Untitled

```
url <- "https://github.com/dataprofessor/data/raw/master/acetylcholinesterase_06_bioactivity_data_3class
download.file(url, destfile = "acetylcholinesterase_06_bioactivity_data_3class_pIC50_pubchem_fp.csv")

df_Ach <- read.csv("acetylcholinesterase_06_bioactivity_data_3class_pIC50_pubchem_fp.csv")

head(df_Ach)
```

```
##   PubchemFP0 PubchemFP1 PubchemFP2 PubchemFP3 PubchemFP4 PubchemFP5 PubchemFP6
## 1          1          1          1          0          0          0          0
## 2          1          1          1          0          0          0          0
## 3          1          1          1          0          0          0          0
## 4          1          1          0          0          0          0          0
## 5          1          1          0          0          0          0          0
## 6          1          1          0          0          0          0          0
##   PubchemFP7 PubchemFP8 PubchemFP9 PubchemFP10 PubchemFP11 PubchemFP12
## 1          0          0          1           1           1           1
## 2          0          0          1           1           1           1
## 3          0          0          1           1           1           1
## 4          0          0          1           1           1           1
## 5          0          0          1           1           1           0
## 6          0          0          1           1           1           0
##   PubchemFP13 PubchemFP14 PubchemFP15 PubchemFP16 PubchemFP17 PubchemFP18
## 1           0           1           1           1           0           1
## 2           0           1           1           0           0           1
## 3           0           1           1           1           0           1
## 4           0           1           1           1           0           1
## 5           0           1           1           1           0           1
## 6           0           1           1           1           0           1
##   PubchemFP19 PubchemFP20 PubchemFP21 PubchemFP22 PubchemFP23 PubchemFP24
## 1           0           0           0           0           0           0
## 2           1           1           0           0           0           0
## 3           0           0           0           0           1           1
## 4           0           0           0           0           1           1
## 5           1           0           0           0           1           1
## 6           0           0           0           0           0           0
##   PubchemFP25 PubchemFP26 PubchemFP27 PubchemFP28 PubchemFP29 PubchemFP30
## 1           0           0           0           0           0           0
## 2           0           0           0           0           0           0
## 3           0           0           0           0           0           0
## 4           0           0           0           0           0           0
## 5           0           0           0           0           0           0
## 6           0           0           0           0           0           0
##   PubchemFP31 PubchemFP32 PubchemFP33 PubchemFP34 PubchemFP35 PubchemFP36
## 1           0           0           1           0           0           0
## 2           0           0           0           0           0           0
```

```
## 3               0               0               1               0               0               0
## 4               0               0               1               0               0               0
## 5               0               0               1               0               0               0
## 6               0               0               1               0               0               0
##     PubchemFP37 PubchemFP38 PubchemFP39 PubchemFP40 PubchemFP41 PubchemFP42
## 1             1           0           0           0           0           0
## 2             0           0           0           0           0           0
## 3             1           0           0           0           0           0
## 4             1           0           0           0           0           0
## 5             0           0           0           0           0           0
## 6             1           0           0           0           0           0
##     PubchemFP43 PubchemFP44 PubchemFP45 PubchemFP46 PubchemFP47 PubchemFP48
## 1             0           0           0           0           0           0
## 2             0           0           0           0           0           0
## 3             0           0           0           0           0           0
## 4             0           0           0           0           0           0
## 5             0           0           0           0           0           0
## 6             0           0           0           0           0           0
##     PubchemFP49 PubchemFP50 PubchemFP51 PubchemFP52 PubchemFP53 PubchemFP54
## 1             0           0           0           0           0           0
## 2             0           0           0           0           0           0
## 3             0           0           0           0           0           0
## 4             0           0           0           0           0           0
## 5             0           0           0           0           0           0
## 6             0           0           0           0           0           0
##     PubchemFP55 PubchemFP56 PubchemFP57 PubchemFP58 PubchemFP59 PubchemFP60
## 1             0           0           0           0           0           0
## 2             0           0           0           0           0           0
## 3             0           0           0           0           0           0
## 4             0           0           0           0           0           0
## 5             0           0           0           0           0           0
## 6             0           0           0           0           0           0
##     PubchemFP61 PubchemFP62 PubchemFP63 PubchemFP64 PubchemFP65 PubchemFP66
## 1             0           0           0           0           0           0
## 2             0           0           0           0           0           0
## 3             0           0           0           0           0           0
## 4             0           0           0           0           0           0
## 5             0           0           0           0           0           0
## 6             0           0           0           0           0           0
##     PubchemFP67 PubchemFP68 PubchemFP69 PubchemFP70 PubchemFP71 PubchemFP72
## 1             0           0           0           0           0           0
## 2             0           0           0           0           0           0
## 3             0           0           0           0           0           0
## 4             0           0           0           0           0           0
## 5             0           0           0           0           0           0
## 6             0           0           0           0           0           0
##     PubchemFP73 PubchemFP74 PubchemFP75 PubchemFP76 PubchemFP77 PubchemFP78
## 1             0           0           0           0           0           0
## 2             0           0           0           0           0           0
## 3             0           0           0           0           0           0
## 4             0           0           0           0           0           0
## 5             0           0           0           0           0           0
## 6             0           0           0           0           0           0
##     PubchemFP79 PubchemFP80 PubchemFP81 PubchemFP82 PubchemFP83 PubchemFP84
```

```
## 1              0           0           0           0           0           0
## 2              0           0           0           0           0           0
## 3              0           0           0           0           0           0
## 4              0           0           0           0           0           0
## 5              0           0           0           0           0           0
## 6              0           0           0           0           0           0
##     PubchemFP85 PubchemFP86 PubchemFP87 PubchemFP88 PubchemFP89 PubchemFP90
## 1              0           0           0           0           0           0
## 2              0           0           0           0           0           0
## 3              0           0           0           0           0           0
## 4              0           0           0           0           0           0
## 5              0           0           0           0           0           0
## 6              0           0           0           0           0           0
##     PubchemFP91 PubchemFP92 PubchemFP93 PubchemFP94 PubchemFP95 PubchemFP96
## 1              0           0           0           0           0           0
## 2              0           0           0           0           0           0
## 3              0           0           0           0           0           0
## 4              0           0           0           0           0           0
## 5              0           0           0           0           0           0
## 6              0           0           0           0           0           0
##     PubchemFP97 PubchemFP98 PubchemFP99 PubchemFP100 PubchemFP101 PubchemFP102
## 1              0           0           0            0            0            0
## 2              0           0           0            0            0            0
## 3              0           0           0            0            0            0
## 4              0           0           0            0            0            0
## 5              0           0           0            0            0            0
## 6              0           0           0            0            0            0
##     PubchemFP103 PubchemFP104 PubchemFP105 PubchemFP106 PubchemFP107 PubchemFP108
## 1              0            0            0            0            0            0
## 2              0            0            0            0            0            0
## 3              0            0            0            0            0            0
## 4              0            0            0            0            0            0
## 5              0            0            0            0            0            0
## 6              0            0            0            0            0            0
##     PubchemFP109 PubchemFP110 PubchemFP111 PubchemFP112 PubchemFP113 PubchemFP114
## 1              0            0            0            0            0            0
## 2              0            0            0            0            0            0
## 3              0            0            0            0            0            0
## 4              0            0            0            0            0            0
## 5              0            0            0            0            0            0
## 6              0            0            0            0            0            0
##     PubchemFP115 PubchemFP116 PubchemFP117 PubchemFP118 PubchemFP119 PubchemFP120
## 1              1            1            0            0            0            0
## 2              0            0            0            0            0            0
## 3              0            0            0            0            0            0
## 4              0            0            0            0            0            0
## 5              0            0            0            0            0            0
## 6              0            0            0            0            0            0
##     PubchemFP121 PubchemFP122 PubchemFP123 PubchemFP124 PubchemFP125 PubchemFP126
## 1              0            0            0            0            0            0
## 2              0            0            0            0            0            0
## 3              0            0            0            0            0            0
## 4              0            0            0            0            0            0
## 5              0            0            0            0            0            0
```

```
## 6               0               0               0               0               0               0
##    PubchemFP127 PubchemFP128 PubchemFP129 PubchemFP130 PubchemFP131 PubchemFP132
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP133 PubchemFP134 PubchemFP135 PubchemFP136 PubchemFP137 PubchemFP138
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP139 PubchemFP140 PubchemFP141 PubchemFP142 PubchemFP143 PubchemFP144
## 1             0            0            0            0            1            0
## 2             0            0            0            0            1            0
## 3             0            0            0            0            1            0
## 4             0            0            0            0            1            0
## 5             0            0            0            0            1            0
## 6             0            0            0            0            1            0
##    PubchemFP145 PubchemFP146 PubchemFP147 PubchemFP148 PubchemFP149 PubchemFP150
## 1             1            1            0            0            0            0
## 2             1            1            0            0            0            0
## 3             1            1            0            0            0            0
## 4             1            1            0            0            0            0
## 5             1            1            0            0            0            0
## 6             1            1            0            0            0            0
##    PubchemFP151 PubchemFP152 PubchemFP153 PubchemFP154 PubchemFP155 PubchemFP156
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP157 PubchemFP158 PubchemFP159 PubchemFP160 PubchemFP161 PubchemFP162
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP163 PubchemFP164 PubchemFP165 PubchemFP166 PubchemFP167 PubchemFP168
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP169 PubchemFP170 PubchemFP171 PubchemFP172 PubchemFP173 PubchemFP174
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
```

```
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##    PubchemFP175 PubchemFP176 PubchemFP177 PubchemFP178 PubchemFP179 PubchemFP180
## 1              0              0              0              1              1              1
## 2              0              0              0              1              1              0
## 3              0              0              0              1              1              1
## 4              0              0              0              1              1              0
## 5              0              0              0              1              1              0
## 6              0              0              0              1              1              0
##    PubchemFP181 PubchemFP182 PubchemFP183 PubchemFP184 PubchemFP185 PubchemFP186
## 1              1              0              0              0              1              0
## 2              0              0              0              0              1              1
## 3              1              0              0              0              1              0
## 4              0              0              0              0              1              1
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##    PubchemFP187 PubchemFP188 PubchemFP189 PubchemFP190 PubchemFP191 PubchemFP192
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##    PubchemFP193 PubchemFP194 PubchemFP195 PubchemFP196 PubchemFP197 PubchemFP198
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##    PubchemFP199 PubchemFP200 PubchemFP201 PubchemFP202 PubchemFP203 PubchemFP204
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##    PubchemFP205 PubchemFP206 PubchemFP207 PubchemFP208 PubchemFP209 PubchemFP210
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##    PubchemFP211 PubchemFP212 PubchemFP213 PubchemFP214 PubchemFP215 PubchemFP216
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##    PubchemFP217 PubchemFP218 PubchemFP219 PubchemFP220 PubchemFP221 PubchemFP222
## 1              0              0              0              0              0              0
```

```
##   2              0           0           0           0           0           0
## 3              0           0           0           0           0           0
## 4              0           0           0           0           0           0
## 5              0           0           0           0           0           0
## 6              0           0           0           0           0           0
##    PubchemFP223 PubchemFP224 PubchemFP225 PubchemFP226 PubchemFP227 PubchemFP228
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP229 PubchemFP230 PubchemFP231 PubchemFP232 PubchemFP233 PubchemFP234
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP235 PubchemFP236 PubchemFP237 PubchemFP238 PubchemFP239 PubchemFP240
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP241 PubchemFP242 PubchemFP243 PubchemFP244 PubchemFP245 PubchemFP246
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP247 PubchemFP248 PubchemFP249 PubchemFP250 PubchemFP251 PubchemFP252
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP253 PubchemFP254 PubchemFP255 PubchemFP256 PubchemFP257 PubchemFP258
## 1             0            0            1            1            1            0
## 2             0            0            1            1            1            0
## 3             0            0            1            1            1            0
## 4             0            0            1            1            1            0
## 5             0            0            1            1            1            0
## 6             0            0            1            1            1            0
##    PubchemFP259 PubchemFP260 PubchemFP261 PubchemFP262 PubchemFP263 PubchemFP264
## 1             0            0            0            0            0            0
## 2             1            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             1            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
```

```
##      PubchemFP265 PubchemFP266 PubchemFP267 PubchemFP268 PubchemFP269 PubchemFP270
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP271 PubchemFP272 PubchemFP273 PubchemFP274 PubchemFP275 PubchemFP276
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP277 PubchemFP278 PubchemFP279 PubchemFP280 PubchemFP281 PubchemFP282
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP283 PubchemFP284 PubchemFP285 PubchemFP286 PubchemFP287 PubchemFP288
## 1               1            1            1            1            0            0
## 2               1            1            1            1            0            0
## 3               1            1            1            1            1            0
## 4               1            1            1            1            1            0
## 5               1            1            1            1            1            0
## 6               1            1            1            1            0            0
##      PubchemFP289 PubchemFP290 PubchemFP291 PubchemFP292 PubchemFP293 PubchemFP294
## 1               0            0            0            0            1            1
## 2               0            0            0            0            0            0
## 3               0            0            0            0            1            1
## 4               0            0            0            0            1            1
## 5               0            0            0            0            1            0
## 6               0            0            0            0            1            1
##      PubchemFP295 PubchemFP296 PubchemFP297 PubchemFP298 PubchemFP299 PubchemFP300
## 1               0            0            0            0            0            1
## 2               0            0            0            0            0            1
## 3               0            0            0            0            0            1
## 4               0            0            0            0            0            1
## 5               0            0            0            0            0            1
## 6               0            0            0            0            0            1
##      PubchemFP301 PubchemFP302 PubchemFP303 PubchemFP304 PubchemFP305 PubchemFP306
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP307 PubchemFP308 PubchemFP309 PubchemFP310 PubchemFP311 PubchemFP312
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
```

```
## 5               0          0          0          0          0          0
## 6               0          0          0          0          0          0
##    PubchemFP313 PubchemFP314 PubchemFP315 PubchemFP316 PubchemFP317 PubchemFP318
## 1             0          0          0          0          0          0
## 2             0          0          0          0          0          0
## 3             0          0          0          0          0          0
## 4             0          0          0          0          0          0
## 5             0          0          0          0          0          0
## 6             0          0          0          0          0          0
##    PubchemFP319 PubchemFP320 PubchemFP321 PubchemFP322 PubchemFP323 PubchemFP324
## 1             0          0          0          0          0          0
## 2             0          0          0          0          0          0
## 3             0          0          0          0          0          0
## 4             0          0          0          0          0          0
## 5             0          0          0          0          0          0
## 6             0          0          0          0          0          0
##    PubchemFP325 PubchemFP326 PubchemFP327 PubchemFP328 PubchemFP329 PubchemFP330
## 1             0          0          0          0          0          0
## 2             0          0          0          0          0          0
## 3             0          0          0          0          0          0
## 4             0          0          0          0          0          0
## 5             0          0          0          0          0          0
## 6             0          0          0          0          0          0
##    PubchemFP331 PubchemFP332 PubchemFP333 PubchemFP334 PubchemFP335 PubchemFP336
## 1             0          1          1          0          1          0
## 2             0          1          1          0          0          0
## 3             0          1          1          0          0          0
## 4             0          1          1          0          0          0
## 5             0          1          1          0          0          0
## 6             0          1          1          0          0          0
##    PubchemFP337 PubchemFP338 PubchemFP339 PubchemFP340 PubchemFP341 PubchemFP342
## 1             0          0          0          0          0          1
## 2             0          0          0          1          1          0
## 3             0          0          0          0          0          1
## 4             0          0          0          1          0          1
## 5             0          0          0          0          1          0
## 6             0          0          0          0          0          1
##    PubchemFP343 PubchemFP344 PubchemFP345 PubchemFP346 PubchemFP347 PubchemFP348
## 1             0          1          1          0          0          0
## 2             0          1          0          1          0          0
## 3             0          1          1          0          0          0
## 4             0          1          0          0          0          0
## 5             0          1          0          0          0          0
## 6             0          1          0          0          0          0
##    PubchemFP349 PubchemFP350 PubchemFP351 PubchemFP352 PubchemFP353 PubchemFP354
## 1             1          0          1          0          1          0
## 2             0          0          1          1          0          0
## 3             1          0          1          0          1          0
## 4             1          0          1          0          1          0
## 5             0          0          1          1          0          0
## 6             0          0          1          0          0          0
##    PubchemFP355 PubchemFP356 PubchemFP357 PubchemFP358 PubchemFP359 PubchemFP360
## 1             1          1          0          1          1          0
## 2             1          1          0          0          0          0
```

8

```
## 3            1            1            0            1            1            0
## 4            1            1            0            1            1            0
## 5            1            1            0            1            1            0
## 6            1            1            0            1            1            0
##   PubchemFP361 PubchemFP362 PubchemFP363 PubchemFP364 PubchemFP365 PubchemFP366
## 1            0            1            0            0            1            0
## 2            0            0            0            0            0            1
## 3            0            1            1            0            1            0
## 4            0            1            1            0            1            0
## 5            0            0            1            0            1            0
## 6            0            1            0            0            1            0
##   PubchemFP367 PubchemFP368 PubchemFP369 PubchemFP370 PubchemFP371 PubchemFP372
## 1            0            1            0            1            1            0
## 2            0            0            0            1            1            0
## 3            0            1            0            1            1            0
## 4            0            1            0            1            1            0
## 5            0            1            0            1            1            0
## 6            0            1            0            1            1            0
##   PubchemFP373 PubchemFP374 PubchemFP375 PubchemFP376 PubchemFP377 PubchemFP378
## 1            0            0            1            1            0            1
## 2            0            1            0            1            1            0
## 3            0            0            1            1            0            1
## 4            0            1            1            1            1            1
## 5            0            1            1            1            0            1
## 6            0            1            1            1            0            1
##   PubchemFP379 PubchemFP380 PubchemFP381 PubchemFP382 PubchemFP383 PubchemFP384
## 1            1            0            0            0            0            1
## 2            0            1            1            1            0            1
## 3            1            0            0            0            0            1
## 4            1            0            0            0            0            1
## 5            1            0            1            1            0            1
## 6            1            0            0            0            0            1
##   PubchemFP385 PubchemFP386 PubchemFP387 PubchemFP388 PubchemFP389 PubchemFP390
## 1            1            0            1            1            1            1
## 2            0            0            0            0            0            1
## 3            1            0            1            1            1            1
## 4            1            0            1            1            1            1
## 5            1            0            1            1            1            1
## 6            1            0            1            1            1            1
##   PubchemFP391 PubchemFP392 PubchemFP393 PubchemFP394 PubchemFP395 PubchemFP396
## 1            1            0            0            0            0            1
## 2            0            0            0            0            0            1
## 3            1            0            0            0            0            1
## 4            1            0            0            0            0            1
## 5            1            0            0            0            0            1
## 6            1            0            0            0            0            1
##   PubchemFP397 PubchemFP398 PubchemFP399 PubchemFP400 PubchemFP401 PubchemFP402
## 1            0            0            0            0            0            0
## 2            0            0            0            0            0            0
## 3            0            0            0            0            0            0
## 4            0            0            0            0            0            0
## 5            0            0            0            0            0            0
## 6            0            0            0            0            0            0
##   PubchemFP403 PubchemFP404 PubchemFP405 PubchemFP406 PubchemFP407 PubchemFP408
```

```
## 1              1             0             0             0             0             0
## 2              0             0             1             0             0             0
## 3              1             0             0             0             0             0
## 4              1             0             0             0             0             0
## 5              1             0             1             0             0             0
## 6              1             0             0             0             0             0
##     PubchemFP409 PubchemFP410 PubchemFP411 PubchemFP412 PubchemFP413 PubchemFP414
## 1              0             0             0             1             0             0
## 2              0             0             0             0             0             0
## 3              0             0             0             1             0             0
## 4              0             0             0             1             0             0
## 5              0             0             0             1             0             0
## 6              0             0             0             1             0             0
##     PubchemFP415 PubchemFP416 PubchemFP417 PubchemFP418 PubchemFP419 PubchemFP420
## 1              0             1             0             1             0             1
## 2              0             1             0             1             0             1
## 3              0             1             0             1             0             1
## 4              0             1             0             1             0             1
## 5              0             1             0             1             0             1
## 6              0             1             0             1             0             1
##     PubchemFP421 PubchemFP422 PubchemFP423 PubchemFP424 PubchemFP425 PubchemFP426
## 1              0             1             0             0             0             0
## 2              0             1             0             0             0             0
## 3              0             1             0             0             0             0
## 4              0             1             0             0             0             0
## 5              0             1             0             0             0             0
## 6              0             1             0             0             0             0
##     PubchemFP427 PubchemFP428 PubchemFP429 PubchemFP430 PubchemFP431 PubchemFP432
## 1              0             0             0             1             0             0
## 2              0             0             0             1             1             0
## 3              0             0             0             1             0             0
## 4              0             0             0             1             0             0
## 5              0             0             0             1             0             0
## 6              0             0             0             1             0             0
##     PubchemFP433 PubchemFP434 PubchemFP435 PubchemFP436 PubchemFP437 PubchemFP438
## 1              0             1             0             0             0             1
## 2              0             1             0             0             1             0
## 3              0             1             0             0             0             1
## 4              0             1             0             0             1             1
## 5              0             1             0             0             0             1
## 6              0             1             0             0             0             1
##     PubchemFP439 PubchemFP440 PubchemFP441 PubchemFP442 PubchemFP443 PubchemFP444
## 1              0             0             1             1             0             0
## 2              0             0             1             1             0             0
## 3              0             0             1             1             0             0
## 4              0             0             1             1             0             0
## 5              0             0             1             1             0             0
## 6              0             0             1             1             0             0
##     PubchemFP445 PubchemFP446 PubchemFP447 PubchemFP448 PubchemFP449 PubchemFP450
## 1              0             1             0             0             1             1
## 2              0             1             0             0             1             0
## 3              0             1             0             0             1             1
## 4              0             1             0             0             1             1
## 5              0             1             0             0             1             1
```

```
## 6               0               1               0               0               1               1
##    PubchemFP451 PubchemFP452 PubchemFP453 PubchemFP454 PubchemFP455 PubchemFP456
## 1             1             0             1             0             0             0
## 2             1             1             1             0             0             0
## 3             1             0             1             0             0             0
## 4             1             0             1             0             0             0
## 5             1             0             1             0             0             0
## 6             1             0             1             0             0             0
##    PubchemFP457 PubchemFP458 PubchemFP459 PubchemFP460 PubchemFP461 PubchemFP462
## 1             0             0             0             0             0             0
## 2             0             0             0             0             0             0
## 3             0             0             0             0             0             0
## 4             0             0             0             0             0             0
## 5             0             0             0             0             0             0
## 6             0             0             0             0             0             0
##    PubchemFP463 PubchemFP464 PubchemFP465 PubchemFP466 PubchemFP467 PubchemFP468
## 1             0             1             0             0             1             0
## 2             0             1             0             0             1             0
## 3             0             1             0             0             1             0
## 4             0             1             0             0             1             0
## 5             0             1             0             0             1             0
## 6             0             1             0             0             1             0
##    PubchemFP469 PubchemFP470 PubchemFP471 PubchemFP472 PubchemFP473 PubchemFP474
## 1             0             1             0             1             1             0
## 2             0             1             0             1             0             0
## 3             0             1             0             1             1             0
## 4             0             1             0             1             1             0
## 5             0             1             0             1             1             0
## 6             0             1             0             1             1             0
##    PubchemFP475 PubchemFP476 PubchemFP477 PubchemFP478 PubchemFP479 PubchemFP480
## 1             1             0             1             0             0             0
## 2             0             1             1             0             0             0
## 3             1             0             1             0             0             0
## 4             1             0             1             0             0             0
## 5             1             1             1             0             0             0
## 6             1             0             1             0             0             0
##    PubchemFP481 PubchemFP482 PubchemFP483 PubchemFP484 PubchemFP485 PubchemFP486
## 1             0             1             0             1             0             1
## 2             0             1             0             0             0             0
## 3             0             1             0             1             0             1
## 4             0             1             0             1             0             1
## 5             0             1             0             1             0             1
## 6             0             1             0             1             0             1
##    PubchemFP487 PubchemFP488 PubchemFP489 PubchemFP490 PubchemFP491 PubchemFP492
## 1             1             0             1             1             0             0
## 2             0             0             0             1             0             1
## 3             1             0             1             1             0             0
## 4             1             0             1             1             0             0
## 5             1             0             1             1             0             0
## 6             1             0             1             1             0             0
##    PubchemFP493 PubchemFP494 PubchemFP495 PubchemFP496 PubchemFP497 PubchemFP498
## 1             0             0             1             0             0             0
## 2             0             0             1             0             0             1
## 3             0             0             1             0             0             0
```

```
## 4            0            0            1            0            0            0
## 5            0            0            1            0            0            1
## 6            0            0            1            0            0            0
##   PubchemFP499 PubchemFP500 PubchemFP501 PubchemFP502 PubchemFP503 PubchemFP504
## 1            0            0            1            0            1            1
## 2            0            0            0            1            0            0
## 3            0            0            1            0            1            1
## 4            0            0            1            1            1            1
## 5            0            0            0            0            0            1
## 6            0            0            1            0            1            1
##   PubchemFP505 PubchemFP506 PubchemFP507 PubchemFP508 PubchemFP509 PubchemFP510
## 1            0            1            1            0            0            0
## 2            0            1            0            0            0            0
## 3            0            1            1            0            0            0
## 4            0            1            1            0            0            0
## 5            0            1            0            0            0            0
## 6            0            1            0            0            0            0
##   PubchemFP511 PubchemFP512 PubchemFP513 PubchemFP514 PubchemFP515 PubchemFP516
## 1            0            0            0            0            1            1
## 2            0            0            0            0            1            1
## 3            0            0            0            0            1            1
## 4            0            0            0            0            1            1
## 5            0            0            0            0            1            1
## 6            0            0            0            0            1            1
##   PubchemFP517 PubchemFP518 PubchemFP519 PubchemFP520 PubchemFP521 PubchemFP522
## 1            1            1            1            1            0            0
## 2            0            1            0            1            0            0
## 3            1            1            1            1            0            0
## 4            1            1            1            1            0            0
## 5            1            1            1            1            0            0
## 6            1            1            1            1            0            0
##   PubchemFP523 PubchemFP524 PubchemFP525 PubchemFP526 PubchemFP527 PubchemFP528
## 1            1            1            0            0            0            0
## 2            1            1            0            0            0            0
## 3            1            1            0            0            0            0
## 4            1            1            0            0            0            0
## 5            1            1            0            0            0            0
## 6            1            1            0            0            0            0
##   PubchemFP529 PubchemFP530 PubchemFP531 PubchemFP532 PubchemFP533 PubchemFP534
## 1            0            1            0            0            0            0
## 2            0            1            0            0            0            0
## 3            0            1            0            0            0            0
## 4            0            1            0            0            0            0
## 5            0            1            0            0            0            0
## 6            0            1            0            0            0            0
##   PubchemFP535 PubchemFP536 PubchemFP537 PubchemFP538 PubchemFP539 PubchemFP540
## 1            0            0            0            1            0            1
## 2            0            0            0            1            1            0
## 3            0            0            0            1            0            1
## 4            0            0            0            1            0            1
## 5            0            0            0            1            0            1
## 6            0            0            0            1            0            1
##   PubchemFP541 PubchemFP542 PubchemFP543 PubchemFP544 PubchemFP545 PubchemFP546
## 1            0            0            0            0            1            0
```

```
## 2             1             1             0             0             1             1
## 3             0             0             0             0             1             0
## 4             0             0             0             0             1             1
## 5             1             1             0             0             1             0
## 6             0             0             0             0             1             0
##    PubchemFP547 PubchemFP548 PubchemFP549 PubchemFP550 PubchemFP551 PubchemFP552
## 1             0             0             1             1             0             1
## 2             0             1             1             0             0             1
## 3             0             0             1             1             0             1
## 4             0             0             1             1             0             1
## 5             0             1             1             0             0             1
## 6             0             0             1             1             0             1
##    PubchemFP553 PubchemFP554 PubchemFP555 PubchemFP556 PubchemFP557 PubchemFP558
## 1             0             0             1             1             0             0
## 2             0             0             1             1             0             0
## 3             0             0             1             1             0             0
## 4             0             0             1             1             0             0
## 5             0             0             1             1             0             0
## 6             0             0             1             1             0             0
##    PubchemFP559 PubchemFP560 PubchemFP561 PubchemFP562 PubchemFP563 PubchemFP564
## 1             0             1             0             0             0             1
## 2             0             0             0             0             1             1
## 3             0             1             0             0             0             1
## 4             0             1             0             0             0             1
## 5             0             1             0             0             0             1
## 6             0             1             0             0             0             1
##    PubchemFP565 PubchemFP566 PubchemFP567 PubchemFP568 PubchemFP569 PubchemFP570
## 1             0             0             0             0             0             1
## 2             1             0             0             0             0             1
## 3             0             0             0             0             0             1
## 4             0             0             0             0             0             1
## 5             1             0             0             0             0             1
## 6             0             0             0             0             0             1
##    PubchemFP571 PubchemFP572 PubchemFP573 PubchemFP574 PubchemFP575 PubchemFP576
## 1             0             1             0             0             0             1
## 2             0             0             1             1             0             1
## 3             0             1             0             0             0             1
## 4             0             1             0             0             0             1
## 5             0             1             1             1             0             1
## 6             0             1             0             0             0             1
##    PubchemFP577 PubchemFP578 PubchemFP579 PubchemFP580 PubchemFP581 PubchemFP582
## 1             0             1             0             0             0             1
## 2             0             1             0             0             0             1
## 3             0             1             0             0             0             1
## 4             0             1             0             0             0             1
## 5             0             1             0             0             0             1
## 6             0             1             0             0             0             1
##    PubchemFP583 PubchemFP584 PubchemFP585 PubchemFP586 PubchemFP587 PubchemFP588
## 1             0             1             1             1             0             0
## 2             0             1             1             0             0             0
## 3             0             1             1             0             0             0
## 4             0             1             1             0             0             0
## 5             0             1             1             0             0             0
## 6             0             1             1             0             0             0
```

```
##    PubchemFP589 PubchemFP590 PubchemFP591 PubchemFP592 PubchemFP593 PubchemFP594
## 1             0            0            1            1            0            0
## 2             1            0            0            1            0            1
## 3             0            0            1            1            0            0
## 4             0            0            1            1            0            0
## 5             1            0            0            1            0            1
## 6             0            0            1            1            0            0
##    PubchemFP595 PubchemFP596 PubchemFP597 PubchemFP598 PubchemFP599 PubchemFP600
## 1             1            1            0            1            1            1
## 2             1            0            0            0            1            1
## 3             1            1            0            1            1            1
## 4             1            1            0            1            1            1
## 5             1            1            0            0            1            1
## 6             1            1            0            1            1            1
##    PubchemFP601 PubchemFP602 PubchemFP603 PubchemFP604 PubchemFP605 PubchemFP606
## 1             0            0            1            0            0            0
## 2             0            0            1            1            0            1
## 3             0            0            1            0            0            0
## 4             0            0            1            0            0            0
## 5             0            0            1            1            0            1
## 6             0            0            1            0            0            0
##    PubchemFP607 PubchemFP608 PubchemFP609 PubchemFP610 PubchemFP611 PubchemFP612
## 1             1            1            0            0            0            0
## 2             1            1            0            1            0            1
## 3             1            1            0            0            0            0
## 4             1            1            0            0            0            0
## 5             1            1            0            0            0            0
## 6             1            1            0            0            0            0
##    PubchemFP613 PubchemFP614 PubchemFP615 PubchemFP616 PubchemFP617 PubchemFP618
## 1             1            0            0            1            0            1
## 2             1            1            0            0            0            1
## 3             1            0            0            1            0            1
## 4             1            0            0            1            0            1
## 5             1            0            0            1            0            1
## 6             1            0            0            1            0            1
##    PubchemFP619 PubchemFP620 PubchemFP621 PubchemFP622 PubchemFP623 PubchemFP624
## 1             0            0            0            0            0            0
## 2             1            1            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             1            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP625 PubchemFP626 PubchemFP627 PubchemFP628 PubchemFP629 PubchemFP630
## 1             0            0            0            1            0            0
## 2             0            1            0            1            0            0
## 3             0            0            0            1            0            0
## 4             0            0            0            1            0            0
## 5             0            1            0            1            0            0
## 6             0            0            0            1            0            0
##    PubchemFP631 PubchemFP632 PubchemFP633 PubchemFP634 PubchemFP635 PubchemFP636
## 1             0            0            1            1            0            1
## 2             0            1            1            1            0            0
## 3             0            0            1            1            0            1
## 4             0            0            1            1            0            1
```

```
## 5             0              0              1              1              0              1
## 6             0              0              1              1              0              1
##    PubchemFP637 PubchemFP638 PubchemFP639 PubchemFP640 PubchemFP641 PubchemFP642
## 1             0              0              0              1              0              0
## 2             1              1              0              1              1              0
## 3             0              0              0              1              0              0
## 4             0              0              0              1              0              0
## 5             1              0              0              1              1              0
## 6             0              0              0              1              0              0
##    PubchemFP643 PubchemFP644 PubchemFP645 PubchemFP646 PubchemFP647 PubchemFP648
## 1             0              1              1              1              1              0
## 2             0              1              1              0              0              0
## 3             0              1              1              1              1              0
## 4             0              1              1              1              1              0
## 5             0              1              0              1              1              0
## 6             0              1              0              1              1              0
##    PubchemFP649 PubchemFP650 PubchemFP651 PubchemFP652 PubchemFP653 PubchemFP654
## 1             0              0              0              0              0              1
## 2             0              0              1              1              0              0
## 3             0              0              0              0              0              1
## 4             0              0              0              0              0              1
## 5             0              0              1              0              0              1
## 6             0              0              0              0              0              1
##    PubchemFP655 PubchemFP656 PubchemFP657 PubchemFP658 PubchemFP659 PubchemFP660
## 1             0              1              1              0              0              1
## 2             1              0              1              0              0              1
## 3             0              1              1              0              0              1
## 4             0              0              1              0              0              1
## 5             1              0              1              0              0              1
## 6             0              0              1              0              0              1
##    PubchemFP661 PubchemFP662 PubchemFP663 PubchemFP664 PubchemFP665 PubchemFP666
## 1             0              0              0              1              1              0
## 2             1              0              0              1              1              1
## 3             0              0              0              1              1              0
## 4             0              0              0              1              1              0
## 5             0              0              0              1              1              1
## 6             0              0              0              1              1              0
##    PubchemFP667 PubchemFP668 PubchemFP669 PubchemFP670 PubchemFP671 PubchemFP672
## 1             0              1              0              0              0              0
## 2             0              1              0              0              0              0
## 3             0              1              0              0              0              0
## 4             0              1              0              0              0              0
## 5             0              1              0              0              0              0
## 6             0              1              0              0              0              0
##    PubchemFP673 PubchemFP674 PubchemFP675 PubchemFP676 PubchemFP677 PubchemFP678
## 1             0              1              0              0              1              1
## 2             0              0              0              0              1              1
## 3             0              1              0              0              1              1
## 4             0              1              0              0              1              1
## 5             0              1              0              0              1              1
## 6             0              1              0              0              1              1
##    PubchemFP679 PubchemFP680 PubchemFP681 PubchemFP682 PubchemFP683 PubchemFP684
## 1             1              0              0              0              1              0
## 2             1              1              0              0              1              0
```

```
## 3            1           0           0           0           1           0
## 4            1           0           0           0           1           0
## 5            1           1           0           0           1           0
## 6            1           0           0           0           1           0
##   PubchemFP685 PubchemFP686 PubchemFP687 PubchemFP688 PubchemFP689 PubchemFP690
## 1            0           0           0           1           0           0
## 2            0           0           0           1           1           0
## 3            0           0           0           1           0           0
## 4            0           0           0           1           0           0
## 5            0           0           0           1           1           0
## 6            0           0           0           1           0           0
##   PubchemFP691 PubchemFP692 PubchemFP693 PubchemFP694 PubchemFP695 PubchemFP696
## 1            0           0           0           0           0           0
## 2            1           0           0           0           0           0
## 3            0           0           0           0           0           0
## 4            0           0           0           0           0           0
## 5            1           0           0           0           0           0
## 6            0           0           0           0           0           0
##   PubchemFP697 PubchemFP698 PubchemFP699 PubchemFP700 PubchemFP701 PubchemFP702
## 1            0           0           0           0           0           0
## 2            0           1           0           0           0           1
## 3            0           0           0           0           0           0
## 4            0           0           0           0           0           0
## 5            0           0           0           0           0           1
## 6            0           0           0           0           0           0
##   PubchemFP703 PubchemFP704 PubchemFP705 PubchemFP706 PubchemFP707 PubchemFP708
## 1            0           0           0           0           0           1
## 2            1           0           0           0           0           1
## 3            0           0           0           0           0           1
## 4            0           0           0           0           0           1
## 5            0           0           0           0           0           1
## 6            0           0           0           0           0           1
##   PubchemFP709 PubchemFP710 PubchemFP711 PubchemFP712 PubchemFP713 PubchemFP714
## 1            1           1           0           0           0           0
## 2            1           1           0           0           0           0
## 3            1           1           0           0           0           0
## 4            1           1           0           0           0           0
## 5            1           1           0           0           0           1
## 6            1           1           0           0           0           0
##   PubchemFP715 PubchemFP716 PubchemFP717 PubchemFP718 PubchemFP719 PubchemFP720
## 1            0           0           1           0           0           0
## 2            0           0           0           0           0           0
## 3            0           0           1           0           0           0
## 4            0           0           1           0           0           0
## 5            0           0           0           0           0           0
## 6            0           0           1           0           0           0
##   PubchemFP721 PubchemFP722 PubchemFP723 PubchemFP724 PubchemFP725 PubchemFP726
## 1            0           0           0           0           0           0
## 2            0           0           0           0           0           0
## 3            0           0           0           0           0           0
## 4            0           0           0           0           0           0
## 5            0           0           0           0           0           0
## 6            0           0           0           0           0           0
##   PubchemFP727 PubchemFP728 PubchemFP729 PubchemFP730 PubchemFP731 PubchemFP732
```

```
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##   PubchemFP733 PubchemFP734 PubchemFP735 PubchemFP736 PubchemFP737 PubchemFP738
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##   PubchemFP739 PubchemFP740 PubchemFP741 PubchemFP742 PubchemFP743 PubchemFP744
## 1              0              0              0              0              0              0
## 2              0              0              0              1              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##   PubchemFP745 PubchemFP746 PubchemFP747 PubchemFP748 PubchemFP749 PubchemFP750
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##   PubchemFP751 PubchemFP752 PubchemFP753 PubchemFP754 PubchemFP755 PubchemFP756
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##   PubchemFP757 PubchemFP758 PubchemFP759 PubchemFP760 PubchemFP761 PubchemFP762
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##   PubchemFP763 PubchemFP764 PubchemFP765 PubchemFP766 PubchemFP767 PubchemFP768
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
## 6              0              0              0              0              0              0
##   PubchemFP769 PubchemFP770 PubchemFP771 PubchemFP772 PubchemFP773 PubchemFP774
## 1              0              0              0              0              0              0
## 2              0              0              0              0              0              0
## 3              0              0              0              0              0              0
## 4              0              0              0              0              0              0
## 5              0              0              0              0              0              0
```

```
## 6               0            0            0            0            0            0
##      PubchemFP775 PubchemFP776 PubchemFP777 PubchemFP778 PubchemFP779 PubchemFP780
## 1               0            0            0            0            0            1
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            1
## 4               0            0            0            0            0            1
## 5               0            0            1            0            0            0
## 6               0            0            0            0            0            1
##      PubchemFP781 PubchemFP782 PubchemFP783 PubchemFP784 PubchemFP785 PubchemFP786
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP787 PubchemFP788 PubchemFP789 PubchemFP790 PubchemFP791 PubchemFP792
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP793 PubchemFP794 PubchemFP795 PubchemFP796 PubchemFP797 PubchemFP798
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP799 PubchemFP800 PubchemFP801 PubchemFP802 PubchemFP803 PubchemFP804
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP805 PubchemFP806 PubchemFP807 PubchemFP808 PubchemFP809 PubchemFP810
## 1               0            0            0            0            0            0
## 2               1            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP811 PubchemFP812 PubchemFP813 PubchemFP814 PubchemFP815 PubchemFP816
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##      PubchemFP817 PubchemFP818 PubchemFP819 PubchemFP820 PubchemFP821 PubchemFP822
## 1               0            0            0            0            0            0
## 2               0            0            0            0            0            0
## 3               0            0            0            0            0            0
```

```
## 4               0            0            0            0            0            0
## 5               0            0            0            0            0            0
## 6               0            0            0            0            0            0
##    PubchemFP823 PubchemFP824 PubchemFP825 PubchemFP826 PubchemFP827 PubchemFP828
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP829 PubchemFP830 PubchemFP831 PubchemFP832 PubchemFP833 PubchemFP834
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP835 PubchemFP836 PubchemFP837 PubchemFP838 PubchemFP839 PubchemFP840
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP841 PubchemFP842 PubchemFP843 PubchemFP844 PubchemFP845 PubchemFP846
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP847 PubchemFP848 PubchemFP849 PubchemFP850 PubchemFP851 PubchemFP852
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP853 PubchemFP854 PubchemFP855 PubchemFP856 PubchemFP857 PubchemFP858
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP859 PubchemFP860 PubchemFP861 PubchemFP862 PubchemFP863 PubchemFP864
## 1             0            0            0            0            0            0
## 2             0            0            0            0            0            0
## 3             0            0            0            0            0            0
## 4             0            0            0            0            0            0
## 5             0            0            0            0            0            0
## 6             0            0            0            0            0            0
##    PubchemFP865 PubchemFP866 PubchemFP867 PubchemFP868 PubchemFP869 PubchemFP870
## 1             0            0            0            0            0            0
```

```
## 2              0           0           0           0           0           0
## 3              0           0           0           0           0           0
## 4              0           0           0           0           0           0
## 5              0           0           0           0           0           0
## 6              0           0           0           0           0           0
##    PubchemFP871 PubchemFP872 PubchemFP873 PubchemFP874 PubchemFP875 PubchemFP876
## 1              0           0           0           0           0           0
## 2              0           0           0           0           0           0
## 3              0           0           0           0           0           0
## 4              0           0           0           0           0           0
## 5              0           0           0           0           0           0
## 6              0           0           0           0           0           0
##    PubchemFP877 PubchemFP878 PubchemFP879 PubchemFP880     pIC50
## 1              0           0           0           0 6.124939
## 2              0           0           0           0 7.000000
## 3              0           0           0           0 4.301030
## 4              0           0           0           0 6.522879
## 5              0           0           0           0 6.096910
## 6              0           0           0           0 5.619789
```

```r
dim(df_Ach)
```

```
## [1] 4695  882
```

```r
X <- df_Ach[, !(names(df_Ach) %in% c("pIC50"))]
Y <- df_Ach$pIC50
```

```r
library(caret)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: lattice
```

```r
# Remove low variance features
nzv <- nearZeroVar(X, saveMetrics = TRUE)
X <- X[, !nzv$nzv]
```

```r
dim(X)
```

```
## [1] 4695  323
```

```r
library(caret)

# Set random seed for reproducibility
set.seed(42)

# Create data partitions
trainIndex <- createDataPartition(Y, p = 0.8, list = FALSE)
X_train <- X[trainIndex, ]
X_test <- X[-trainIndex, ]
Y_train <- Y[trainIndex]
Y_test <- Y[-trainIndex]
```

```r
list.of.packages <- c("nnls", "quadprog", "SuperLearner", "ggplot2", "raster", "sp", "rgdal", "rgeos",
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages, repos = "http://cran.us.r-project.org")

invisible(lapply(list.of.packages, library, character.only = TRUE))
```

```
## Loading required package: gam

## Loading required package: splines

## Loading required package: foreach

## Loaded gam 1.22-2

## Super Learner

## Version: 2.0-28

## Package created on 2021-05-04

## Loading required package: sp

## Please note that rgdal will be retired during 2023,
## plan transition to sf/stars/terra functions using GDAL and PROJ
## at your earliest convenience.
## See https://r-spatial.org/r/2022/04/12/evolution.html and https://github.com/r-spatial/evolution
## rgdal: version: 1.6-5, (SVN revision 1199)
## Geospatial Data Abstraction Library extensions to R successfully loaded
## Loaded GDAL runtime: GDAL 3.5.3, released 2022/10/21
## Path to GDAL shared files: /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/library/rgdal
##  GDAL does not use iconv for recoding strings.
## GDAL binary built with GEOS: TRUE
## Loaded PROJ runtime: Rel. 9.1.0, September 1st, 2022, [PJ_VERSION: 910]
## Path to PROJ shared files: /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/library/rgdal
## PROJ CDN enabled: FALSE
## Linking to sp version:1.6-0
## To mute warnings of possible GDAL/OSR exportToProj4() degradation,
## use options("rgdal_show_exportToProj4_warnings"="none") before loading sp or rgdal.

## rgeos version: 0.6-2, (SVN revision 693)
##  GEOS runtime version: 3.11.0-CAPI-1.17.0
##  Please note that rgeos will be retired during 2023,
## plan transition to sf functions using GEOS at your earliest convenience.
##  GEOS using OverlayNG
##  Linking to sp version: 1.6-0
##  Polygon checking: TRUE

## Loading required package: Matrix

## Loaded glmnet 4.1-7
```

```
## randomForest 4.7-1.1
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:ggplot2':
##
##     margin
```

```r
mtry = create.Learner("SL.randomForest", tune = list(mtry = c(2,3)), name_prefix="mtry")
sl_libraries <- mtry$names
sl_mtry <- SuperLearner(Y = Y_train,
                        X = X_train,
                        family = gaussian(),
                        SL.library = sl_libraries,
                        cvControl = list(V=5))
sl_mtry
```

```
##
## Call:
## SuperLearner(Y = Y_train, X = X_train, family = gaussian(), SL.library = sl_libraries,
##     cvControl = list(V = 5))
##
##
##                  Risk Coef
## mtry_1_All 1.562131    0
## mtry_2_All 1.375657    1
```

Predicted values

```r
pred <- predict(sl_mtry, X_test, onlySL = T)
```

```r
cv_sl_2 = CV.SuperLearner(Y = Y_train,
                   X = X_train, family = gaussian(), SL.library = c("SL.lm", "SL.kernelKnn"),
                   cvControl = list(V=5))
```

```
## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading
```

```
## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading
```

```
## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading
```

```
## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading
```

```
## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
```

```
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
```

```
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
```

```
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(fit, newdata = newX, type = "response"): prediction from
## a rank-deficient fit may be misleading
```

```
summary(cv_sl_2)
```

```
##
## Call:
## CV.SuperLearner(Y = Y_train, X = X_train, family = gaussian(), SL.library = c("SL.lm",
##      "SL.kernelKnn"), cvControl = list(V = 5))
##
## Risk is based on: Mean Squared Error
##
## All risk estimates are based on V =  5
```

```
##
##            Algorithm    Ave       se     Min     Max
##       Super Learner 1.3069 0.042570 1.2458 1.4044
##         Discrete SL 1.3688 0.045212 1.2946 1.4373
##           SL.lm_All 1.5541 0.047724 1.4743 1.6801
##   SL.kernelKnn_All 1.3688 0.045212 1.2946 1.4373
```

```
plot(cv_sl_2)
```
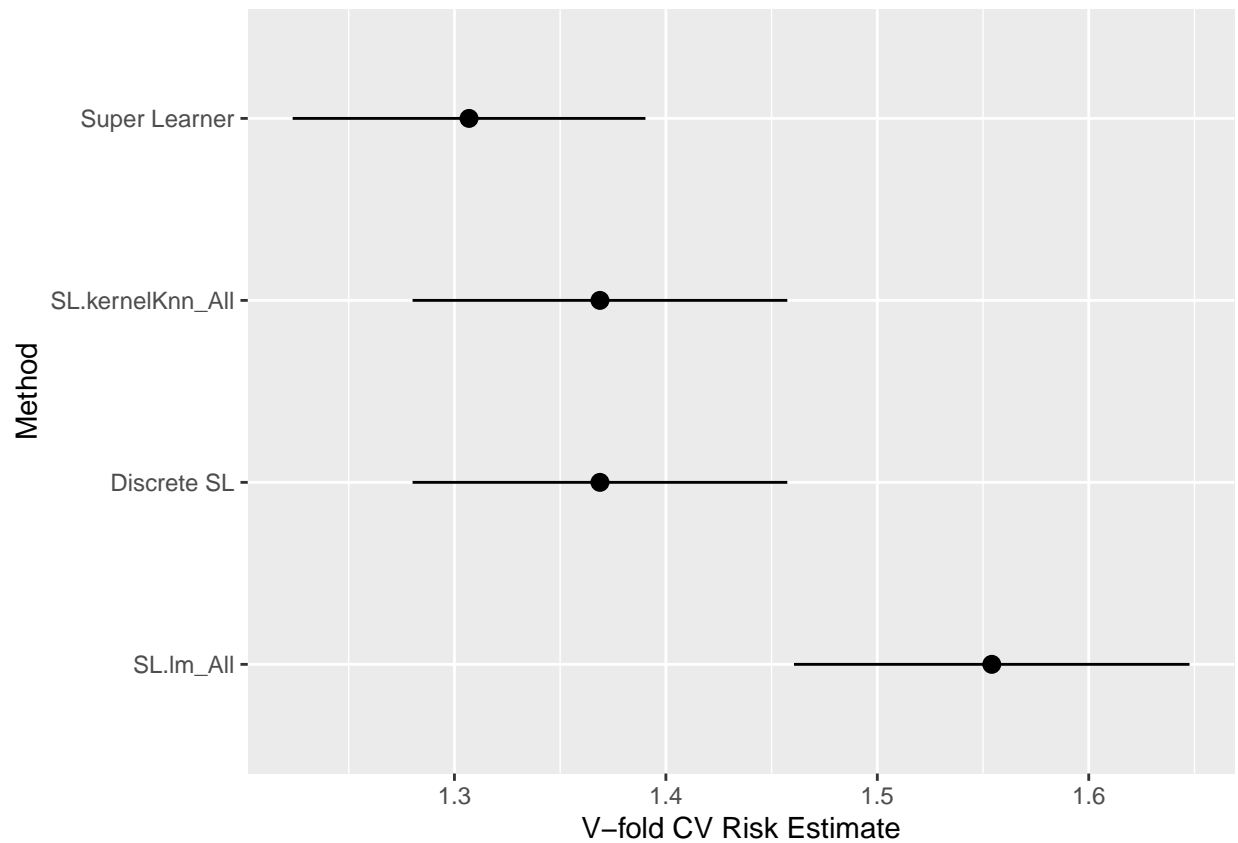


```
ggplot(data.frame(Y_test, pred$pred), aes(pred$pred , Y_test)) + geom_point() + geom_abline(slope=1, in
```
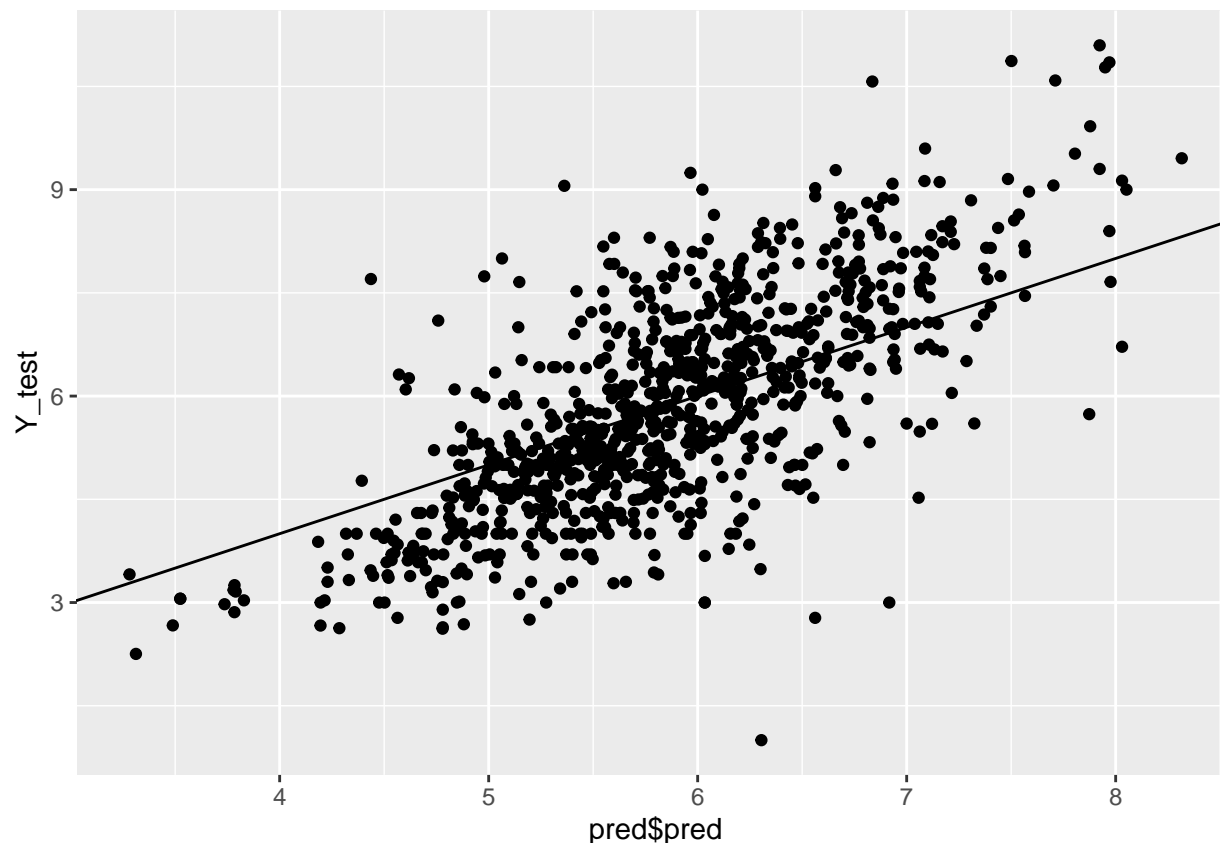
```r
mse = mean((Y_test-pred$pred)^2)
print(paste("Mean squared error:", round(mse, digits=4)))
```

```
## [1] "Mean squared error: 1.2466"
```

```r
print(paste("Correlation:", round(cor(Y_test,pred$pred), digits=4)))
```

```
## [1] "Correlation: 0.7197"
```

```r
library(caret)

# Define your SuperLearner object
sl_libraries <- c("SL.glm", "SL.randomForest", "SL.knn")
sl_mtry1 <- SuperLearner(Y = Y_train, X = X_train, family = gaussian(),
                         SL.library = sl_libraries, cvControl = list(V = 5))
```

```
## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type == :
## prediction from a rank-deficient fit may be misleading
```

```
## Error in (function (Y, X, newX, family, k = 10, ...)  :
##    SL.knn only available for family = binomial()
```

```
## Warning in FUN(X[[i]], ...): Error in algorithm SL.knn
```

```
##    The Algorithm will be removed from the Super Learner (i.e. given weight 0)

## Warning in FUN(X[[i]], ...): prediction from a rank-deficient fit may be misleading

## Error in (function (Y, X, newX, family, k = 10, ...)  :
##    SL.knn only available for family = binomial()

## Warning in FUN(X[[i]], ...): Error in algorithm SL.knn
##    The Algorithm will be removed from the Super Learner (i.e. given weight 0)

## Warning in FUN(X[[i]], ...): prediction from a rank-deficient fit may be misleading

## Error in (function (Y, X, newX, family, k = 10, ...)  :
##    SL.knn only available for family = binomial()

## Warning in FUN(X[[i]], ...): Error in algorithm SL.knn
##    The Algorithm will be removed from the Super Learner (i.e. given weight 0)

## Warning in FUN(X[[i]], ...): prediction from a rank-deficient fit may be misleading

## Error in (function (Y, X, newX, family, k = 10, ...)  :
##    SL.knn only available for family = binomial()

## Warning in FUN(X[[i]], ...): Error in algorithm SL.knn
##    The Algorithm will be removed from the Super Learner (i.e. given weight 0)

## Warning in FUN(X[[i]], ...): prediction from a rank-deficient fit may be misleading

## Error in (function (Y, X, newX, family, k = 10, ...)  :
##    SL.knn only available for family = binomial()

## Warning in FUN(X[[i]], ...): Error in algorithm SL.knn
##    The Algorithm will be removed from the Super Learner (i.e. given weight 0)

## Warning in FUN(X[[i]], ...): prediction from a rank-deficient fit may be misleading

## Error in (function (Y, X, newX, family, k = 10, ...)  :
##    SL.knn only available for family = binomial()

## Warning in FUN(X[[i]], ...): Error in algorithm SL.knn  on full data
##    The Algorithm will be removed from the Super Learner (i.e. given weight 0)

## Warning in SuperLearner(Y = Y_train, X = X_train, family = gaussian(),
## SL.library = sl_libraries, : Coefficients already 0 for all failed algorithm(s)
```

```
sl_mtry1
```

```
##
## Call:
## SuperLearner(Y = Y_train, X = X_train, family = gaussian(), SL.library = sl_libraries,
```

```
##      cvControl = list(V = 5))
##
##
##                        Risk      Coef
## SL.glm_All            1.598625 0.1361646
## SL.randomForest_All 1.171995 0.8638354
## SL.knn_All                NA 0.0000000
```

```
summary(sl_mtry1)
```

```
##                   Length Class  Mode
## call                   6 -none- call
## libraryNames           3 -none- character
## SL.library             2 -none- list
## SL.predict          3758 -none- numeric
## coef                   3 -none- numeric
## library.predict    11274 -none- numeric
## Z                  11274 -none- numeric
## cvRisk                 3 -none- numeric
## family                11 family list
## fitLibrary             3 -none- list
## cvFitLibrary           0 -none- NULL
## varNames             323 -none- character
## validRows              5 -none- list
## method                 3 -none- list
## whichScreen          323 -none- logical
## control                3 -none- list
## cvControl              4 -none- list
## errorsInCVLibrary      3 -none- logical
## errorsInLibrary        3 -none- logical
## metaOptimizer          8 nnls   list
## env                   22 -none- environment
## times                  3 -none- list
```

```
pred1 <- predict(sl_mtry1, X_test, onlySL = T)
```

```
## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type == :
## prediction from a rank-deficient fit may be misleading
```

```
ggplot(data.frame(Y_test, pred1$pred), aes(pred1$pred , Y_test)) + geom_point() + geom_abline(slope=1,
```
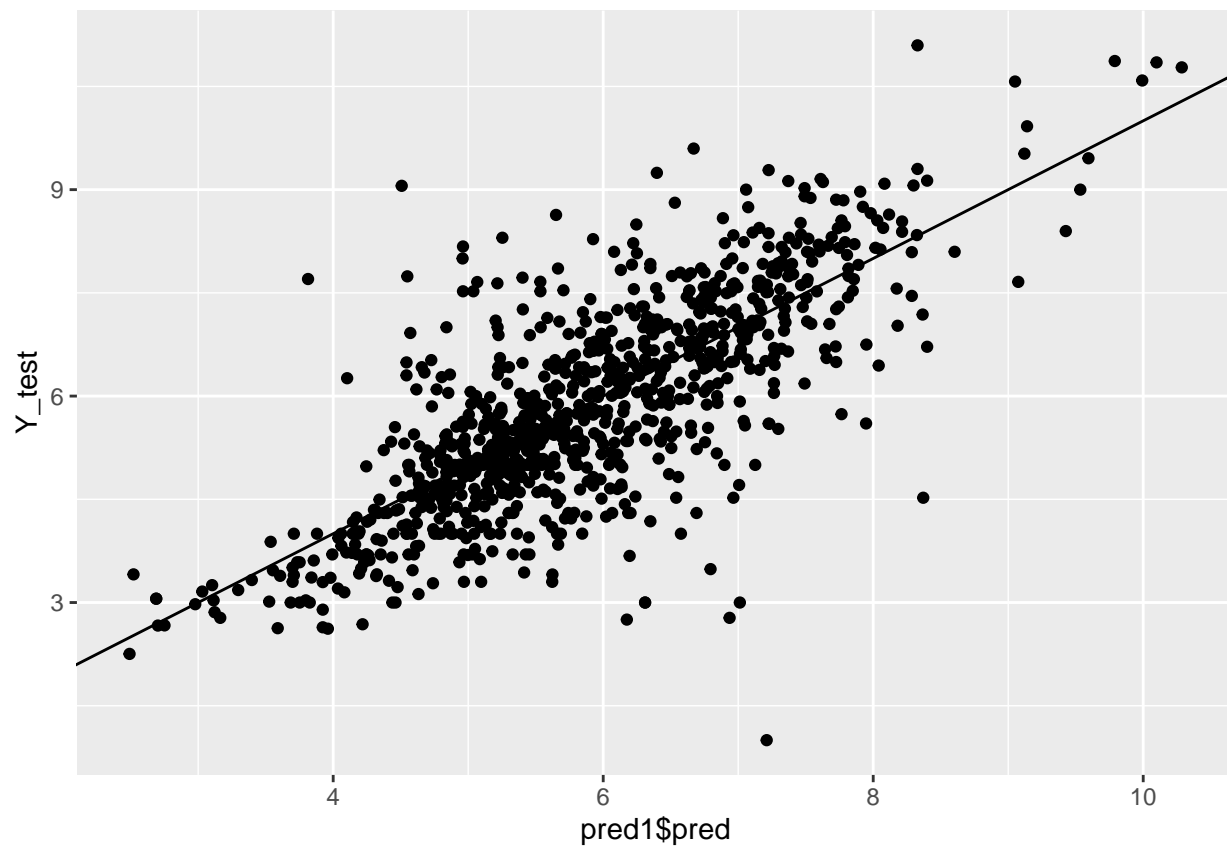
```
print(paste("Correlation:", round(cor(Y_test,pred1$pred), digits=4)))
```

```
## [1] "Correlation: 0.7679"
```

```
ggplot(data.frame(Y_test, pred1$pred), aes(pred1$pred, Y_test)) +
  geom_point() +
  geom_abline(slope = 1, intercept = 0, color = "red") +
  labs(x = "Experimental pIC50", y = "Predicted pIC50")
```
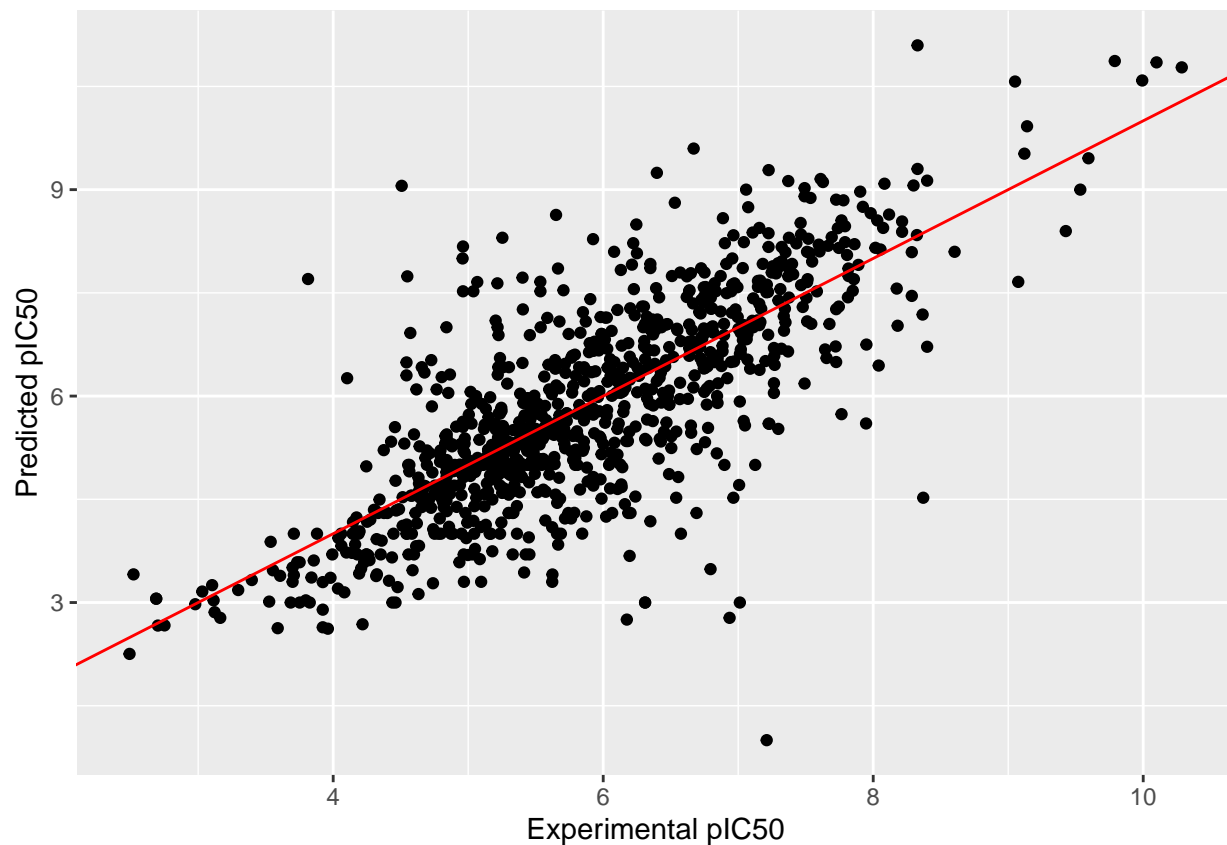
```
importances <- c(251.06030, 192.32849, 158.69406, 158.42195, 154.25947,
                 110.67713, 108.31098, 107.41293, 98.17413, 97.12171,
                 92.67365, 92.42779, 86.46519, 83.93010, 81.44788,
                 81.11657, 79.32465, 76.00375, 73.36621, 72.87880)
names(importances) <- c("PubchemFP3", "PubchemFP528", "PubchemFP559", "PubchemFP821", "PubchemFP601",
                        "PubchemFP758", "PubchemFP335", "PubchemFP308", "PubchemFP813", "PubchemFP750",
                        "PubchemFP193", "PubchemFP493", "PubchemFP621", "PubchemFP391", "PubchemFP180",
                        "PubchemFP672", "PubchemFP258", "PubchemFP623", "PubchemFP261", "PubchemFP688")

barplot(importances, horiz=TRUE, las=1, main="Feature Importances")
```

Feature Importances
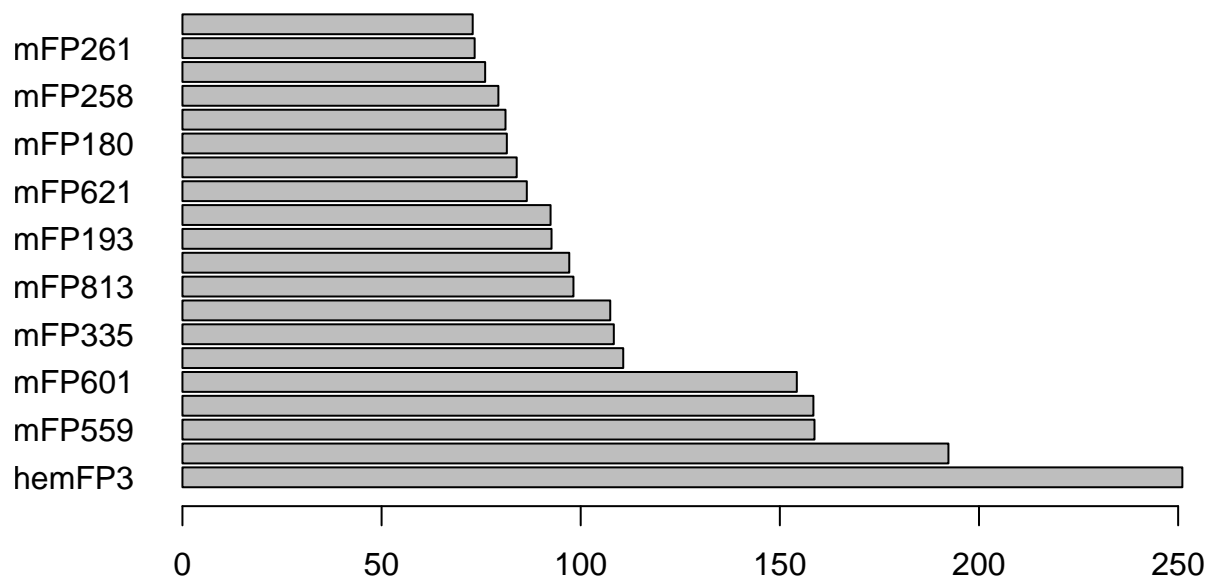

```
# Set the colors for the bars
#bar_colors <- rainbow(length(importances))

# Create the histogram
par(mar=c(10, 5, 4, 2))
barplot(importances, col= "#50C878", xlab="", ylab="Importance", las=2, cex.names=0.8)
mtext("Feature", side = 1, line = 6, cex = 0.9)
title("Top20 Molecular Descriptors")
```

Top20 Molecular Descriptors
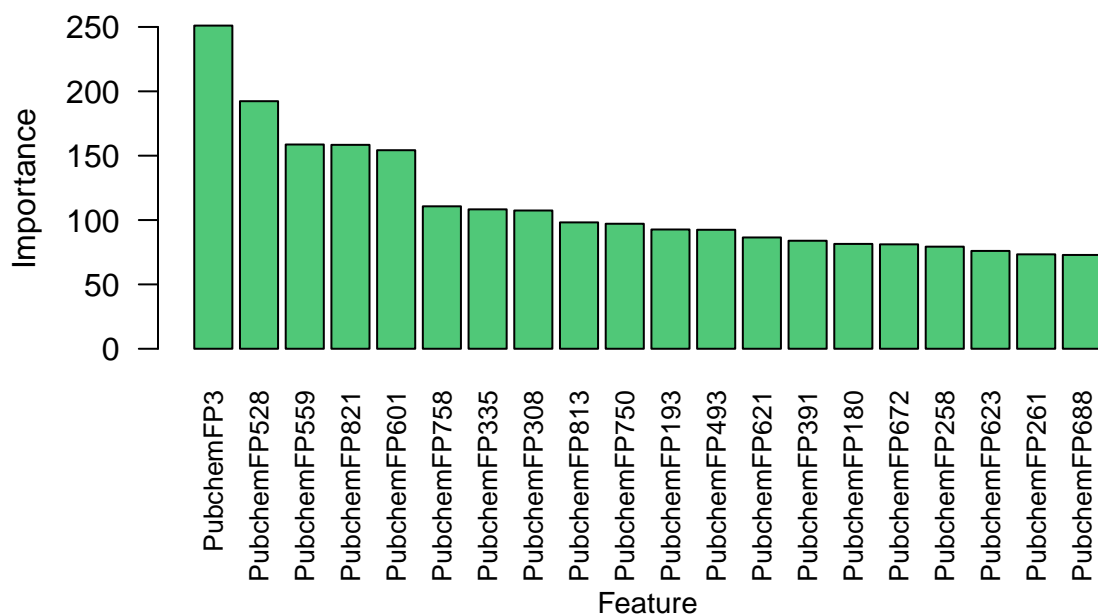

```r
# Compute RMSE for a set of predicted and true values
rmse <- function(pred, true) {
  sqrt(mean((pred - true)^2))
}

rmse_val <- rmse(pred1$pred, Y_test)
print(paste0("RMSE = ", rmse_val))
```

```
## [1] "RMSE = 0.98464020422991"
```

```r
mse1 = mean((Y_test-pred1$pred)^2)
print(paste("Mean squared error:", round(mse1, digits=4)))
```

```
## [1] "Mean squared error: 0.9695"
```