Machine Learning I – STA314

Tahir Muhammad

04/10/2020

Tutorial 1 - KNN

Today we will conduct some basic data analysis on an different methods of Advertising data from a Company X. Afterwards, we will go through the K-Nearest Neighbours Algorithm in R, train the model, and illustrate it's MSE. Let's Get Started!

```
 #------------------------------------------------------------------
 # Data Analysis: loading data sets into R and making plots
 #------------------------------------------------------------------

# read in the dataset
getwd()
```

```
## [1] "/home/tahir/Github/R/ML"
```

```
df <- read.csv("/home/tahir/Downloads/Advertising.csv")
# Look at the top few rows
head(df)
```

```
##   X    TV radio newspaper sales
## 1 1 230.1  37.8      69.2  22.1
## 2 2  44.5  39.3      45.1  10.4
## 3 3  17.2  45.9      69.3   9.3
## 4 4 151.5  41.3      58.5  18.5
## 5 5 180.8  10.8      58.4  12.9
## 6 6   8.7  48.9      75.0   7.2
```

Hi

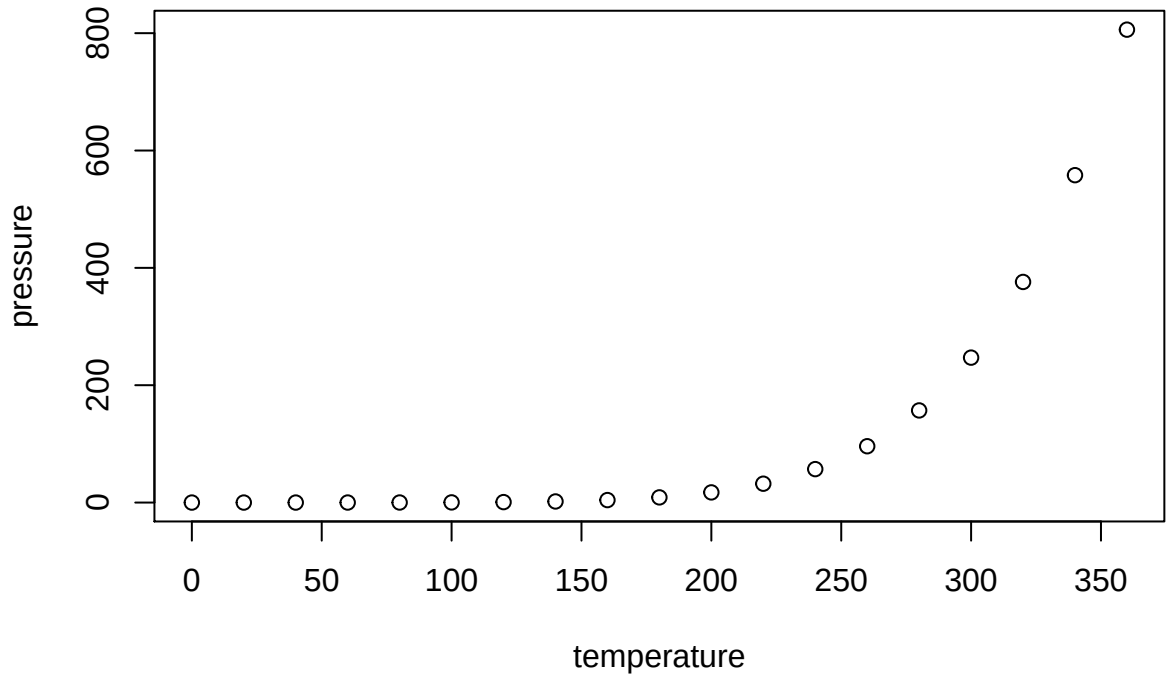
```
 #fix(df) # take a look at data in table form
# names(df) # names of entries in df
 #$TV # values of column TV
 #TV # does not work, not in memory
 #attach(df) # make elements of res available in workspace
 #TV # now works
```

```
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.