

5.

Dataset Info:

Data Info

Data Set Name

bupa

Data Set Size

Rows: 345  
Columns: 7

Features

Categorical: -  
Numeric: 6

Targets

Categorical outcome with 2 values

Meta Attributes

None

Location

Data is stored in memory

Data Attributes

?

345

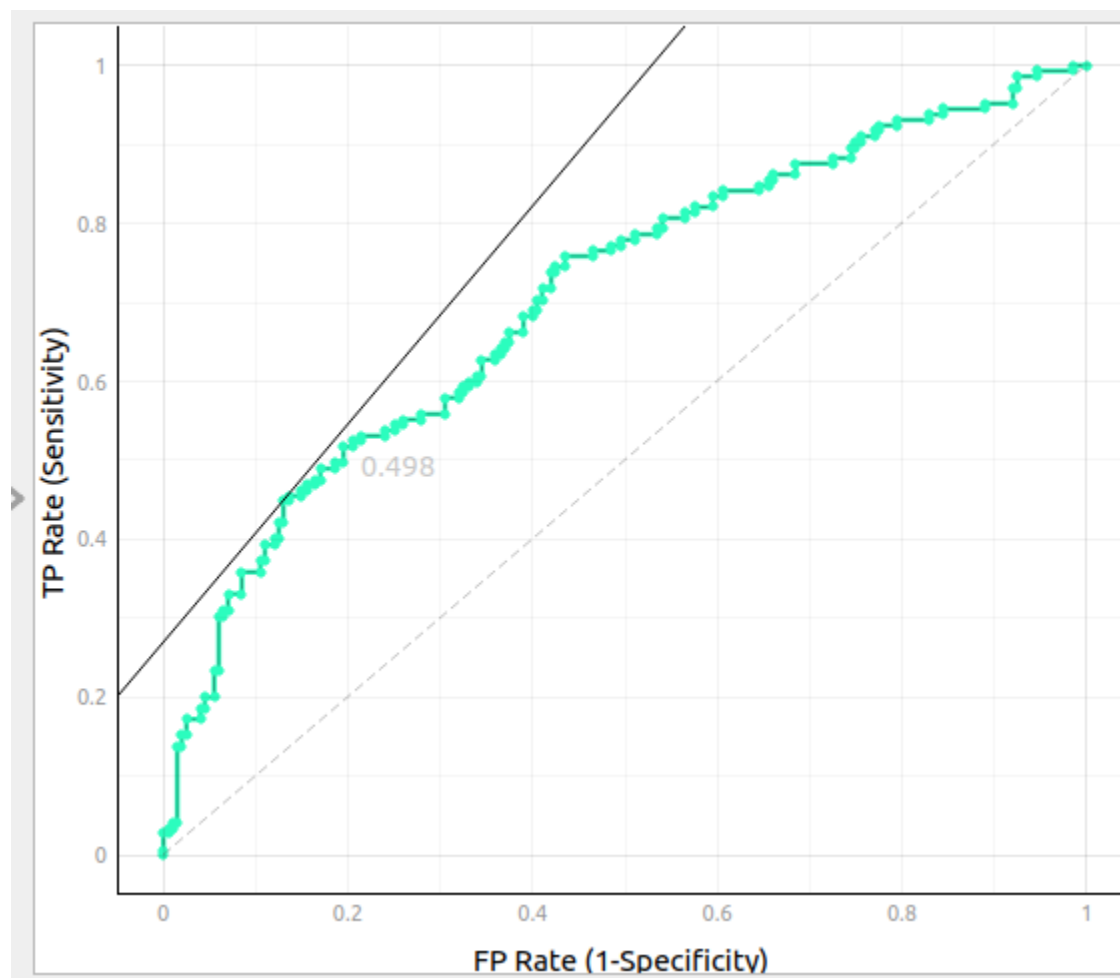
The BUPA dataset contains 345 single male patients with 6 numeric attributes. Five of these attributes are blood tests which are thought to be sensitive to liver disorders that might arise from excessive alcohol consumption. Each line in the dataset constitutes the record of a single male individual.

Target Variable : Liver Disorder Prediction (Y/N) [Binary]

Logistic Regression:

Evaluation Results						
Model	AUC	CA	F1	Precision	Recall	
Logistic Regression	0.706	0.678	0.671	0.674	0.678	

		Predicted		
		1	2	$\Sigma$
Actual	1	75	70	145
	2	41	159	200
	$\Sigma$	116	229	345



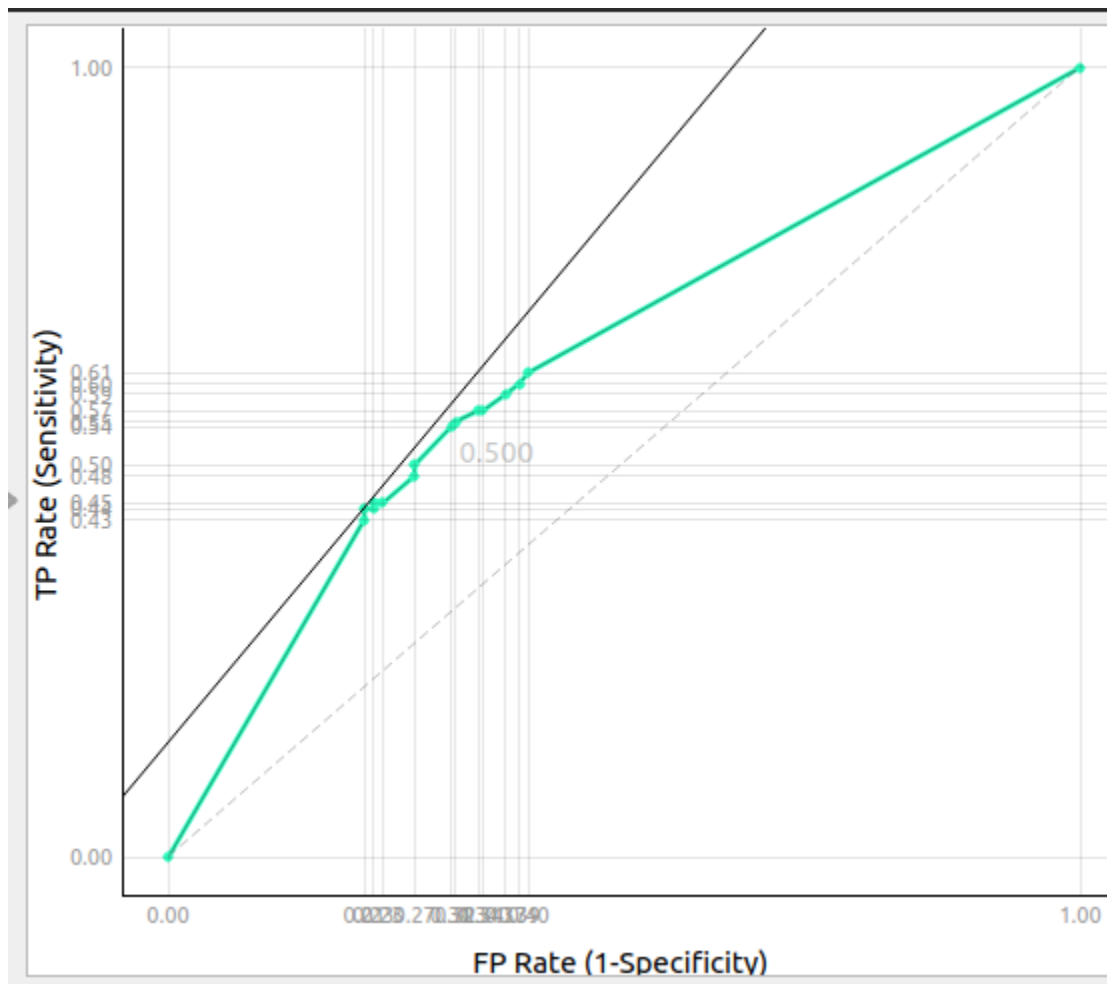
Decision Tree:

#### Evaluation Results

Model	AUC	CA	F1	Precision	Recall
Random Forest	0.629	0.629	0.628	0.628	0.629

Random Forest with 1 tree = Decision Tree

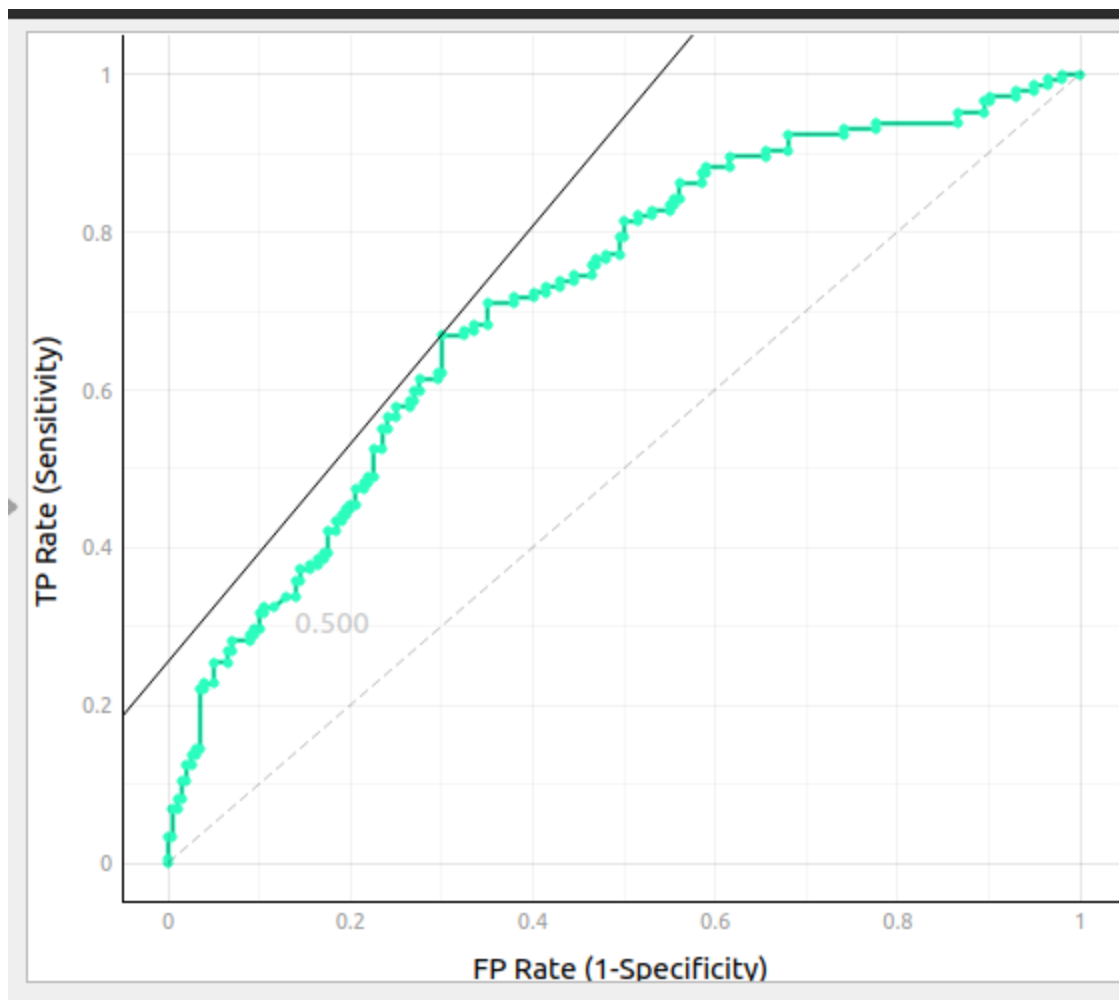
Actual	Predicted		$\Sigma$
	1	2	
	1	2	
1	79	66	145
2	62	138	200
$\Sigma$	141	204	345



Support Vector Machine :

Model ^	AUC	CA	F1	Precision	Recall
SVM	0.718	0.649	0.623	0.649	0.649

		Predicted		
		1	2	$\Sigma$
Actual	1	52	93	145
	2	28	172	200
	$\Sigma$	80	265	345



Naive Bayes:

Evaluation Results

Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.731	0.667	0.663	0.663	0.667

Model Confusion by AUC

Show: Number of instances

		Predicted		
		1	2	$\Sigma$
Actual	1	79	66	145
	2	49	151	200
	$\Sigma$	128	217	345

Select Correct

Select Misclassified

Clear Selection

