# anomaly

*Matous Dzivjak*

*12/2/2019*

```
library('FNN')
library('AUC')
```

```
## AUC 0.3.0
```

```
## Type AUCNews() to see the change log and ?AUC to get an overview.
```

```
read_dir <- function(dirName) {
  anomalFile = paste(dirName, 'anomalous.csv', sep='')
  normalFile = paste(dirName, 'normal.csv', sep='')
  data_anomalous = as.matrix(read.csv(anomalFile, header=TRUE))
  data_anomalous = cbind(data_anomalous, rep(1,nrow(data_anomalous)))
  data_normal = as.matrix(read.csv(normalFile, header=TRUE))
  data_normal = cbind(data_normal, rep(0,nrow(data_anomalous)))
  problem = rbind(data_anomalous, data_normal)
  colnames(problem) = c("x","y","anomalous")
  return(data.frame(problem))
}

knn_eval <- function(data, k) {
  knn <- get.knn(data.frame(data$x, data$y), k)
  dists <- knn$nn.dist
  sorted_desc <- order(dists[,k], decreasing = TRUE)
  anomalous_count = length(which(data$anomalous==1))
  normal_count = length(which(data$anomalous==0))
  p <- anomalous_count/normal_count
  p_count <- round(p*nrow(data))
  # mark first 'p' (those with longest distances) as anomalous
  first_p <- sorted_desc[1:p_count]
  anomalous_labels = rep(0,nrow(dists))
  anomalous_labels[first_p] = 1
  # calculate auc for given k
  auc_val <- auc(roc(anomalous_labels, factor(data$anomalous)))
  return(auc_val)
}

bloodData <- read_dir('problems/blood-transfusion/')
```

```
## Warning in cbind(data_normal, rep(0, nrow(data_anomalous))): number of rows of
## result is not a multiple of vector length (arg 2)
```

```
print(bloodData)
```

```
##                x             y anomalous
## 1   -0.256510252 -2.0103353652         1
## 2   -0.178006419 -2.3449379285         1
## 3   -0.245764696 -2.0170966523         1
## 4   -0.223348797 -2.0344945068         1
## 5   -0.182220317 -2.2919111998         1
```

```
## 6   -0.764882655 -1.5612132523         0
## 7   -0.213193279 -2.0702533973         0
## 8   -2.054388754  0.1036452460         0
## 9   -0.571026432 -1.6305229683         0
## 10  -0.955679343 -1.4640008997         0
## 11   0.231954437  0.0220609349         0
## 12  -0.637241157 -1.7118538805         0
## 13  -1.836978021 -0.5306836855         0
## 14  -1.194612230 -1.1235409717         0
## 15  -0.512987688 -1.6709995913         0
## 16  -0.287153098 -1.7892984566         0
## 17  -1.855788825 -0.5017484368         0
## 18  -0.990246280 -1.0899078910         0
## 19  -2.083929361  0.2186450597         0
## 20  -0.686718205 -1.4046241597         0
## 21  -1.423123190 -0.7625203450         0
## 22  -0.583674149 -1.5311954194         0
## 23  -0.221130651 -1.6879901500         0
## 24  -1.441397469 -0.7442013210         0
## 25  -1.193018408 -0.9145376418         0
## 26  -0.693966851 -1.2150096133         0
## 27  -0.239558520 -1.7381423698         0
## 28  -2.079837779  0.1592210008         0
## 29  -1.216325736 -1.0463312208         0
## 30  -0.697096579 -1.0187323976         0
## 31  -1.873099398 -0.2416058249         0
## 32  -2.048724663  0.1725450129         0
## 33  -1.335954573 -0.8157005283         0
## 34  -2.018142280  0.3090877228         0
## 35   0.287303143 -0.0834003738         0
## 36  -1.222839092 -0.9669067714         0
## 37  -0.210950345 -1.6908173400         0
## 38  -0.928286457 -1.0225324984         0
## 39  -1.426923343 -0.5513133724         0
## 40  -0.232197519 -1.8010481223         0
## 41  -0.674335043 -0.7697508404         0
## 42  -1.415285432 -0.4966875322         0
## 43  -1.887422019 -0.0687894663         0
## 44  -0.459818209 -1.0958643157         0
## 45  -0.786564887 -0.9383491889         0
## 46  -1.429870060 -0.5153032940         0
## 47  -0.283004760 -1.2950270961         0
## 48  -0.196766258 -2.1034698953         0
## 49  -0.464556123 -1.1508651687         0
## 50  -0.693256349 -0.7107570735         0
## 51  -1.463367422 -0.2684205705         0
## 52  -1.111254422 -0.3709060991         0
## 53  -1.389721040 -0.4466379102         0
## 54  -0.248101683 -1.3414349364         0
## 55  -1.955182612  0.1255406013         0
## 56  -1.069555132 -0.5375378413         0
## 57  -0.259066621 -1.3251563715         0
## 58  -1.628262189 -0.0663249066         0
## 59  -1.054912357 -0.3864027976         0
```

```
## 60  -0.236938444 -1.3705249664         0
## 61   0.084750861  1.8461725322         0
## 62  -1.955071415  0.1262879736         0
## 63  -0.422387026 -0.9862898059         0
## 64  -0.187148423 -1.4339998716         0
## 65  -1.856036702 -0.1658434998         0
## 66  -1.861934769 -0.0741846983         0
## 67  -0.979854006 -0.4613938699         0
## 68  -0.335490146 -1.1781009193         0
## 69  -2.086910132  0.2741553641         0
## 70   0.326715020 -0.1373889575         0
## 71  -1.544269633 -0.0479899622         0
## 72   0.057521759  1.8733273543         0
## 73  -1.131121133 -0.1889143081         0
## 74  -0.248956973 -1.3107201457         0
## 75   0.332666833  0.5611626938         0
## 76   0.332396486 -0.0524741961         0
## 77  -1.103931182 -0.3534043201         0
## 78  -1.439010443 -0.2558098789         0
## 79   0.062697741 -1.5039464018         0
## 80  -0.741030926 -0.6554842477         0
## 81   0.347994163 -0.2195923936         0
## 82  -1.928449726  0.0877505386         0
## 83  -0.853812816 -0.4083317626         0
## 84   0.288631250  0.6478401925         0
## 85   0.405244336 -0.0901753701         0
## 86  -0.762373776 -0.5080727986         0
## 87  -0.687589942 -0.6849444259         0
## 88   0.355144122 -0.0417766007         0
## 89   0.266623536  0.7672769211         0
## 90  -1.879986886  0.0704663378         0
## 91  -0.468753341 -0.7740068825         0
## 92   0.267457513  0.7538565025         0
## 93   0.022596675  1.9150046375         0
## 94  -2.084909515  0.3201691756         0
## 95  -1.471947419 -0.0808124314         0
## 96   0.374407358  0.0444365543         0
## 97   0.350053622  0.6132753287         0
## 98   0.545954710 -0.3152043082         0
## 99   0.022119455  1.9030718878         0
## 100 -1.917001313  0.1171262505         0
## 101 -0.985810452 -0.1375136669         0
## 102  0.544926139 -0.2305301211         0
## 103 -0.101094092 -1.9918832293         0
## 104 -0.070473485  1.9805761653         0
## 105 -0.882393887 -0.3240958275         0
## 106  0.419159466 -0.5006601035         0
## 107 -0.246366085 -1.3192242696         0
## 108 -0.028687554  2.0381163152         0
## 109 -0.482646464 -0.7177071937         0
## 110  0.405265649 -0.5615483929         0
## 111 -1.187667059 -0.0413290618         0
## 112  0.381832761  0.6020293655         0
## 113  0.356399952  0.2433171905         0
```

```
## 114 -1.417059949 -0.0837842313        0
## 115  0.433874055 -0.5141442036        0
## 116 -1.014768897 -0.1828716482        0
## 117 -1.515761072 -0.0194894277        0
## 118 -0.689627856 -0.5374082673        0
## 119 -1.492133358  0.0024970274        0
## 120  0.472674350  0.0212635903        0
## 121 -0.460015584 -0.7488810215        0
## 122  0.262894966  1.7128404276        0
## 123  0.348631924 -0.2787966631        0
## 124  0.395061482  0.6597707310        0
## 125 -0.143552069  1.9400495701        0
## 126 -0.949204442 -0.1560405400        0
## 127  0.409653292 -0.5500936961        0
## 128 -1.224043725 -0.0611021450        0
## 129  0.166068438 -1.4646275426        0
## 130  0.494316504  0.0091742904        0
## 131  0.203966416 -1.3845393028        0
## 132  0.349916016  1.4843399673        0
## 133 -0.080944524  1.9771784958        0
## 134  0.228743859  1.0110693264        0
## 135 -0.018249643  2.0292274493        0
## 136 -1.467307035 -0.0018045646        0
## 137  0.406049586 -0.4838393620        0
## 138  0.174950983  1.3913504894        0
## 139  0.216308342  0.9958149711        0
## 140 -0.524770155 -0.6110682209        0
## 141  0.366357863  1.4536070450        0
## 142  0.176003878  2.0811101235        0
## 143 -0.971655554 -0.0234681903        0
## 144  0.418600933  0.7308648537        0
## 145 -0.504223290 -0.6680626563        0
## 146  0.084641285  1.5760429920        0
## 147  0.346490689  1.7400120546        0
## 148  0.322874327  1.9638460814        0
## 149  0.220718920  0.9842461785        0
## 150  0.001791277  1.9955072489        0
## 151  0.587196703 -0.1969735466        0
## 152 -0.063723835  2.0041144130        0
## 153 -1.380861197 -0.0087064816        0
## 154  0.013816987  1.9059035002        0
## 155  0.353397868  1.6075471672        0
## 156  0.231683163  0.9828619919        0
## 157  0.175002643 -1.4550709061        0
## 158  0.231939379  1.0985189624        0
## 159  0.475874272  0.5850237709        0
## 160  0.202155529  2.0527267805        0
## 161 -1.087411033  0.0177522672        0
## 162  0.160087667 -1.4755736724        0
## 163  0.478256665  0.6036328992        0
## 164  0.549603357  0.2223023194        0
## 165  0.475075508  0.7090406978        0
## 166  0.242961769  2.0663705264        0
## 167  0.107940968  1.4890939628        0
```

```
## 168  0.761429458 -0.1878127079        0
## 169  0.264588170  1.2228083043        0
## 170  0.388990365  1.5572565286        0
## 171 -0.171517618 -2.4210580194        0
## 172  0.763067990  0.0908594305        0
## 173 -0.030441221  2.0497160642        0
## 174  0.037169951  1.6963372088        0
## 175  0.728303679  0.3271670569        0
## 176  0.383445811  1.5374785808        0
## 177 -1.086861303  0.0259715936        0
## 178  0.492831100  0.6912718106        0
## 179  0.388156157  1.7282442574        0
## 180  0.058259828  1.6074220155        0
## 181  0.251753271 -1.2954967466        0
## 182  0.627009584  0.2391350012        0
## 183  0.052781529  1.6032959897        0
## 184 -0.527765984 -0.6067781269        0
## 185  0.296114800 -1.2108897788        0
## 186  0.834841401 -0.1583901194        0
## 187  0.278093957 -1.2601615144        0
## 188  0.527941975  1.2907705390        0
## 189  0.057030407  1.6027277214        0
## 190  0.470989486  1.5778764865        0
## 191 -0.983438484  0.0043950877        0
## 192  0.860776218  0.2235153622        0
## 193  0.563595630  1.2271775578        0
## 194  0.675005628  0.3237034949        0
## 195  0.186913449  2.0787990184        0
## 196  0.841178464 -0.2566165103        0
## 197  0.778239414  0.3843556323        0
## 198  0.839709044 -0.2365135219        0
## 199  0.809607124  0.3065944180        0
## 200  0.817769901  0.2742601731        0
## 201 -0.954453629  0.0022763535        0
## 202  0.819055615  0.2734535417        0
## 203  0.578107517  1.1903953421        0
## 204  0.945187153 -0.3299275427        0
## 205  1.763502533  0.2743238545        0
## 206  0.878977756 -0.2741846242        0
## 207  0.399042794  1.9083937093        0
## 208  1.617667373  0.0447611526        0
## 209  1.396482665 -0.1589725391        0
## 210  0.509990167  1.4278297760        0
## 211  0.836658080  0.0929273080        0
## 212  0.554572467  1.2693917109        0
## 213  0.235710343  2.0545629284        0
## 214  0.610160015  1.0547454989        0
## 215  0.389752063  1.8831802863        0
## 216  1.826768919  0.3575545960        0
## 217  0.409851361  1.8487878931        0
## 218  0.802469212  0.2474286227        0
## 219  0.805645272  0.1129206229        0
## 220  1.789110246  0.3948126534        0
## 221  0.912602068 -0.0511563443        0
```

```
## 222  1.705079002  0.2121805103        0
## 223  0.565093081  1.2689888374        0
## 224  0.771700111  0.3949868929        0
## 225  0.856465022  0.1599803134        0
## 226  0.490696815  1.4836248780        0
## 227  0.975167682 -0.3075846485        0
## 228  1.777865926  0.3216524222        0
## 229  0.761072470  0.5157242262        0
## 230  0.226762701 -1.4082252560        0
## 231  0.978220268 -0.3153812010        0
## 232  0.629980162  1.0838138595        0
## 233  0.940050320  0.0489428997        0
## 234  0.759920656  0.4376981148        0
## 235  0.259143230 -1.3586187506        0
## 236  1.663554152  0.0353274567        0
## 237  1.725612199  0.3054791085        0
## 238  1.258234607 -0.2447315217        0
## 239  0.666868566  0.9356647671        0
## 240  1.673182874  0.1152839114        0
## 241  1.640514855  0.0074637365        0
## 242  1.740226016  0.2796094130        0
## 243  1.796025301  0.4041676320        0
## 244  1.354921213 -0.1708318783        0
## 245  1.507570888 -0.1012692055        0
## 246  1.414420806 -0.1898049636        0
## 247  0.933007852  0.1038424816        0
## 248  1.377359120 -0.1490055112        0
## 249  1.607959655 -0.0253025692        0
## 250  1.591148078 -0.0068314173        0
## 251  0.962993942 -0.0210359135        0
## 252  1.688757386  0.0664818262        0
## 253  1.259217310 -0.2141821257        0
## 254  1.793194414  0.4242949433        0
## 255  1.390812783 -0.1720360764        0
## 256  1.072455678 -0.2621077067        0
## 257  1.499094906 -0.1100837765        0
## 258  1.638647675 -0.0201956725        0
## 259  1.310049338 -0.1705943996        0
## 260  0.972465598  0.0109626772        0
## 261  1.277879619 -0.1746425926        0
## 262  1.111428560 -0.2098230435        0
## 263  1.118455044 -0.2232310797        0
## 264  1.127589172 -0.1841576593        0
## 265  1.474006112  0.0681968025        0
## 266  1.504801159  0.1058625980        0
## 267  1.498457841  0.0923948605        0
## 268  1.410508992  0.0007648041        0
## 269 -0.166825574 -2.4420100936        0
## 270 -0.966336637 -1.4627790123        0
## 271 -0.966432312 -1.4443326254        0
## 272 -0.952683514 -1.4608278841        0
## 273 -0.676497359 -1.7116194975        0
## 274 -1.075994358 -1.2640266233        0
## 275 -1.907581780 -0.2321283480        0
```

```
## 276 -0.334799104 -1.7586172755         0
## 277 -0.162616540 -2.1065385441         0
## 278 -1.145803753 -1.0892905350         0
## 279 -1.872018935 -0.2785830650         0
## 280 -1.350495185 -0.7730790795         0
## 281 -0.679817049 -1.1345410538         0
## 282 -0.220935361 -1.4227172031         0
## 283 -0.614291976 -1.7262507477         0
## 284 -1.939160996 -0.0178580334         0
## 285 -1.864433687 -0.2665729246         0
## 286 -1.519043927 -0.1896515091         0
## 287 -1.845420412 -0.2919075192         0
## 288 -1.301631110 -0.7518289409         0
## 289 -1.269421547 -0.7222621847         0
## 290  0.195214991  1.7826432325         0
## 291 -0.736409306 -0.8164302113         0
## 292 -0.709266613 -0.9295147176         0
## 293  0.406470976 -0.2146071108         0
## 294 -0.208352972 -1.4578012474         0
## 295 -0.414254828 -0.9982119421         0
## 296 -0.694536735 -0.9283706630         0
## 297 -0.863802465 -0.4588165410         0
## 298 -0.848715830 -0.4221802619         0
## 299 -1.970261371  0.2648273616         0
## 300  0.359668677 -0.0354924820         0
## 301  0.105773556  1.8259404120         0
## 302 -1.574222132 -0.0552142664         0
## 303  0.336324515 -0.2049558366         0
## 304 -0.348273153 -1.1667886212         0
## 305 -1.025878854 -0.3350030409         0
## 306 -0.920103757 -0.4354750866         0
## 307 -1.506227561 -0.0738561199         0
## 308  0.340606521 -0.2171658659         0
## 309 -0.880319834 -0.4234773288         0
## 310 -0.034842764  1.9531272588         0
## 311  0.347190303  0.5986816682         0
## 312 -1.938647869  0.1540827056         0
## 313 -0.250323304 -1.3497090997         0
## 314 -0.465576354 -0.7691423290         0
## 315 -0.660191264 -0.6920502119         0
## 316  0.262673268  0.7483949356         0
## 317 -1.096215509 -0.3902572929         0
## 318  0.522176417 -0.1887701404         0
## 319 -0.601322714 -0.6205456978         0
## 320 -0.267468935 -1.3012780473         0
## 321  0.588061258 -0.2485187977         0
## 322 -1.527885469 -0.0166445484         0
## 323 -1.440546350 -0.0738623111         0
## 324  0.171678088 -1.4286183487         0
## 325  0.140847134  1.9614084978         0
## 326  0.158725506 -1.4735376357         0
## 327  0.427922243  0.5949249062         0
## 328  0.162936392  1.4379887327         0
## 329 -0.757926364 -0.4181855790         0
```

```
## 330  0.239131037  0.9342796961        0
## 331  0.223504633  0.9873948713        0
## 332 -1.266045795 -0.0031440857        0
## 333  0.525334640  0.1619818560        0
## 334  0.360612692  1.4786130612        0
## 335 -1.373035023 -0.0094432403        0
## 336  0.279669709  1.9198868803        0
## 337  0.195786496 -1.4247028254        0
## 338 -1.336012024 -0.0036738798        0
## 339  0.539785600  0.1312122285        0
## 340  0.092790394  1.5233214034        0
## 341  0.535068073  0.2261033046        0
## 342  0.422436994  1.2874069214        0
## 343  0.340499263  1.9304659575        0
## 344  0.806700484 -0.2544822980        0
## 345 -0.527202123 -0.6128309581        0
## 346  0.185371056 -1.4268197906        0
## 347  0.187674221 -1.4316887664        0
## 348  0.542850927 -0.3849872126        0
## 349  0.358737171  1.8842018426        0
## 350  0.234748722 -1.3465281239        0
## 351  0.837628736 -0.2445984098        0
## 352  0.414633522  1.6749406913        0
## 353 -1.006608437  0.0215567902        0
## 354  0.061131719  1.5915613599        0
## 355  0.322209926  1.1868725209        0
## 356  0.491607934  1.3779831386        0
## 357  1.424735010 -0.1526375638        0
## 358  0.854644174 -0.2677722582        0
## 359  0.613503797  0.2464835726        0
## 360  0.502450324  1.5442248321        0
## 361  0.514997269  1.4223116035        0
## 362  0.792881725 -0.1328781897        0
## 363  1.813771164  0.3606351850        0
## 364  0.522393019  1.3971715910        0
## 365  0.460626864  1.6356085651        0
## 366  0.927837197  0.0644227934        0
## 367  1.410786058 -0.1721289452        0
## 368  1.428277789 -0.1353073708        0
## 369  0.961821278 -0.2318024229        0
## 370  0.786085974  0.2259648860        0
## 371  0.909645853 -0.0701015690        0
## 372  1.766297281  0.2911003729        0
## 373  1.229727436 -0.2325422773        0
## 374  1.643478020  0.0600871496        0
## 375  1.624167062 -0.0045362325        0
## 376  1.400820038 -0.1372001243        0
## 377  1.617470230  0.0129124787        0
## 378  1.681080169  0.0721689315        0
## 379  0.959584599  0.0112762198        0
## 380  1.414397640 -0.1539297661        0
## 381  1.073145330 -0.2688318464        0
## 382  1.007822960 -0.0680575721        0
## 383  1.643250762 -0.0174564866        0
```

```
## 384  1.486740471 -0.1601329562         0
## 385  1.352048627 -0.1531881428         0
## 386  1.499119926  0.0922918204         0
## 387  1.413966057  0.0147733915         0
```

```
k_values = c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,25,50)
for(k in k_values) {
  aucval <- knn_eval(bloodData, k)
  print(aucval)
}
```

```
## [1] 0.4934555
## [1] 0.4934555
## [1] 0.4934555
## [1] 0.4934555
## [1] 0.4934555
## [1] 0.4934555
## [1] 0.4934555
## [1] 0.4934555
## [1] 0.4934555
## [1] 0.4934555
## [1] 0.6960733
## [1] 0.6960733
## [1] 0.6960733
## [1] 0.6960733
## [1] 0.6960733
## [1] 0.6960733
## [1] 0.6960733
## [1] 0.6960733
## [1] 0.6960733
## [1] 0.6960733
## [1] 0.6960733
```