



UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES
CURSO INTERSEMESTRAL: MACHINE Y DEEP LEARNING

ACTIVIDAD 1 06. KNN EN R(PRÁCTICA).

DOCENTE:

FRANCISCO JAVIER LUNA ROSAS

ALUMNO:

EMILIO LUNA PÉREZ, ID:244182

JUNIO-AGOSTO DE 2021

Evidencias de la Practica

Defina el algoritmo KNN, explique el funcionamiento del algoritmo KNN e implemente el algoritmo KNN en R, el dataset que se utilizará, es el dataset visto en clase de la flor iris.

K Nearest Neighbours (KNN): Vecinos cercanos

Los k vecinos más cercanos, se trata de un algoritmo no paramétrico que se utiliza para clasificación, pero, a su vez para regresión. En ambos casos, se basa en elegir de entre un conjunto de k elementos, los que se encuentran más cerca para catalogar o para clasificar el objeto en sí mismo o bien hacer una regresión.

Por ende, la salida va a depender de si se utiliza para clasificar o para regresión. En este caso, lo usaremos para la clasificación, por lo tanto, la salida tiene que ser un miembro de la clase a clasificar, deberá ser una variable categórica de modo que el objeto se va a clasificar por mayoría, el voto de mayoría por sus vecinos.

Cuantos más vecinos tenga cerca el objeto, la votación indicará precisamente que pertenece a dicha categoría. k <- indica cuántas familias o cuántos vecinos existen dentro de la clasificación.

Algoritmo de Knn en R

```
install.packages('kknn',dependencies=TRUE)
library(kknn)
setwd("C:/Users/emili/Desktop/VACACIONES 2021/MACHINE LEARNING")
rm(list = ls(all=TRUE)) #BORRA TODAS LAS VARIABLES DE MEMORIA
datos<-read.csv("iris.csv",sep = ";",dec = '.',header = T)
datos
dim(datos)
head(datos)
summary(datos)
muestra<-sample(1:150,50)
muestra
ttesting<-datos[muestra,]
dim(ttesting)
taprendizaje <- datos[-muestra,]
dim(taprendizaje)
#train.kknn escoje el k usando leave-one-out crossvalidation

modelo<-train.kknn(tipo~., data = taprendizaje, kmax = 9)
modelo

#Que tan bien predice el modelo
prediccion<-predict(modelo,ttesting[,5])
prediccion

#Matriz de confusion
MC<-table(ttesting[,5],prediccion)
MC

#Porcentaje de buena clasificacion
acierto<-(sum(diag(MC)))/sum(MC)
acierto
error<-1-acierto
error
```

```
> modelo<-train.kknn(tipo~.,data=taprendizaje,kmax=9)
> modelo

Call:
train.kknn(formula = tipo ~ ., data = taprendizaje, kmax = 9)

Type of response variable: nominal
Minimal misclassification: 0.04
Best kernel: optimal
Best k: 6
```

```

call:
train.kknn(formula = tipo ~ ., data = taprendizaje, kmax = 9)

Type of response variable: nominal
Minimal misclassification: 0.06
Best kernel: optimal
Best k: 1

```

```
> ## Matriz de Confusion
```

```
>
```

```
> MC<-table(ttesting[,5],prediccion)
```

```
> MC
```

	prediccion		
	setosa	versicolor	virginica
setosa	17	0	0
versicolor	0	15	0
virginica	0	2	16

```
> MC<-table(ttesting[,5],prediccion)
```

```
> MC
```

	prediccion		
	setosa	versicolor	virginica
setosa	17	0	0
versicolor	0	15	0
virginica	0	2	16

```
> acierto<-(sum(diag(MC)))/sum(MC)
```

```
> acierto
```

```
[1] 0.96
```

```
Levels: setosa versicolor virginica
```

```
> MC<-table(ttesting[,5],prediccion)
```

```
> MC
```

	prediccion		
	setosa	versicolor	virginica
setosa	16	1	0
versicolor	0	13	4
virginica	0	3	13

```
> acierto<-(sum(diag(MC)))/sum(MC)
```

```
> acierto
```

```
[1] 0.84
```

```
>
```

Conclusiones

Hemos realizado una práctica teórica y didáctica del algoritmo de aprendizaje supervisado KNN con un ejemplo clásico del uso de este algoritmo es decir con los datos de flores de iris, pusimos al algoritmo a predecir la clasificación de las flores, también puedo agregar que matemáticamente es un algoritmo complejo, sin embargo, con el uso de las librerías de r se nos facilitó el trabajo, y simplemente tuvimos que usar funciones sencillas.

Referencias

RPubs (2020). KNN. Consultado en agosto 1, 2021, de Rpubs.com Sitio web: <https://rpubs.com/JairoAyala/601703>

Download R-4.1.0 for Windows. (2021). The R-project for statistical computing. Consultado en julio 28, 2021, de R-project.org Sitio web: <https://cran.r-project.org/bin/windows/base/>

Download the RStudio IDE. (2021). Consultado en julio 28, 2021, de Rstudio.com Sitio web: <https://www.rstudio.com/products/rstudio/download/>