Class 8: Tree-based Models and Their Application to Targeted Marketing

Dr Wei Miao

UCL School of Management

October 25, 2023

Section 1

Decision Tree

Introduction to Decision Tree

- A **decision tree** is a flowchart-like tree structure.
- Used in classification and regression.
- Consists of nodes representing decisions and leaves representing outcomes.

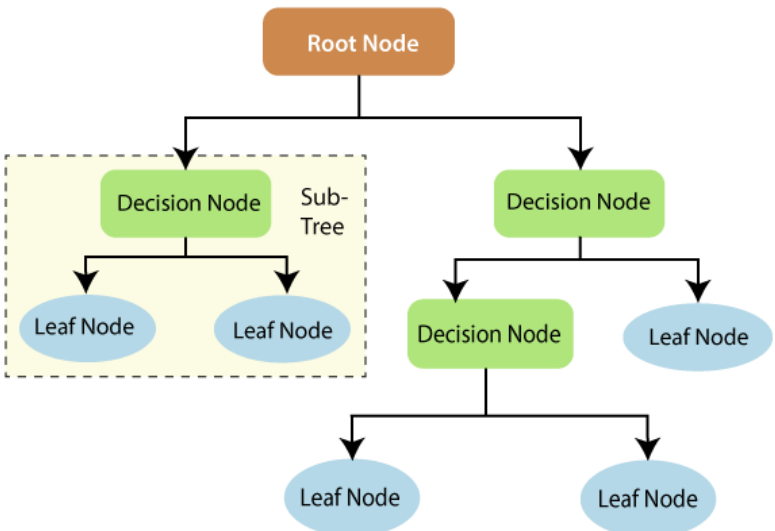


Figure 1: How a decision tree looks like. Source: Medium.

Example: Predict Customer Response to Marketing Offers

- Tesco made marketing offers to customers in the data, and the variable `Response` represents whether or not customers responded to our offer.
- **Business objective**: From `data_full`, we want to train a decision tree model to predict the outcome variable `Response` based on `Recency` and `totalspending` (for simplicity).
- **Data collection and cleaning:**

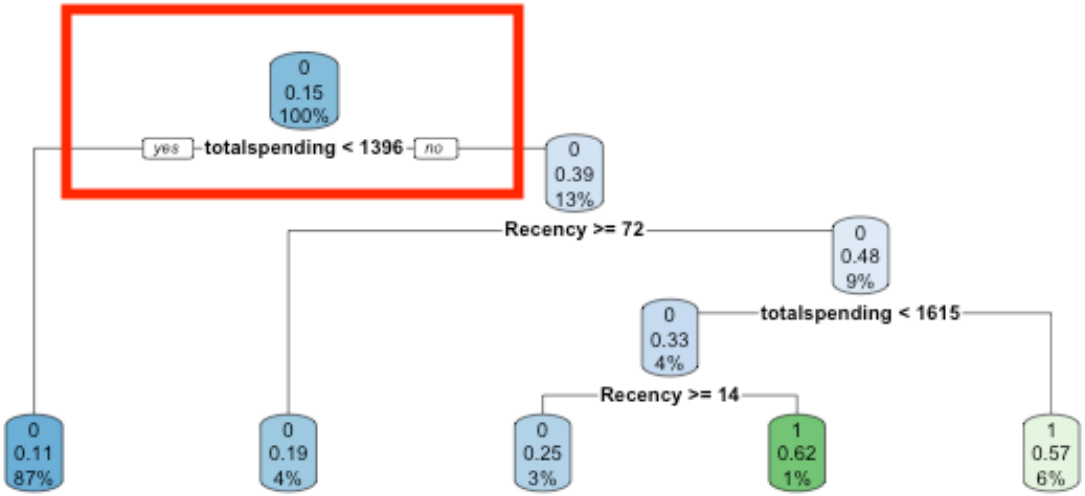
```
pacman::p_load(dplyr,modelsummary)
data_demo <- read.csv(file = "https://www.dropbox.com/s/a0v38lpydls2emy
                      header = T)
data_purchase <- read.csv(file = "https://www.dropbox.com/s/de435r8zdxy

data_full <- data_purchase %>%
  left_join(data_demo, by = c('ID' = 'ID')) %>%
  mutate(totalspending = MntWines + MntFruits +
           MntMeatProducts + MntFishProducts +
           MntSweetProducts + MntGoldProds)
```

Implementation of Decision Tree in R

- Package `rpart` provides efficient implementation of decision trees in R
- Package `rpart.plot` provides nice visualizations of decision trees

```
# Load the necessary packages
pacman::p_load(rpart,rpart.plot)

# Below example shows how to train a decision tree
tree1 <- rpart(
  formula = Response ~ Recency + totalspending,
  data    = data_full,
  method  = "class" # classification task; or 'anova' for regression
  )

# visualize the tree
rpart.plot(tree1)
```

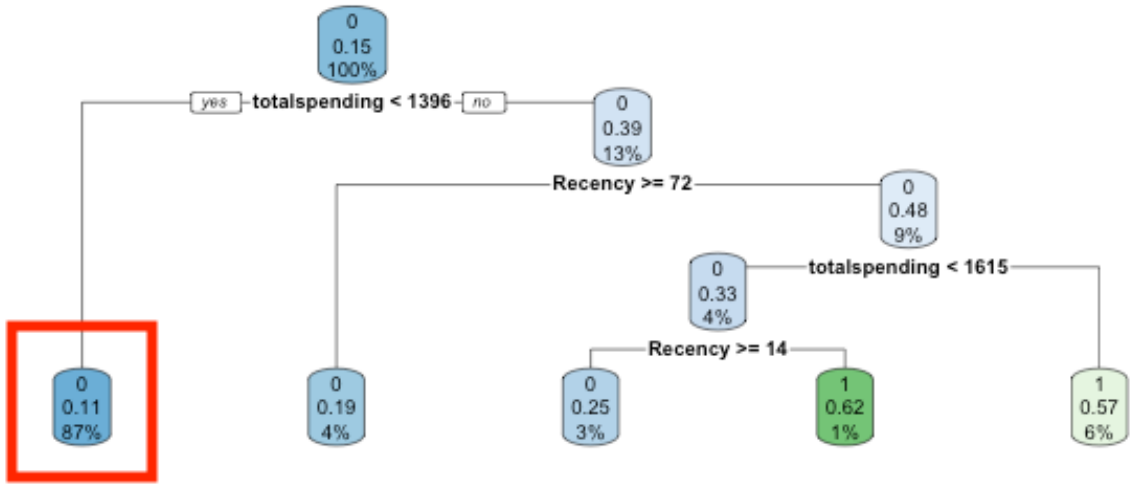
How Decision Tree Works: Step 1



1. Decision tree (DT) will try to split customers into 2 groups based on each unique value of each variable, and see which split can lead to customers being most differentiated in terms of `Response`.
   - After this step, DT finds that total spending is the best variable and 1396 is the best cutoff.
   - DT therefore splits customers into 2 groups based on 1396.[1]

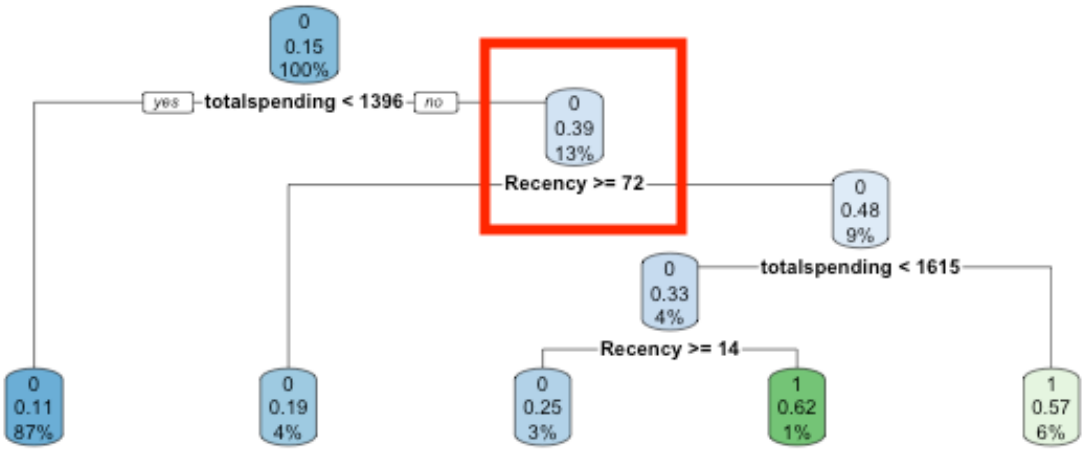
[1] In each node, the 3 numbers are (1) predicted outcome (2) predicted probability of outcome being 1, and (3) percentage of customers in the node

How Decision Tree Works: Step 2



2. For customers in the left branch (`totalspending` < 1396), DT will continue to split based on each unique value of each variable, and see which split can result in the customers to be most different in terms of Response.
    - However, DT couldn't find a cutoff that sufficiently differentiate customers, so DT stops in the left branch.

How Decision Tree Works: Step 3 ...



3. For customers in the right branch (`totalspending` $>=$ 1396), DT will continue to split based on each unique value of each variable, and see which split can result in the customers to be most different in terms of Response.
   - After this step, DT finds `Recency` is the best variable and 72 is the best cutoff. DT further splits customers into 2 groups.
4. This process continues until DT determines that there is no need to further split customers.

Advantages of Decision Trees

- They are very interpretable.
- Making predictions is fast.
- It's easy to understand what variables are important in making the prediction. The internal nodes (splits) are those variables that most largely reduce the SSE (criteria for split).

Section 2

Random Forest

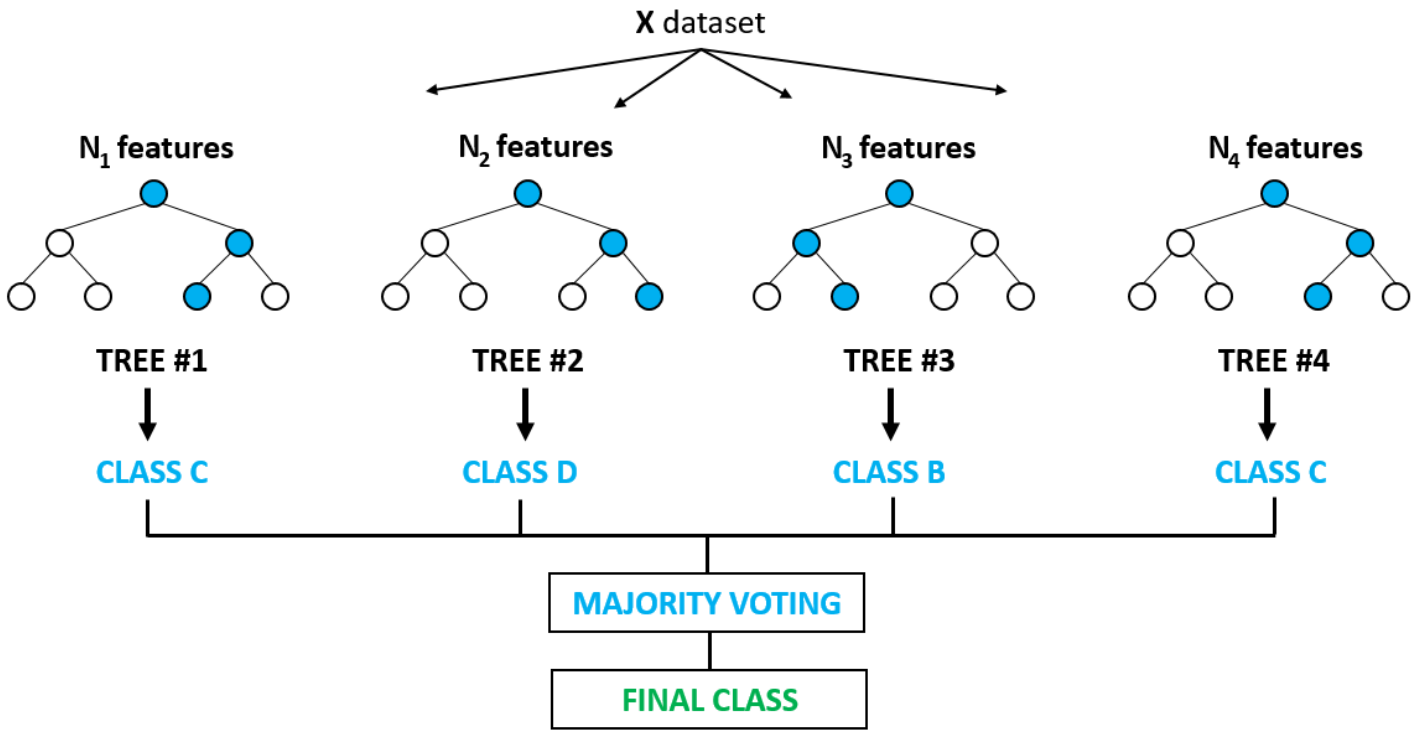
Disadvantages of Decision Trees

- Single regression trees tend to have high variance (overfitting), resulting in unstable predictions.
- Due to the high variance, single regression trees tend to have poor predictive accuracy.

Random Forest

- To overcome the overfitting tendency of a single decision tree, random forest has been developed by (Breiman 2001).
  - Instead of using all customers, each tree is grown to a random subsample of customers
  - Instead of using all features for splitting, each treen is grown to a random subset of features

Visualization of Random Forest



For a new customer,

- Each tree gives a prediction of the outcome
- Random forest takes the average of all trees' predictions as the final prediction

Implementation of Random Forest in R

- Package `ranger` provides implementation of random forest in R.
- `ranger()` is the function in the package to train a random forest; refer to its help function for more details.
- The following code shows how to train a random forest consisting of 500 decision trees, where the outcome variable is mpg, and the predictors are 5 car attribute variables.

```
pacman::p_load(ranger)
randomforest1 <- ranger(
    formula   = Response ~ totalspending + Recency,
    data      = data_full, # dataset to train the model
    num.trees = 500, # 500 decision trees
    seed = 888, # make sure of replication
    probability  = TRUE
  )
```

Make Predictions from Random Forest

- After we train the predictive model, we can use `predict()` function to make predictions
  - The 1st argument is the trained model object
  - The 2nd argument is the dataset to make predictions on

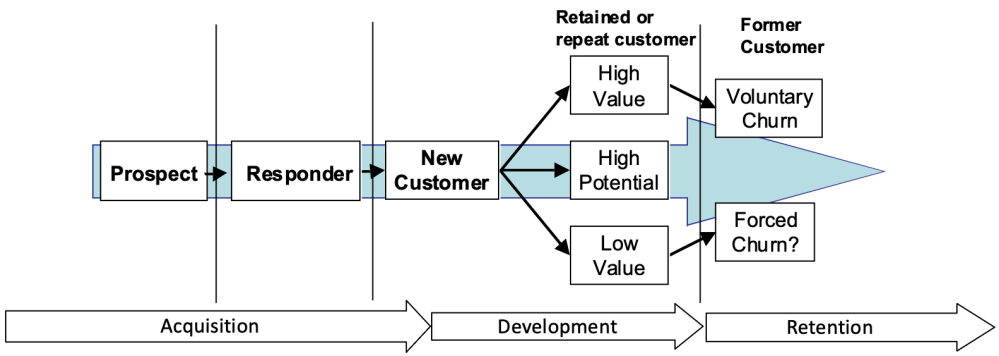
```
# Make predictions on the mtcars
prediction_rf <- predict(randomforest1,
                         data = data_full)

# Because prediction_rf is a list object
# Need to use $ to extract the predicted value as a numeric vector
prediction_rf$predictions
```

Section 3

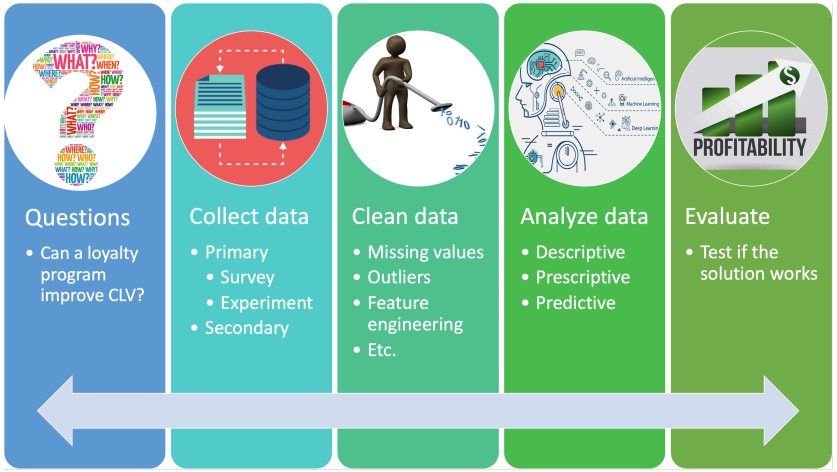
Improve Marketing Efficiency Using Supervised Learning

Customer Life Cycle

- Acquisition (Tesco Case Study)
  - Use predictive analytics to target responsive customers to reduce marketing costs
- Development
  - Use predictive analytics to recommend products to customers (personalized recommendation system); for each customer, promote the item with the highest purchase probability
- Retention
  - Use predictive analytics to find valuable customers who are likely to churn and conduct targeted churn management



Workflow



1. Define a business objective: target responsive customers in acquisition stage to reduce customer acquisition costs
2. Collect data
3. Clean and prepare data
4. Analyze data using predictive analytics
5. Conduct break-even analyses to show the profitability of the proposed marketing compaign

After-Class Reading

- (optional) Varian, Hal R. "Big data: New tricks for econometrics." Journal of Economic Perspectives 28, no. 2 (2014): 3-28
- (next week) Predictive Analytics for Tesco
- (recommended) Decision tree in R
- (recommended) Random forest in R