

Project STAT 652

Kotomi Oda

[illegible]

```
Warning: There was 1 warning in `mutate()``.
i In argument: `emp_length = as.numeric(gsub("[0-9]+.*$", "\\1",
  emp_length))`.
Caused by warning:
! NAs introduced by coercion
```

```
data$emp_length <- data$emp_length %>% replace(is.na(.), 0)
data$loan_status <- factor(data$loan_status)
```

```
set.seed(123)
sample <- sample.split(data$loan_status, SplitRatio = 0.75)
train_data <- subset(data, sample == T)
test_data <- subset(data, sample == F)
```

C50

```
doParallel::registerDoParallel(cores = 8)
tree_mod <- C5.0(train_data[, -18], train_data$loan_status)
```

```
predictions <- predict(tree_mod, test_data)
table(test_data$loan_status, predictions)
```

```
predictions
  0    1
0 1879 147
1   223 180
```

```
mean(predictions == test$loan_status)
```

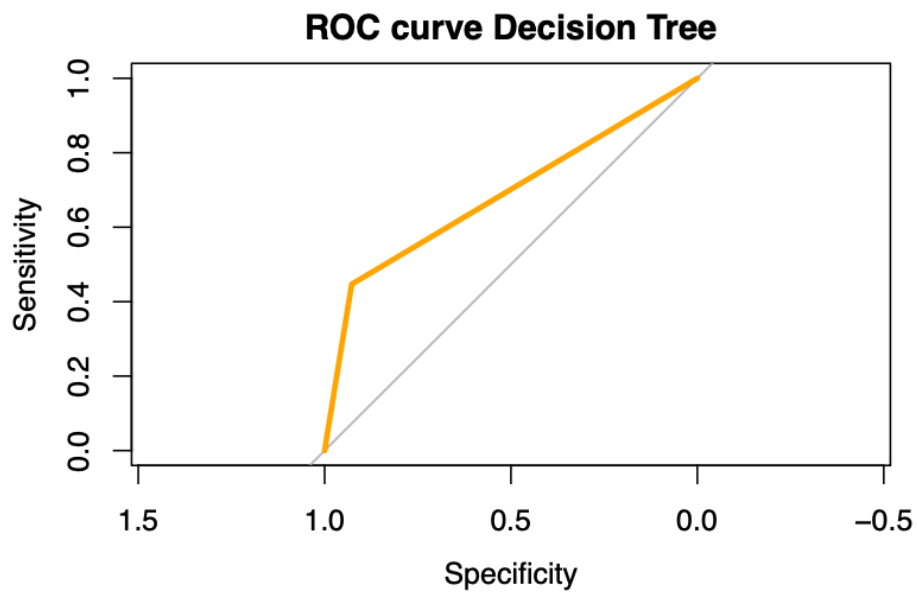
Error in mean(predictions == test\$loan_status): object 'test' not found

```
roc1 <- roc(as.numeric(test_data$loan_status), as.numeric(predictions))
```

Setting levels: control = 1, case = 2

Setting direction: controls < cases

```
plot(roc1, col="Orange", lwd=3, main="ROC curve Decision Tree")
```



Random Forest

```
rf_model <- randomForest(loan_status ~ ., data = train_data, na.action=na.roughfix, import
rf_model$confusion
```

```
      0      1 class.error
0 5889 188  0.03093632
1  767 443  0.63388430
```

```
rf_pred <- predict(rf_model, na.roughfix(test_data))
table(test_data$loan_status, rf_pred)
```

```
rf_pred
      0      1
0 1961    65
1  254   149
```

```
mean(rf_pred == test_data$loan_status)
```

```
[1] 0.8686702
```