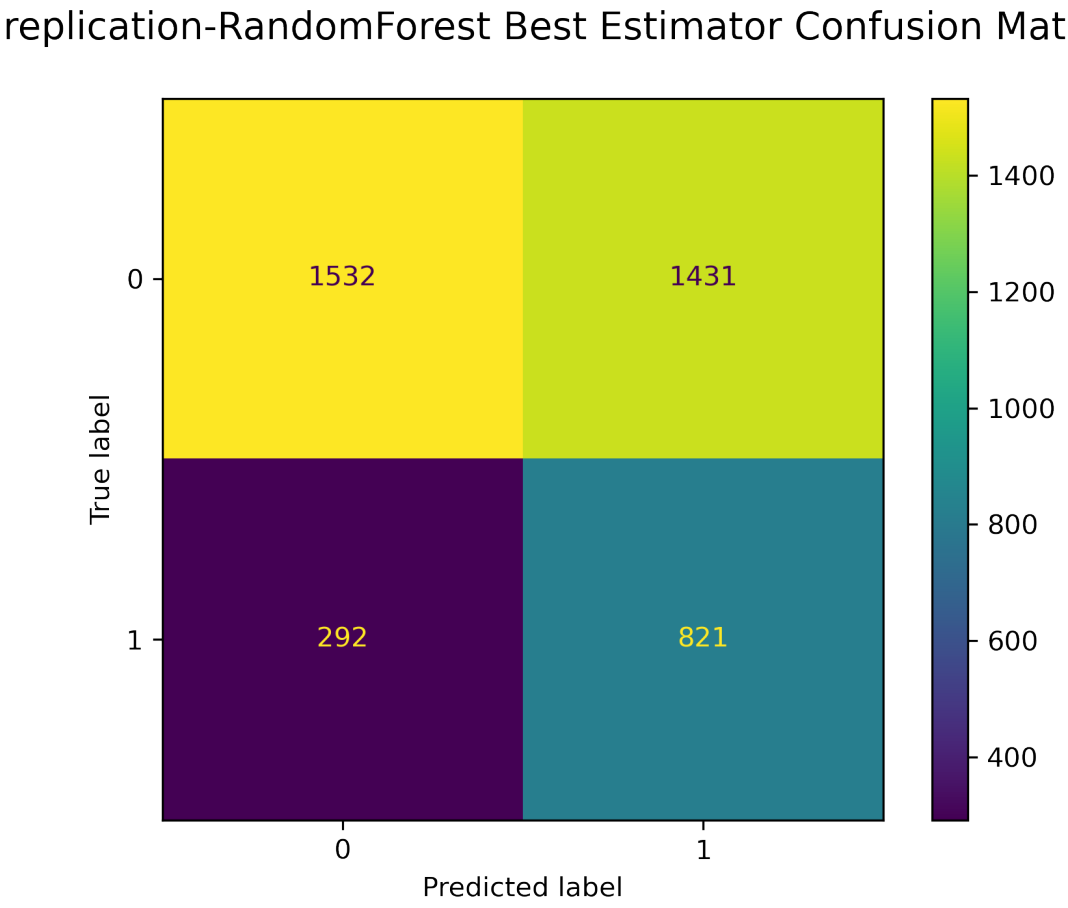

Figure 15: Random Forest best estimator final confusion matrix.

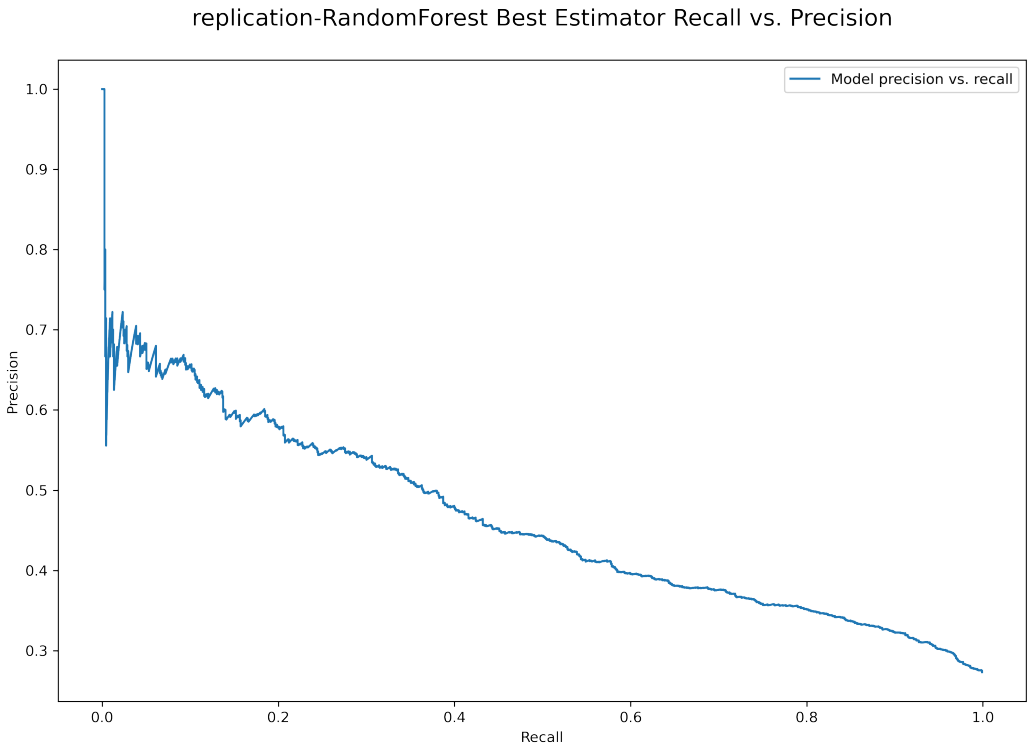

Figure 16: Random Forest best estimator recall vs. precision.

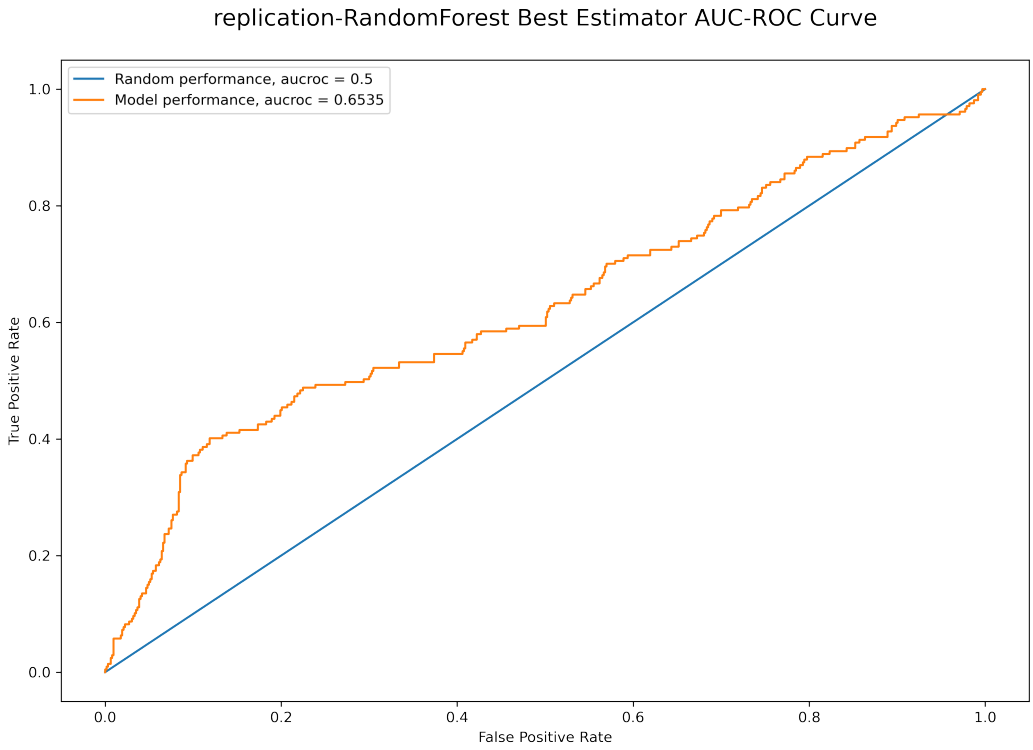

Figure 17: Random Forest best estimator auc-roc curve.

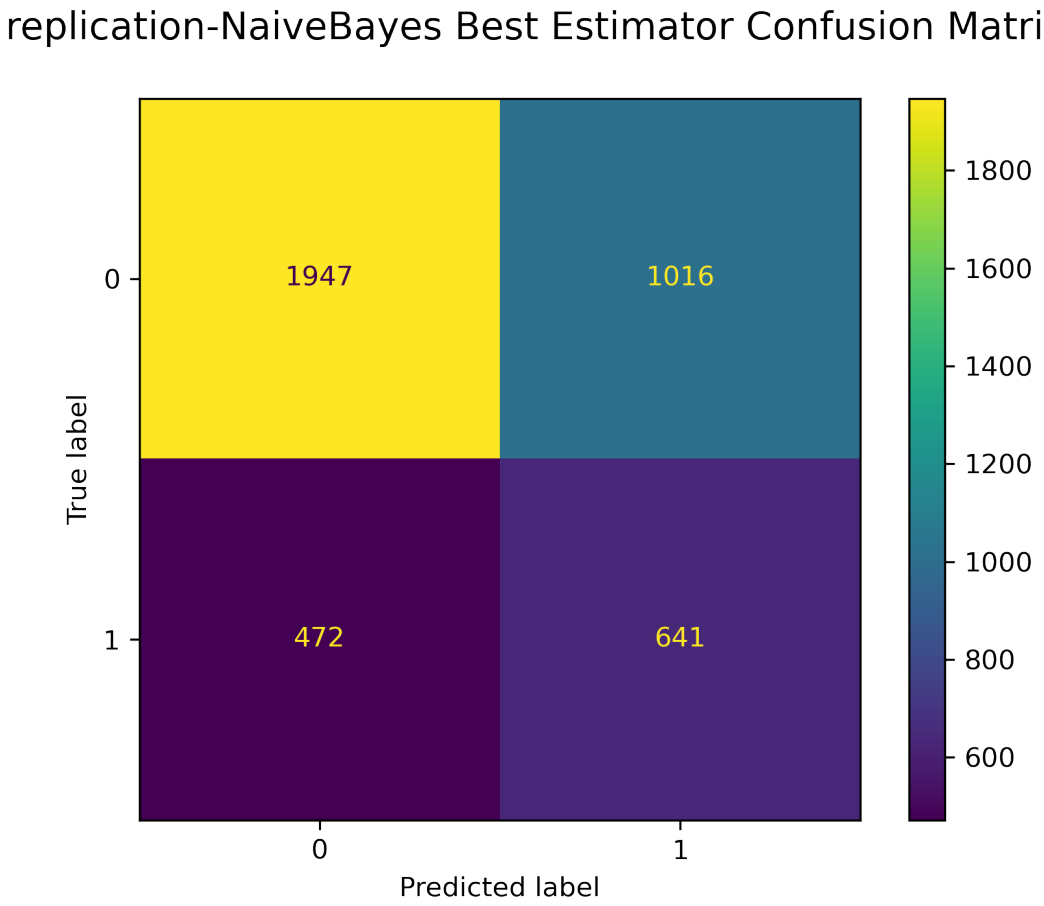

Figure 18: Naive Bayes best estimator final confusion matrix.

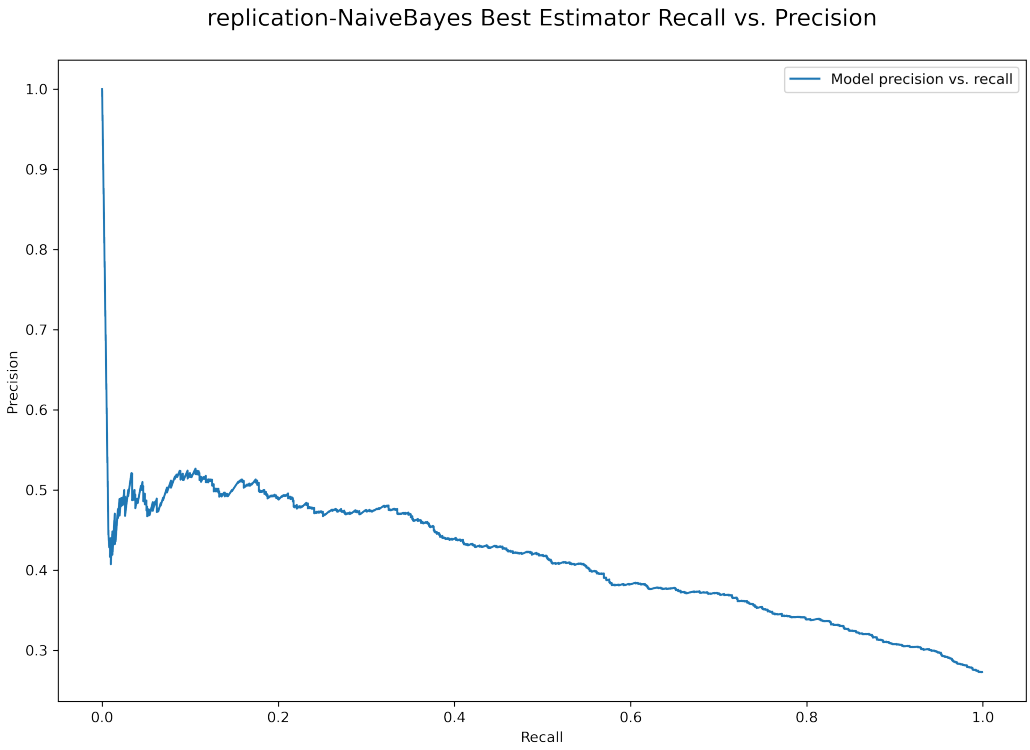

Figure 19: Naive Bayes best estimator recall vs. precision.

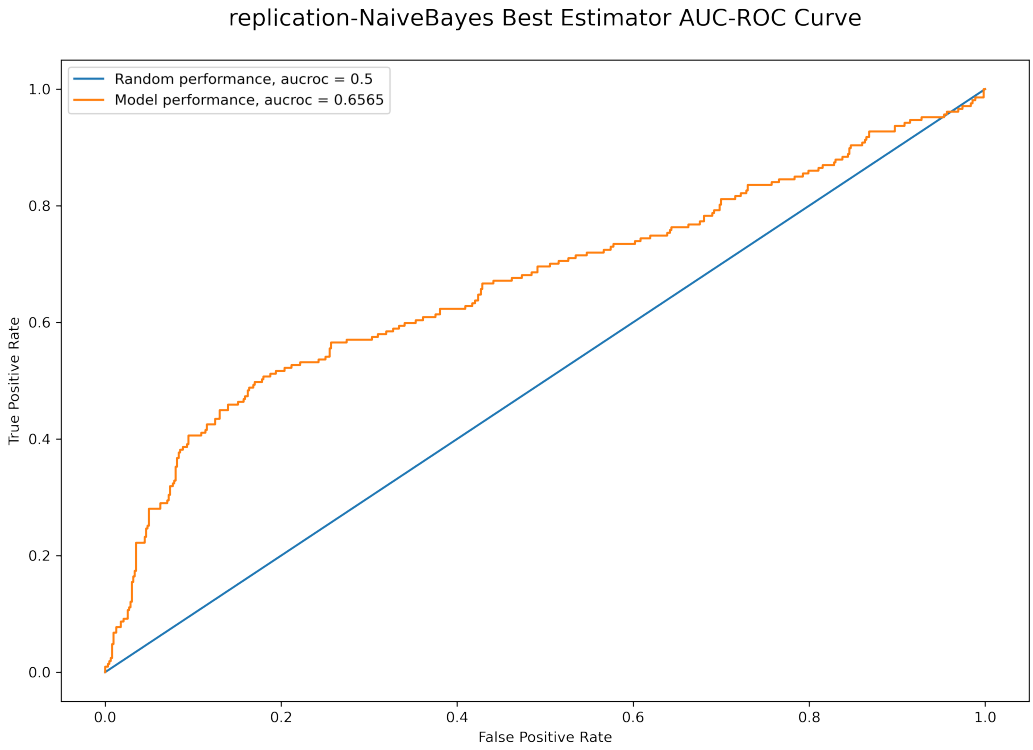

Figure 20: Naive Bayes best estimator auc-roc curve.