Data table properties

Name: titanic
Size: 2201 rows, 4 columns
Features: 3 categorical
Targets: categorical outcome with 2 classes

Additional attributes

Name: Titanic dataset
Description: Data from the real Titanic passengers and their survival.
**Author:** Kaggle
Year: 1912

Settings

Sampling type: Stratified 10-fold Cross validation
Target class: None, show average over classes

Scores

Model	AUC	CA	F1	Prec	Recall	MCC
Logistic Regression	0.743	0.770	0.756	0.764	0.770	0.448
Naive Bayes	0.707	0.772	0.758	0.765	0.772	0.451
Random Forest	0.749	0.783	0.751	0.810	0.783	0.489

Test and Scores on 70% training set.

Target class: no
Costs: FP = 500, FN = 500
Target probability: 67.0 %

| Logistic Regression | Naive Bayes | Random Forest |
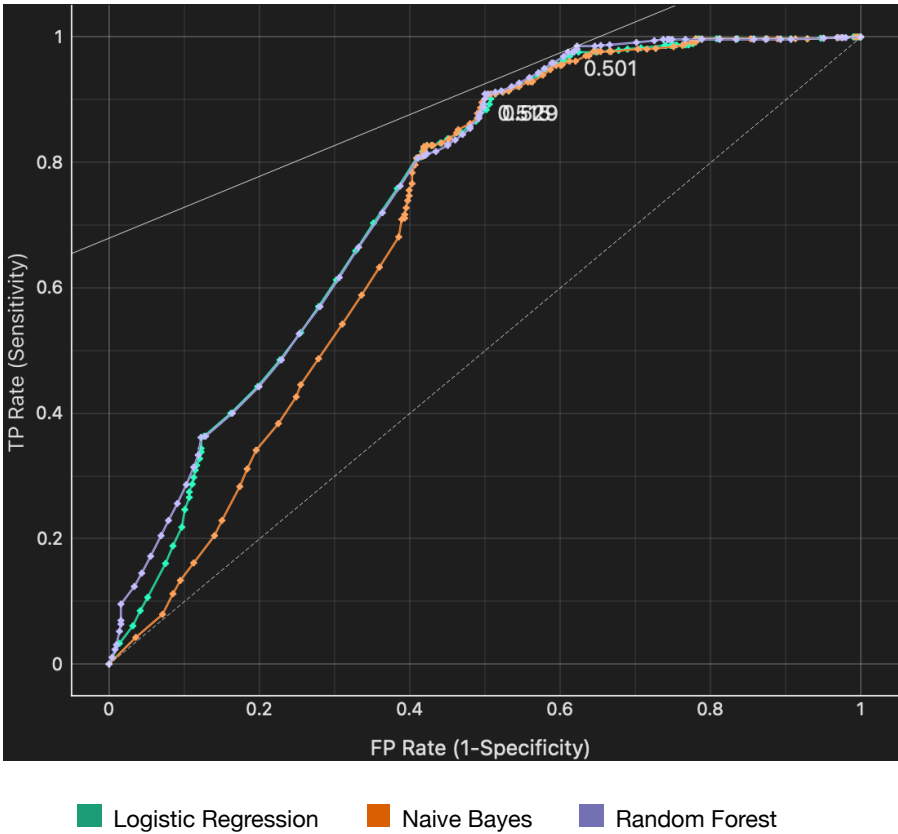
Naive Bayes Performs the worst.

Confusion matrix for Logistic Regression (showing number of instances)

Predicted

		no	yes	Σ
Actual	**no**	941	94	**1,035**
	yes	260	246	**506**
	Σ	**1,201**	**340**	**1,541**

Out of Logistic Regression, Naive Bayes and Random Forest.
L.R. performs relatively the based on ROC Analysis.