

hw7__ANN.R

lukemcevoy

2021-11-15

```
# clear the environment
rm(list=ls())

# select the data
filename<-' /Users/lukemcevoy/Develop/stevens/f21/dataMining/week10/hw7/wisc_bc_ContinuousVar.csv'
cancer<-read.csv(filename)
View(cancer)
# cancer_df<-data.frame(lapply(cancer[,-1], as.numeric))
cancer$diagnosis[cancer$diagnosis=="M"]<-0
cancer$diagnosis[cancer$diagnosis=="B"]<-1
cancer_df<-data.frame(lapply(na.omit(cancer),as.numeric))
cancer_df<-cancer_df[-1]
View(cancer_df)

normalized_cancer_df<-as.data.frame(apply(cancer_df[,1:ncol(cancer_df)], 2, function(x) (x-min(x))/(max(x)-min(x))))
View(normalized_cancer_df)

# split data
index<-sort(sample(nrow(normalized_cancer_df), round(.3*nrow(normalized_cancer_df))))
training<-normalized_cancer_df[-index,]
test<-normalized_cancer_df[index,]

library("neuralnet")
net_cancer<-neuralnet(diagnosis~., training, hidden=5, threshold=0.01)

plot(net_cancer)

pred<-predict(net_cancer, test)
pred

##           [,1]
## 3    -0.0004583046
## 13    1.4159669215
## 18   -0.0004544585
## 21    0.9978323736
## 27   -0.0124109843
## 28   -0.0004448290
## 29   -0.0004476881
## 32   -0.0053930363
## 33   -0.0004585165
## 36   -0.0004595520
```

37 -0.0123989287
38 0.9945123410
39 2.6695922912
40 -0.8532229734
41 0.9926352679
44 -0.0014469292
46 -0.0004570614
51 0.9980643918
54 -0.0004585838
55 -0.0753585439
57 -0.0004479411
59 0.9973642655
60 0.9977783895
62 0.9973087635
64 0.9973087625
70 1.0005386628
72 0.9973085069
82 0.7502591113
83 0.0138405280
85 0.9983127877
92 -0.0747428122
103 1.0167883101
105 0.9973087257
106 -0.0004811463
111 0.9973144740
114 0.9973085774
116 0.9979020581
120 -0.0050041885
128 -0.0023938459
129 0.9898347225
131 0.9973173897
132 0.0008219911
134 0.8629482335
137 0.5133290098
140 0.9256325369
145 0.9981922036
147 -0.2298383850
149 0.9466156418
159 1.0124160953
160 0.9982087072
165 -0.0004584707
171 1.0619500223
172 -0.0026708367
176 0.9973094690
181 -0.0004286497
183 -0.0011258929
184 0.9976103140
187 0.0452902651
190 0.9974167746
191 0.9973067230
198 0.6739041564
207 0.9974524820
209 1.0719776207
212 0.9975387008

214 0.9485670545
217 0.9973303423
219 -0.0004584442
222 1.0027685648
223 0.9973188246
226 0.9819928985
227 0.9970045994
232 0.9975525156
233 1.0018629031
236 1.0073851429
238 -0.0004900342
242 0.9979449124
244 0.9974921541
250 1.0118974803
251 -0.0004585656
259 -0.0004604439
260 0.2525588972
262 0.0102225497
263 -0.0585830127
264 0.6042973864
267 0.9973227537
268 0.9975569520
269 0.9973704615
271 0.9977994321
273 -0.0004585495
283 0.0009609133
290 0.9973318928
295 0.9175914538
301 -0.0004591618
302 0.9973132806
305 0.9973046709
309 0.9990415188
317 0.9978062191
318 0.0047762635
320 0.9973504168
321 0.9973094125
324 -0.0004585720
325 0.9956197621
327 1.0101009497
332 0.9974624615
337 0.9973162981
338 -0.0004588220
340 -0.0004572845
345 0.9976745214
347 0.9979362157
349 0.9986009141
352 -0.0006702175
356 0.9973318979
357 0.9973079646
358 1.0023536241
363 0.9975447700
367 -0.0004585828
369 -0.0004515801
372 0.9996289164

374 0.0007014207
379 0.9973332177
381 0.9902569006
385 1.0011448239
386 -0.0113532730
389 0.9976104742
390 -0.0016993622
391 0.9975271672
394 -0.0004584211
396 1.0049065935
402 1.6513539307
406 1.0016840013
412 0.9993028986
419 0.9978280321
423 0.9935495981
424 1.0129547628
427 0.9973205703
428 0.9973730745
433 -0.0003037298
440 0.8244095275
447 0.0251724043
449 1.4073669842
470 0.7869131842
476 0.9993470400
477 0.9948736231
480 -0.0373231390
483 1.0024582031
485 0.2286383125
486 0.9973085114
487 1.2234813145
490 1.0010883457
491 1.0130966771
492 1.4873113996
494 0.9974211360
506 0.9973119430
508 0.9973103451
510 -0.0004427468
511 0.9973107572
521 0.9973093236
527 0.8447722446
528 1.1486569749
533 1.0054281674
534 -0.0004585763
537 0.0790841878
538 0.9981230611
540 0.9973084197
547 0.9974146182
548 0.9973192590
549 0.9973259404
557 0.9973112477
559 1.0537666625
561 1.0108496339
566 -0.0003858969

```
ann<-compute(net_cancer, test)
ann$net.result
```

```
##           [,1]
## 3  -0.0004583046
## 13  1.4159669215
## 18  -0.0004544585
## 21  0.9978323736
## 27  -0.0124109843
## 28  -0.0004448290
## 29  -0.0004476881
## 32  -0.0053930363
## 33  -0.0004585165
## 36  -0.0004595520
## 37  -0.0123989287
## 38  0.9945123410
## 39  2.6695922912
## 40  -0.8532229734
## 41  0.9926352679
## 44  -0.0014469292
## 46  -0.0004570614
## 51  0.9980643918
## 54  -0.0004585838
## 55  -0.0753585439
## 57  -0.0004479411
## 59  0.9973642655
## 60  0.9977783895
## 62  0.9973087635
## 64  0.9973087625
## 70  1.0005386628
## 72  0.9973085069
## 82  0.7502591113
## 83  0.0138405280
## 85  0.9983127877
## 92  -0.0747428122
## 103 1.0167883101
## 105 0.9973087257
## 106 -0.0004811463
## 111 0.9973144740
## 114 0.9973085774
## 116 0.9979020581
## 120 -0.0050041885
## 128 -0.0023938459
## 129 0.9898347225
## 131 0.9973173897
## 132 0.0008219911
## 134 0.8629482335
## 137 0.5133290098
## 140 0.9256325369
## 145 0.9981922036
## 147 -0.2298383850
## 149 0.9466156418
## 159 1.0124160953
```

160 0.9982087072
165 -0.0004584707
171 1.0619500223
172 -0.0026708367
176 0.9973094690
181 -0.0004286497
183 -0.0011258929
184 0.9976103140
187 0.0452902651
190 0.9974167746
191 0.9973067230
198 0.6739041564
207 0.9974524820
209 1.0719776207
212 0.9975387008
214 0.9485670545
217 0.9973303423
219 -0.0004584442
222 1.0027685648
223 0.9973188246
226 0.9819928985
227 0.9970045994
232 0.9975525156
233 1.0018629031
236 1.0073851429
238 -0.0004900342
242 0.9979449124
244 0.9974921541
250 1.0118974803
251 -0.0004585656
259 -0.0004604439
260 0.2525588972
262 0.0102225497
263 -0.0585830127
264 0.6042973864
267 0.9973227537
268 0.9975569520
269 0.9973704615
271 0.9977994321
273 -0.0004585495
283 0.0009609133
290 0.9973318928
295 0.9175914538
301 -0.0004591618
302 0.9973132806
305 0.9973046709
309 0.9990415188
317 0.9978062191
318 0.0047762635
320 0.9973504168
321 0.9973094125
324 -0.0004585720
325 0.9956197621
327 1.0101009497

332 0.9974624615
337 0.9973162981
338 -0.0004588220
340 -0.0004572845
345 0.9976745214
347 0.9979362157
349 0.9986009141
352 -0.0006702175
356 0.9973318979
357 0.9973079646
358 1.0023536241
363 0.9975447700
367 -0.0004585828
369 -0.0004515801
372 0.9996289164
374 0.0007014207
379 0.9973332177
381 0.9902569006
385 1.0011448239
386 -0.0113532730
389 0.9976104742
390 -0.0016993622
391 0.9975271672
394 -0.0004584211
396 1.0049065935
402 1.6513539307
406 1.0016840013
412 0.9993028986
419 0.9978280321
423 0.9935495981
424 1.0129547628
427 0.9973205703
428 0.9973730745
433 -0.0003037298
440 0.8244095275
447 0.0251724043
449 1.4073669842
470 0.7869131842
476 0.9993470400
477 0.9948736231
480 -0.0373231390
483 1.0024582031
485 0.2286383125
486 0.9973085114
487 1.2234813145
490 1.0010883457
491 1.0130966771
492 1.4873113996
494 0.9974211360
506 0.9973119430
508 0.9973103451
510 -0.0004427468
511 0.9973107572
521 0.9973093236

```
## 527 0.8447722446
## 528 1.1486569749
## 533 1.0054281674
## 534 -0.0004585763
## 537 0.0790841878
## 538 0.9981230611
## 540 0.9973084197
## 547 0.9974146182
## 548 0.9973192590
## 549 0.9973259404
## 557 0.9973112477
## 559 1.0537666625
## 561 1.0108496339
## 566 -0.0003858969
```

```
length(ann$net.result)
```

```
## [1] 171
```

```
ann_cat<-ifelse(ann$net.result<1.5,1,2)
ann_cat<-ifelse(ann$net.result <0.5,0,1)
length(ann_cat)
```

```
## [1] 171
```

```
table(Actual=test$diagnosis, predict=ann_cat)
```

```
##      predict
## Actual    0    1
##      0  55    8
##      1   1 107
```

```
wrong<-(test$diagnosis!=ann_cat)
error_rate<-sum(wrong)/length(wrong)
error_rate
```

```
## [1] 0.05263158
```