

Homework 3 Report

Classification of liver malfunction severity (LDA)

Ian Effendi \ iae2784@rit.edu

October 11, 2021

Contents

Certification	1
Overview	1
ELT	4
Session Information	5

Certification

I certify that I indeed finished reading Ch. 4 from *An Introduction to Statistical Learning*, by James Gareth, Daniela Witten, Trevor Hastie, Robert Tibshirani.

Overview

```
# Comments that begin with '## ----' can be used
# to separate content. R Studio provides them as
# little bookmarks that you can use.

## ---- Hidden Variables ----
.VERBOSE = TRUE
.FILEPATHS = list(
  utils      = list(
    .          = "R/utils.R",
    dependencies = "R/utils/dependencies.R",
    paths       = "R/utils/paths.R",
    printf      = "R/utils/printf.R"
  ),
  packages   = "R/packages.R",
  build      = list(
    .          = "R/build.R",
    dataset    = "R/build/dataset.R",
    report     = "R/build/report.R"
  ),
  analysis   = list(
    .          = "R/analysis.R",
```

```

    setup      = "R/analysis/setup.R",
    eda        = "R/analysis/eda.R",
    model      = "R/analysis/model.R",
    metrics    = "R/analysis/metrics.R",
    validation  = "R/analysis/validation.R"
  )
)

## ---- Default Settings (Knitr) ----
knitr::opts_knit$set(
  root.dir = here::here()
)

## ---- Default Settings (Chunk) ----
knitr::opts_chunk$set(

  ## ---- Options (Execution) ----
  eval      = TRUE,    # By default, turn off chunk evaluation.
  cache     = FALSE,   # By default, cache all chunk results.
  error     = FALSE,   # Stop R when an error is raised.

  ## ---- Options (Figure) ----
  fig.path  = 'figure/',
  fig.keep  = 'all',
  fig.align = 'center',
  fig.width = 8,
  fig.height = 6,
  dpi       = 600,

  ## ---- Options (Formatting) ----
  results   = 'hold',
  collapse  = FALSE,
  strip.white = FALSE, # Remove blank lines at beginning and end of the source code block
  tidy      = FALSE,
  size      = 'tiny',
  R.options = list(width = 60),

  ## ---- Options (Display) ----
  include   = TRUE,
  echo      = TRUE,
  message   = TRUE,
  warning   = TRUE
)

## ---- Utility Functions ----

#' Use relative path to import all
#' chunks from an external script,
#' without executing the code inside
#' of the document.
#'

```

```

#' @param path Relative path to file.
#' @param ... Additional parameters passed to the `read_chunk` call.
#' @param roxygen_comments FALSE; should comments be stripped out from external code?
#' @param verbose Print messages to stderr.
import.chunks <- function(path, ..., verbose = .VERBOSE) {
  if (verbose) { message(sprintf("Importing chunks from external script: '%s'...", basename(
    file = here::here(path)
    knitr::read_chunk(file, ...)
  })

  #' Use relative path to source an
  #' external script in its entirety.
  #'
  #' @param path Relative path to file.
  #' @param ... Additional parameters passed to the `source` call.
  #' @param verbose Print messages to stderr.
import.source <- function(path, ..., verbose = .VERBOSE) {
  if (verbose) { message(sprintf("Sourcing external script: '%s'...", basename(path))) }
  source(here::here(path), ..., echo = .VERBOSE)
}

```

```

## ---- Import (R/utils.R) ----
import.chunks(.FILEPATHS$utils$.)

```

```
## Importing chunks from external script: 'utils.R'...
```

```

## ---- Import (R/packages.R) ----
import.chunks(.FILEPATHS$packages)

```

```
## Importing chunks from external script: 'packages.R'...
```

```

## ---- Import (R/build/*) ----
import.chunks(.FILEPATHS$build$dataset)

```

```
## Importing chunks from external script: 'dataset.R'...
```

```
import.chunks(.FILEPATHS$build$report)
```

```
## Importing chunks from external script: 'report.R'...
```

```
## ---- Import (R/analysis/*) ----
```

```

## ---- Source (R/utils.R) ----
source.utils(local = FALSE, force = TRUE, verbose = .VERBOSE)

```

```
## Loading 3 utilit(y/ies)...\n\n## Loading utility: 'printf'\n\n## Loading utility: 'paths'\n\n## Loading utility: 'dependencies'
```

```
## ---- Source (R/packages.R) ----\ninstall.deps(file = "requirements.txt")
```

```
## * The project is already synchronized with the lockfile.
```

In this assignment we will:

- Perform exploratory data analysis (*EDA*) on the dataset.
- Fit and analyze a linear discriminant analysis (*LDA*) model on the dataset.
- Perform multiple cross-validation tasks at different k -fold values ($k = 3, k = 10$).

ELT

Much of the *extract*, *load*, and *transform* (*ELT*) process from the previous report has been revised for this assignment. Notably, a `make.dataset()` function streamlines the process of parsing the source `Liver.txt` file into a compatible `data.frame`.

Session Information

*This document was generated from an **R Markdown** Notebook (See the `vignettes/HW3_report.Rmd` in the project's sub-directory). The setup chunk for this document sets the root directory to the project root directory using the `here` package; all file paths are relative to the project root.*

```
## R version 4.1.1 (2021-08-10)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19042)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] stats      graphics  grDevices datasets  utils
## [6] methods    base
##
## other attached packages:
## [1] dplyr_1.0.7      tidyr_1.1.4      forcats_0.5.1
## [4] ggplot2_3.3.5    foreach_1.5.1    magrittr_2.0.1
## [7] mime_0.12        markdown_1.1     rmarkdown_2.11
## [10] knitr_1.36
##
## loaded via a namespace (and not attached):
## [1] compiler_4.1.1   pillar_1.6.3     jquerylib_0.1.4
## [4] iterators_1.0.13 tools_4.1.1       bit_4.0.4
## [7] digest_0.6.28    corrplot_0.90     jsonlite_1.7.2
## [10] evaluate_0.14    lifecycle_1.0.1   tibble_3.1.5
## [13] gtable_0.3.0     pkgconfig_2.0.3   rlang_0.4.11
## [16] parallel_4.1.1   yaml_2.2.1        xfun_0.26
## [19] fastmap_1.1.0    withr_2.4.2       stringr_1.4.0
## [22] hms_1.1.1        generics_0.1.0    vctrs_0.3.8
## [25] bit64_4.0.5      tidyselect_1.1.1  rprojroot_2.0.2
## [28] grid_4.1.1       glue_1.4.2        here_1.0.1
## [31] R6_2.5.1         fansi_0.5.0       vroom_1.5.5
## [34] tzdb_0.1.2       readr_2.0.2       purrr_0.3.4
## [37] scales_1.1.1     codetools_0.2-18  htmltools_0.5.2
## [40] ellipsis_0.3.2   colorspace_2.0-2  renv_0.14.0
## [43] utf8_1.2.2       stringi_1.7.5     munsell_0.5.0
## [46] crayon_1.4.1
```