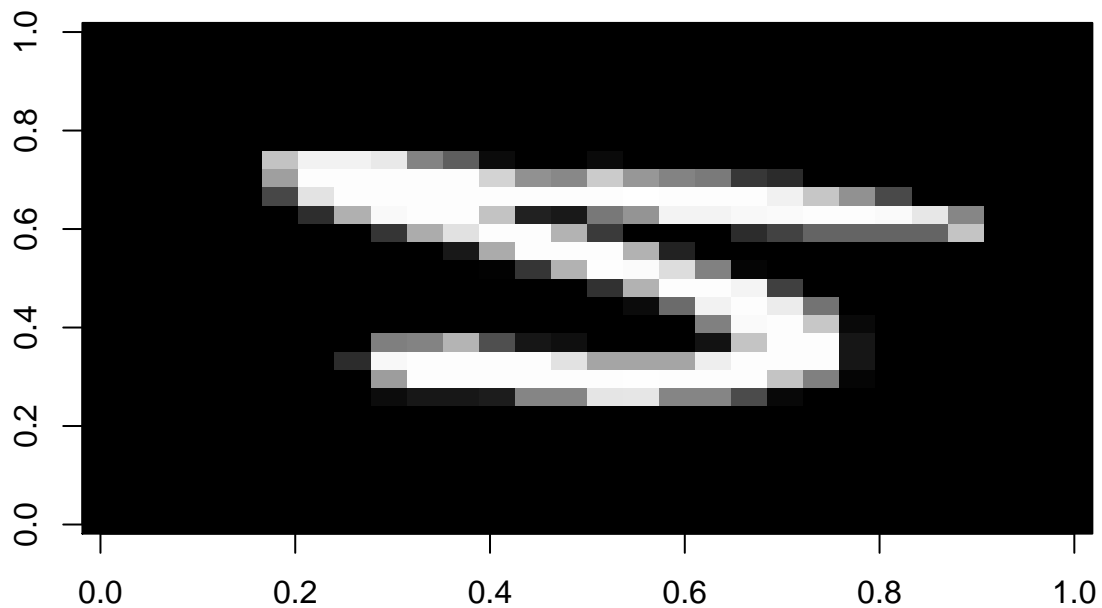# Task 3 - recognition

Tobias Machnitzki, Finn Burgemeister

26 November 2017

NR1: a linear model

```r
library(softmaxreg)
library(ggplot2)
x = load_image_file("train-images-idx3-ubyte")
y = load_label_file("train-labels-idx1-ubyte")
image(matrix(x[200,], nrow=28), col=grey(seq(0, 1, length = 256)))
```
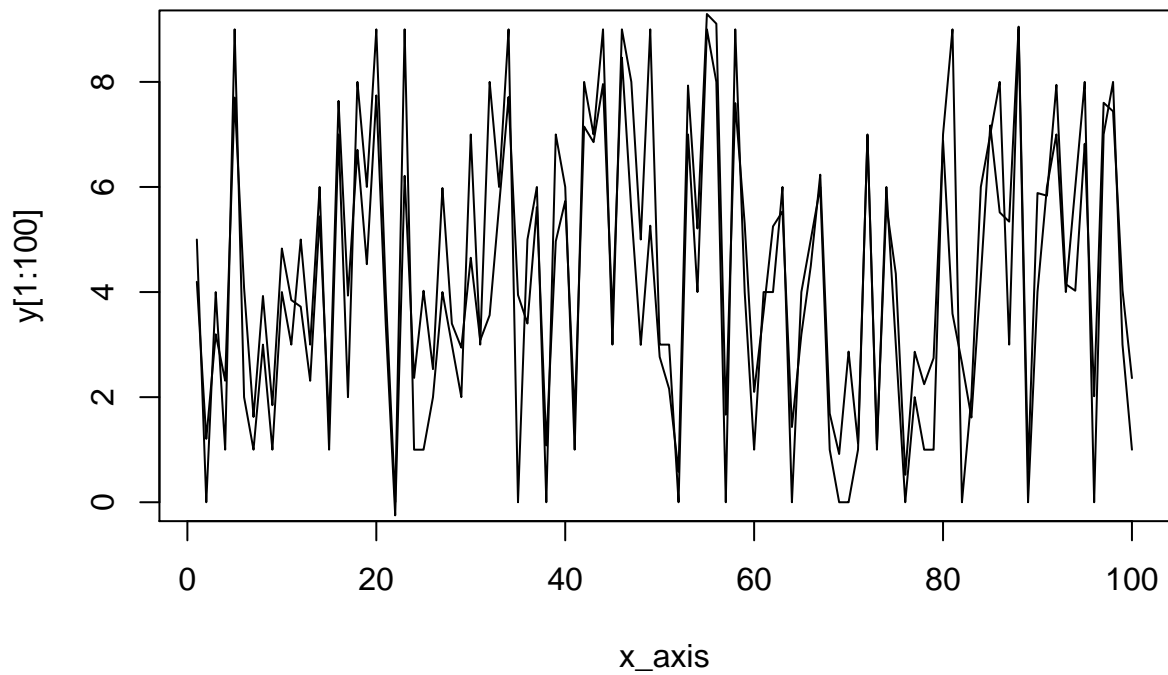


```r
m = lm(y~x)
df = data.frame(x)
pr = predict(m,df)
```

```
## Warning in predict.lm(m, df): prediction from a rank-deficient fit may be
## misleading
```

```r
x_axis = seq(1,length(pr[1:100]))
```

```r
plot(x=x_axis,y=y[1:100],type="l")
lines(x=x_axis,y=pr[1:100])
```
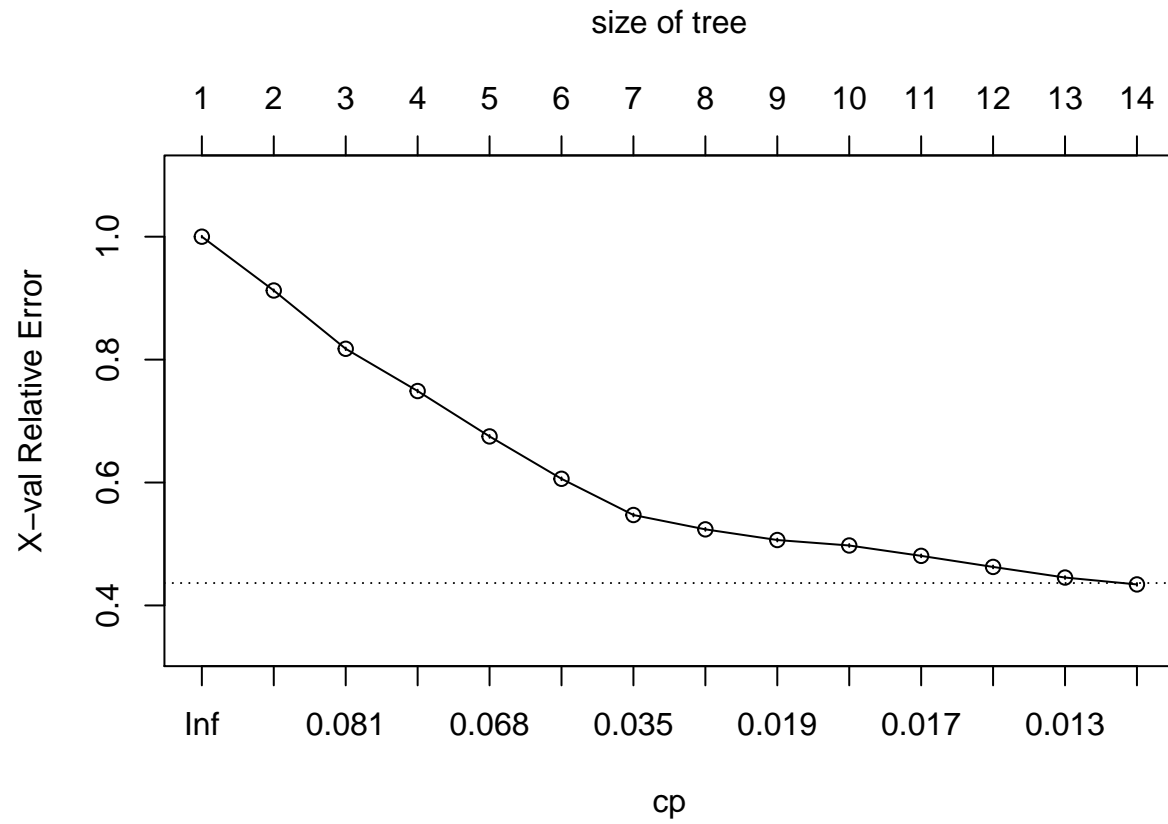
What I find really suspicious is, that the 7 never gets mistaken for a 1 and vice versa. I probably set up the model not right...

NR2 A classification tree

```r
library(rpart)

# grow tree
fit = rpart(y ~ x,method="class")
plotcp(fit)
```

2

size of tree


```
plot(fit, uniform=TRUE,
    main="Classification Tree for MNIST Data")
text(fit, use.n=TRUE, all=TRUE, cex=.8)
```

Classification Tree for MNIST Data

x351>=131.5

5923/6742/5958/6131/5842/5421/5918/6265/5851/5949

x490>=26.5   x569>=0.5

6/4819/489/2155/1143/4542/3762/614772/1312/5353/3266/4775/5773/3081/4772

x235< 0.5 x291< 34.5   x436< 0.5   x431>=0.5

5273/2864/436/2090/1862/1494/220/1698/2004/265/1642/902/847/5133/1668/2771/5568/1

  1   8   3   5   0   x347< 0.5   x212< 28.5   x406>=2.5

                                    2   4   7

                                x656< 0.5   x99< 0.5   x157>=0.5

                                2   6   4   9   8   7

12/191/3128/732/157/93/523/408/369/423/189/272/286/025/650/021

                                   6   8   4   6   6   9