Homework 1

Riley Payung

2/24/2020

PART A

Read homework 1 data

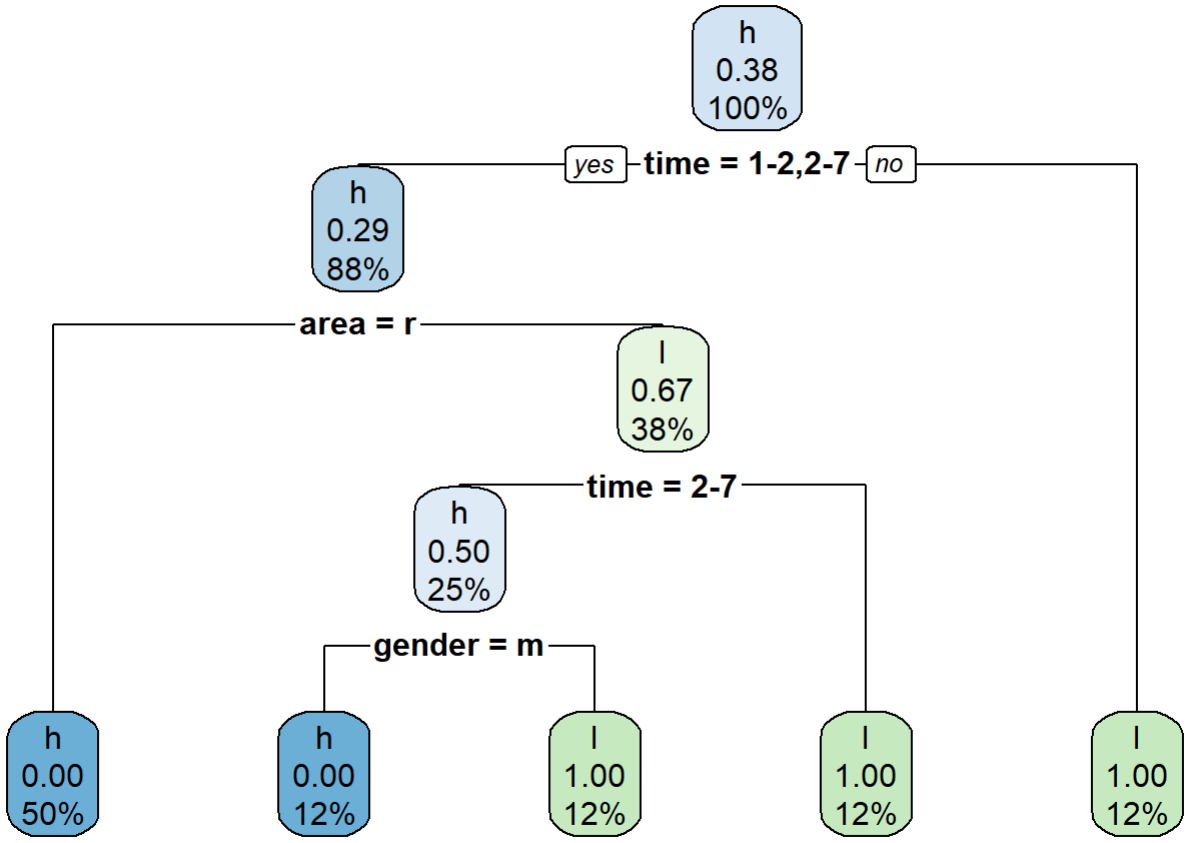
```
data <- read.csv(file="hw1data.csv")
```

Create Model

```
model <- rpart(risk~time+gender+area,data=data,method='class',control=rpart.control(minsplit=2))
model
```

```
## n= 8
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##  1) root 8 3 h (0.6250000 0.3750000)
##    2) time=1-2,2-7 7 2 h (0.7142857 0.2857143)
##      4) area=r 4 0 h (1.0000000 0.0000000) *
##      5) area=u 3 1 l (0.3333333 0.6666667)
##       10) time=2-7 2 1 h (0.5000000 0.5000000)
##         20) gender=m 1 0 h (1.0000000 0.0000000) *
##         21) gender=f 1 0 l (0.0000000 1.0000000) *
##       11) time=1-2 1 0 l (0.0000000 1.0000000) *
##    3) time=>7 1 0 l (0.0000000 1.0000000) *
```

```
rpart.plot(model)
```



PART B

Load Test Data:

```
test <- read.csv(file="predict.csv")
test <- test %>%
  rename(X = "ï..X")
head(test)
```

	X <int>	person <int>	time <fctr>	gender <fctr>	area <fctr>	risk <lgl>
1	1	1	1-2	f	r	*NA*
2	2	2	2-7	m	u	*NA*
3	3	3	1-2	f	u	*NA*

3 rows

Predict

```
predicted <- predict(model, test, type='class')
test$risk = predicted
head(test)
```

	X	person	time	gender	area	risk
	<int>	<int>	<fctr>	<fctr>	<fctr>	<fctr>
1	1	1	1-2	f	r	h
2	2	2	2-7	m	u	h
3	3	3	1-2	f	u	l

3 rows