Decision Tree

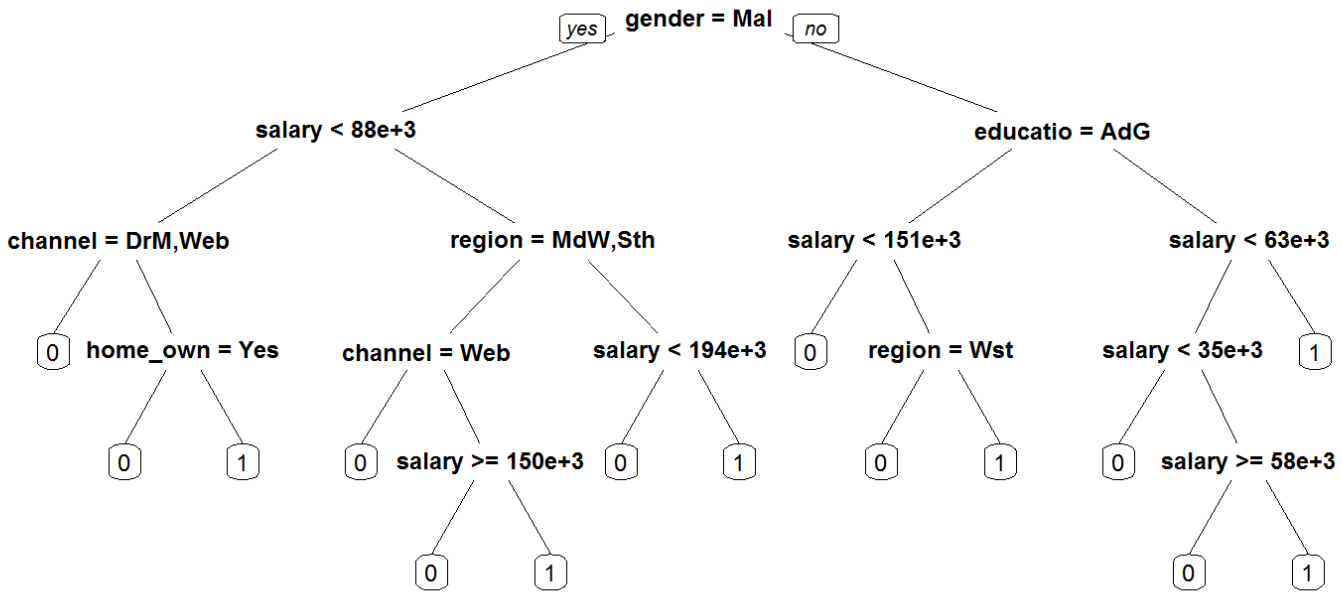
Decision trees are commonly used in operations research, specifically in **decision analysis**, to help identify a strategy most likely to reach a goal, but are also a popular tool in **machine learning**.

Split data into training and testing chunks

```
[1] 975   8
```

```
[1] 325   8
```

Create and plot the decision tree



Generate predictions on the training set and check the accuracy with a confusion matrix

```
Confusion Matrix and Statistics

          Reference
Prediction   0   1
         0 551 180
         1  38 206

               Accuracy : 0.7764
                 95% CI : (0.7489, 0.8022)
    No Information Rate : 0.6041
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.5009
 Mcnemar's Test P-Value : < 2.2e-16

            Sensitivity : 0.9355
            Specificity : 0.5337
         Pos Pred Value : 0.7538
         Neg Pred Value : 0.8443
             Prevalence : 0.6041
         Detection Rate : 0.5651
   Detection Prevalence : 0.7497
      Balanced Accuracy : 0.7346

       'Positive' Class : 0
```

Validate the accuracy with the testing dataset

```
Confusion Matrix and Statistics

          Reference
Prediction   0   1
         0 189  69
         1  22  45

               Accuracy : 0.72
                 95% CI : (0.6678, 0.7682)
    No Information Rate : 0.6492
    P-Value [Acc > NIR] : 0.003966

                  Kappa : 0.3209
 Mcnemar's Test P-Value : 1.42e-06

            Sensitivity : 0.8957
            Specificity : 0.3947
         Pos Pred Value : 0.7326
         Neg Pred Value : 0.6716
             Prevalence : 0.6492
         Detection Rate : 0.5815
   Detection Prevalence : 0.7938
      Balanced Accuracy : 0.6452

       'Positive' Class : 0
```

Create cross validation, trainControl object and view model

```
CART

975 samples
  7 predictor
  2 classes: '0', '1'

No pre-processing
Resampling: Cross-Validated (10 fold, repeated 2 times)
Summary of sample sizes: 878, 877, 879, 877, 878, 877, ...
Resampling results

  Accuracy   Kappa      Accuracy SD  Kappa SD
  0.7215037  0.3989268  0.04869316   0.1054461

Tuning parameter 'cp' was held constant at a value of 0.001
```

Fancy decision tree