

# Blood Donation Prediction

## ETM 538

Peter Boss  
Mengyu Li  
Jordan Hilton  
Andey Nunes, MS

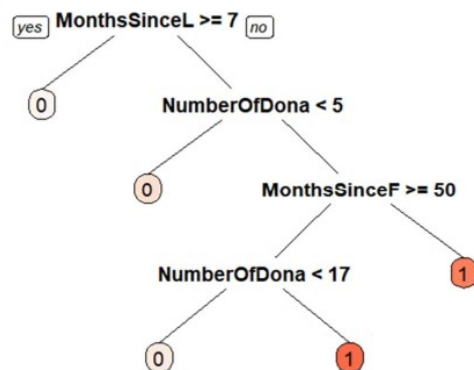
## Decision Tree – By Mengyu

Check the structure of data frame

```
'data.frame': 576 obs. of 6 variables:
 $ X                : int  619 664 441 160 358 335 47 164 736 436 ...
 $ Months.since.Last.Donation : int  2 0 1 2 1 4 2 1 5 0 ...
 $ Number.of.Donations      : int  50 13 16 20 24 4 7 12 46 3 ...
 $ Total.volume.Donated..c.c.: int 12500 3250 4000 5000 6000 1000 1750 3000 11500
 750 ...
 $ Months.since.First.Donation: int  98 28 35 45 77 4 14 35 98 4 ...
 $ Made.Donation.in.March.2007: int  1 1 1 0 0 1 0 1 0 ...
```

## DT: Trained Classifier Result

- Split data into train and test data,  $p = 0.8$
- Cross validation into 10 folds
- $C_p = 0.04337671$



## DT: Confusion Matrix

- Accuracy = 0.7193
- The error rate = 0.2807
- Kappa = 0.1817, kappa coefficient is slight

Confusion Matrix and Statistics

	Reference	
Prediction	0	1
0	73	18
1	14	9

Accuracy : 0.7193