AGENDA - MACHINE LEARNING

Machine Learning (DSC 305) Spring 2023

- Download clean.R from GitHub (bit.ly/nb_clean)
- Run the file on the shell (M-x eshell) as a batch job $\frac{1}{2}$:

```
R CMD BATCH clean.R
ls -al .RData
```

• Load the .RData file in your current R session and check that packages and user-defined objects were loaded:

```
load(".RData")
search()
ls()
```

```
[1] ".GlobalEnv"
                           "ESSR"
                                                 "package:stats"
 [4] "package:graphics"
                           "package:grDevices" "package:utils"
[7] "package:datasets"
[10] "package:methods"
[1] "all_msg"
                                                "package:httr"
                           "package:stringr"
                                                 "package:base"
                           "Autoloads"
                                 "all_msg_chars"
 [3] "api_key"
                                 "ask_chatgpt"
                                 "chardonnay_df"
 [5] "chardonnay corpus"
                                 "chardonnay_src"
 [7] "chardonnay m"
 [9] "chardonnay_tdm"
                                 "chardonnay vec"
[11] "clean chardonnay"
                                 "clean chardonnay corpus"
[13] "clean coffee"
                                 "clean coffee corpus"
[15] "coffee corpus"
                                 "coffee df"
                                 "coffee_vec"
[17] "coffee_src"
[19] "color_pal"
                                 "ham"
[21] "load packages"
                                 "m"
[23] "sms corpus"
                                 "sms_corpus_clean"
[25] "sms_corpus_no_numbers"
                                 "sms_dtm"
[27] "sms_dtm_test"
                                 "sms_dtm_train"
[29] "sms dtm2"
                                  "sms_raw"
[31] "sms_test_labels"
                                 "sms_train_labels"
[33] "spam"
                                 "spam_ham"
[35] "spam_ham_prop"
                                 "spam_ham_prop_2"
[37] "stops"
                                 "string"
[39] "stringX"
                                 "term_frequency"
[41] "terms_vec"
                                  "tokens"
[43] "word freq"
```

• If we don't finish with a session, save your data from now on:

```
save.image(file=".RData")
shell("ls -al .RData")
```

```
-rwx----+ 1 Birkenkrahe LYONNET+Group(513) 3246618 Mar 31 10:47 .RData
```

• Now: continue with your practice file 5_naive_bayes_practice.org after adding the next (and final) chapters from here: bit.ly/nb final

Footnotes:

 $\frac{1}{2}$ It is possible that the existence of an .RData file converts warnings about packages you're trying to load into errors, which stops the packages from being loaded. In this case *erase* the .RData file and restart R.

Author: MARCUS BIRKENKRAHE (pledged)

Created: 2023-03-31 Fri 10:47