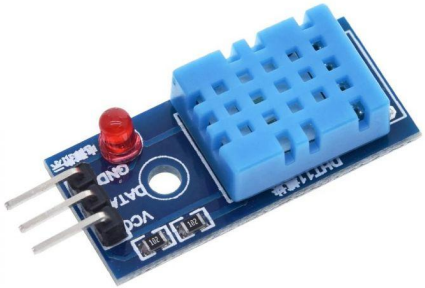
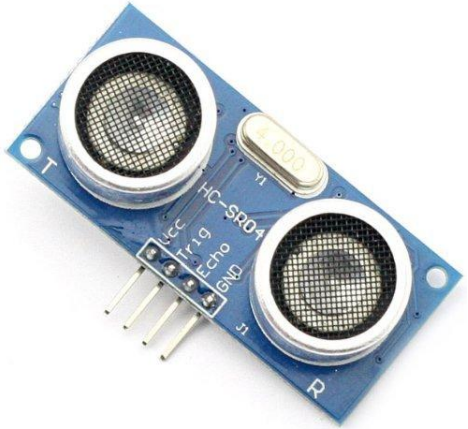
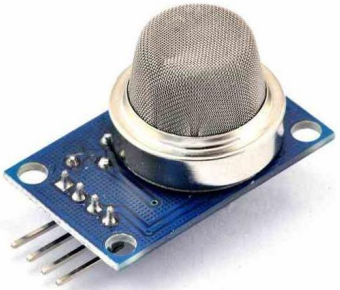
Assignment 3
Supervised Methods in Machine Learning
3rd report guideline: “Supervised learning final project”

Santiago Meza
Alejandro Pachon
Alexander Morgan

Sensores



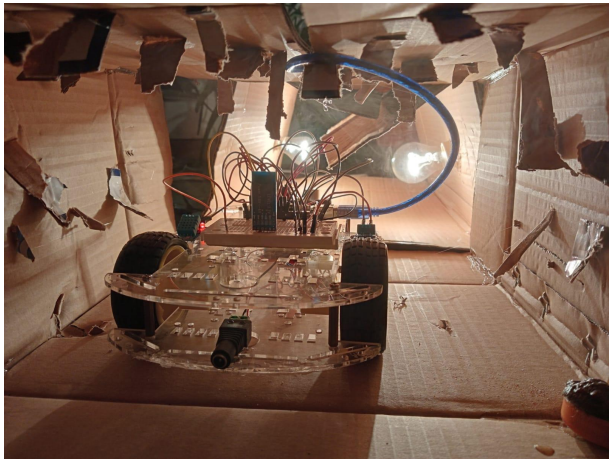




Ambientes

Desierto

Jungla



Cueva Volcánica

Modelos de Predicción

```
##-----------------------Arbol de decision-----------------
predictors <- colnames(DataSet)[-7]
data.samples <- sample(1:nrow(DataSet),
                       nrow(DataSet) *0.7, replace = FALSE)
sample(DataSet)
training.data <- DataSet[data.samples,c(predictors,"AMBIENTE") ]
test.data <- DataSet[-data.samples,c(predictors,"AMBIENTE") ]

fit.rf <- randomForest(AMBIENTE ~ PPM + TEMP + HUMEDAD + DIS_UP, data = training.data)


prediction.rf <- predict(fit.rf, test.data)
output <- data.frame(test.data$Mileage, prediction.rf)
RMSE = sqrt(sum((output$test.data.Mileage - output$prediction.rf)^2)/
                nrow(output))

RMSE
RMSEmodelo2 = data.frame(prediccion = prediction.rf
                         ,ahora = test.data$Mileage
                         ,RSE = sqrt((prediction.rf-test.data$mileage)^2)
                         )
```

Modelos de Predicción

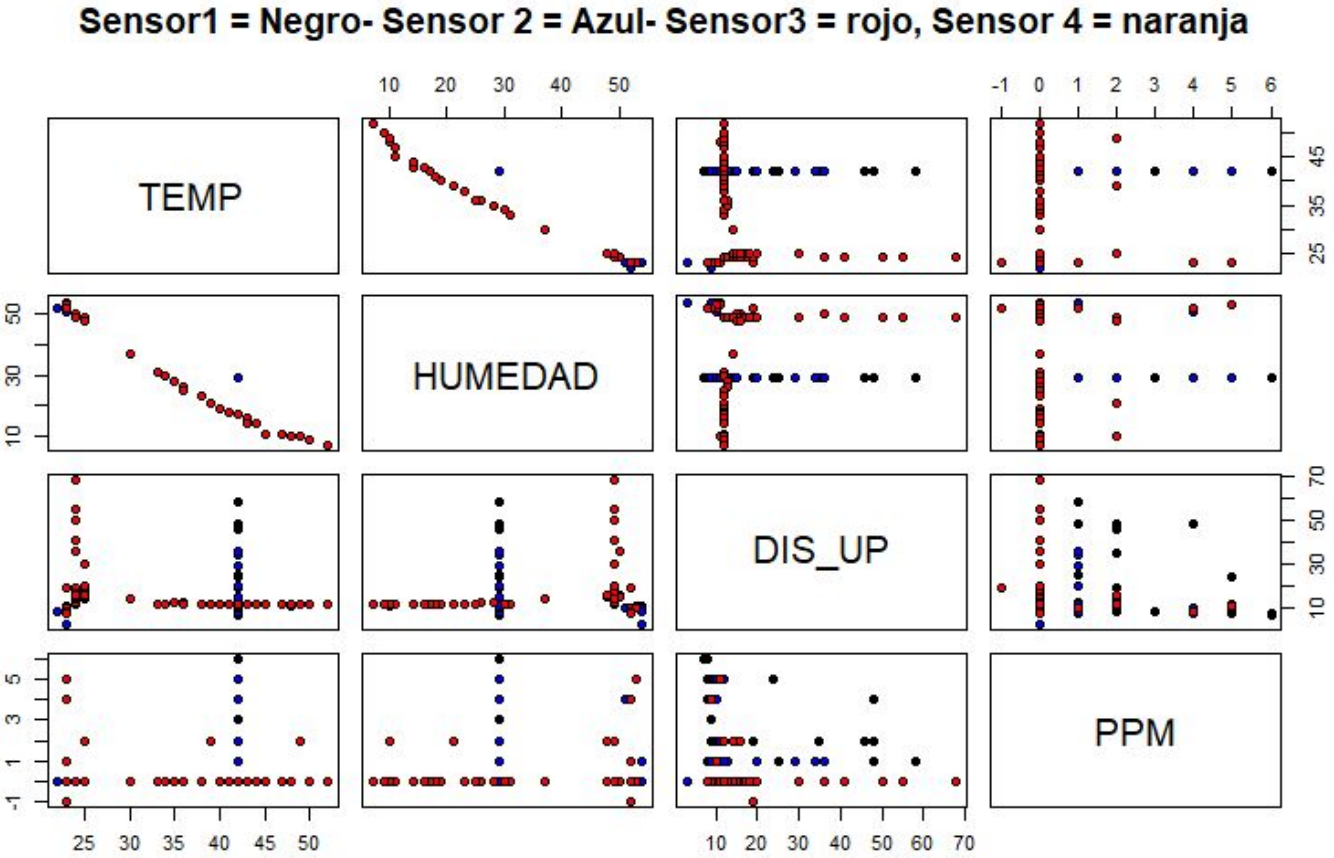
```
#-------------------------------- Modelo Lineal--------------------------
modelo_multilineal <- lm(DataSet$AMBIENTE ~ DataSet$TEMP + DataSet$HUMEDAD + DataSet$DIS_UP +
                           DataSet$PPM, data = training.data)
summary(modelo_multilineal)
prediccionlm <- predict(modelo_multilineal,test.data)
prediccionlm
RMS1 = data.frame(prediccion = prediccionlm
                          ,actual =DataSet
                          ,RSE = sqrt((prediccionlm-DataSet$AMBIENTE)^2)
)
View(RMS1)
```

Modelos de Predicción

```
##-------------------Modelo GLM ---------------------------------------------
Modelo2 <- glm(AMBIENTE ~ TEMP + HUMEDAD + DIS_UP + PPM, data = training.data)
summary(Modelo2)
modeloPredictivo2 <- predict(Modelo2,test.data)
RMS2 = data.frame(prediccion = modeloPredictivo2
                  ,actual =test.data$AMBIENTE
                  ,RSE = sqrt((modeloPredictivo2-test.data$AMBIENTE)^2)
)
View(RMS2)
```

Modelos de Predicción

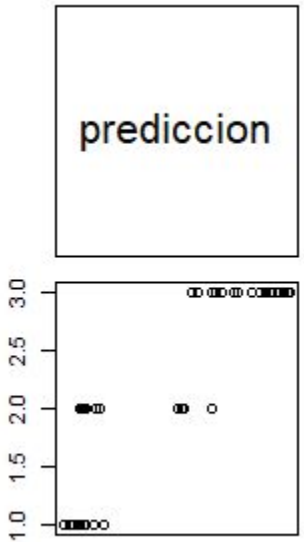
```
##--------------------- Modelo KNN ----------------------------------------------
DataSet$AMBIENTE <- as.factor(DataSet$AMBIENTE)
data.samples<- sample(1:nrow(DataSet)
                      , nrow(DataSet)*0.7
                      , replace = F)

predictors <- c("TEMP","HUMEDAD","DIS_UP","PPM")

training.data <-
  DataSet[data.samples,c(predictors,"AMBIENTE"),drop=F]
test.data <-
  DataSet[-data.samples,c(predictors,"AMBIENTE"),drop=F]

ctrl <- trainControl(method = "cv",p=7)
knnFit <- train(AMBIENTE ~ TEMP + HUMEDAD + DIS_UP + PPM
                , data = training.data
                , method = "knn",trControl=ctrl
                , preProcess = c("center","scale")
                , tuneLength = 20)
knnFit
plot(knnFit)

knnPredict <- predict(knnFit,newdata = test.data)
caret::confusionMatrix(knnPredict,test.data$AMBIENTE)
```


Sensor1 = Negro- Sensor 2 = Azul- Sensor3 = rojo, Sensor 4 = naranja
TEMP
HUMEDAD
DIS_UP
PPM

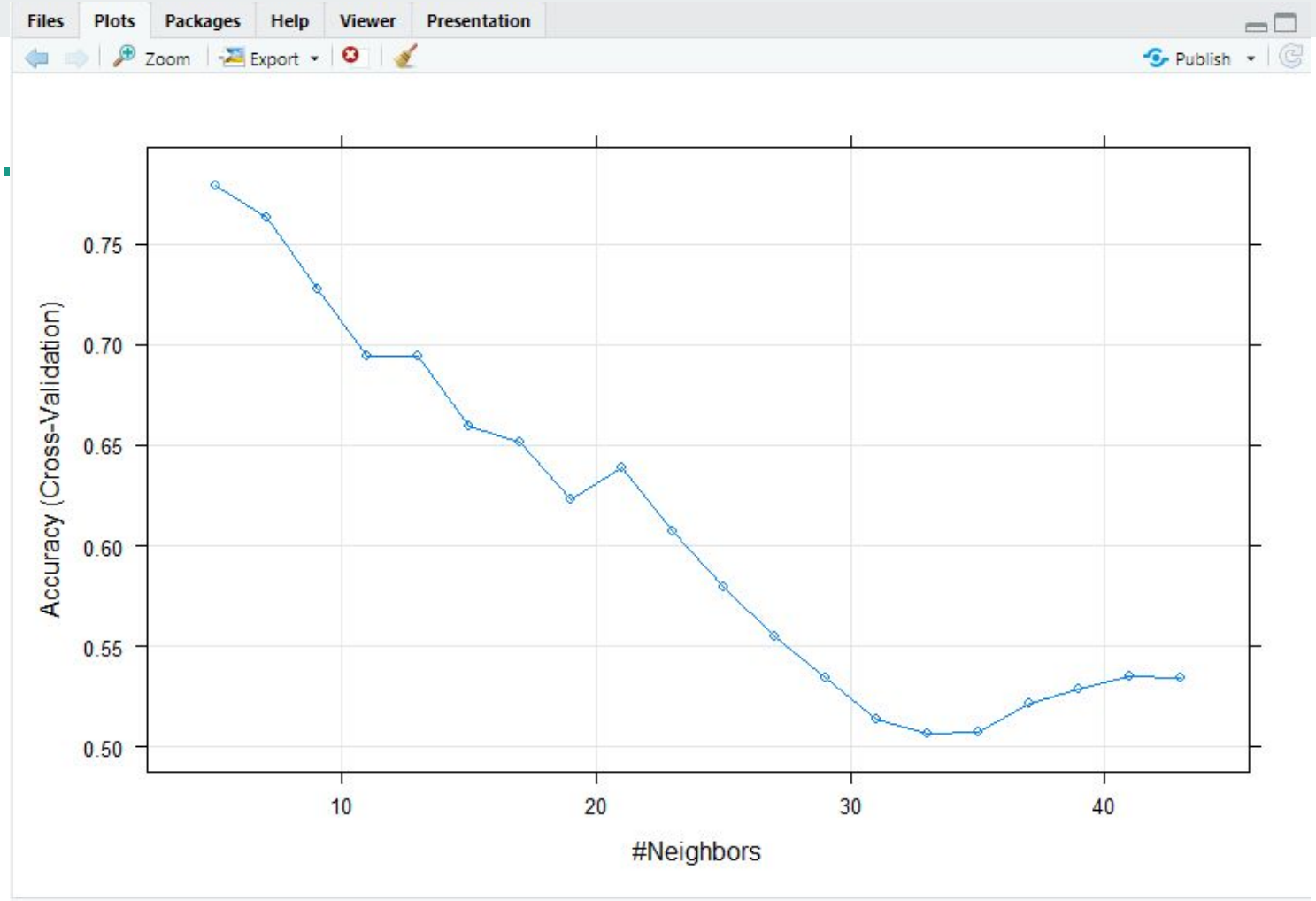
Modelo Lineal

Modelo GLM

	prediccion	actual	RSE
1	1.287688	1	NA
2	1.282117	1	NA
3	1.287688	1	NA
4	1.287688	1	NA
5	1.380586	1	NA
6	1.334137	1	NA
7	1.386158	1	NA
8	1.386158	1	NA
9	1.438178	1	NA
10	1.438178	1	NA
11	1.427035	1	NA
12	1.427035	1	NA
13	1.421464	1	NA
14	1.235668	1	NA
15	1.386158	1	NA
16	1.597867	1	NA
17	1.504969	1	NA
18	1.566318	2	NA



Modelo KNN


Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish
Accuracy (Cross-Validation)
0.75
0.70
0.65
0.60
0.55
0.50
10
20
30
40
#Neighbors