Project

Someone

2025-04-15

```r
###Extracting features using grid partitioning


##Loading necessary libraries

library(jpeg)
```

```
## Warning: package 'jpeg' was built under R version 4.4.3
```

```r
##Defining Grid Feature Extraction Function

get_grid_features <- function(img,grid_size=10) {
  #Get image dimensions
  h <- dim(img)[1]
  w <- dim(img)[2]
  c <- dim(img)[3]

  #Size of each grid cell
  h_step <- floor(h / grid_size)
  w_step <- floor(w / grid_size)

  features <- c()
  for (i in 0:(grid_size - 1)){
    for (j in 0:(grid_size - 1)){
      for (k in 1:c) {
        patch <- img[(i * h_step + 1):min((i + 1) * h_step,h),
                     (j * w_step + 1):min((j + 1) * w_step, w), k]
        features <- c(features, median(patch))
      }
    }
  }
  return(features)
}
```

```r
##Apply to all images
photometadata <- read.csv("D:\\MATH 3333\\Winter 25\\Final Project\\photoMetaData.csv")
n <- nrow(photometadata)
grid_size <- 10
X_grid <- matrix (NA, nrow = n, ncol = grid_size * grid_size * 3)

for (j in 1:n) {
  img <- readJPEG(paste0("D:\\MATH 3333\\Winter 25\\Final Project\\columbiaImages\\columbiaImages\\", pl
  X_grid[j,] <- get_grid_features(img, grid_size)
  print(sprintf("%03d / %03d", j, n))
}
```

```
## [1] "001 / 800"
## [1] "002 / 800"
## [1] "003 / 800"
## [1] "004 / 800"
## [1] "005 / 800"
## [1] "006 / 800"
## [1] "007 / 800"
## [1] "008 / 800"
## [1] "009 / 800"
## [1] "010 / 800"
## [1] "011 / 800"
## [1] "012 / 800"
## [1] "013 / 800"
## [1] "014 / 800"
## [1] "015 / 800"
## [1] "016 / 800"
## [1] "017 / 800"
## [1] "018 / 800"
## [1] "019 / 800"
## [1] "020 / 800"
## [1] "021 / 800"
## [1] "022 / 800"
## [1] "023 / 800"
## [1] "024 / 800"
## [1] "025 / 800"
## [1] "026 / 800"
## [1] "027 / 800"
## [1] "028 / 800"
## [1] "029 / 800"
## [1] "030 / 800"
## [1] "031 / 800"
## [1] "032 / 800"
## [1] "033 / 800"
## [1] "034 / 800"
## [1] "035 / 800"
## [1] "036 / 800"
## [1] "037 / 800"
## [1] "038 / 800"
## [1] "039 / 800"
## [1] "040 / 800"
## [1] "041 / 800"
## [1] "042 / 800"
## [1] "043 / 800"
## [1] "044 / 800"
## [1] "045 / 800"
## [1] "046 / 800"
## [1] "047 / 800"
## [1] "048 / 800"
## [1] "049 / 800"
## [1] "050 / 800"
## [1] "051 / 800"
## [1] "052 / 800"
## [1] "053 / 800"
## [1] "054 / 800"
## [1] "055 / 800"
## [1] "056 / 800"
## [1] "057 / 800"
## [1] "058 / 800"
## [1] "059 / 800"
## [1] "060 / 800"
## [1] "061 / 800"
## [1] "062 / 800"
## [1] "063 / 800"
## [1] "064 / 800"
## [1] "065 / 800"
## [1] "066 / 800"
## [1] "067 / 800"
## [1] "068 / 800"
## [1] "069 / 800"
## [1] "070 / 800"
## [1] "071 / 800"
## [1] "072 / 800"
## [1] "073 / 800"
## [1] "074 / 800"
## [1] "075 / 800"
## [1] "076 / 800"
## [1] "077 / 800"
## [1] "078 / 800"
## [1] "079 / 800"
## [1] "080 / 800"
## [1] "081 / 800"
## [1] "082 / 800"
## [1] "083 / 800"
## [1] "084 / 800"
## [1] "085 / 800"
## [1] "086 / 800"
## [1] "087 / 800"
## [1] "088 / 800"
## [1] "089 / 800"
## [1] "090 / 800"
## [1] "091 / 800"
## [1] "092 / 800"
## [1] "093 / 800"
## [1] "094 / 800"
## [1] "095 / 800"
## [1] "096 / 800"
## [1] "097 / 800"
## [1] "098 / 800"
## [1] "099 / 800"
## [1] "100 / 800"
## [1] "101 / 800"
## [1] "102 / 800"
## [1] "103 / 800"
## [1] "104 / 800"
## [1] "105 / 800"
## [1] "106 / 800"
## [1] "107 / 800"
## [1] "108 / 800"
## [1] "109 / 800"
## [1] "110 / 800"
## [1] "111 / 800"
## [1] "112 / 800"
## [1] "113 / 800"
## [1] "114 / 800"
## [1] "115 / 800"
## [1] "116 / 800"
## [1] "117 / 800"
## [1] "118 / 800"
## [1] "119 / 800"
## [1] "120 / 800"
## [1] "121 / 800"
## [1] "122 / 800"
## [1] "123 / 800"
## [1] "124 / 800"
## [1] "125 / 800"
## [1] "126 / 800"
## [1] "127 / 800"
## [1] "128 / 800"
## [1] "129 / 800"
## [1] "130 / 800"
## [1] "131 / 800"
## [1] "132 / 800"
## [1] "133 / 800"
## [1] "134 / 800"
## [1] "135 / 800"
## [1] "136 / 800"
## [1] "137 / 800"
## [1] "138 / 800"
## [1] "139 / 800"
## [1] "140 / 800"
## [1] "141 / 800"
## [1] "142 / 800"
## [1] "143 / 800"
## [1] "144 / 800"
## [1] "145 / 800"
## [1] "146 / 800"
## [1] "147 / 800"
## [1] "148 / 800"
## [1] "149 / 800"
## [1] "150 / 800"
## [1] "151 / 800"
## [1] "152 / 800"
## [1] "153 / 800"
## [1] "154 / 800"
## [1] "155 / 800"
## [1] "156 / 800"
## [1] "157 / 800"
## [1] "158 / 800"
## [1] "159 / 800"
## [1] "160 / 800"
## [1] "161 / 800"
## [1] "162 / 800"
## [1] "163 / 800"
## [1] "164 / 800"
## [1] "165 / 800"
## [1] "166 / 800"
## [1] "167 / 800"
## [1] "168 / 800"
## [1] "169 / 800"
## [1] "170 / 800"
## [1] "171 / 800"
## [1] "172 / 800"
## [1] "173 / 800"
## [1] "174 / 800"
## [1] "175 / 800"
## [1] "176 / 800"
## [1] "177 / 800"
## [1] "178 / 800"
## [1] "179 / 800"
## [1] "180 / 800"
## [1] "181 / 800"
## [1] "182 / 800"
## [1] "183 / 800"
## [1] "184 / 800"
## [1] "185 / 800"
## [1] "186 / 800"
## [1] "187 / 800"
## [1] "188 / 800"
## [1] "189 / 800"
## [1] "190 / 800"
## [1] "191 / 800"
## [1] "192 / 800"
## [1] "193 / 800"
## [1] "194 / 800"
## [1] "195 / 800"
## [1] "196 / 800"
## [1] "197 / 800"
## [1] "198 / 800"
## [1] "199 / 800"
## [1] "200 / 800"
## [1] "201 / 800"
## [1] "202 / 800"
## [1] "203 / 800"
## [1] "204 / 800"
## [1] "205 / 800"
## [1] "206 / 800"
## [1] "207 / 800"
## [1] "208 / 800"
## [1] "209 / 800"
## [1] "210 / 800"
## [1] "211 / 800"
## [1] "212 / 800"
## [1] "213 / 800"
## [1] "214 / 800"
## [1] "215 / 800"
## [1] "216 / 800"
## [1] "217 / 800"
## [1] "218 / 800"
## [1] "219 / 800"
## [1] "220 / 800"
## [1] "221 / 800"
## [1] "222 / 800"
## [1] "223 / 800"
## [1] "224 / 800"
## [1] "225 / 800"
## [1] "226 / 800"
## [1] "227 / 800"
## [1] "228 / 800"
## [1] "229 / 800"
## [1] "230 / 800"
## [1] "231 / 800"
## [1] "232 / 800"
## [1] "233 / 800"
## [1] "234 / 800"
## [1] "235 / 800"
## [1] "236 / 800"
## [1] "237 / 800"
## [1] "238 / 800"
## [1] "239 / 800"
## [1] "240 / 800"
## [1] "241 / 800"
## [1] "242 / 800"
## [1] "243 / 800"
## [1] "244 / 800"
## [1] "245 / 800"
## [1] "246 / 800"
## [1] "247 / 800"
## [1] "248 / 800"
## [1] "249 / 800"
## [1] "250 / 800"
## [1] "251 / 800"
## [1] "252 / 800"
## [1] "253 / 800"
## [1] "254 / 800"
## [1] "255 / 800"
## [1] "256 / 800"
## [1] "257 / 800"
## [1] "258 / 800"
## [1] "259 / 800"
## [1] "260 / 800"
## [1] "261 / 800"
## [1] "262 / 800"
## [1] "263 / 800"
## [1] "264 / 800"
## [1] "265 / 800"
## [1] "266 / 800"
## [1] "267 / 800"
## [1] "268 / 800"
## [1] "269 / 800"
## [1] "270 / 800"
## [1] "271 / 800"
## [1] "272 / 800"
## [1] "273 / 800"
## [1] "274 / 800"
## [1] "275 / 800"
## [1] "276 / 800"
## [1] "277 / 800"
## [1] "278 / 800"
## [1] "279 / 800"
## [1] "280 / 800"
## [1] "281 / 800"
## [1] "282 / 800"
## [1] "283 / 800"
## [1] "284 / 800"
## [1] "285 / 800"
## [1] "286 / 800"
## [1] "287 / 800"
## [1] "288 / 800"
## [1] "289 / 800"
## [1] "290 / 800"
## [1] "291 / 800"
## [1] "292 / 800"
## [1] "293 / 800"
## [1] "294 / 800"
## [1] "295 / 800"
## [1] "296 / 800"
## [1] "297 / 800"
## [1] "298 / 800"
## [1] "299 / 800"
## [1] "300 / 800"
## [1] "301 / 800"
## [1] "302 / 800"
## [1] "303 / 800"
## [1] "304 / 800"
## [1] "305 / 800"
## [1] "306 / 800"
## [1] "307 / 800"
## [1] "308 / 800"
## [1] "309 / 800"
## [1] "310 / 800"
## [1] "311 / 800"
## [1] "312 / 800"
## [1] "313 / 800"
## [1] "314 / 800"
## [1] "315 / 800"
## [1] "316 / 800"
## [1] "317 / 800"
## [1] "318 / 800"
## [1] "319 / 800"
## [1] "320 / 800"
## [1] "321 / 800"
## [1] "322 / 800"
## [1] "323 / 800"
## [1] "324 / 800"
## [1] "325 / 800"
## [1] "326 / 800"
## [1] "327 / 800"
## [1] "328 / 800"
## [1] "329 / 800"
## [1] "330 / 800"
## [1] "331 / 800"
## [1] "332 / 800"
## [1] "333 / 800"
## [1] "334 / 800"
## [1] "335 / 800"
## [1] "336 / 800"
## [1] "337 / 800"
## [1] "338 / 800"
## [1] "339 / 800"
## [1] "340 / 800"
## [1] "341 / 800"
## [1] "342 / 800"
## [1] "343 / 800"
## [1] "344 / 800"
## [1] "345 / 800"
## [1] "346 / 800"
## [1] "347 / 800"
## [1] "348 / 800"
## [1] "349 / 800"
## [1] "350 / 800"
## [1] "351 / 800"
## [1] "352 / 800"
## [1] "353 / 800"
## [1] "354 / 800"
## [1] "355 / 800"
## [1] "356 / 800"
## [1] "357 / 800"
## [1] "358 / 800"
## [1] "359 / 800"
## [1] "360 / 800"
## [1] "361 / 800"
## [1] "362 / 800"
## [1] "363 / 800"
## [1] "364 / 800"
## [1] "365 / 800"
## [1] "366 / 800"
## [1] "367 / 800"
## [1] "368 / 800"
## [1] "369 / 800"
## [1] "370 / 800"
## [1] "371 / 800"
## [1] "372 / 800"
## [1] "373 / 800"
## [1] "374 / 800"
## [1] "375 / 800"
## [1] "376 / 800"
## [1] "377 / 800"
## [1] "378 / 800"
## [1] "379 / 800"
## [1] "380 / 800"
## [1] "381 / 800"
## [1] "382 / 800"
## [1] "383 / 800"
## [1] "384 / 800"
## [1] "385 / 800"
## [1] "386 / 800"
## [1] "387 / 800"
## [1] "388 / 800"
## [1] "389 / 800"
## [1] "390 / 800"
## [1] "391 / 800"
## [1] "392 / 800"
## [1] "393 / 800"
## [1] "394 / 800"
## [1] "395 / 800"
## [1] "396 / 800"
## [1] "397 / 800"
## [1] "398 / 800"
## [1] "399 / 800"
## [1] "400 / 800"
## [1] "401 / 800"
## [1] "402 / 800"
## [1] "403 / 800"
## [1] "404 / 800"
## [1] "405 / 800"
## [1] "406 / 800"
## [1] "407 / 800"
## [1] "408 / 800"
## [1] "409 / 800"
## [1] "410 / 800"
## [1] "411 / 800"
## [1] "412 / 800"
## [1] "413 / 800"
## [1] "414 / 800"
## [1] "415 / 800"
## [1] "416 / 800"
## [1] "417 / 800"
## [1] "418 / 800"
## [1] "419 / 800"
## [1] "420 / 800"
## [1] "421 / 800"
## [1] "422 / 800"
## [1] "423 / 800"
## [1] "424 / 800"
## [1] "425 / 800"
## [1] "426 / 800"
## [1] "427 / 800"
## [1] "428 / 800"
## [1] "429 / 800"
## [1] "430 / 800"
## [1] "431 / 800"
## [1] "432 / 800"
## [1] "433 / 800"
## [1] "434 / 800"
## [1] "435 / 800"
## [1] "436 / 800"
## [1] "437 / 800"
## [1] "438 / 800"
## [1] "439 / 800"
## [1] "440 / 800"
## [1] "441 / 800"
## [1] "442 / 800"
## [1] "443 / 800"
## [1] "444 / 800"
## [1] "445 / 800"
## [1] "446 / 800"
## [1] "447 / 800"
## [1] "448 / 800"
## [1] "449 / 800"
## [1] "450 / 800"
## [1] "451 / 800"
## [1] "452 / 800"
## [1] "453 / 800"
## [1] "454 / 800"
## [1] "455 / 800"
## [1] "456 / 800"
## [1] "457 / 800"
## [1] "458 / 800"
## [1] "459 / 800"
## [1] "460 / 800"
## [1] "461 / 800"
## [1] "462 / 800"
## [1] "463 / 800"
## [1] "464 / 800"
## [1] "465 / 800"
## [1] "466 / 800"
## [1] "467 / 800"
## [1] "468 / 800"
## [1] "469 / 800"
## [1] "470 / 800"
## [1] "471 / 800"
## [1] "472 / 800"
## [1] "473 / 800"
## [1] "474 / 800"
## [1] "475 / 800"
## [1] "476 / 800"
## [1] "477 / 800"
## [1] "478 / 800"
## [1] "479 / 800"
## [1] "480 / 800"
## [1] "481 / 800"
## [1] "482 / 800"
## [1] "483 / 800"
## [1] "484 / 800"
## [1] "485 / 800"
## [1] "486 / 800"
## [1] "487 / 800"
## [1] "488 / 800"
## [1] "489 / 800"
## [1] "490 / 800"
## [1] "491 / 800"
## [1] "492 / 800"
## [1] "493 / 800"
## [1] "494 / 800"
## [1] "495 / 800"
## [1] "496 / 800"
## [1] "497 / 800"
## [1] "498 / 800"
## [1] "499 / 800"
## [1] "500 / 800"
## [1] "501 / 800"
## [1] "502 / 800"
## [1] "503 / 800"
## [1] "504 / 800"
## [1] "505 / 800"
## [1] "506 / 800"
## [1] "507 / 800"
## [1] "508 / 800"
## [1] "509 / 800"
## [1] "510 / 800"
## [1] "511 / 800"
## [1] "512 / 800"
## [1] "513 / 800"
## [1] "514 / 800"
## [1] "515 / 800"
## [1] "516 / 800"
## [1] "517 / 800"
## [1] "518 / 800"
## [1] "519 / 800"
## [1] "520 / 800"
## [1] "521 / 800"
## [1] "522 / 800"
## [1] "523 / 800"
## [1] "524 / 800"
## [1] "525 / 800"
## [1] "526 / 800"
## [1] "527 / 800"
## [1] "528 / 800"
## [1] "529 / 800"
## [1] "530 / 800"
## [1] "531 / 800"
## [1] "532 / 800"
## [1] "533 / 800"
## [1] "534 / 800"
## [1] "535 / 800"
## [1] "536 / 800"
## [1] "537 / 800"
## [1] "538 / 800"
## [1] "539 / 800"
## [1] "540 / 800"
## [1] "541 / 800"
## [1] "542 / 800"
## [1] "543 / 800"
## [1] "544 / 800"
## [1] "545 / 800"
## [1] "546 / 800"
## [1] "547 / 800"
## [1] "548 / 800"
## [1] "549 / 800"
## [1] "550 / 800"
## [1] "551 / 800"
## [1] "552 / 800"
## [1] "553 / 800"
## [1] "554 / 800"
## [1] "555 / 800"
## [1] "556 / 800"
## [1] "557 / 800"
## [1] "558 / 800"
## [1] "559 / 800"
## [1] "560 / 800"
## [1] "561 / 800"
## [1] "562 / 800"
## [1] "563 / 800"
## [1] "564 / 800"
## [1] "565 / 800"
## [1] "566 / 800"
## [1] "567 / 800"
## [1] "568 / 800"
## [1] "569 / 800"
## [1] "570 / 800"
## [1] "571 / 800"
## [1] "572 / 800"
## [1] "573 / 800"
## [1] "574 / 800"
## [1] "575 / 800"
## [1] "576 / 800"
## [1] "577 / 800"
## [1] "578 / 800"
## [1] "579 / 800"
## [1] "580 / 800"
## [1] "581 / 800"
## [1] "582 / 800"
## [1] "583 / 800"
## [1] "584 / 800"
## [1] "585 / 800"
## [1] "586 / 800"
## [1] "587 / 800"
## [1] "588 / 800"
## [1] "589 / 800"
## [1] "590 / 800"
## [1] "591 / 800"
## [1] "592 / 800"
## [1] "593 / 800"
## [1] "594 / 800"
## [1] "595 / 800"
## [1] "596 / 800"
## [1] "597 / 800"
## [1] "598 / 800"
## [1] "599 / 800"
## [1] "600 / 800"
## [1] "601 / 800"
## [1] "602 / 800"
## [1] "603 / 800"
## [1] "604 / 800"
## [1] "605 / 800"
## [1] "606 / 800"
## [1] "607 / 800"
## [1] "608 / 800"
## [1] "609 / 800"
## [1] "610 / 800"
## [1] "611 / 800"
## [1] "612 / 800"
## [1] "613 / 800"
## [1] "614 / 800"
## [1] "615 / 800"
## [1] "616 / 800"
## [1] "617 / 800"
## [1] "618 / 800"
## [1] "619 / 800"
## [1] "620 / 800"
## [1] "621 / 800"
## [1] "622 / 800"
## [1] "623 / 800"
## [1] "624 / 800"
## [1] "625 / 800"
## [1] "626 / 800"
## [1] "627 / 800"
## [1] "628 / 800"
## [1] "629 / 800"
## [1] "630 / 800"
## [1] "631 / 800"
## [1] "632 / 800"
## [1] "633 / 800"
## [1] "634 / 800"
## [1] "635 / 800"
## [1] "636 / 800"
## [1] "637 / 800"
## [1] "638 / 800"
## [1] "639 / 800"
## [1] "640 / 800"
## [1] "641 / 800"
## [1] "642 / 800"
## [1] "643 / 800"
## [1] "644 / 800"
## [1] "645 / 800"
## [1] "646 / 800"
## [1] "647 / 800"
## [1] "648 / 800"
## [1] "649 / 800"
## [1] "650 / 800"
## [1] "651 / 800"
## [1] "652 / 800"
## [1] "653 / 800"
## [1] "654 / 800"
## [1] "655 / 800"
## [1] "656 / 800"
## [1] "657 / 800"
## [1] "658 / 800"
## [1] "659 / 800"
## [1] "660 / 800"
## [1] "661 / 800"
## [1] "662 / 800"
## [1] "663 / 800"
## [1] "664 / 800"
## [1] "665 / 800"
## [1] "666 / 800"
## [1] "667 / 800"
## [1] "668 / 800"
## [1] "669 / 800"
## [1] "670 / 800"
## [1] "671 / 800"
## [1] "672 / 800"
## [1] "673 / 800"
## [1] "674 / 800"
## [1] "675 / 800"
## [1] "676 / 800"
## [1] "677 / 800"
## [1] "678 / 800"
## [1] "679 / 800"
## [1] "680 / 800"
## [1] "681 / 800"
## [1] "682 / 800"
## [1] "683 / 800"
## [1] "684 / 800"
## [1] "685 / 800"
## [1] "686 / 800"
## [1] "687 / 800"
## [1] "688 / 800"
## [1] "689 / 800"
## [1] "690 / 800"
## [1] "691 / 800"
## [1] "692 / 800"
## [1] "693 / 800"
## [1] "694 / 800"
## [1] "695 / 800"
## [1] "696 / 800"
## [1] "697 / 800"
## [1] "698 / 800"
## [1] "699 / 800"
## [1] "700 / 800"
## [1] "701 / 800"
## [1] "702 / 800"
## [1] "703 / 800"
## [1] "704 / 800"
## [1] "705 / 800"
## [1] "706 / 800"
## [1] "707 / 800"
## [1] "708 / 800"
## [1] "709 / 800"
## [1] "710 / 800"
## [1] "711 / 800"
## [1] "712 / 800"
## [1] "713 / 800"
## [1] "714 / 800"
## [1] "715 / 800"
## [1] "716 / 800"
## [1] "717 / 800"
## [1] "718 / 800"
## [1] "719 / 800"
## [1] "720 / 800"
## [1] "721 / 800"
## [1] "722 / 800"
## [1] "723 / 800"
## [1] "724 / 800"
## [1] "725 / 800"
## [1] "726 / 800"
## [1] "727 / 800"
## [1] "728 / 800"
## [1] "729 / 800"
## [1] "730 / 800"
## [1] "731 / 800"
## [1] "732 / 800"
## [1] "733 / 800"
## [1] "734 / 800"
## [1] "735 / 800"
## [1] "736 / 800"
## [1] "737 / 800"
## [1] "738 / 800"
## [1] "739 / 800"
## [1] "740 / 800"
## [1] "741 / 800"
## [1] "742 / 800"
## [1] "743 / 800"
## [1] "744 / 800"
## [1] "745 / 800"
## [1] "746 / 800"
## [1] "747 / 800"
## [1] "748 / 800"
## [1] "749 / 800"
## [1] "750 / 800"
## [1] "751 / 800"
## [1] "752 / 800"
## [1] "753 / 800"
## [1] "754 / 800"
## [1] "755 / 800"
## [1] "756 / 800"
## [1] "757 / 800"
## [1] "758 / 800"
## [1] "759 / 800"
## [1] "760 / 800"
## [1] "761 / 800"
## [1] "762 / 800"
## [1] "763 / 800"
## [1] "764 / 800"
## [1] "765 / 800"
## [1] "766 / 800"
## [1] "767 / 800"
## [1] "768 / 800"
## [1] "769 / 800"
## [1] "770 / 800"
## [1] "771 / 800"
## [1] "772 / 800"
## [1] "773 / 800"
## [1] "774 / 800"
## [1] "775 / 800"
## [1] "776 / 800"
## [1] "777 / 800"
## [1] "778 / 800"
## [1] "779 / 800"
## [1] "780 / 800"
## [1] "781 / 800"
## [1] "782 / 800"
## [1] "783 / 800"
## [1] "784 / 800"
## [1] "785 / 800"
## [1] "786 / 800"
## [1] "787 / 800"
## [1] "788 / 800"
## [1] "789 / 800"
## [1] "790 / 800"
## [1] "791 / 800"
## [1] "792 / 800"
## [1] "793 / 800"
## [1] "794 / 800"
## [1] "795 / 800"
## [1] "796 / 800"
## [1] "797 / 800"
## [1] "798 / 800"
## [1] "799 / 800"
## [1] "800 / 800"
```

```
##Using Random Forest

#Label and Split
```

```r
y <- as.numeric(photometadata$category == "outdoor-day")
set.seed(42)
trainFlag <- runif(length(y)) > 0.5

#Train a Random Forest
library(randomForest)
```

```
## Warning: package 'randomForest' was built under R version 4.4.3
```

```
## randomForest 4.7-1.2
```

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

```r
rf_model <- randomForest(x = X_grid[trainFlag, ],
                         y = as.factor(y[trainFlag]),
                         ntree = 200)

rf_model
```

```
##
## Call:
##  randomForest(x = X_grid[trainFlag, ], y = as.factor(y[trainFlag]),      ntree = 200)
##                Type of random forest: classification
##                      Number of trees: 200
## No. of variables tried at each split: 17
##
##         OOB estimate of  error rate: 22.31%
## Confusion matrix:
##     0  1 class.error
## 0 220 26   0.1056911
## 1  59 76   0.4370370
```

```r
##Predict and Evaluate

rf_pred_probs <- predict(rf_model, X_grid[!trainFlag, ], type = "prob")[,2]
rf_pred_class <- as.numeric(rf_pred_probs > 0.5)
y_test <- y[!trainFlag]

#Accuracy
rf_acc <- mean(rf_pred_class == y_test)

#Misclassification Error
rf_misc <- 1 - rf_acc

#Sensitivity (True Positive Rate)
rf_sens <- sum(rf_pred_class == 1 & y_test == 1) / sum(y_test == 1)

#Specificity (True Negative Rate)
rf_spec <- sum(rf_pred_class == 0 & y_test == 0) / sum(y_test == 0)

rf_acc
```

```
## [1] 0.7995227
```

```
rf_misc
```

```
## [1] 0.2004773
```

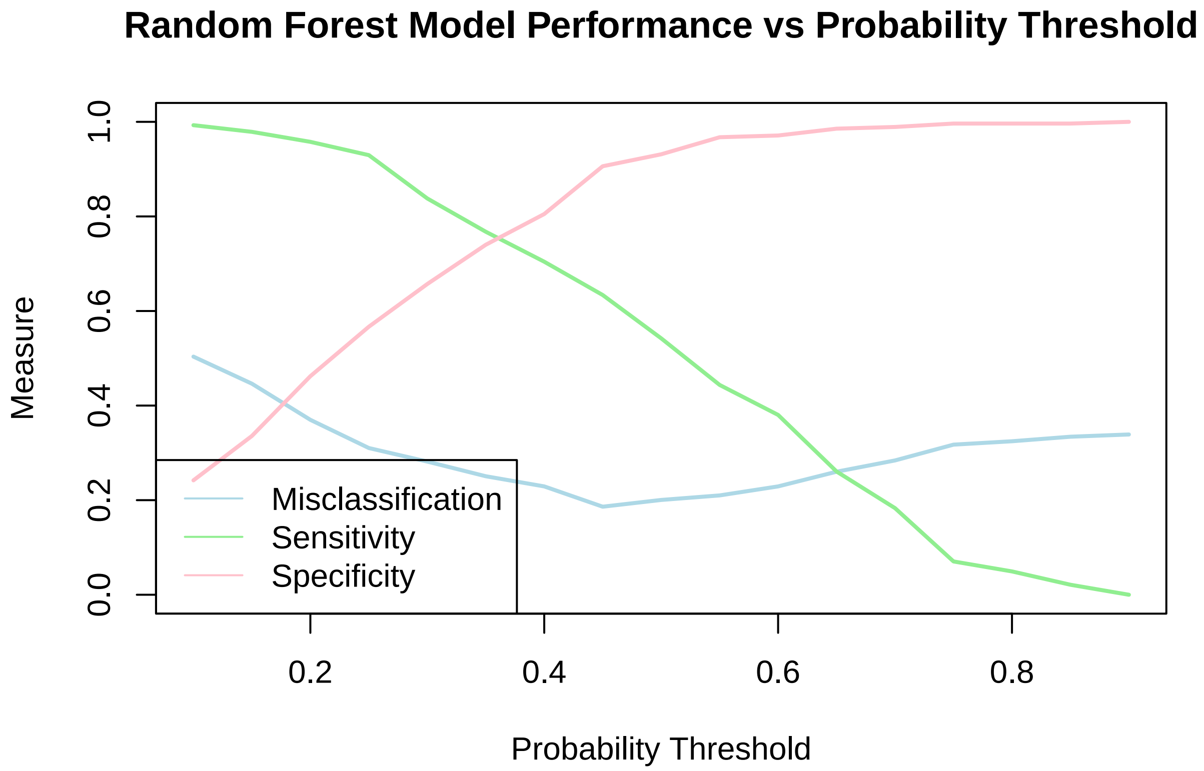
```
rf_sens
```

```
## [1] 0.5422535
```

```
rf_spec
```

```
## [1] 0.9314079
```

```
thresholds  <- seq(0.1,0.9, by = 0.05)

accuracy <- c()
misclassification <- c()
sensitivity <- c()
specificity <- c()

for (t in thresholds) {
  preds <- as.numeric(rf_pred_probs > t)

  acc <- mean(preds == y_test)
  misc <- 1 - acc
  sens <- sum(preds == 1 & y_test == 1) / sum(y_test == 1)
  spec <- sum(preds == 0 & y_test == 0) / sum(y_test == 0)

  accuracy <- c(accuracy, acc)
  misclassification <- c(misclassification, misc)
  sensitivity <- c(sensitivity, sens)
  specificity <- c(specificity, spec)
}

#Base plot
plot(thresholds, misclassification, type="l", col="lightblue", lwd=2, ylim=c(0,1),
     xlab="Probability Threshold", ylab="Measure",
     main="Random Forest Model Performance vs Probability Threshold")
lines(thresholds, sensitivity, col="lightgreen", lwd=2)
lines(thresholds, specificity, col="pink", lwd=2)

legend("bottomleft", legend=c("Misclassification", "Sensitivity", "Specificity"),
       col=c("lightblue", "lightgreen", "pink"), lwd=1)
```



```r
#Finding where misclassification is lowest for Random Forest

rf_best_index <- which.min(misclassification)

rf_best_threshold <- thresholds[rf_best_index]
rf_best_misclassification <- misclassification[rf_best_index]
rf_best_sensitivity <- sensitivity[rf_best_index]
rf_best_specificity <- specificity[rf_best_index]


cat("Best Probability Threshold:", rf_best_threshold, "\n")
```

```
## Best Probability Threshold: 0.45
```

```r
cat("Misclassification at best probability threshold is", rf_best_misclassification, "\n")
```

```
## Misclassification at best probability threshold is 0.1861575
```

```r
cat("Sensitivity at best probability threshold is", rf_best_sensitivity, "\n")
```

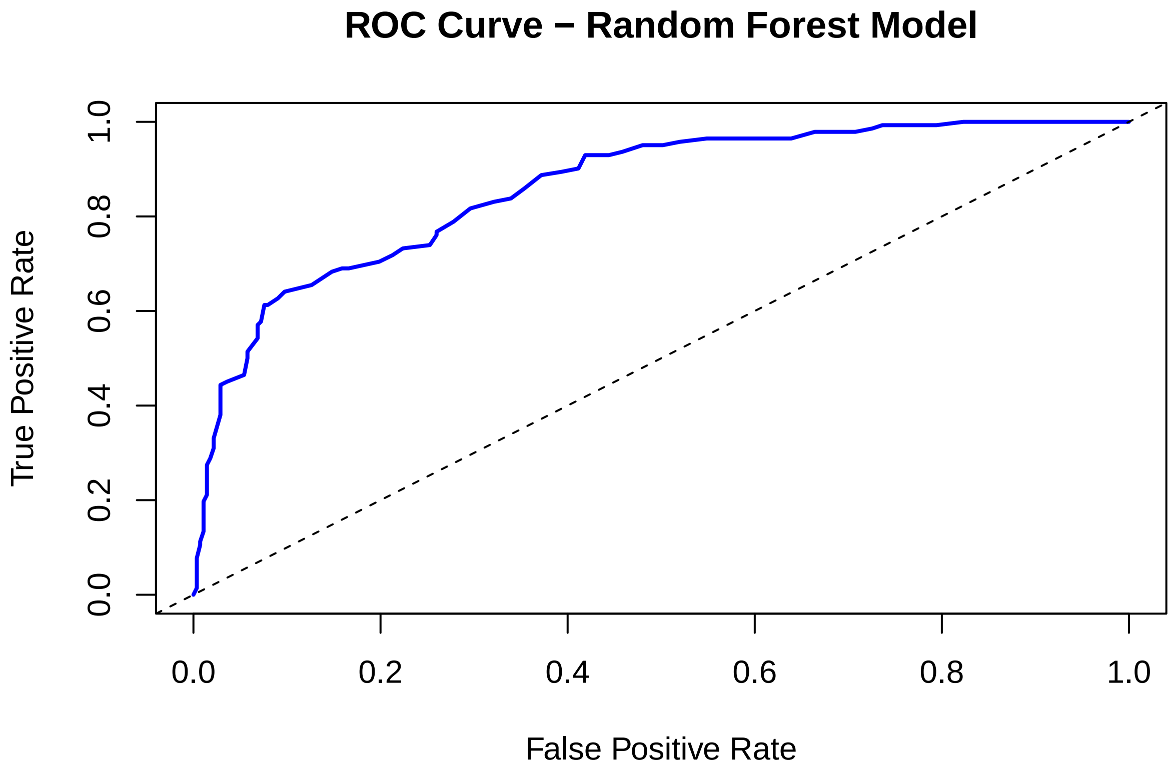
```
## Sensitivity at best probability threshold is 0.6338028
```

```r
cat("Specificity at best probability threshold is", rf_best_specificity, "\n")
```

```
## Specificity at best probability threshold is 0.9061372
```

```r
get_roc_data <- function(y_true, probs) {
  thresholds <- seq(0, 1, by = 0.01)
  tpr <- c()  # Sensitivity
  fpr <- c()  # 1 - Specificity

  for (t in thresholds) {
    pred_class <- as.numeric(probs >= t)

    tp <- sum(pred_class == 1 & y_true == 1)
    fn <- sum(pred_class == 0 & y_true == 1)
    fp <- sum(pred_class == 1 & y_true == 0)
    tn <- sum(pred_class == 0 & y_true == 0)

    tpr <- c(tpr, tp / (tp + fn))
    fpr <- c(fpr, fp / (fp + tn))
  }

  return(list(fpr = fpr, tpr = tpr))
}
```

```r
rf_roc <- get_roc_data(y_test, rf_pred_probs)

# Plot ROC Curve
plot(rf_roc$fpr, rf_roc$tpr, type = "l", col = "blue", lwd = 2,
     xlab = "False Positive Rate", ylab = "True Positive Rate",
     main = "ROC Curve - Random Forest Model")
abline(0, 1, lty = 2, col = "black")
```



```
get_auc <- function(fpr, tpr) {
  # Order by increasing FPR
  ord <- order(fpr)
  fpr <- fpr[ord]
  tpr <- tpr[ord]

  auc <- sum(diff(fpr) * (tpr[-1] + tpr[-length(tpr)]) / 2)  #Trapezoid method
  return(auc)
}

rf_auc <- get_auc(rf_roc$fpr, rf_roc$tpr)

cat("AUC for Random Forest Model:", round(rf_auc, 4), "\n")
```

```
## AUC for Random Forest Model: 0.859
```

```
library(MASS)
```

```
## Warning: package 'MASS' was built under R version 4.4.3
```

```
#Train LDA model
lda_model <- lda(X_grid[trainFlag, ], grouping = as.factor(y[trainFlag]))

#Predict on test set
```

```
lda_pred <- predict(lda_model, X_grid[!trainFlag, ])
lda_probs <- lda_pred$posterior[,2]

#Using Threshold to get predicted classes
lda_class <- as.numeric(lda_probs > 0.5)
y_test <- y[!trainFlag]

#Evaluate the performance of the model

#Accuracy
lda_acc <- mean(lda_class == y_test)
#Misclassification
lda_misc <- 1 - lda_acc
#Sensitivity
lda_sens <- sum(lda_class == 1 & y_test == 1) / sum(y_test == 1)
#Specificity
lda_spec <- sum(lda_class == 0 & y_test == 0) / sum(y_test == 0)

cat("LDA Misclassification:", lda_misc,"\n")
```

```
## LDA Misclassification: 0.3627685
```

```
cat("LDA Sensitivity:", lda_sens,"\n")
```

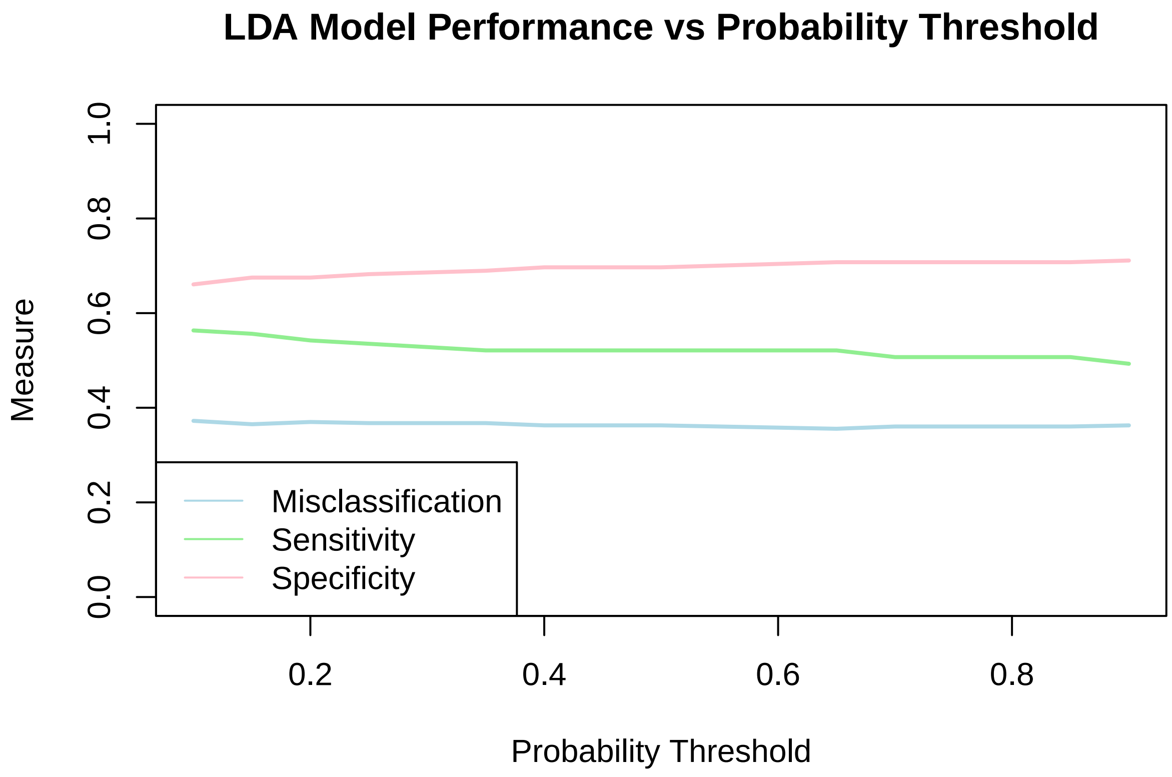
```
## LDA Sensitivity: 0.5211268
```

```
cat("LDA Specificity:", lda_spec,"\n")
```

```
## LDA Specificity: 0.6967509
```

```
thresholds  <- seq(0.1,0.9, by = 0.05)

lda_accuracy <- c()
lda_misclassification <- c()
lda_sensitivity <- c()
lda_specificity <- c()

for (t in thresholds) {
  lda_preds <- as.numeric(lda_probs > t)

  acc <- mean(lda_preds == y_test)
  misc <- 1 - acc
  sens <- sum(lda_preds == 1 & y_test == 1) / sum(y_test == 1)
  spec <- sum(lda_preds == 0 & y_test == 0) / sum(y_test == 0)

  lda_accuracy <- c(lda_accuracy, acc)
  lda_misclassification <- c(lda_misclassification, misc)
  lda_sensitivity <- c(lda_sensitivity, sens)
  lda_specificity <- c(lda_specificity, spec)
}
```

```r
#Base plot
plot(thresholds, lda_misclassification, type="l", col="lightblue", lwd=2, ylim=c(0,1), xlab="Probabilit
     main="LDA Model Performance vs Probability Threshold")
lines(thresholds, lda_sensitivity, col="lightgreen", lwd=2)
lines(thresholds, lda_specificity, col="pink", lwd=2)

legend("bottomleft", legend=c("Misclassification", "Sensitivity", "Specificity"),
       col=c("lightblue", "lightgreen", "pink"), lwd=1)
```



```r
#Finding where misclassification is lowest for LDA

lda_best_index <- which.min(lda_misclassification)

lda_best_threshold <- thresholds[lda_best_index]
lda_best_misclassification <- lda_misclassification[lda_best_index]
lda_best_sensitivity <- lda_sensitivity[lda_best_index]
lda_best_specificity <- lda_specificity[lda_best_index]


cat("Best Probability Threshold:", lda_best_threshold, "\n")
```

```
## Best Probability Threshold: 0.65
```

```
cat("Misclassification at best probability threshold is", lda_best_misclassification, "\n")
```

```
## Misclassification at best probability threshold is 0.3556086
```

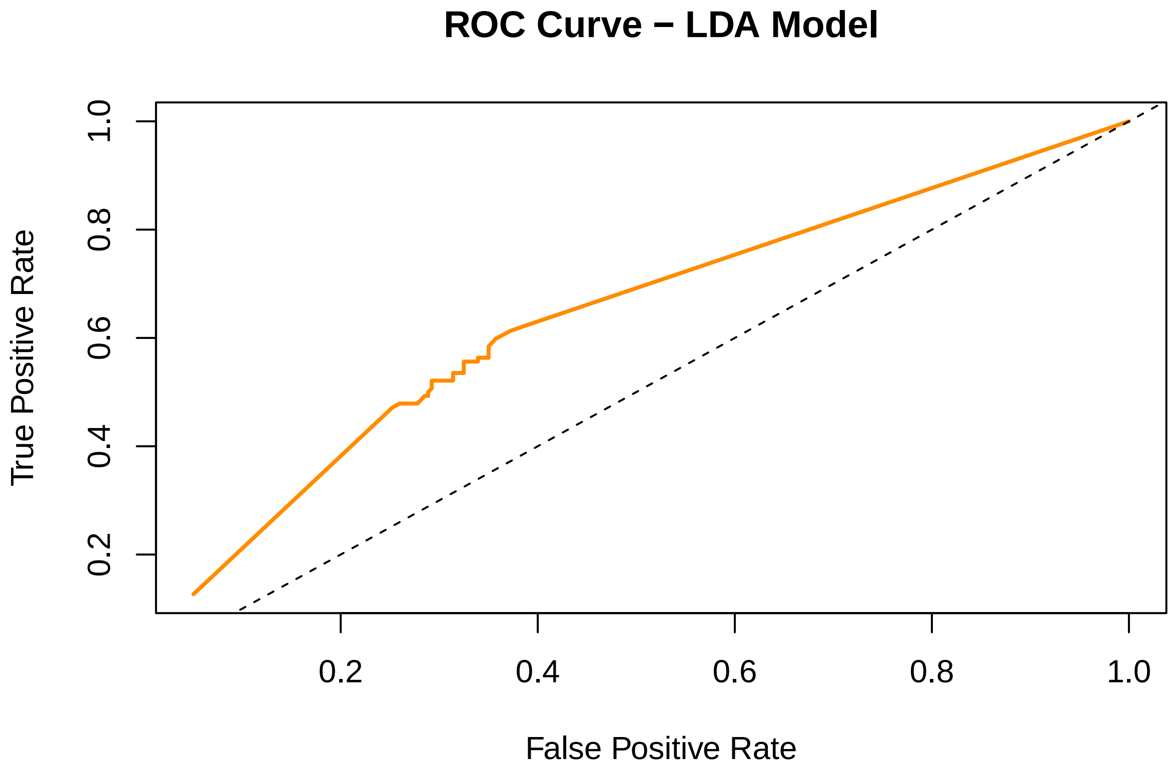
```
cat("Sensitivity at best probability threshold is", lda_best_sensitivity, "\n")
```

```
## Sensitivity at best probability threshold is 0.5211268
```

```
cat("Specificity at best probability threshold is", lda_best_specificity, "\n")
```

```
## Specificity at best probability threshold is 0.7075812
```

```
lda_roc <- get_roc_data(y_test, lda_probs)

# Plot ROC Curve for LDA
plot(lda_roc$fpr, lda_roc$tpr, type = "l", col = "darkorange", lwd = 2,
     xlab = "False Positive Rate", ylab = "True Positive Rate",
     main = "ROC Curve - LDA Model")
abline(0, 1, lty = 2, col = "black")
```



```
#AUC

lda_auc <- get_auc(lda_roc$fpr, lda_roc$tpr)
cat("AUC for LDA Model:", round(lda_auc, 4), "\n")
```

```
## AUC for LDA Model: 0.6305
```

```
##KNN Model

library(class)
```

```
## Warning: package 'class' was built under R version 4.4.3
```

```
#Create test and train data
X_train <- X_grid[trainFlag, ]
X_test <- X_grid[!trainFlag, ]
y_train <- as.factor(y[trainFlag])
y_test <- y[!trainFlag]

#Run K-nearest neighbor, starting with k=5
knn_pred <- knn(train = X_train, test = X_test, cl = y_train, k = 5, prob = TRUE)

#Convert predictions to numeric
knn_class <- as.numeric(as.character(knn_pred))

#Accuracy
knn_acc <- mean(knn_class == y_test)

#Misclassification
knn_misc <- 1 - knn_acc

# Sensitivity
knn_sens <- sum(knn_class == 1 & y_test == 1) / sum(y_test == 1)

# Specificity
knn_spec <- sum(knn_class == 0 & y_test == 0) / sum(y_test == 0)

# Print
cat("KNN (k=5) Misclassification:", knn_misc, "\n")
```

```
## KNN (k=5) Misclassification: 0.2696897
```

```
cat("Sensitivity:", knn_sens, "\n")
```

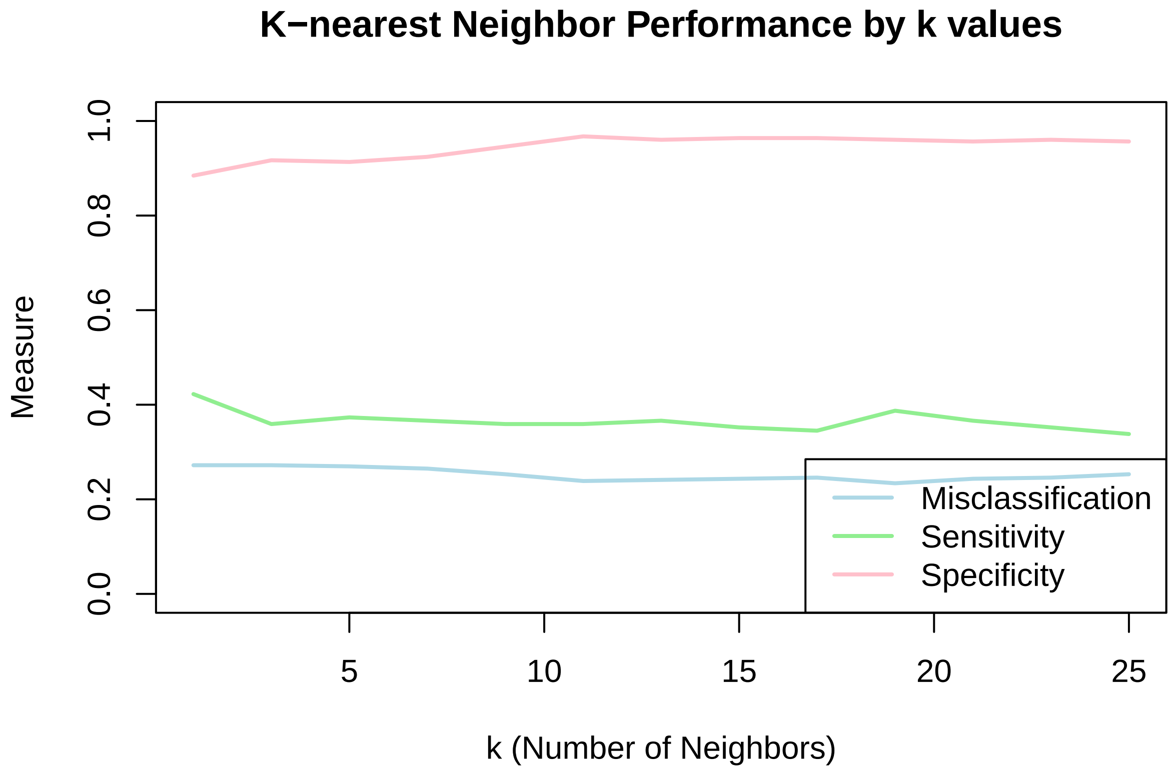
```
## Sensitivity: 0.3732394
```

```
cat("Specificity:", knn_spec, "\n")
```

```
## Specificity: 0.9133574
```

```
k_vals <- seq(1, 25, by=2)
acc_vec <- c()
misc_vec <- c()
sens_vec <- c()
spec_vec <- c()
```

```r
for (k in k_vals) {
  pred <- knn(train = X_train, test = X_test, cl = y_train, k = k)
  pred_num <- as.numeric(as.character(pred))

  acc <- mean(pred_num == y_test)
  misc <- 1 - acc
  sens <- sum(pred_num == 1 & y_test == 1) / sum(y_test == 1)
  spec <- sum(pred_num == 0 & y_test == 0) / sum(y_test == 0)

  acc_vec <- c(acc_vec, acc)
  misc_vec <- c(misc_vec, misc)
  sens_vec <- c(sens_vec, sens)
  spec_vec <- c(spec_vec, spec)
}

# Plot
plot(k_vals, misc_vec, type="l", col="lightblue", ylim=c(0,1), lwd=2,
     xlab="k (Number of Neighbors)", ylab="Measure",
     main="K-nearest Neighbor Performance by k values")
lines(k_vals, sens_vec, col="lightgreen", lwd=2)
lines(k_vals, spec_vec, col="pink", lwd=2)
legend("bottomright", legend=c("Misclassification", "Sensitivity", "Specificity"),
       col=c("lightblue", "lightgreen", "pink"), lwd=2)
```



```
knn_best_index <- which.min(misc_vec)

best_kvalue <- k_vals[knn_best_index]
knn_best_misclassification <- misc_vec[knn_best_index]
knn_best_sensitivity <- sens_vec[knn_best_index]
knn_best_specificity <- spec_vec[knn_best_index]


cat("Best Performing K value:", best_kvalue, "\n")
```

```
## Best Performing K value: 19
```

```
cat("Misclassification at best K value is", knn_best_misclassification, "\n")
```

```
## Misclassification at best K value is 0.2338902
```

```
cat("Sensitivity at best K value is", knn_best_sensitivity, "\n")
```

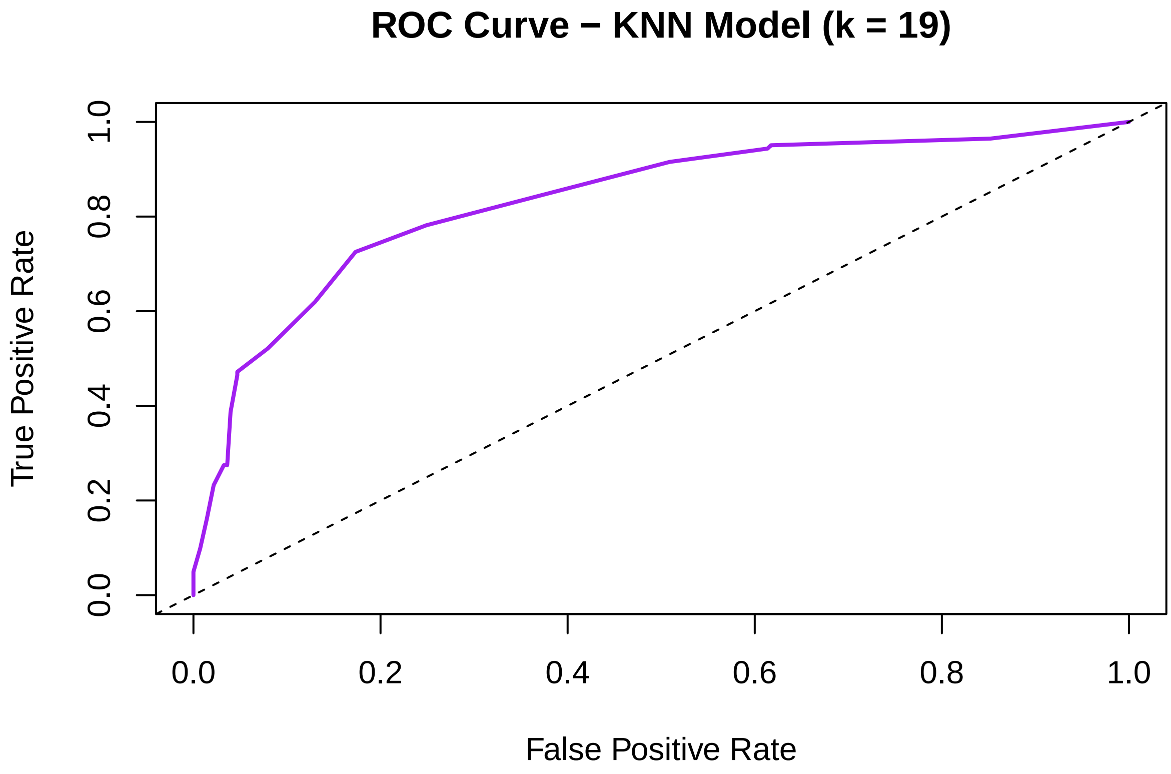
```
## Sensitivity at best K value is 0.3873239
```

```
cat("Specificity at best K value is", knn_best_specificity, "\n")
```

```
## Specificity at best K value is 0.9602888
```

```
knn_pred <- knn(train = X_train, test = X_test, cl = y_train, k = best_kvalue, prob = TRUE)

# Get vote proportions
knn_vote_probs <- attr(knn_pred, "prob")

# Predicted class labels (as numeric)
knn_class <- as.numeric(as.character(knn_pred))

# Convert to class-1 probability (outdoor-day)
# If predicted class is 1, use prob directly; otherwise, use 1 - prob
knn_probs <- ifelse(knn_class == 1, knn_vote_probs, 1 - knn_vote_probs)
```

```
knn_roc <- get_roc_data(y_test, knn_probs)

# Plot ROC
plot(knn_roc$fpr, knn_roc$tpr, type = "l", col = "purple", lwd = 2,
     xlab = "False Positive Rate", ylab = "True Positive Rate",
     main = "ROC Curve - KNN Model (k = 19)")
abline(0, 1, lty = 2, col = "black")
```



```
knn_auc <- get_auc(knn_roc$fpr, knn_roc$tpr)
cat("AUC for KNN Model (k = 19):", round(knn_auc, 4), "\n")
```

```
## AUC for KNN Model (k = 19): 0.8334
```