

## Aufgabe 3 - Classification

Das MNIST Datenset ist sozusagen das HelloWorld Datenset wenn es um Classification geht. Wir missbrauchen hier dieses Datenset für traditionelle Classification.

1. Führen Sie u.g. Source Code aus und beantworten Sie folgende Fragen
  - a. Wieviele Bilder sind in der Matrix `mnist_matrix` encodiert
  - b. Da es sich um einen Supervised Machine Learning task handelt muss ein Label (Target Variable) bereitgestellt sein - welche Spalte der Matrix enthält das Label?
  - c. Wieviele Pixel haben die Bilder?
  - d. Wie hoch/breit sind die Bilder?

```
mnist_matrix = read.csv( '{see material in the folder}' )  
dim(mnist_matrix)  
sort(unique(mnist_matrix[,1]))
```

```
par( mfrow = c(10,10), mai = c(0,0,0,0))  
for(i in 1:100){  
  y = as.matrix(mnist_matrix[i, 2:785])  
  dim(y) = c(28, 28)  
  image( y[,nrow(y):1], axes = FALSE, col = gray(255:0 / 255))  
  text( 0.2, 0, mnist_matrix[i,1], cex = 3, col = 2, pos = c(3,4))  
}
```

2. Nehmen Sie einen Classifier Ihrer Wahl und trainieren Sie Ihn mit der bereitgestellten Matrix.
  - a. Teilen Sie die Matrix in ein sinnvolles Training und Test set auf (Ein Validation Set wird hier nicht benötigt da nicht erwartet wird Parameter des Classifiers zu tunen)
  - b. Verwenden Sie nun das Training Set um einen Classifier Ihrer Wahl zu trainieren
  - c. Berechnen Sie den Prozentsatz der richtig klassifizierten Daten indem Sie Ihren trainierten Classifier auf das Test Set anwenden