Quant Strategy

读文件，确定自变量和因变量

涨跌幅低于-1%记为 0;涨跌幅高于-1%，低于 1%，记为 1;涨跌幅高于 1%记为 2

```
dataTrain <- read.csv("./data/hsp.csv")
dataTrain <- na.omit(dataTrain)
x <- dataTrain[,-16]
y <- dataTrain[,16]
```

计算 SVM 在 2 种分类机，4 种核函数下模型的错误次数

```
type=c("C-classification","nu-classification")
kernel=c("linear","polynomial","radial","sigmoid")
accuracy=matrix(0,2,4)
for (i in 1:2)
{
  for ( j in 1:4)
  {
    model <- svm(x,y,type=type[i],kernel = kernel[j])
    pred_temp=predict(model,x)
    accuracy[i,j]=sum(pred_temp!=y)
  }
}
dimnames(accuracy)=list(type,kernel)
accuracy
```

```
##                   linear polynomial radial sigmoid
## C-classification    1014       1023    967    1509
## nu-classification    987       1018    926    1346
```

由以上结果可知，使用 SVM 进行实验，type="nu-classification",kernel = "radial"的模型最优。

实验 1 用训练数据的前 2666 条作为训练集，后 200 条作为测试集，看看预测结果

```
model1 <- svm(x[1:2666,],y[1:2666],type="nu-classification",kernel = "radial")
pred1 <- predict(model1,x[2667:2866,])

table(pred1,y[2667:2866])
```

```
##
## pred1   0   1   2
##     0   1   0   0
##     1   7 179  12
##     2   0   0   1
```

实验 2 使用全部训练样本展示预测结果，并与真实情况的比较。

```
model_fitted <- svm(x,y,type="nu-classification",kernel = "radial")
summary(model_fitted)
```

```
##
## Call:
## svm.default(x = x, y = y, type = "nu-classification", kernel = "radi
al")
##
##
## Parameters:
##    SVM-Type:  nu-classification
##  SVM-Kernel:  radial
##       gamma:  0.06666667
##          nu:  0.5
##
## Number of Support Vectors:  2138
##
##  ( 960 551 627 )
##
##
## Number of Classes:  3
##
## Levels:
##  0 1 2
```

```
pred <- predict(model_fitted,x)
```

```
table(pred,y)
```

```
##     y
## pred    0    1    2
##    0  139   39   33
##    1  399 1465  319
##    2   28  108  336
```