

HW3

Sijia Yue

11/5/2019

```
df_train = read_table("synth.tr.txt") %>%  
  mutate(yc = as.factor(yc)) %>%  
  as.data.frame()
```

```
## Parsed with column specification:  
## cols(  
##   xs = col_double(),  
##   ys = col_double(),  
##   yc = col_double()  
## )
```

```
df_test = read_table("synth.te.txt") %>%  
  mutate(yc = as.factor(yc)) %>%  
  as.data.frame()
```

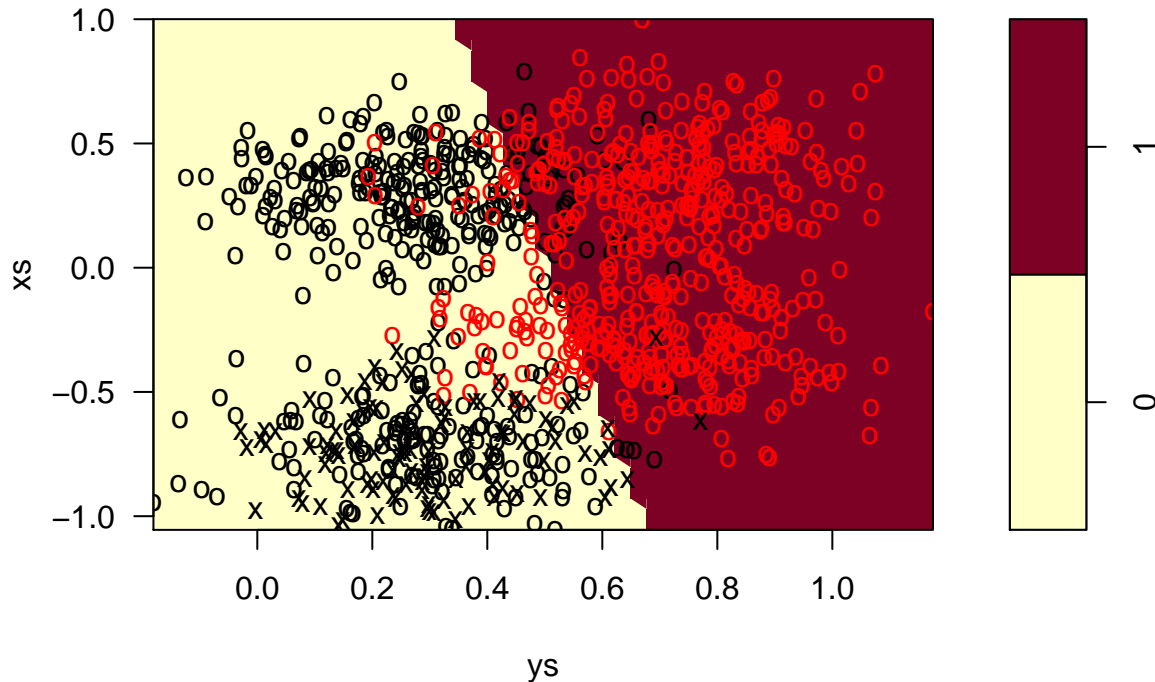
```
## Parsed with column specification:  
## cols(  
##   xs = col_double(),  
##   ys = col_double(),  
##   yc = col_double()  
## )
```

Construct a linear support vector classifier

```
set.seed(313)  
# tune the lowest cost  
tune_cost = tune(svm, yc ~., data = df_train, kernel = "linear", range = list(cost = c(1:10)))  
tune_cost  
  
##  
## Parameter tuning of 'svm':  
##  
## - sampling method: 10-fold cross validation  
##  
## - best parameters:  
##   cost  
##     1  
##  
## - best performance: 0.148  
  
# train the model  
svm_Linear <- svm(yc~., data = df_train, kernel = "linear", scale = FALSE, cost = 2)  
# predict on test set  
test_pred <- predict(svm_Linear, df_test)  
# get the test error  
test_error = mean(test_pred != df_test$yc)  
# get the standard error of the test error  
se = sd(test_pred != df_test$yc)/sqrt(nrow(df_test))
```

```
# visualize the prediction
plot(svm_Linear, df_test)
```

SVM classification plot

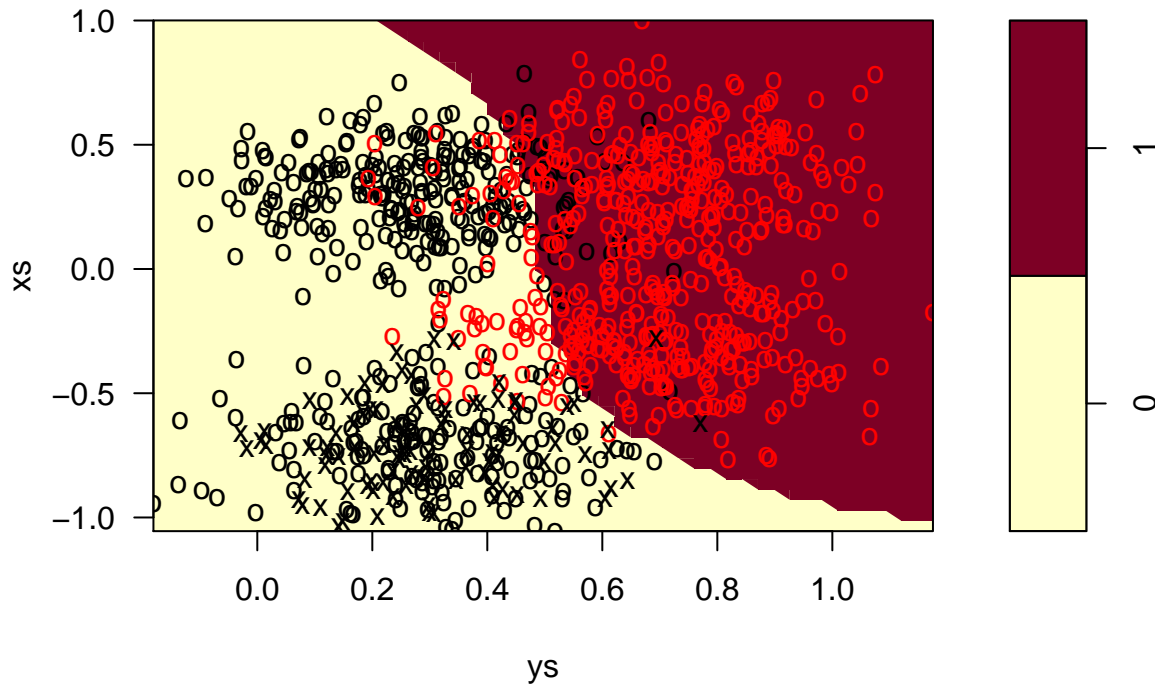


The test error is 0.103 and the standard error is 0.0096.

Construct a support vector classifier with Radial kernel.

```
set.seed(333)
# tune the lowest cost
tune_cost = tune(svm, yc ~., data = df_train, kernel = "radial", range = list(cost = c(1:10)))
# train the model
svm_radial <- svm(yc~., data = df_train, kernel = "radial", scale = FALSE, cost = 8)
# predict on test set
test_pred <- predict(svm_radial, df_test)
# get the test error
test_error = mean(test_pred != df_test$yc)
# get the standard error of the test error
se = sd(test_pred != df_test$yc)/sqrt(nrow(df_test))
# visualize the prediction
plot(svm_radial, df_test)
```

SVM classification plot

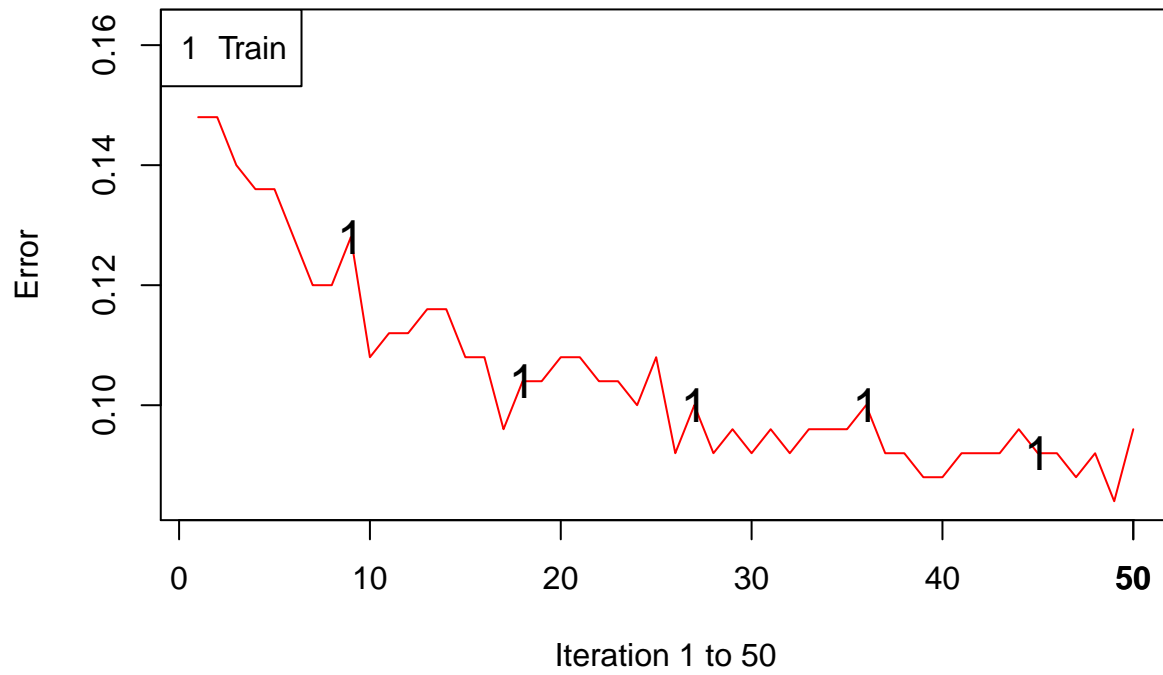


The test error is 0.096 and the standard error is 0.0093.

Construct a classifier using AdaBoost algorithm (with 50 boosting iterations) with decision stumps as weak learners.

```
set.seed(333)
# train the model
ada_fit <- ada(yc ~., data = df_train, iter = 50)
# predict on test set
test_pred <- predict(ada_fit, df_test)
# get the test error
test_error = mean(test_pred != df_test$yc)
# get the standard error of the test error
se = sd(test_pred != df_test$yc)/sqrt(nrow(df_test))
# visualize the prediction
plot(ada_fit)
```

Training Error



The test error is 0.102 and the standard error is 0.0096.