

Introducción al paquete caret: “classification and regression training in R.”

Laura Rodriguez Navas

7/2/2020

El paquete **caret** [2, 1] se está convirtiendo durante los últimos años en un referente clave para realizar tareas de clasificación y regresión en **R**. También podríamos mencionar otro paquete, **rattle**, con su útil interfaz gráfica, en muchas cuestiones similar a **WEKA**. **caret** por otro lado trabaja en línea de comandos, con una amplísima variedad de funciones que puedes consultar con el comando `ls(“package:caret”)`. Entre estas, hay que destacar las relacionadas con la construcción de modelos y su evaluación, así como algoritmos para selección de variables.

Esta breve introducción nos dibujará los ejes principales del software. Se mostrarán referencias para ampliar los conocimientos sobre el paquete.

Empecemos con ejemplos sobre la construcción de modelos de clasificación supervisada. Este paso final de la construcción nos lo ofrece la función **train**: sus opciones, amplísimas, las puedes encontrar en <https://www.rdocumentation.org/packages/caret/versions/6.0-84/topics/train>. Recuerda que la carga de paquetes en **R** se realiza desde la línea de comandos `install.packages(“caret”)`. Partiremos de la base de datos sonar, con 208 casos, 60 predictoras y dos valores en la variable clase. Puedes consultar en numerosos enlaces de la web acerca de la naturaleza de este dataset. Cargaremos la librería **mlbench**, que recoge numerosos datasets benchmark para tareas de aprendizaje automático. Es interesante localizar estos datasets en la propia instalación del paquete sobre el disco duro del ordenador.

Se pueden encontrar cientos de manuales, para diferentes niveles de conocimiento, para introducirse en

Dependiendo los paquetes de R que tengamos instalados y los que se necesiten para seguir el flujo de la

La ayuda de cada función en R se obtiene mediante los comandos `nombreFuncion` or `help(nombreFuncion)`. Un

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

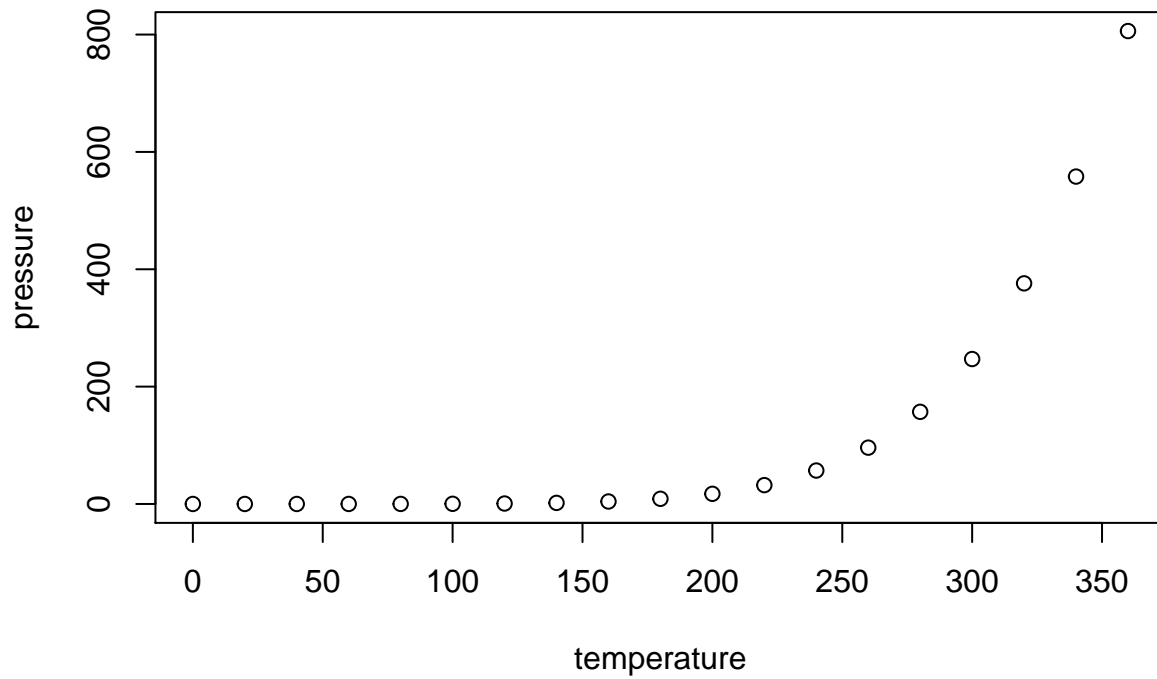
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed          dist
##  Min.   : 4.0      Min.   :  2.00
##  1st Qu.:12.0      1st Qu.: 26.00
##  Median :15.0      Median : 36.00
##  Mean   :15.4      Mean   : 42.98
##  3rd Qu.:19.0      3rd Qu.: 56.00
##  Max.   :25.0      Max.    :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.