

R Notebook

Paschalis Agapitos

Introduction

Aim of this notebook is to validate the methods used in the paper *A Stylometric Analysis of Seneca's Disputed Plays: Authorship Verification of Octavia and Hercules Oetaeus*.

Our dataset includes the following authors and texts (we have removed the disputed plays by Seneca since these will part of the main analysis and not of the validation of the methods used):

- Marcus Annaeus Lucanus (i.e., Lucan):
 - *Pharsalia* (each book splitted into a separate txt file (10 in total))
- Marcus Manilius (i.e., Manilius):
 - *Astronomica* (each book splitted into a separate txt file (5 in total))
- Publius Ovidius Naso (i.e., Ovid):
 - *Amores*
 - *Ars Amatoria*
 - *Epistulae*
 - *Fasti*
 - *Ibis*
 - *Medicamina faciei femineae*
 - *Ex Ponto*
 - *Remedia Amoris*
 - *Tristia*
- Aulus Persius Flaccus (i.e., Persius):
 - *Satires* (each satire splitted into a separate txt file (6 in total))
- Gaius Julius Phaedrus (i.e., Phaedrus):
 - *Fabulae* (each book splitted into a separate txt file (6 in total))
- Lucius Annaeus Seneca Minor (i.e., Seneca the Younger):
 - *Agamemnon*
 - *Hercules Furens*
 - *Medea*
 - *Oedipus*
 - *Phaedra*
 - *Phoenissae*
 - *Thyestes*
 - *Troades*
- Silius Italicus (i.e., Silius):
 - *Punica* (each book splitted into a separate txt file (17 in total))
- Publius Papinius Statius (i.e., Statius):
 - *Achilleid*
 - *Silvae* (each book splitted into a separate txt file (5 in total))
 - *Thebaid* (each book splitted into a separate txt file (12 in total))
- Valerius Flaccus:
 - *Argonautica* (each book splitted into a separate txt file (8 in total))

In this notebook we will apply the `imposters()` method to every text of the dataset. To demonstrate that

the algorithm is not affected by the author name mentioned in the filename, we have manually changed some of the filenames; we also include two “tricky” texts to test the limits of this method. The texts from which the name have been manually changed are the following: * *Medicamina faciei femineae* by Ovid (i.e., `unknown_medicam.txt`) * *Satire 4* by Persius (i.e., `unknown_sat_4.txt`) * *Punica* book 6 by Silius Italicus (i.e., `unknown_pun_6.txt`)

The first to “unknown” texts are the two shortest compared to the other texts of the corpus. These are the two “tricky” cases in an effort to test the limits of the `imposters()` method.

The same pre-processing step will be applied to every text that is used here and the results will be generated using Most Frequent Characters (henceforward MFCs) 4grams and pentagrams; these number of n-grams will be applied to each one of the aforementioned methods and their variations.

```
set.seed(100) # random seed for reproducibility
```

Importing the libraries

```
library(stylo)

##
## ### stylo version: 0.7.4 ###
##
## If you plan to cite this software (please do!), use the following reference:
##   Eder, M., Rybicki, J. and Kestemont, M. (2016). Stylometry with R:
##   a package for computational text analysis. R Journal 8(1): 107-121.
##   <https://journal.r-project.org/archive/2016/RJ-2016-007/index.html>
##
## To get full BibTeX entry, type: citation("stylo")
library(writexl)
```

Setting working directory

The first step that we need to take is to set the working directory to the folder where this notebook and our dataset are saved. In our case `Analysis` is our working directory.

```
setwd("../validation_imposters/")
getwd()
```

```
## [1] "/Users/paschalis/Documents/MA_DH/Thesis/seneca_stylometry/validation/validation_imposters"
```

Preparation of the dataset

The “preparation of the dataset” step consists of six steps.

- 1) importing the corpus; at this step we need to make sure that we point out to the correct directory.
- 2) tokenizing the corpus, lowering the case of the letters and removing punctuation marks. A very crucial detail on which we should pay attention is the the aspect of the language. Our dataset is written in Latin, so we could either use the `Latin` or `Latin.corr` in the parameter `lang`. Since a lot of manuscripts in Latin do not distinguish the letter `v` from the letter `u`, the second option suits better in this case.
- 3) removing the pronouns; this step is optional but since some studies have shown that pronouns are related to the genre or content of a text, it was decided for them to be removed.
- 4) extracting the features that we want to use; in our case character tetra-grams and penta-grams will be used.
- 5) creating the frequency list

6) creating the table with frequencies using the frequency list that was created in the previous step.

```
# step 1
raw.corpus <- load.corpus(
  files = "all",
  corpus.dir = "validation_imp_corpus/",
  encoding = "UTF-8"
)

# step 2
tokenized.corpus <- txt.to.words.ext(
  raw.corpus,
  corpus.lang = "Latin.corr",
  preserve.case = F,
)

# step 3
corpus.no.pronouns <- delete.stop.words(
  tokenized.corpus,
  stop.words = stylo.pronouns(corpus.lang = "Latin.corr")
)

# step 4
corpus.char.tetragrams <- txt.to.features(
  tokenized.text = corpus.no.pronouns,
  features = "c",
  ngram.size = 4
)

# step 5
features_char_tetragrams <- make.frequency.list(
  corpus.char.tetragrams,
  head = 3000
)

# step 6
data = make.table.of.frequencies(
  corpus = corpus.char.tetragrams,
  features = features_char_tetragrams,
  relative = T
)

## processing 88 text samples
## .....
## combining frequencies into a table...
```

Implementation of imposters method

In order to implement `imposters` to the data that we have prepared in the previous steps we first need to find in which rows the disputed texts are. This will help us when we need to split the dataset into reference and test set; we could do that just by referring to the number of the row where the texts under investigation are. Since our dataset contains more than 999 rows we need to set the option for printing more than 999 rows. We have 1814 text samples, so if we set `max.option = 2000` it will output all the rows.

```
options(max.print=100)
rownames(data)
```

```
## [1] "luc_phars_1.txt"      "luc_phars_10.txt"     "luc_phars_2.txt"
## [4] "luc_phars_3.txt"      "luc_phars_4.txt"      "luc_phars_5.txt"
## [7] "luc_phars_6.txt"      "luc_phars_7.txt"      "luc_phars_8.txt"
## [10] "luc_phars_9.txt"      "manil_astro_1.txt"     "manil_astro_2.txt"
## [13] "manil_astro_3.txt"     "manil_astro_4.txt"     "manil_astro_5.txt"
## [16] "ovid_am.txt"           "ovid_ars.txt"          "ovid_epist.txt"
## [19] "ovid_fasti.txt"        "ovid_ibis.txt"         "ovid_meta.txt"
## [22] "ovid_ponto.txt"        "ovid_remed.txt"        "ovid_tristia.txt"
## [25] "persius_sati_1.txt"     "persius_sati_2.txt"     "persius_sati_3.txt"
## [28] "persius_sati_5.txt"     "persius_sati_6.txt"     "phaed_fables_1.txt"
## [31] "phaed_fables_2.txt"     "phaed_fables_3.txt"     "phaed_fables_4.txt"
## [34] "phaed_fables_5.txt"     "phaed_fables_6.txt"     "sen_ag.txt"
## [37] "sen_her_f.txt"          "sen_med.txt"           "sen_oed.txt"
## [40] "sen_phaed.txt"          "sen_phoen.txt"         "sen_thy.txt"
## [43] "sen_tro.txt"           "sil.ita_pun_1.txt"      "sil.ita_pun_10.txt"
## [46] "sil.ita_pun_11.txt"     "sil.ita_pun_12.txt"     "sil.ita_pun_13.txt"
## [49] "sil.ita_pun_14.txt"     "sil.ita_pun_15.txt"     "sil.ita_pun_16.txt"
## [52] "sil.ita_pun_17.txt"     "sil.ita_pun_2.txt"      "sil.ita_pun_3.txt"
## [55] "sil.ita_pun_4.txt"      "sil.ita_pun_5.txt"      "sil.ita_pun_7.txt"
## [58] "sil.ita_pun_8.txt"      "sil.ita_pun_9.txt"      "stat_achill.txt"
## [61] "stat_silv_1.txt"        "stat_silv_2.txt"        "stat_silv_3.txt"
## [64] "stat_silv_4.txt"        "stat_silv_5.txt"        "stat_theb_1.txt"
## [67] "stat_theb_10.txt"       "stat_theb_11.txt"       "stat_theb_12.txt"
## [70] "stat_theb_2.txt"        "stat_theb_3.txt"        "stat_theb_4.txt"
## [73] "stat_theb_5.txt"        "stat_theb_6.txt"        "stat_theb_7.txt"
## [76] "stat_theb_8.txt"        "stat_theb_9.txt"        "unknown_medicam.txt"
## [79] "unknown_pun_6.txt"      "unknown_sati_4.txt"     "valflac_argon_1.txt"
## [82] "valflac_argon_2.txt"    "valflac_argon_3.txt"    "valflac_argon_4.txt"
## [85] "valflac_argon_5.txt"    "valflac_argon_6.txt"    "valflac_argon_7.txt"
## [88] "valflac_argon_8.txt"
```

```
# create an empty dataframe where we are going to save the results
results.imposters <- data.frame()

# define the test text as the current row
# apply the imposters method to compare the test text against the reference set
# every time imposters is applied to a different text in the corpus
# the disputed Senecan plays (i.e., Octavia and Hercules Oetaeus have been manually removed)
for (n in 1:nrow(data)) {
  test <- data[n, 1:1500]

  # define the reference set as all rows except for the current row
  reference.set <- data[-n, 1:1500]

  # set a random seed
  set.seed(n)

  # apply the imposters method to compare the test text against the reference set
  # and extract the maximum result for the current test text
  result <- max(summary(imposters(reference.set = reference.set,
                                   test = test,
                                   iterations = 100,
                                   distance = 'eder'))))
```

```

# append the result to the results dataframe
results.imposters <- rbind(results.imposters,
                           data.frame(document = rownames(data)[n],
                                       result = result))
}

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    1
## manil   0
## ovid    0.03
## persius 0
## phaed   0
## sen     0
## sil.ita 0.01
## stat    0
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    1
## manil   0
## ovid    0.04
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...

```

```

## luc    1
## manil   0
## ovid    0.02
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0
## unknown 0
## valflac 0

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    1
## manil   0
## ovid    0.08
## persius 0
## phaed   0
## sen     0
## sil.ita 0.01
## stat    0
## unknown 0
## valflac 0

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    1
## manil   0
## ovid    0.02
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0

```

```

## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  1
## manil  0
## ovid  0.06
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0.88
## manil  0
## ovid  0.55
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0.02
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  1
## manil  0

```

```

## ovid      0.04
## persius   0
## phaed     0
## sen       0
## sil.ita   0
## stat      0
## unknown   0
## valflac   0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##

## Testing a given candidate against imposters...
## luc  1
## manil  0
## ovid   0.09
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0
## unknown 0
## valflac 0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##

## Testing a given candidate against imposters...
## luc  0.8
## manil  0
## ovid   0.69
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0.01
## unknown 0
## valflac 0

```



```

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc   0.03
## manil   1
## ovid   0.12
## persius  0
## phaed   0
## sen    0
## sil.ita  0
## stat   0.01
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc   0
## manil   1
## ovid   0.05
## persius  0
## phaed   0
## sen    0
## sil.ita  0
## stat   0.02
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc   0.02
## manil   1
## ovid   0.14
## persius  0

```

```

## phaed      0
## sen        0
## sil.ita    0
## stat       0
## unknown    0
## valflac    0.01
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc      0.05
## manil     0.98
## ovid      0.32
## persius   0
## phaed     0
## sen        0
## sil.ita   0
## stat      0.03
## unknown   0
## valflac   0.02
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc      0.07
## manil     1
## ovid      0.5
## persius   0
## phaed     0
## sen        0
## sil.ita   0.01
## stat      0.1
## unknown   0
## valflac   0.04
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac

```

```

##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    1
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    1
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0.01
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    1
## persius 0
## phaed   0
## sen     0

```

```

## sil.ita  0
## stat     0
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  1
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0.01
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  1
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0.01
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...

```

```

## luc    0.06
## manil   0
## ovid    1
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0.05
## unknown 0
## valflac 0

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    0.01
## manil   0
## ovid    1
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0
## unknown 0
## valflac 0

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc     0
## manil   0
## ovid    1
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0

```

```

## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  1
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.69
## persius  0.27
## phaed  0.01
## sen  0.03
## sil.ita  0
## stat  0.44
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0

```

```

## ovid      0.09
## persius   0.17
## phaed     0
## sen       0.12
## sil.ita   0
## stat      0.95
## unknown   0
## valflac   0.01
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0.04
## manil  0
## ovid  0.43
## persius  0.03
## phaed  0
## sen  0.28
## sil.ita  0.01
## stat  0.86
## unknown  0
## valflac  0.01
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0.13
## manil  0
## ovid  0.62
## persius  0
## phaed  0
## sen  0.06
## sil.ita  0
## stat  0.82
## unknown  0
## valflac  0

```

```

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc   0
## manil   0
## ovid    0.06
## persius 0.32
## phaed   0
## sen    0.59
## sil.ita 0
## stat    0.65
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc   0
## manil   0
## ovid    0.11
## persius 0
## phaed   1
## sen    0.03
## sil.ita 0
## stat    0
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc   0
## manil   0
## ovid    0.05
## persius 0

```



```

## phaed      1
## sen        0
## sil.ita    0
## stat       0
## unknown    0
## valflac    0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc        0
## manil      0
## ovid       0.09
## persius    0
## phaed      1
## sen        0.02
## sil.ita    0
## stat       0
## unknown    0
## valflac    0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc        0.01
## manil      0
## ovid       0.17
## persius    0
## phaed      1
## sen        0.04
## sil.ita    0
## stat       0.01
## unknown    0
## valflac    0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac

```

```

##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.02
## persius 0
## phaed   1
## sen     0.03
## sil.ita 0
## stat    0
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.17
## persius 0
## phaed   1
## sen     0.05
## sil.ita 0
## stat    0.01
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.02
## persius 0
## phaed   0
## sen     1

```

```

## sil.ita  0
## stat     0
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.04
## persius  0
## phaed  0
## sen  1
## sil.ita  0
## stat  0.01
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0.01
## manil  0
## ovid  0
## persius  0
## phaed  0
## sen  1
## sil.ita  0
## stat  0
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...

```

```

## luc    0
## manil   0
## ovid    0.03
## persius 0
## phaed   0
## sen     1
## sil.ita 0
## stat    0.01
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0.01
## manil   0
## ovid    0
## persius 0
## phaed   0
## sen     1
## sil.ita 0
## stat    0.01
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.02
## persius 0
## phaed   0
## sen     1
## sil.ita 0
## stat    0

```

```

## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.01
## persius  0
## phaed  0
## sen  1
## sil.ita  0
## stat  0.01
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0
## persius  0
## phaed  0
## sen  1
## sil.ita  0
## stat  0
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0

```

```

## ovid      0
## persius   0
## phaed     0
## sen       0
## sil.ita   1
## stat      0
## unknown   0.04
## valflac   0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc      0
## manil     0
## ovid      0
## persius   0
## phaed     0
## sen       0
## sil.ita   1
## stat      0
## unknown   0
## valflac   0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc      0
## manil     0
## ovid      0
## persius   0
## phaed     0
## sen       0
## sil.ita   0.99
## stat      0
## unknown   0.04
## valflac   0

```

```

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0
## persius 0
## phaed   0
## sen     0
## sil.ita 1
## stat    0
## unknown 0.01
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.28
## persius 0
## phaed   0
## sen     0
## sil.ita 0.89
## stat    0
## unknown 0.13
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0
## persius 0

```

```

## phaed      0
## sen        0
## sil.ita    1
## stat       0
## unknown    0
## valflac    0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc      0
## manil    0
## ovid     0
## persius  0
## phaed    0
## sen      0
## sil.ita  0.99
## stat     0
## unknown  0.01
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc      0
## manil    0
## ovid     0
## persius  0
## phaed    0
## sen      0
## sil.ita  1
## stat     0
## unknown  0.02
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac

```



```

##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0
## persius 0
## phaed   0
## sen     0
## sil.ita 1
## stat    0
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0
## persius 0
## phaed   0
## sen     0
## sil.ita 1
## stat    0
## unknown 0.08
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0
## persius 0
## phaed   0
## sen     0

```

```

## sil.ita  1
## stat      0.01
## unknown   0.02
## valflac   0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0
## persius  0
## phaed  0
## sen  0
## sil.ita  1
## stat  0
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0
## persius  0
## phaed  0
## sen  0
## sil.ita  1
## stat  0
## unknown  0.03
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...

```

```

## luc    0
## manil   0
## ovid    0
## persius 0
## phaed   0
## sen     0
## sil.ita 0.99
## stat    0
## unknown 0.11
## valflac 0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##

## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0
## persius 0
## phaed   0
## sen     0
## sil.ita 0.98
## stat    0.01
## unknown 0.01
## valflac 0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##

## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0
## persius 0
## phaed   0
## sen     0
## sil.ita 1
## stat    0

```

```

## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.84
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0.57
## unknown  0
## valflac  0.18
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.29
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0.98
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0

```

```

## ovid      0.09
## persius   0
## phaed     0
## sen       0
## sil.ita   0
## stat      1
## unknown   0
## valflac   0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc      0
## manil     0
## ovid      0.03
## persius   0
## phaed     0
## sen       0
## sil.ita   0
## stat      1
## unknown   0
## valflac   0.03

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc      0
## manil     0
## ovid      0.03
## persius   0
## phaed     0
## sen       0
## sil.ita   0
## stat      1
## unknown   0
## valflac   0.01

```

```

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.44
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0.93
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.05
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    1
## unknown 0
## valflac 0.01
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.01
## persius 0

```

```

## phaed      0
## sen        0
## sil.ita    0
## stat       1
## unknown    0
## valflac    0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc      0
## manil     0
## ovid      0.01
## persius   0
## phaed     0
## sen       0
## sil.ita   0
## stat      1
## unknown   0
## valflac   0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc      0
## manil     0
## ovid      0.07
## persius   0
## phaed     0
## sen       0
## sil.ita   0
## stat      0.99
## unknown   0
## valflac   0
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac

```

```

##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.02
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    1
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    1
## unknown 0
## valflac 0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.03
## persius 0
## phaed   0
## sen     0

```



```

## sil.ita  0
## stat     1
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.04
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  1
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.08
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0.99
## unknown  0
## valflac  0.01
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...

```

```

## luc    0
## manil   0
## ovid    0.03
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    1
## unknown 0
## valflac 0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##

## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.02
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    1
## unknown 0
## valflac 0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##

## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.19
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0.97

```

```

## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  1
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.01
## persius  0
## phaed  0
## sen  0
## sil.ita  1
## stat  0
## unknown  0
## valflac  0
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0

```

```

## ovid      0.09
## persius   0.99
## phaed     0.18
## sen       0.01
## sil.ita   0
## stat      0.02
## unknown   0
## valflac   0

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.54
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0.84
## unknown  0
## valflac  0.48

## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc  0
## manil  0
## ovid  0.09
## persius  0
## phaed  0
## sen  0
## sil.ita  0
## stat  0.12
## unknown  0
## valflac  1

```

```

## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc   0
## manil   0
## ovid    0.06
## persius  0
## phaed   0
## sen     0
## sil.ita  0
## stat    0.33
## unknown  0
## valflac  0.92
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc   0
## manil   0
## ovid    0.11
## persius  0
## phaed   0
## sen     0
## sil.ita  0
## stat    0.22
## unknown  0
## valflac  0.98
## No candidate set specified; testing the following classes (one at a time):
##   luc   manil   ovid   persius   phaed   sen   sil.ita   stat   unknown   valflac
##
## Testing a given candidate against imposters...
## luc   0
## manil   0
## ovid    0.04
## persius  0

```

```

## phaed      0
## sen        0
## sil.ita    0
## stat       0.75
## unknown    0
## valflac    0.62
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc        0
## manil      0
## ovid       0.24
## persius    0
## phaed      0
## sen        0
## sil.ita    0
## stat       0.43
## unknown    0
## valflac    0.75
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac
##
## Testing a given candidate against imposters...
## luc        0
## manil      0
## ovid       0.75
## persius    0
## phaed      0
## sen        0
## sil.ita    0
## stat       0.06
## unknown    0
## valflac    0.63
## No candidate set specified; testing the following classes (one at a time):
##  luc  manil  ovid  persius  phaed  sen  sil.ita  stat  unknown  valflac

```

```
##
## Testing a given candidate against imposters...
## luc    0
## manil   0
## ovid    0.7
## persius 0
## phaed   0
## sen     0
## sil.ita 0
## stat    0.22
## unknown 0
## valflac 0.66
```

```
# check if the dataframe is empty
if (nrow(results.imposters) == 0) {
  print("No results found.")
} else {
  # view the results data frame
  print(results.imposters)
}
```

```
##          document result
## 1    luc_phars_1.txt  1.00
## 2    luc_phars_10.txt 1.00
## 3    luc_phars_2.txt  1.00
## 4    luc_phars_3.txt  1.00
## 5    luc_phars_4.txt  1.00
## 6    luc_phars_5.txt  1.00
## 7    luc_phars_6.txt  0.88
## 8    luc_phars_7.txt  1.00
## 9    luc_phars_8.txt  1.00
## 10   luc_phars_9.txt  0.80
## 11   manil_astro_1.txt 1.00
## 12   manil_astro_2.txt 1.00
## 13   manil_astro_3.txt 1.00
## 14   manil_astro_4.txt 0.98
## 15   manil_astro_5.txt 1.00
## 16       ovid_am.txt  1.00
## 17       ovid_ars.txt  1.00
## 18       ovid_epist.txt 1.00
## 19       ovid_fasti.txt 1.00
## 20       ovid_ibis.txt  1.00
## 21       ovid_meta.txt  1.00
## 22       ovid_ponto.txt 1.00
## 23       ovid_remed.txt 1.00
## 24       ovid_tristia.txt 1.00
## 25   persius_sati_1.txt 0.69
## 26   persius_sati_2.txt 0.95
## 27   persius_sati_3.txt 0.86
```

```
## 28 persius_sati_5.txt 0.82
## 29 persius_sati_6.txt 0.65
## 30 phaed_fables_1.txt 1.00
## 31 phaed_fables_2.txt 1.00
## 32 phaed_fables_3.txt 1.00
## 33 phaed_fables_4.txt 1.00
## 34 phaed_fables_5.txt 1.00
## 35 phaed_fables_6.txt 1.00
## 36      sen_ag.txt 1.00
## 37      sen_her_f.txt 1.00
## 38      sen_med.txt 1.00
## 39      sen_oed.txt 1.00
## 40      sen_phaed.txt 1.00
## 41      sen_phoen.txt 1.00
## 42      sen_thy.txt 1.00
## 43      sen_tro.txt 1.00
## 44 sil.ita_pun_1.txt 1.00
## 45 sil.ita_pun_10.txt 1.00
## 46 sil.ita_pun_11.txt 0.99
## 47 sil.ita_pun_12.txt 1.00
## 48 sil.ita_pun_13.txt 0.89
## 49 sil.ita_pun_14.txt 1.00
## 50 sil.ita_pun_15.txt 0.99
## [ reached 'max' / getOption("max.print") -- omitted 38 rows ]

# write the data frame into an excel spreadsheet
# write_xlsx(results imposters, "../validation imposters/results imposters_test.xlsx")
```

Optimize

The `imposters` method comes with an in-built function that optimizes the hyperparameters used in the General Imposters method. In other words it tries to define the grey area where these result could be valid or not.

```
# help("imposters.optimize")
set.seed(1)
imposters.optimize(data)
```

```
## luc_phars_1.txt ...
## luc_phars_10.txt ...
## luc_phars_2.txt ...
## luc_phars_3.txt ...
## luc_phars_4.txt ...
## luc_phars_5.txt ...
## luc_phars_6.txt ...
## luc_phars_7.txt ...
## luc_phars_8.txt ...
## luc_phars_9.txt ...
## manil_astro_1.txt ...
```


manil_astro_2.txt ...
manil_astro_3.txt ...
manil_astro_4.txt ...
manil_astro_5.txt ...
ovid_am.txt ...
ovid_ars.txt ...
ovid_epist.txt ...
ovid_fasti.txt ...
ovid_ibis.txt ...
ovid_meta.txt ...
ovid_ponto.txt ...
ovid_remed.txt ...
ovid_tristia.txt ...
persius_sati_1.txt ...
persius_sati_2.txt ...
persius_sati_3.txt ...
persius_sati_5.txt ...
persius_sati_6.txt ...
phaed_fables_1.txt ...
phaed_fables_2.txt ...
phaed_fables_3.txt ...
phaed_fables_4.txt ...
phaed_fables_5.txt ...
phaed_fables_6.txt ...
sen_ag.txt ...
sen_her_f.txt ...
sen_med.txt ...
sen_oed.txt ...
sen_phaed.txt ...
sen_phoen.txt ...
sen_thy.txt ...
sen_tro.txt ...
sil.ita_pun_1.txt ...
sil.ita_pun_10.txt ...
sil.ita_pun_11.txt ...
sil.ita_pun_12.txt ...

sil.ita_pun_13.txt ...
sil.ita_pun_14.txt ...
sil.ita_pun_15.txt ...
sil.ita_pun_16.txt ...
sil.ita_pun_17.txt ...
sil.ita_pun_2.txt ...
sil.ita_pun_3.txt ...
sil.ita_pun_4.txt ...
sil.ita_pun_5.txt ...
sil.ita_pun_7.txt ...
sil.ita_pun_8.txt ...
sil.ita_pun_9.txt ...
stat_achill.txt ...
stat_silv_1.txt ...
stat_silv_2.txt ...
stat_silv_3.txt ...
stat_silv_4.txt ...
stat_silv_5.txt ...
stat_theb_1.txt ...
stat_theb_10.txt ...
stat_theb_11.txt ...
stat_theb_12.txt ...
stat_theb_2.txt ...
stat_theb_3.txt ...
stat_theb_4.txt ...
stat_theb_5.txt ...
stat_theb_6.txt ...
stat_theb_7.txt ...
stat_theb_8.txt ...
stat_theb_9.txt ...
unknown_medicam.txt ...
unknown_pun_6.txt ...
unknown_sati_4.txt ...
valflac_argon_1.txt ...
valflac_argon_2.txt ...
valflac_argon_3.txt ...

```
## valflac_argon_4.txt ...
## valflac_argon_5.txt ...
## valflac_argon_6.txt ...
## valflac_argon_7.txt ...
## valflac_argon_8.txt ...
## [1] 0.34 0.64
```