

# Final Project

Jarod Wright

4/26/2021

```
library(ggplot2)
library(caret)
```

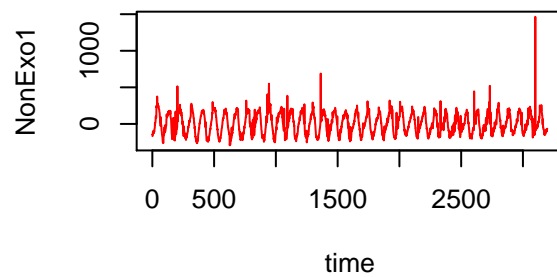
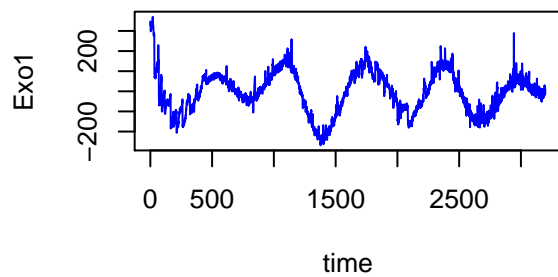
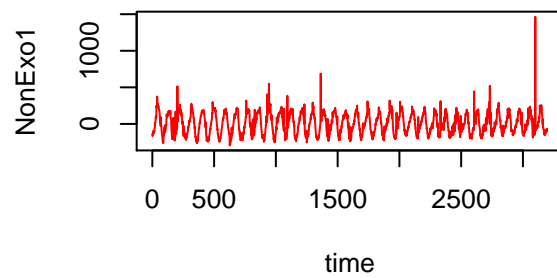
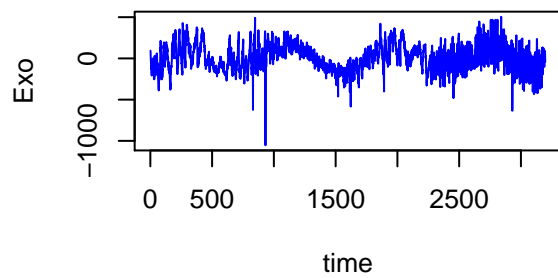
```
## Loading required package: lattice
```

```
t1<- file.choose()
t2<-file.choose()
exo.Train = read.csv(t1) # preprocessed exoplanet data
exo.Test = read.csv(t2)

exo.Train$LABEL <- as.factor(exo.Train$LABEL)
exo.Test$LABEL <- as.factor(exo.Test$LABEL)

exo.Train.true=subset(exo.Train, LABEL==2)
exo.Train.false=subset(exo.Train, LABEL==1)
#-----
Exo=c(exo.Train.true[1,2:3198])
Exo1=c(exo.Train.true[4,2:3198])
NonExo=c(exo.Train.false[1,2:3198]) # arbitrary star time series
NonExo1=c(exo.Train.false[4,2:3198])

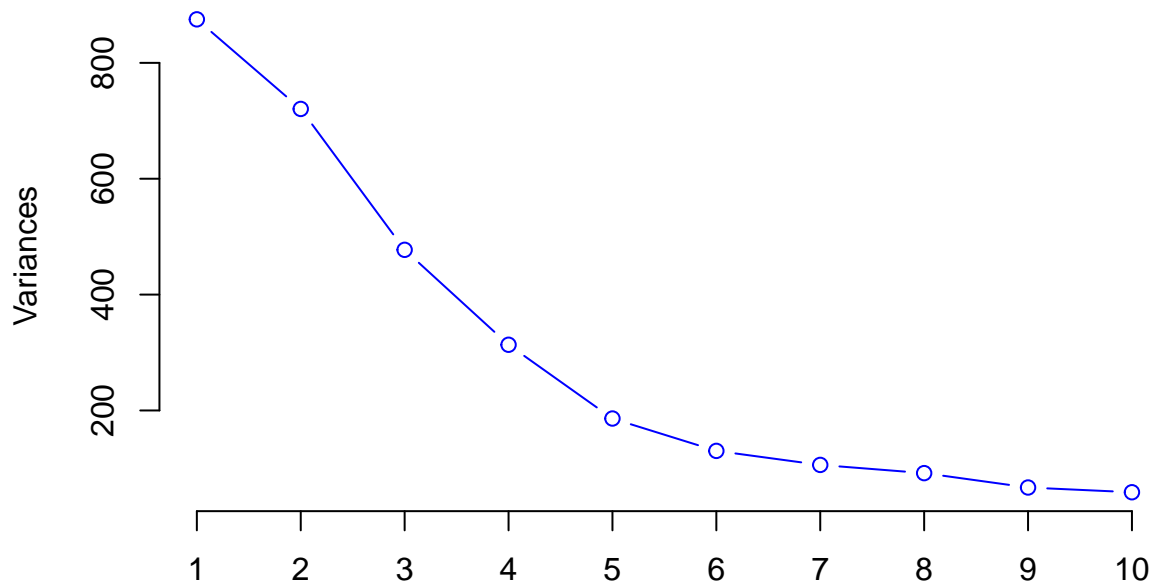
time <- c(1:3197)
par(mfrow=c(2,2))
plot(time,Exo,type = "l",col="blue")
plot(time,NonExo1,type="l",col="red")
plot(time,Exo1,type = "l",col="blue")
plot(time,NonExo1,type="l",col="red")
```



```
#plots of exoplanet vs. non exoplanet
#-----
#normalize
#PCA
#logistic regression
par(mfrow=c(1,1))
exo.PCA = prcomp(exo.Train[,c(-1)],scale=TRUE, center = TRUE)

plot(exo.PCA,type="l", col="blue")#spree
```

## exo.PCA



```
NewData <- data.frame(LABEL = exo.Train[, "LABEL"], exo.PCA$x[, 1:6])
head(NewData)
```

```
##   LABEL      PC1      PC2      PC3      PC4      PC5      PC6
## 1     2 -0.8775396 -0.207663548 -0.37778196 -0.03944037 0.04909936 0.3034687
## 2     2 -0.8909994 -0.157259659 -0.41408001 -0.10507538 0.08607877 0.2307515
## 3     2 -0.9823419  0.078045508 -0.56160362 -0.37241677 0.13519529 0.2048472
## 4     2 -0.9194254 -0.135489501 -0.44157165 -0.18354111 0.10914435 0.2195717
## 5     2 -1.3856596 -0.309331578 -0.06462467  0.47519107 0.03329498 0.4534340
## 6     2 -0.7588780  0.004431154 -0.25531724  0.21644911 0.14697202 0.2997393
```

```
logRegPCA <- glm(LABEL~., data=NewData, family = binomial)
```

```
test.p <- predict(exo.PCA, exo.Test[, 2:3198])
```

```
pred <- predict(logRegPCA, data.frame(test.p[, 1:6]), type = "response")
```

```
predLabel <- factor(ifelse(pred >= 0.5, "2", "1"))
table(exo.Test$LABEL, predLabel)
```

```
##   predLabel
##      1
## 1 565
## 2   5
```

```

##started CNN attempt
##library(kiras)
##library(tensorflow)

##t1<- file.choose()
##t2<-file.choose()
##exo.Train = read.csv(t1)
##exo.Test = read.csv(t2)
##exo.Train$LABEL <- ifelse(exo.Train$LABEL==2, 1, 0)

##normalize <- function(x) { return ((x - min(x)) / (max(x) - min(x)))}
##maxmindf.train <- as.data.frame(lapply(exo.Train, normalize))
##maxmindf.test <- as.data.frame(lapply(exo.Test, normalize))

##nn <- neuralnet(LABEL~.,data=exo.Train,hidden = )

```