04_figure1_flowchart

randy

2023-10-24

flowchart for data cleaning

```
flowchart1 <- DiagrammeR::grViz("
digraph graph2 {

graph [layout = dot]

# node definitions with substituted label text
node [shape = rectangle, width = 4, fillcolor = Biege]
a [label = 'Raw Data: \n total number of children \n registered in the EPIC dataset \n(n = 1772)']
b [label = 'Eligible Subjects: \n  (n = 1370)']
c [label = 'Working Data: \n with all eligible subjects \n randomly split dataset into two subsets']
d [label = 'Training Data:\n serving as a pool for selecting \n the set of information donors \n (n = 9:
e [label = 'Testing Data:\n serving as targets for prediction validation \n and evaluationfor model per:

 c -> {d e}

# These are the main nodes at top of graph
  a -> b

  [label =
  <
        Subjects Excluded:<br ALIGN = 'LEFT'/>
          &#8226; No missing baseline variables<br ALIGN = 'LEFT'/>
           &#8226; Minimal 10 times of visits<br ALIGN = 'LEFT'/>
          &#8226; Minimal 5 years of following up<br ALIGN = 'LEFT'/>
           &#8226; Minimal age of 4 registerd in the study<br ALIGN = 'LEFT'/>
          (n = 402) <br ALIGN = 'LEFT'/>
    >
  ]

 b -> c

 [label =
 <
        Data Cleaning:<br ALIGN = 'LEFT'/>
               &#8226; Remove the time points before age 3 <br ALIGN = 'LEFT'/>
                &#8226; Realign with the earliest age <br ALIGN = 'LEFT'/>
                &#8226; Covert age into time variable<br ALIGN = 'LEFT'/>
  >
 ]
```
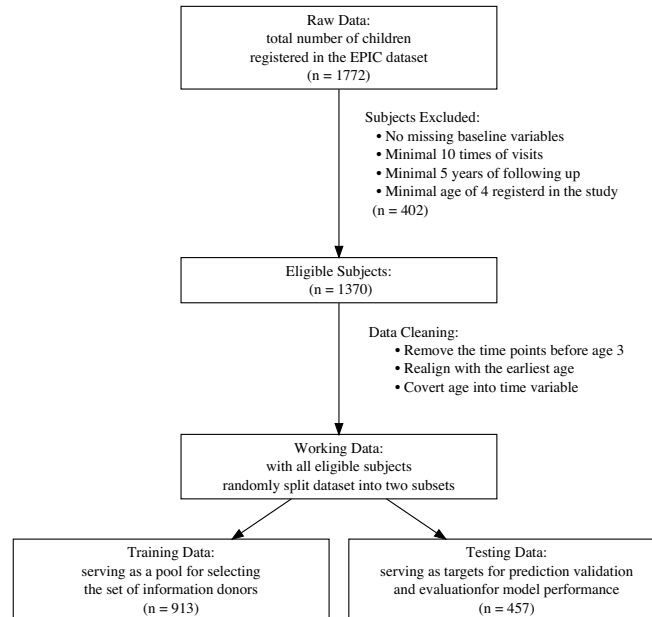
```
}

")
```

```
## see the flowchart
flowchart1
```



```
# flowchart1 %>%
#   export_svg() %>%
#   read_xml() %>%
#   write_xml(paste0("~/Desktop/project/plmlmm/paper/figure/00_plmlmm_data_cleaning", Sys.Date(), ".svg

## save the flowchart
flowchart1 %>%
  export_svg() %>%
  charToRaw() %>%
  rsvg_png(paste0("figure/S04_plmlmm_data_cleaning_", Sys.Date(), ".png"))
# export_graph(flowchart,
# file_name = "final/flowchart_graph.png",
# file_type = "png")
```

flowchart for algorithm

```
flowchart2 <- DiagrammeR::grViz(
"digraph graph2 {
```

```
graph [layout = dot]

# node definitions with substituted label text
node [shape = rectangle, width = 4, fillcolor = Biege]
a [label =
  <
    <b> </b> <br ALIGN = 'CENTER'/>
    <b>Imputing: </b> <br ALIGN = 'CENTER'/>
      &#8226; Brokenstick model prediction <br ALIGN = 'LEFT'/>
    >
  ]

b [label =
  <
    <b> </b> <br ALIGN = 'CENTER'/>
    <b>Linear Model: </b> <br ALIGN = 'CENTER'/>
      &#8226; Linear model prediction with the imputes<br ALIGN = 'LEFT'/>
    >
  ]

c [label =
 <
   <b> </b> <br ALIGN = 'CENTER'/>
   <b>Matching: </b> <br ALIGN = 'CENTER'/>
    &#8226; Single-time matching <br ALIGN = 'LEFT'/>
      &#8226; Multiple-time matching<br ALIGN = 'LEFT'/>
      &#8226; Matching criterion<br ALIGN = 'LEFT'/>
  >
 ]


d [label =
 <
   <b> </b> <br ALIGN = 'CENTER'/>
   <b>Predicting: </b> <br ALIGN = 'CENTER'/>
    &#8226; Median for predictive trajectory <br ALIGN = 'LEFT'/>
      &#8226; Centiles for predictive intervals <br ALIGN = 'LEFT'/>
  >
 ]

# These are the main nodes at top of graph

  a -> b -> c -> d
}
")


## see the flowchart
flowchart2
```
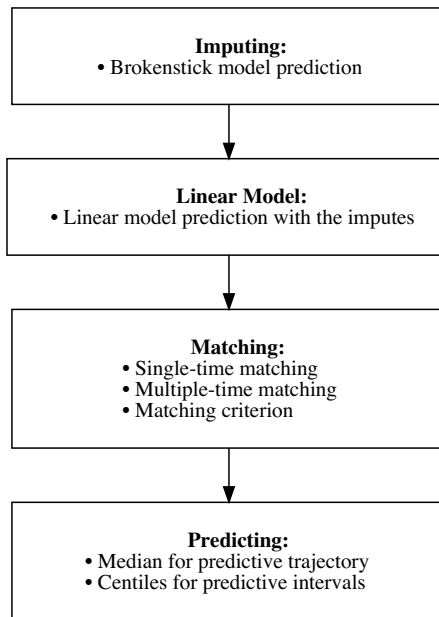
```
Imputing:
• Brokenstick model prediction
```

```
Linear Model:
• Linear model prediction with the imputes
```

```
Matching:
• Single-time matching
• Multiple-time matching
• Matching criterion
```

```
Predicting:
• Median for predictive trajectory
• Centiles for predictive intervals
```

```r
# flowchart1 %>%
#   export_svg() %>%
#   read_xml() %>%
#   write_xml(paste0("~/Desktop/project/plmlmm/paper/figure/00_plmlmm_data_cleaning", Sys.Date(), ".svg

## save the flowchart
flowchart2 %>%
  export_svg() %>%
  charToRaw() %>%
  rsvg_png(paste0("figure/S04_plmlmm_step_algorithm_", Sys.Date(), ".png"))
# export_graph(flowchart,
# file_name = "final/flowchart_graph.png",
# file_type = "png")
```

flowchart for upgraded phd work

```r
# install.packages("DiagrammeR")
flowchart3 <- DiagrammeR::grViz(
"digraph graph2 {

graph [layout = dot]
edge [color = gray]

# node definitions with substituted label text
graph [nodesep = 0.2, ranksep = 0.3]
node [shape = rectangle, width = 3, fillcolor = Biege]
imputation [label =
  <
```

```
    <b> </b> <br ALIGN = 'CENTER'/>
    <b>1. Imputing: </b> <br ALIGN = 'LEFT'/>
      &#8226; Brokenstick model prediction <br ALIGN = 'LEFT'/>
       &#8226; Linear model prediction<br ALIGN = 'LEFT'/>
    >;
    fontcolor = '#5d5d5d'
  ]
  imputation [color = gray]

distance [label =
 <
   <b> </b> <br ALIGN = 'CENTER'/>
   <b>2. Matching: </b> <br ALIGN = 'LEFT'/>
     &#8226; Single-time matching <br ALIGN = 'LEFT'/>
      &#8226; Multiple-time matching<br ALIGN = 'LEFT'/>
      &#8226; Matching criterion<br ALIGN = 'LEFT'/>
  >;
  fontcolor = '#5d5d5d'
 ]
 distance [color = gray]

prediction [label =
 <
   <b> </b> <br ALIGN = 'CENTER'/>
   <b>3. Predicting: </b> <br ALIGN = 'LEFT'/>
     &#8226; Median for predictive trajectory <br ALIGN = 'LEFT'/>
      &#8226; Centiles for predictive intervals <br ALIGN = 'LEFT'/>
  >;
  fontcolor = '#5d5d5d'
 ]
 prediction [color = gray]

 update_imputation [label =
 <
   <b> </b> <br ALIGN = 'LEFT'/>
   <b>INLA Smoothing: </b> <br ALIGN = 'LEFT'/>
     Random Walk (RW)
     <br ALIGN = 'LEFT'/>
     Autoregression (AR) <br ALIGN = 'LEFT'/>
   <br ALIGN = 'LEFT'/>
   <b>  Kernel Smoothing: </b> <br ALIGN = 'LEFT'/>
     Kernel regression
     <br ALIGN = 'LEFT'/>
   <br ALIGN = 'LEFT'/>
   <b>  Multiple LM: </b> <br ALIGN = 'LEFT'/>
     LM at each anchor time point \n<br ALIGN = 'LEFT'/>
  >
 ]

 update_distance [label =
 <
   <b> </b> <br ALIGN = 'CENTER'/>
   <b>Inversed Distance Weights: </b> <br ALIGN = 'LEFT'/>
```

```
      For Euclidean distance and single predictive matching
    <br ALIGN = 'LEFT'/>
    <br ALIGN = 'LEFT'/>
    <b>  Chi-squared Weights: </b> <br ALIGN = 'LEFT'/>
      For Mahalanobis distance matching <br ALIGN = 'LEFT'/>
  >
 ]

 update_model [label =
 <
   <b> </b> <br ALIGN = 'CENTER'/>
   <b>Weighted Regression: </b> <br ALIGN = 'LEFT'/>
     Using inversed distance weights or Chi-squared p-values
   <br ALIGN = 'LEFT'/>
   <br ALIGN = 'LEFT'/>
   <b>  Penalized Regression: </b> <br ALIGN = 'LEFT'/>
     Lasso or Ridge regression Penalization on the Distance
   <br ALIGN = 'LEFT'/>
  >
 ]

# These are the main nodes at top of graph


  imputation -> update_imputation
  {rank = same; imputation; update_imputation}


  distance -> update_distance
  {rank = same; distance; update_distance}


  prediction -> update_model
  {rank = same; prediction; update_model}

  update_imputation -> update_distance -> update_model [color = black]
}
")
## see the flowchart
flowchart3
```
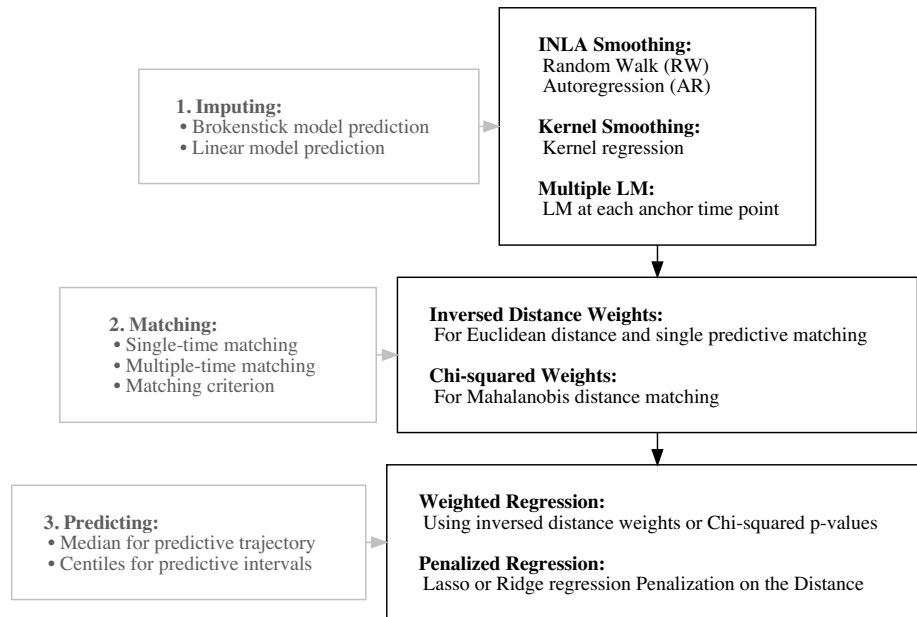
```
# flowchart1 %>%
#   export_svg() %>%
#   read_xml() %>%
#   write_xml(paste0("~/Desktop/project/plmlmm/paper/figure/00_plmlmm_data_cleaning", Sys.Date(), ".svg

## save the flowchart
flowchart3 %>%
  export_svg() %>%
  charToRaw() %>%
  rsvg_png(paste0("figure/S04_plmlmm_updated_work_", Sys.Date(), ".png"))
# export_graph(flowchart,
# file_name = "final/flowchart_graph.png",
# file_type = "png")
```

```
## R version 4.2.2 (2022-10-31)
## Platform: aarch64-apple-darwin20 (64-bit)
## Running under: macOS 14.0
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] DiagrammeRsvg_0.1 DiagrammeR_1.0.10 xml2_1.3.5        rsvg_2.4.0
```

```
##  [5] lubridate_1.9.2    forcats_1.0.0     stringr_1.5.0       dplyr_1.1.2
##  [9] purrr_1.0.1        readr_2.1.4       tidyr_1.3.0         tibble_3.2.1
## [13] ggplot2_3.4.3      tidyverse_2.0.0   here_1.0.1
##
## loaded via a namespace (and not attached):
##  [1] tidyselect_1.2.0   xfun_0.39         V8_4.3.2            colorspace_2.1-0
##  [5] vctrs_0.6.3        generics_0.1.3    htmltools_0.5.5     yaml_2.3.7
##  [9] utf8_1.2.3         rlang_1.1.1       pillar_1.9.0        glue_1.6.2
## [13] withr_2.5.0        RColorBrewer_1.1-3 lifecycle_1.0.3    munsell_0.5.0
## [17] gtable_0.3.3       visNetwork_2.1.2  htmlwidgets_1.6.2   evaluate_0.21
## [21] knitr_1.43         tzdb_0.4.0        callr_3.7.3         fastmap_1.1.1
## [25] ps_1.7.5           curl_5.0.1        fansi_1.0.4         highr_0.10
## [29] Rcpp_1.0.11        scales_1.2.1      webshot_0.5.5       jsonlite_1.8.7
## [33] freshr_1.0.2       hms_1.1.3         digest_0.6.33       stringi_1.7.12
## [37] processx_3.8.2     grid_4.2.2        rprojroot_2.0.3     cli_3.6.1
## [41] tools_4.2.2        magrittr_2.0.3    pkgconfig_2.0.3     ellipsis_0.3.2
## [45] timechange_0.2.0   rmarkdown_2.23    rstudioapi_0.15.0   R6_2.5.1
## [49] compiler_4.2.2
```