

Boston <- read.csv("D:/Rlearn/big data/Boston.csv")

library(rpart)

split <- sample(nrow(`Boston`), floor(0.75\*nrow(`Boston`)))

traindata <- `Boston`[split,]

fit <- rpart(crim~., traindata, method="anova")

library(rpart.plot)

rpart.plot(fit, branch=1, branch.type=2, type=1, extra=101,

shadow.col="gray", box.col="green",

border.col="blue", split.col="red",

split.cex=1.2, main="BostonTM")

Pro:

1.Pretty intelligible and readable, it is good for presentation or business need;

2. It can deal with big data and also can output a effective results;

Con:

1. If the dataset included many variables, it might make mistake;

2. If it is time series, before we process tree method, we may need to do a lot of pre-work, such as transfer some data.