



# Machine Learning, Visualization, and Agentic AI for all with rtemis

Warsaw, 2026-07-09

**Stathis Gennatas, MBBS AICSM PhD**

**Assoc. Professor**

Div. Bioinformatics, Dept. Epidemiology & Biostatistics

Div. Clinical Informatics, Dept. Medicine

UCSF & UC Berkeley Computational Precision Health

University of California, San Francisco

**Co-Director**

UCSF Scientific Software Core

[Stathis.Gennatas@ucsf.edu](mailto:Stathis.Gennatas@ucsf.edu) | [rtemis.org](https://rtemis.org)



Bio





# Bio

- **Medical School;**  
Imperial College London



# Bio

- **Medical School;**  
Imperial College London
- **Asst. Specialist, Neuroimaging;**  
Memory & Aging Center, UCSF





# Bio

- **Medical School;**  
Imperial College London
- **Asst. Specialist, Neuroimaging;**  
Memory & Aging Center, UCSF
- **Neuroscience PhD;**  
University of Pennsylvania



# Bio

- Medical School;  
Imperial College London

- Asst. Specialist, Neuroimaging;  
Memory & Aging Center, UCSF

- Neuroscience PhD;  
University of Pennsylvania



Many large studies, one consistent bottleneck:  
Application of predictive modeling  
requires quantitative + technical expertise;  
existing packages are not user-friendly,  
lack features  
→ rtemis 0.0.1



Bio





# Bio

- **Medical School;**  
Imperial College London
- **Asst. Specialist, Neuroimaging;**  
Memory & Aging Center, UCSF
- **Neuroscience PhD;**  
University of Pennsylvania
- **Asst. Prof. Researcher, ML for Medicine;**  
Medical Physics, Radiation Oncology, UCSF





# Bio

- **Medical School;**  
Imperial College London
- **Asst. Specialist, Neuroimaging;**  
Memory & Aging Center, UCSF
- **Neuroscience PhD;**  
University of Pennsylvania
- **Asst. Prof. Researcher, ML for Medicine;**  
Medical Physics, Radiation Oncology, UCSF
- **Research Scientist, ML for Medicine;**  
Laboratory of Artificial Intelligence for  
Medicine & Biomedical Physics, Stanford



# Bio

- **Medical School;**  
Imperial College London
- **Asst. Specialist, Neuroimaging;**  
Memory & Aging Center, UCSF
- **Neuroscience PhD;**  
University of Pennsylvania
- **Asst. Prof. Researcher, ML for Medicine;**  
Medical Physics, Radiation Oncology, UCSF
- **Research Scientist, ML for Medicine;**  
Laboratory of Artificial Intelligence for  
Medicine & Biomedical Physics, Stanford
- **Assoc. Professor, UCSF**
  - Div. Bioinformatics, Dept. Epidemiology & Biostatistics
  - Div. Clinical Informatics & Digital Transformation, Dept. Medicine
  - UCSF-UC Berkeley Computational Precision Health
- **Co-Director**
  - UCSF Scientific Software Core & OSPO
- **Affiliate**
  - Bakar Computational Health Sciences Institute
  - Center for Intelligent Imaging
  - Institute for Global Health Sciences, UCSF
  - Center for Targeted Machine Learning & Causal Inference, UC Berkeley
- **Co-lead**
  - FTD Center Without Walls Genomics & Transcriptomics Core
- **Instructor**
  - Datasci 213 & 214 Programming for Health Data Science in R I & II



# Bio

- **Medical School;**  
Imperial College London
- **Asst. Specialist, Neuroimaging;**  
Memory & Aging Center, UCSF
- **Neuroscience PhD;**  
University of Pennsylvania
- **Asst. Prof. Researcher, ML for Medicine;**  
Medical Physics, Radiation Oncology, UCSF
- **Research Scientist, ML for Medicine;**  
Laboratory of Artificial Intelligence for  
Medicine & Biomedical Physics, Stanford



Health sciences require  
validated + observable + auditable pipelines;  
low code + no-code access to  
advanced data/stat/ML/AI/visualization pipelines  
→ rtemis 1.0.0 + friends

- **Assoc. Professor, UCSF**
  - Div. Bioinformatics, Dept. Epidemiology & Biostatistics
  - Div. Clinical Informatics & Digital Transformation, Dept. Medicine
  - UCSF-UC Berkeley Computational Precision Health
- **Co-Director**
  - UCSF Scientific Software Core
- **Affiliate**
  - Bakar Computational Health Sciences Institute
  - Center for Intelligent Imaging
  - Institute for Global Health Sciences, UCSF
  - Center for Targeted Machine Learning & Causal Inference, UC Berkeley
- **Co-lead**
  - FTD Center Without Walls Genomics & Transcriptomics Core
- **Instructor**
  - Datasci 213 & 214 Programming for Health Data Science I & II





# Health Data Science





## Domain Expertise

Biomedical Science  
Clinical Medicine



Math  
Statistics  
Machine Learning  
**Quantitative Expertise**



Programming  
Databases  
Sys admin  
Web dev  
Cloud

## Technical Expertise

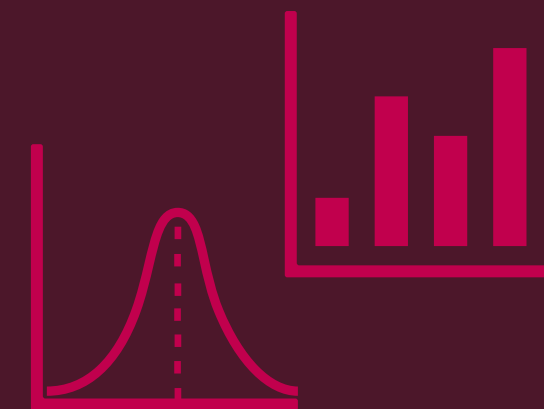


Biostatistics  
Epidemiology  
Public Health/  
Global Health



## Domain Expertise

Biomedical Science  
Clinical Medicine



Math  
Statistics  
Machine Learning  
**Quantitative Expertise**



Programming  
Databases  
Sys admin  
Web dev  
Cloud

## Technical Expertise





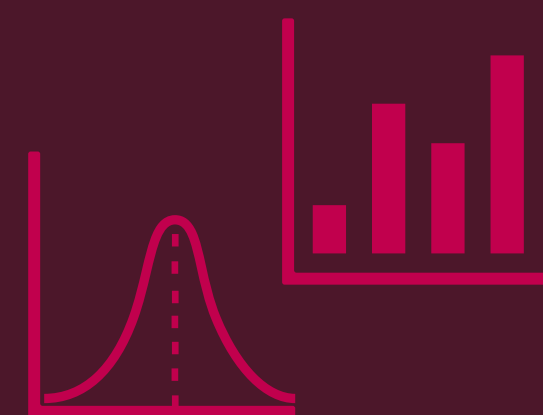
Biostatistics  
Epidemiology  
Public Health/  
Global Health



## Domain Expertise

Biomedical Science  
Clinical Medicine

Math  
Statistics  
Machine Learning  
**Quantitative Expertise**



Programming  
Databases  
Sys admin  
Web dev  
Cloud



## Technical Expertise





Biostatistics  
Epidemiology  
Public Health/  
Global Health

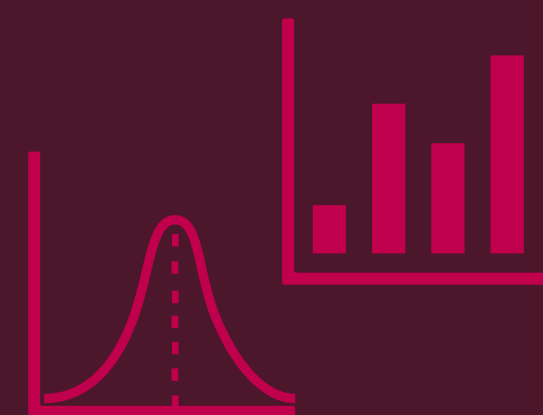
## Domain Expertise

Biomedical Science  
Clinical Medicine



\* Health Data Science

Math  
Statistics  
Machine Learning  
**Quantitative Expertise**



Programming  
Databases  
Sys admin  
Web dev  
Cloud

## Technical Expertise







Quantitative pipelines are often  
challenging:

- Inefficient +/- non-reproducible +/-  
not grounded on domain knowledge  
+/- incorrect







Quantitative pipelines are often  
challenging:

- Inefficient +/- non-reproducible +/-  
not grounded on domain knowledge  
+/- incorrect
- Domain expertise is essential but ML  
software is not accessible to most  
scientists/clinicians/public health  
practitioners.







Purpose-built health data science  
software incorporating  
**domain + quantitative + technical**  
expertise catalyzes collaboration,  
increases efficiency & reproducibility,  
improves results.



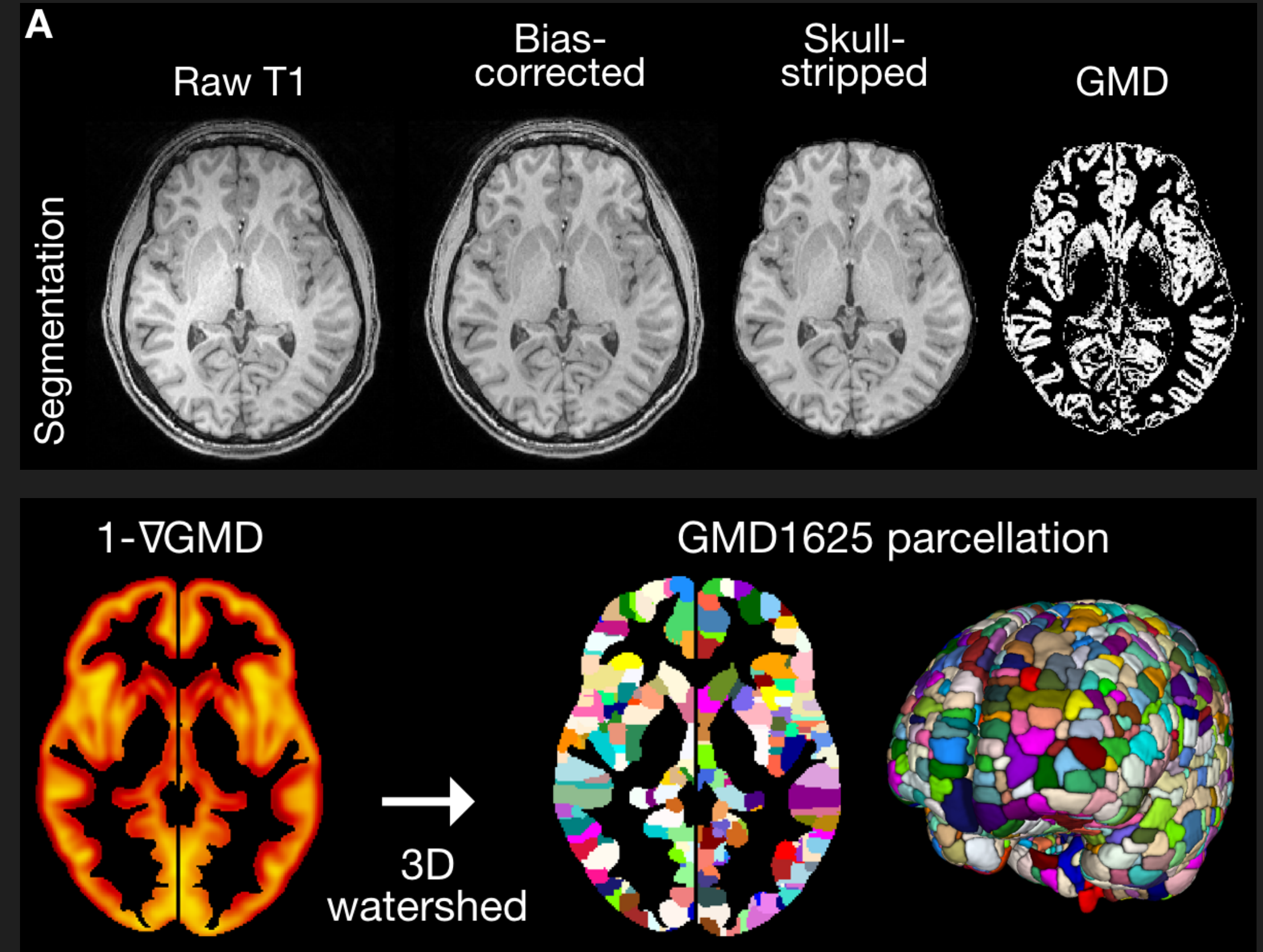


# Initial motivation (~2015)



# Initial motivation (~2015)

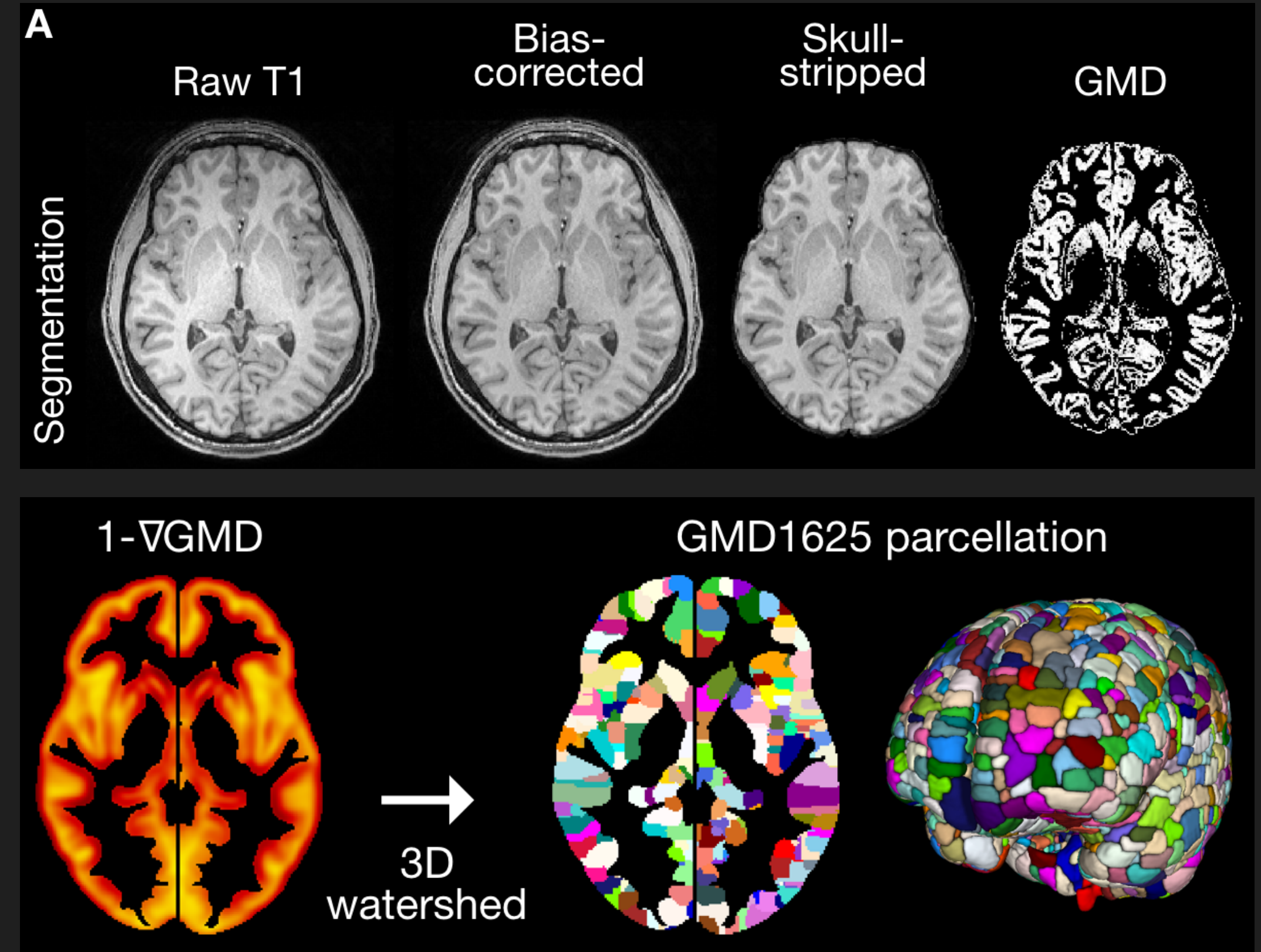
- Neuroimaging of brain development:  
~2k cases \* (1625 ROIs \* 4 modalities +  
demographic + neurocognitive data)





# Initial motivation (~2015)

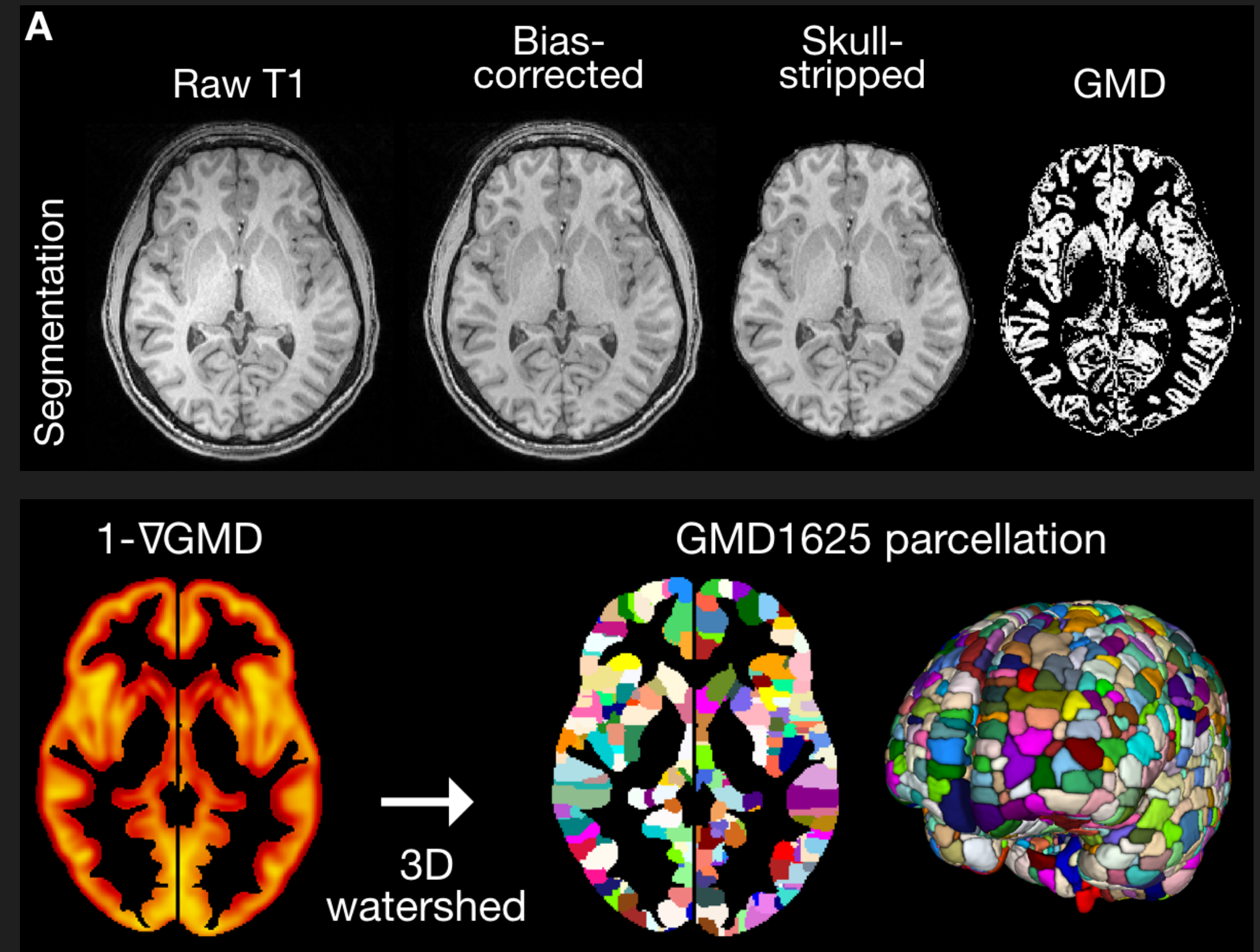
- Neuroimaging of brain development:  
~2k cases \* (1625 ROIs \* 4 modalities +  
demographic + neurocognitive data)
- Started with 2 goals:  
tune gradient boosting & sparse  
decompositions





# Initial motivation (~2015)

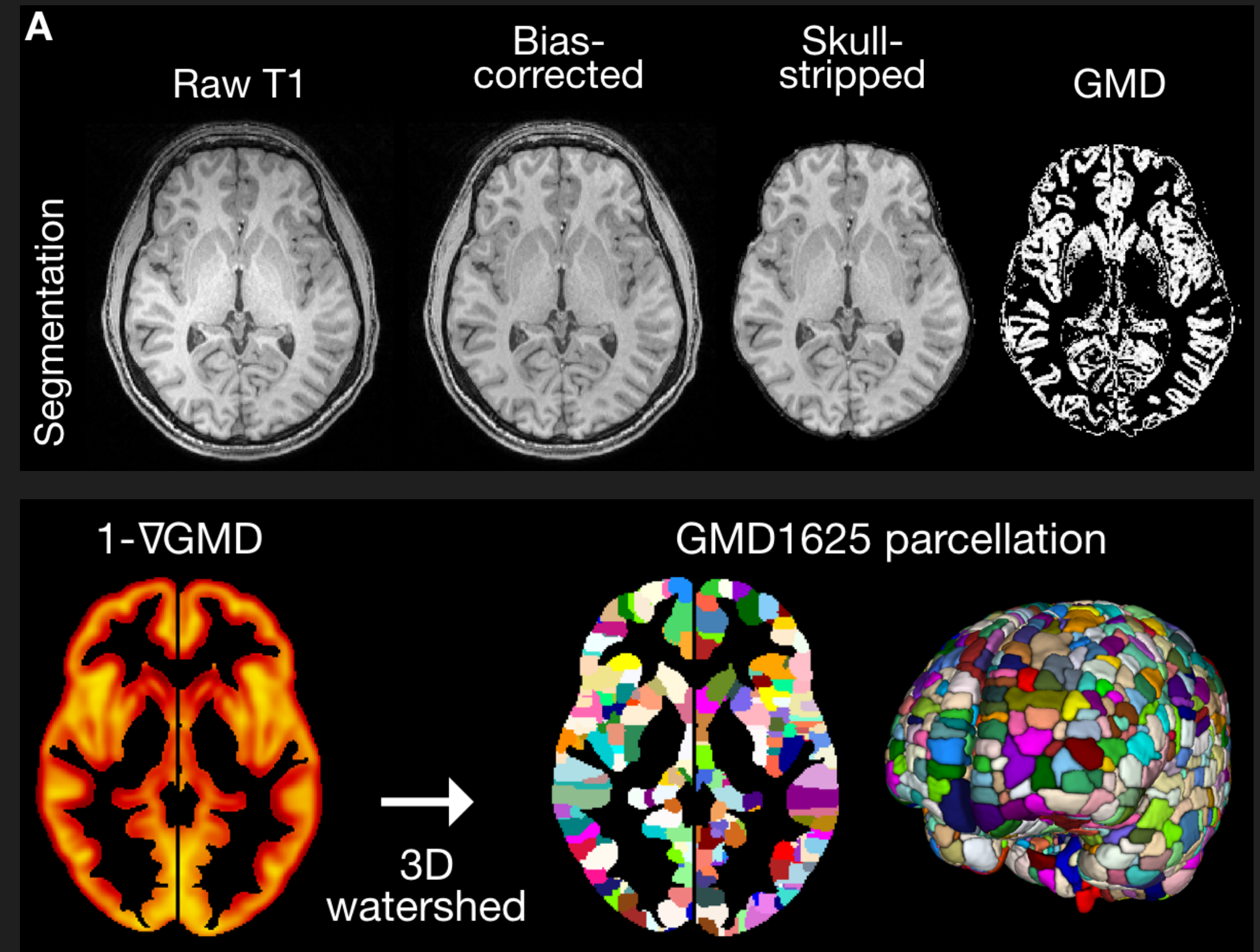
- Neuroimaging of brain development:  
~2k cases \* (1625 ROIs \* 4 modalities +  
demographic + neurocognitive data)
- Started with 2 goals:  
tune gradient boosting & sparse  
decompositions
- Existing packages did not support early  
stopping, performed suboptimal and highly  
inefficient tuning for gradient boosting





# Initial motivation (~2015)

- Neuroimaging of brain development:  
~2k cases \* (1625 ROIs \* 4 modalities +  
demographic + neurocognitive data)
- Started with 2 goals:  
tune gradient boosting & sparse  
decompositions
- Existing packages did not support early  
stopping, performed suboptimal and highly  
inefficient tuning for gradient boosting
- 2 functions → ~600 functions;  
1 package → 12 packages



# rtemis







# rtemis Mk I



# rtemis Mk I (< v0.9)





# rtemis Mk I (< v0.9)

- Learning R & machine learning





# rtemis Mk I (< v0.9)

- Learning R & machine learning
- Integrating functionality for supervised learning, decomposition, clustering and visualization, obsessing over console pretty printing



# rtemis Mk I (< v0.9)

- Learning R & machine learning
- Integrating functionality for supervised learning, decomposition, clustering and visualization, obsessing over console pretty printing
- Organizing growing number of scripts into an R package



# rtemis Mk I (< v0.9)

- Learning R & machine learning
- Integrating functionality for supervised learning, decomposition, clustering and visualization, obsessing over console pretty printing
- Organizing growing number of scripts into an R package
- Writing tutorials in Rmarkdown, later Quarto





# rtemis Mk I (< v0.9)

- Learning R & machine learning
- Integrating functionality for supervised learning, decomposition, clustering and visualization, obsessing over console pretty printing
- Organizing growing number of scripts into an R package
- Writing tutorials in Rmarkdown, later Quarto
- Tried all class systems, both to learn and to choose the best fit:  
S3, S4, RC, R6 - settled on R6



# rtemis Mk I (< v0.9)

- Learning R & machine learning
- Integrating functionality for supervised learning, decomposition, clustering and visualization, obsessing over console pretty printing
- Organizing growing number of scripts into an R package
- Writing tutorials in Rmarkdown, later Quarto
- Tried all class systems, both to learn and to choose the best fit: S3, S4, RC, R6 - settled on R6
- static (base graphics) and interactive (plotly) graphics systems



# rtemis Mk I (< v0.9)

- Learning R & machine learning
- Integrating functionality for supervised learning, decomposition, clustering and visualization, obsessing over console pretty printing
- Organizing growing number of scripts into an R package
- Writing tutorials in Rmarkdown, later Quarto
- Tried all class systems, both to learn and to choose the best fit: S3, S4, RC, R6 - settled on R6
- static (base graphics) and interactive (plotly) graphics systems
- First public release 0.79 on GitHub





The extensive low- to high-level tooling available in **rtemis**,  
makes it an ideal platform for developing & testing novel algorithms.



# Expert-augmented machine learning

Efstathios D. Gennatas<sup>a,1,2</sup>, Jerome H. Friedman<sup>b</sup>, Lyle H. Ungar<sup>c</sup>, Romain Pirracchio<sup>d</sup>, Eric Eaton<sup>c</sup>, Lara G. Reichmann<sup>e</sup>, Yannet Interian<sup>e</sup>, José Marcio Luna<sup>f</sup>, Charles B. Simone II<sup>g</sup>, Andrew Auerbach<sup>h</sup>, Elier Delgado<sup>i</sup>, Mark J. van der Laan<sup>j</sup>, Timothy D. Solberg<sup>a</sup>, and Gilmer Valdes<sup>a</sup>

<sup>a</sup>Department of Radiation Oncology, University of California, San Francisco, CA 94143; <sup>b</sup>Department of Statistics, Stanford University, Stanford, CA 94305; <sup>c</sup>Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104; <sup>d</sup>Department of Anesthesia and Perioperative Care, University of California, San Francisco, CA 94143; <sup>e</sup>Data Institute, University of San Francisco, CA 94105; <sup>f</sup>Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA 19104; <sup>g</sup>Department of Radiation Oncology, New York Proton Center, New York, NY 10035; <sup>h</sup>Division of Hospital Medicine, University of California, San Francisco, CA 94143; <sup>i</sup>Innova Montreal, Inc., Montreal, QC J4W 2P2, Canada; and <sup>j</sup>Division of Biostatistics, University of California, Berkeley, CA 94720

PNAS

## Building more accurate decision trees with the additive tree

José Marcio Luna<sup>a,1,2</sup>, Efstathios D. Gennatas<sup>b,1</sup>, Lyle H. Ungar<sup>c</sup>, Eric Eaton<sup>c</sup>, Eric S. Diffenderfer<sup>a</sup>, Shane T. Jensen<sup>d</sup>, Charles B. Simone II<sup>e</sup>, Jerome H. Friedman<sup>f,2</sup>, Timothy D. Solberg<sup>b</sup>, and Gilmer Valdes<sup>b,2</sup>

<sup>a</sup>Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA 19104; <sup>b</sup>Department of Radiation Oncology, University of California, San Francisco, CA 94115; <sup>c</sup>Department of Computing and Information Science, University of Pennsylvania, Philadelphia, PA 19104; <sup>d</sup>Department of Statistics, University of Pennsylvania, Philadelphia, PA 19104; <sup>e</sup>Department of Radiation Oncology, New York Proton Center, New York, NY 10035; and <sup>f</sup>Department of Statistics, Stanford University, Stanford, CA 94305

## Representational Gradient Boosting: Shallow learning with backpropagation in the space of functions.

Gilmer Valdes, Jerome H. Friedman, Fei Jiang, and Efstathios D. Gennatas

## The Conditional Super Learner

Gilmer Valdes\*, Yannet Interian\*, Efstathios Gennatas and Mark Van der Laan

## The Linear Additive Tree

**Efstathios D. Gennatas**  
*Department of Epidemiology and Biostatistics  
University of California  
San Francisco, CA 94143, USA*

EFSTATHIOS.GENNATAS@UCSF.EDU

**Jerome H. Friedman**  
*Department of Statistics  
Stanford University  
Palo Alto, CA 94305, USA*

JHF@STAT.STANFORD.EDU

**Eric Eaton**  
*Department of Computer and Information Science  
University of Pennsylvania  
Philadelphia, PA 19104, USA*

EEATON@CIS.UPENN.EDU

**Charles B. Simone, II**  
*Department of Radiation Oncology  
New York Proton Center  
New York, NY 10035, USA*

CSIMONE@NYPCC.ORG

**Lyle H. Ungar**  
*Department of Computer and Information Science  
University of Pennsylvania  
Philadelphia, PA 19104, USA*

UNGA@CIS.UPENN.EDU

**Lei Xing**  
*Department of Radiation Oncology  
Stanford University  
Palo Alto, CA 94305, USA*

LEI@STAT.STANFORD.EDU

**Gilmer Valdes**  
*Department of Radiation Oncology  
Department of Epidemiology and Biostatistics  
University of California  
San Francisco, CA 94143, USA*

GILMER.VALDES@UCSF.EDU



# rtemis Mk II





# rtemis Mk II



# rtemis Mk II

- January 2025: began a complete rewrite of backend + streamline public API



# rtemis Mk II

- January 2025: began a complete rewrite of backend + streamline public API
- Switch to S7 classes and methods throughout





# rtemis Mk II

- January 2025: began a complete rewrite of backend + streamline public API
- Switch to S7 classes and methods throughout
- Trimmed down considerably + broke into separate packages



# rtemis Mk II

- January 2025: began a complete rewrite of backend + streamline public API
- Switch to S7 classes and methods throughout
- Trimmed down considerably + broke into separate packages
- **rtemis** 1.0.0 on CRAN 2026-03-26: ML & visualization



# rtemis Mk II

- January 2025: began a complete rewrite of backend + streamline public API
- Switch to S7 classes and methods throughout
- Trimmed down considerably + broke into separate packages
- **rtemis** 1.0.0 on CRAN 2026-03-26: ML & visualization
- **rtemis.a3** 0.5.3 on CRAN 2026-04-28: Amino Acid Annotation (A3) format





# rtemis Mk II

- January 2025: began a complete rewrite of backend + streamline public API
- Switch to S7 classes and methods throughout
- Trimmed down considerably + broke into separate packages
- **rtemis** 1.0.0 on CRAN 2026-03-26: ML & visualization
- **rtemis.a3** 0.5.3 on CRAN 2026-04-28: Amino Acid Annotation (A3) format
- **rtemis.llm** 0.8.1 on CRAN 2026-05-04: LLMs and agents



# rtemis Mk II

- January 2025: began a complete rewrite of backend + streamline public API
- Switch to S7 classes and methods throughout
- Trimmed down considerably + broke into separate packages
- **rtemis** 1.0.0 on CRAN 2026-03-26: ML & visualization
- **rtemis.a3** 0.5.3 on CRAN 2026-04-28: Amino Acid Annotation (A3) format
- **rtemis.llm** 0.8.1 on CRAN 2026-05-04: LLMs and agents
- **rtemis.core** on CRAN 2026-05-09: Shared utilities: logging system, formatting system, progress reporting, custom S7 types, check\_\* functions, ...



# rtemis Mk II

- January 2025: began a complete rewrite of backend + streamline public API
- Switch to S7 classes and methods throughout
- Trimmed down considerably + broke into separate packages
- **rtemis** 1.0.0 on CRAN 2026-03-26: ML & visualization
- **rtemis.a3** 0.5.3 on CRAN 2026-04-28: Amino Acid Annotation (A3) format
- **rtemis.llm** 0.8.1 on CRAN 2026-05-04: LLMs and agents
- **rtemis.core** on CRAN 2026-05-09: Shared utilities: logging system, formatting system, progress reporting, custom S7 types, check\_\* functions, ...
- **rtemis.draw** on R-universe (coming soon to CRAN): high performance interactive graphics using JS libraries





## **rtemis Goal**

Make data science accessible to  
domain experts  
by minimizing technical hurdles  
& promote correctness,  
reproducibility, auditability



Quick dive: **Supervised pipeline**



|                       |                                                                                               |                |
|-----------------------|-----------------------------------------------------------------------------------------------|----------------|
| Data Input            | Input from all popular formats                                                                | > read()       |
| Data check            | Summary of types, issues, Preproc. recommendations                                            | > check_data() |
| Visualization         | Interactive graphics                                                                          | > draw()       |
| Data cleaning         | Feature names; categorical labels                                                             | > preprocess() |
| Variable coding       | Type conversion<br>categorical encoding                                                       |                |
| Handling missing data | Up/down-sampling<br>case weighting                                                            |                |
| Preproc +/- decomp    | Training set processing, & test set application                                               | > train()      |
| Tuning                | Inner resampling / validation set                                                             |                |
| Evaluation            | Outer resampling / test set                                                                   |                |
| Summary               | Plain language output                                                                         | > describe()   |
| Presentation          | Confusion matrix<br>ROC curve<br>PR curve<br>Performance metrics<br>Variable importance plots | > present()    |





```
rtemis::train()
```



```
dat ← read("path/to/data.csv")
```



```
dat ← read("path/to/data.csv")
```

*Reads delimited file formats,  
.arff, .parquet, .rds, .xlsx, .dta, .fasta  
Produces data.table, tibble, or data.frame*





```
mod ← train(dat)
```





```
mod ← train(dat)
```

*Train supervised model: last column is outcome, rest are features*  
*Defaults to ranger random forest*



```
mod ← train(  
  dat,  
  algorithm = "lightgbm"  
)
```



```
mod ← train(  
  dat,  
  algorithm = "lightgbm"  
)  
  
# same as  
  
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM()  
)
```





```
mod ← train(  
  dat,  
  algorithm = "lightgbm"  
)
```

# same as

```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM()  
)
```



```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM()  
)
```



```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM()  
)
```

*Automatic tuning of number of trees  
using early stopping*





```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM()  
)
```

*Automatic tuning of number of trees  
using early stopping*

...





```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM()  
)
```

*Automatic tuning of number of trees  
using early stopping*

...  
*(mission accomplished!)*



```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM()  
)
```



```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM()  
)
```

*Get metrics on train & test set*





```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM(max_depth = 3L)  
)
```



```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM(max_depth = 3L)  
)
```

*Define hyperparameters:*

*setup\_\* functions provide docs + tab-completion*



```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM(max_depth = 2:4)  
)
```





```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM(max_depth = 2:4)  
)
```

*Vector of values  $\Rightarrow$  automatic grid search*



```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM(max_depth = 2:4)  
)
```

*Vector of values  $\Rightarrow$  automatic grid search*

*$(3*5 + 1) = 16$  models*





```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01)  
  )  
)
```



```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01)  
  )  
)
```

*$(3*3*5 + 1) = 46$  models*



```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  )  
)
```





```
mod ← train(  
  dat_train,  
  test = dat_test,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  )  
)
```

*$(3*3*2*5 + 1) = 91$  models*



```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler()  
)
```



```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler()  
)
```

*Estimate generalizability using 10-fold stratified crossvalidation*





```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler()  
)
```

*Estimate generalizability using 10-fold stratified crossvalidation*

*$(3*3*2*5 + 1) * 10 = 910$  models*



```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(),  
  outdir = "path/to/outdir"  
)
```





```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(),  
  outdir = "path/to/outdir"  
)
```

*Save model RDS and log TXT to disk*



```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(),  
  outdir = "path/to/outdir"  
)
```

A very common setup:  
nested resampling + save to disk



```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(  
    n_resamples = 25L, type = "StratSub"),  
  outdir = "path/to/outdir"  
)
```





```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(  
    n_resamples = 25L, type = "StratSub"),  
  outdir = "path/to/outdir"  
)
```

*Estimate generalizability using 25 stratified subsamples*



```
mod ← train(  
  dat,  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(  
    n_resamples = 25L, type = "StratSub"),  
  outdir = "path/to/outdir"  
)
```

*Estimate generalizability using 25 stratified subsamples*

*$(3*3*2*5 + 1) * 25 = 2,275$  models*



```
mod ← train(  
  dat,  
  decomposition_config = setup_ICA(k = 3L),  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(  
    n_resamples = 25L, type = "StratSub"),  
  outdir = "path/to/outdir"  
)
```





```
mod ← train(  
  dat,  
  decomposition_config = setup_ICA(k = 3L),  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(  
    n_resamples = 25L, type = "StratSub"),  
  outdir = "path/to/outdir"  
)
```

*Decompose numeric variables to 3 ICA projections*



```
mod ← train(  
  dat,  
  preprocessor_config = setup_Preprocessor(  
    impute = TRUE  
  ),  
  decomposition_config = setup_ICA(k = 3L),  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(  
    n_resamples = 25L, type = "StratSub"),  
  outdir = "path/to/outdir"  
)
```





```
mod ← train(  
  dat,  
  preprocessor_config = setup_Preprocessor(  
    impute = TRUE  
  ),  
  decomposition_config = setup_ICA(k = 3L),  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(  
    n_resamples = 25L, type = "StratSub"),  
  outdir = "path/to/outdir"  
)
```



```
mod ← train(  
  dat,  
  preprocessor_config = setup_Preprocessor(  
    impute = TRUE  
  ),  
  decomposition_config = setup_ICA(k = 3L),  
  hyperparameters = setup_LightGBM(  
    max_depth = 2:4,  
    learning_rate = c(.001, .005, .01),  
    subsample = c(.5, .9)  
  ),  
  outer_resampling_config = setup_Resampler(  
    n_resamples = 25L, type = "StratSub"),  
  outdir = "path/to/outdir"  
)
```

*Impute missing data using iterative RF*



```
mod ← train(  
  dat          = x,  
  preprocessor_config = setup_Preprocessor(),  
  decomposition_config = setup_{Decomp_name}(),  
  hyperparameters   = setup_{Alg_name}(),  
  tuner_config       = setup_GridSearch(),  
  outer_resampling_config = setup_Resampler()  
)
```





```

mod ← train(
  dat                = x                : class_tabular,
  preprocessor_config = setup_Preprocessor() : PreprocessorConfig,
  decomposition_config = setup_{Decomp_name}() : DecompositionConfig,
  hyperparameters    = setup_{Alg_name}()   : Hyperparameters,
  tuner_config        = setup_GridSearch()   : TunerConfig,
  outer_resampling_config = setup_Resampler() : ResamplerConfig
)

```

What the use doesn't need to worry about: **S7 types**, **S7 classes** and their constructors.

Used internally throughout to perform type checking and validation of all inputs and fail fast with corrective error messages



```
● ● ● R
> library(rtemis)

rtemis
.:rtemis v.1.2.9 🌊 aarch64-apple-darwin23 (10 cores available)

Defaults
├ Theme: darkgraygrid
├ Font: Geist
└ Palette: rtms

Resources
├ Docs: https://docs.rtemis.org/r/ml
├ Learn R: https://pdsr.rtemis.org
└ Cite: > citation(rtemis)

PSA: Do not throw data at algorithms. Compute responsibly!

>
```



```
● ● ● R
> dat ← read("~/icloud/Data/iris.parquet")
2026-07-08 17:54:36 ► Reading iris.parquet using arrow... [read]
2026-07-08 17:54:36 Read in 150 x 5 [read]
2026-07-08 17:54:36 Done in 0.15 seconds. [read]
>
```





```
● ● ● R
> mod_lightgbm <- train(
+   dat,
+   hyperparameters = setup_LightGBM(
+     max_depth = c(2:4),
+     learning_rate = c(.001, .005, .01)
+   ),
+   outer_resampling_config = setup_Resampler(),
+   outdir = "./mod_lightgbm"
+ )
2026-07-09 03:33:06 Checking data is ready for training...
2026-07-09 03:33:06 ✖ rtemis_outcome_error [check_supervised]
Error in check_supervised() :
  Outcome must be integer, numeric, or factor; was character.
>
```



```
● ● ● R
> dat <- read(
+   "~/icloud/Data/iris.parquet",
+   remove_duplicates = TRUE,
+   character2factor = TRUE
+ )
2026-07-08 22:26:00 ► Reading iris.parquet using arrow... [read]
2026-07-08 22:26:00 Read in 150 x 5 [read]
2026-07-08 22:26:00 Removing 1 duplicate case... [preprocess]
2026-07-08 22:26:00 Converting 1 character feature to a factor... [preprocess]
2026-07-08 22:26:00 Preprocessing done. [preprocess]
2026-07-08 22:26:00 Done in 0.11 seconds. [read]
>
```





```
● ● ● ■ R
> mod_lightgbm <- train(
+   dat,
+   hyperparameters = setup_LightGBM(
+     max_depth = c(2:4),
+     learning_rate = c(.001, .005, .01)
+   ),
+   outer_resampling_config = setup_Resampler(),
+   outdir = "./mod_lightgbm"
+ )
2026-07-08 22:33:37 Checking data is ready for training... ✓ [check_supervised]
2026-07-08 22:33:37 ► [train]
2026-07-08 22:33:37 Training set: 149 cases x 4 features. [summarize_supervised]
2026-07-08 22:33:37 // Max workers: 7 { Algorithm: 7; Tuning: 1; Outer Resampling: 1 } [get_n_workers]
2026-07-08 22:33:37 ◀ Training LightGBM Classification using 10 independent folds... [train]
2026-07-08 22:33:37 Input contains more than one column; stratifying on last. [resample]
2026-07-08 22:33:37 Using max n bins possible = 3. [kfold]
⚡ Outer resamples 2/10 > Tuning 10/45 ETA 0:05
```



```
● ● ● R
> mod_lightgbm <- train(
+   dat,
+   hyperparameters = setup_LightGBM(
+     max_depth = c(2:4),
+     learning_rate = c(.001, .005, .01)
+   ),
+   outer_resampling_config = setup_Resampler(),
+   outdir = "./mod_lightgbm"
+ )
2026-07-09 03:45:38 Checking data is ready for training... ✓ [check_supervised]
2026-07-09 03:45:38 ► [train]
2026-07-09 03:45:38 Training set: 149 cases x 4 features. [summarize_supervised]
2026-07-09 03:45:38 // Max workers: 7 { Algorithm: 7; Tuning: 1; Outer Resampling: 1 } [get_n_workers]
2026-07-09 03:45:38 < Training LightGBM Classification using 10 independent folds... [train]
2026-07-09 03:45:38 Input contains more than one column; stratifying on last. [resample]
2026-07-09 03:45:38 Using max n bins possible = 3. [kfold]
2026-07-09 03:46:29 ✓ Outer resamples 10/10 × Tuning 45/45 done in 0:50
2026-07-09 03:46:29 <➤ Outer resampling done. [train]

<Resampled Classification Model>
LightGBM (Gradient Boosting)
• Tuned using exhaustive grid search.
c Tested using 10 independent folds.

<Resampled Classification Training Metrics>
Aggregate Confusion Matrix across resamples.

      Predicted
Reference setosa versicolor virginica
setosa      450          0          0
versicolor   1         428         21
virginica    0          7        434

Showing mean (sd) across resamples.
Balanced Accuracy: 0.978 (0.015)
F1: 0.978 (0.015)
Accuracy: 0.978 (0.015)

<Resampled Classification Test Metrics>
Aggregate Confusion Matrix across resamples.

      Predicted
Reference setosa versicolor virginica
setosa      50          0          0
versicolor   0         46          4
virginica    0          4         45

Showing mean (sd) across resamples.
Balanced Accuracy: 0.945 (0.053)
F1: 0.945 (0.053)
Accuracy: 0.946 (0.053)

2026-07-09 03:46:29 Writing data to ./mod_lightgbm... ✓ 0.44 secs [rt_save]
2026-07-09 03:46:29 Reload with: > obj <- readRDS('/Users/sdg/Code/rtemis/mod_lightgbm/train_LightGBM.rds') [rt_save]
2026-07-09 03:46:29 Done in 50.84 seconds. [train]
> █
```



● ● ● R

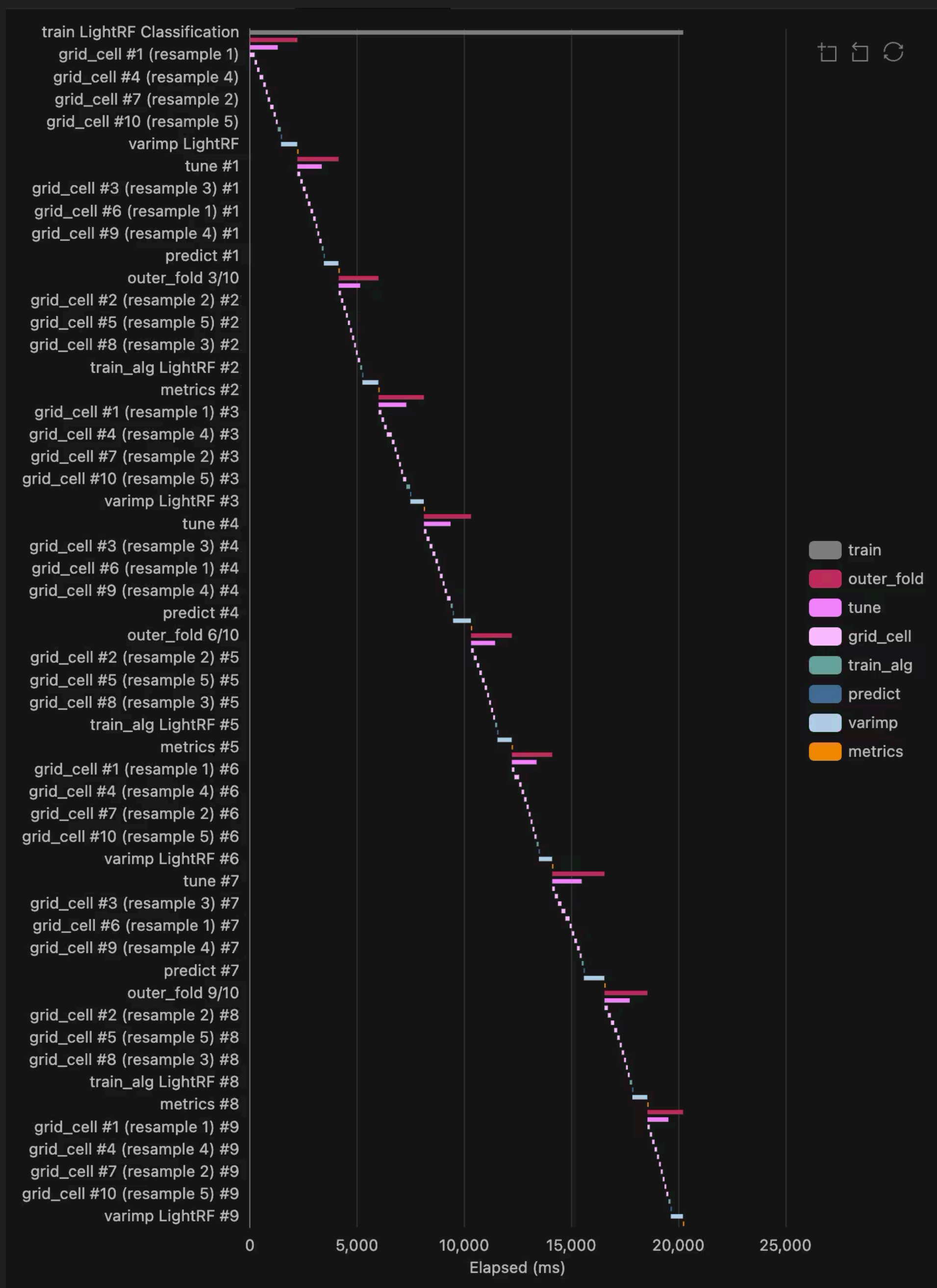
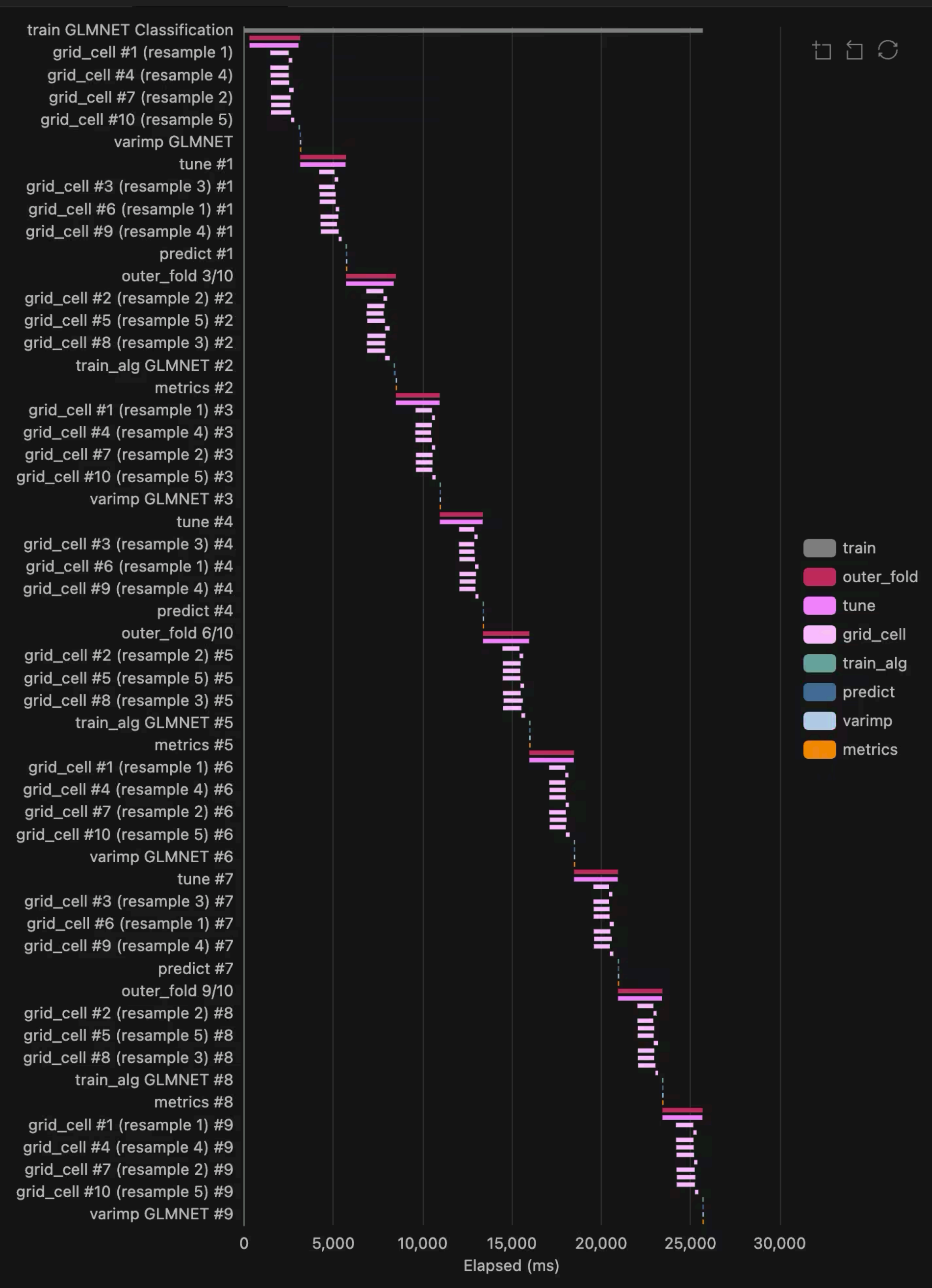
```
> describe(mod_lightgbm)
```

Gradient Boosting was used for classification. Hyperparameter tuning was performed using exhaustive grid search. Mean balanced accuracy was 0.98 in the training set and 0.92 in the test set across 10 independent folds.

```
>
```







plot\_session(mod)

Gantt chart of supervised pipeline





# Complex pipelines you can **trust**: validated, observable, auditable

- **Validated specification** (*what will happen*)  
Custom S7 classes & types throughout + function argument checking:  
fail fast with corrective messages  
`abort()`, `warn()`, `info()`, `check_*`
- **Observable execution** (*what is happening*)  
Custom logging & nested progress reporting `msg()`, `fmt()`, `progress_lapply()`
- **Auditable record** (*what happened*)  
Every step recorded (`SupervisedSession`), pretty printed (`repr()`), plotted (`plot_session()`)



# rtemis schemas



# rtemis schemas

- JSON schemas built from S7 classes
- Hosted on [schema.rtemis.org](https://schema.rtemis.org)
- `write_config()` S7  $\rightarrow$  JSON; `read_config()` JSON  $\rightarrow$  S7
- Used in R, rtemis CLI, rtemislive





# rtemis CLI





# rtemis-cli

- Rust binary
- Validates & describes rtemis config JSONs natively
- Executes by passing data and config to `rtemis::train`
  - `rtemis train dat.{csv,parquet} supervised.json`
  - `rtemis decompose dat.{csv,parquet} decomp.json`
  - `rtemis cluster dat.{csv,parquet} cluster.json`



```
sdg ~ % rtemis

RTEMIS

The rtemis command-line interface.

Usage: rtemis <COMMAND>

Commands:
  validate  Validate one or more config JSON files (use '-' or no path for stdin)
  describe  Pretty-print a config's plan (use '-' or no path for stdin)
  train     Train a supervised model from data and a SuperConfig (R backend)
  decompose Run a decomposition from data and a DecomposeConfig (R backend)
  cluster   Cluster data from a ClusterConfig (R backend)
  schemas   List the known config schemas
  help      Print this message or the help of the given subcommand(s)

Options:
  -h, --help      Print help
  -V, --version    Print version

sdg ~ %
```



```
sdg ~ % rtemis validate supervised.json
✓ valid SuperConfig v1 https://schema.rtemis.org/supervised/v1/schema.json
supervised.json
sdg ~ % |
```





```
sdg ~ % rtemis describe supervised.json

✓ valid SuperConfig v1 https://schema.rtemis.org/supervised/v1/schema.json
supervised.json

Data
└─ positive class case

Preprocessing
└─ remove_duplicates true

Decomposition PCA
├─ center true
├─ k 2
└─ scale true

Algorithm LightGBM
├─ learning_rate 0.1
└─ num_leaves 31

Tuning -

Outer resampling KFold
└─ n 5

Output -

sdg ~ %
```



**rtemis.llm**







 llm





# rtemis.llm

- rtemis-standard S7 backend with functional API
- Support for
  - Ollama, OpenAI-compatible, Anthropic-compatible providers
  - Capturing **reasoning** output
  - Defining **schemas** + generating **structured output**
  - Custom **tools**
  - **Memory** on/off



| Feature           | LLM | Agent |
|-------------------|-----|-------|
| Reasoning         | ✓   | ✓     |
| Structured Output | ✓   | ✓     |
| Tool use          | x   | ✓     |
| Memory            | x   | ✓     |
| Batch generation  | ✓   | ✓     |

Batch generation





# rtemis.llm

| Feature           | LLM | Agent |
|-------------------|-----|-------|
| Reasoning         | ✓   | ✓     |
| Structured Output | ✓   | ✓     |
| Tool use          | x   | ✓     |
| Memory            | x   | ✓     |
| Batch generation  | ✓   | ✓     |

Batch generation

```
llmapply(x, llm)
```



# rtemis.llm

| Feature           | LLM | Agent |
|-------------------|-----|-------|
| Reasoning         | ✓   | ✓     |
| Structured Output | ✓   | ✓     |
| Tool use          | x   | ✓     |
| Memory            | x   | ✓     |
| Batch generation  | ✓   | ✓     |

Batch generation

```
llmapply(x, llm)
```

```
agentapply(x, agent)
```



# rtemis.llm

| Feature           | LLM | Agent |
|-------------------|-----|-------|
| Reasoning         | ✓   | ✓     |
| Structured Output | ✓   | ✓     |
| Tool use          | x   | ✓     |
| Memory            | x   | ✓     |
| Batch generation  | ✓   | ✓     |

Batch generation

```
llmapply(x, llm)
```

```
agentapply(x, agent)
```

```
map(x, llm)
```





# rtemis.llm

| Feature           | LLM | Agent |
|-------------------|-----|-------|
| Reasoning         | ✓   | ✓     |
| Structured Output | ✓   | ✓     |
| Tool use          | x   | ✓     |
| Memory            | x   | ✓     |
| Batch generation  | ✓   | ✓     |

## Batch generation

```
llmapply(x, llm)
```

```
agentapply(x, agent)
```

```
map(x, llm)
```

```
map(x, agent)
```



**rtemis.a3**





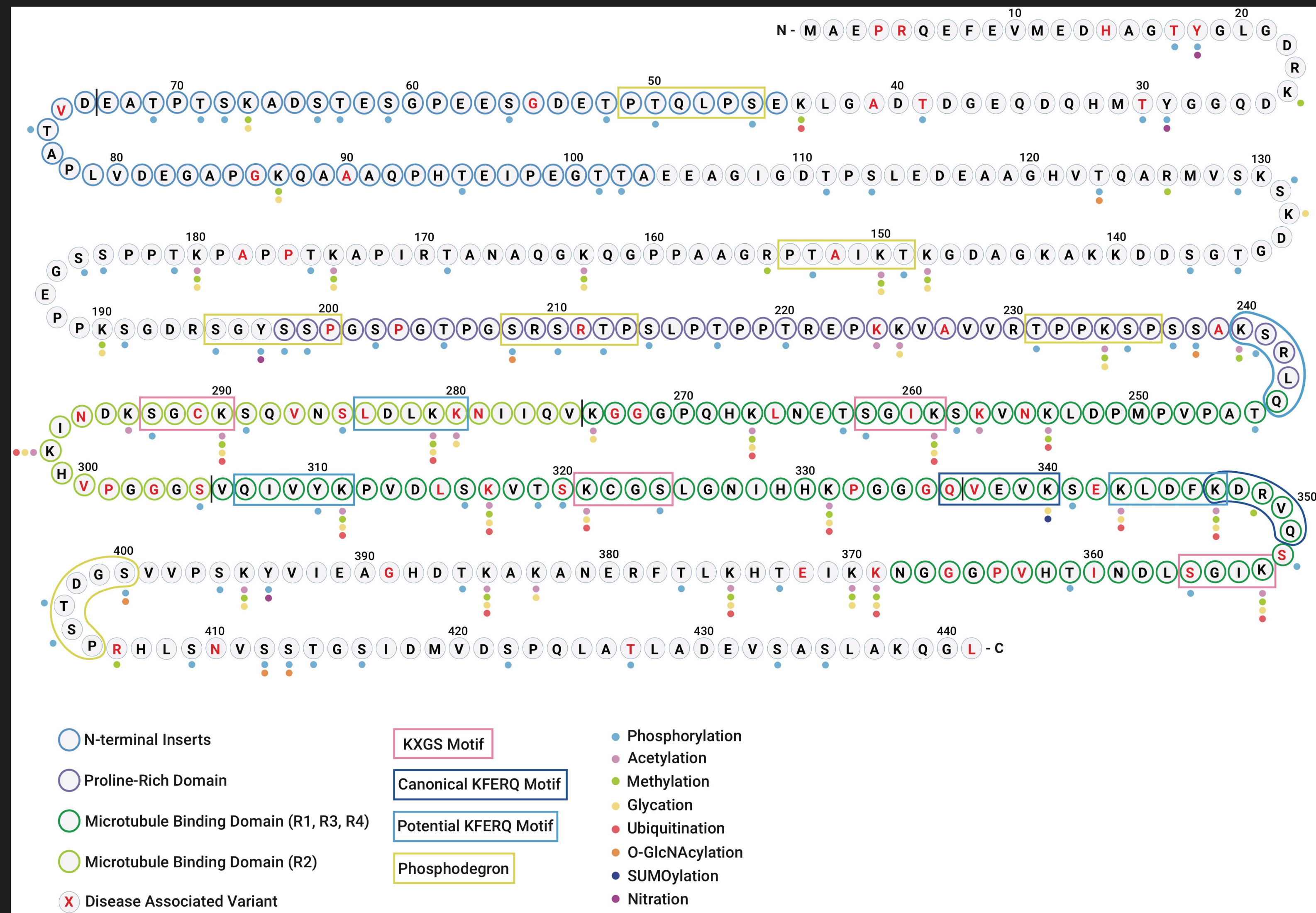


{A3}



Goal:  
Automate  
visualization of  
protein primary  
structure &  
annotations





Goal:  
Automate  
visualization of  
protein primary  
structure &  
annotations





# Amino Acid Annotation (A3) format

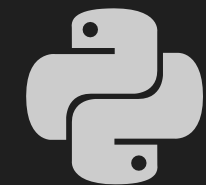
- rtemis infrastructure makes it easy to build domain-specific tools
- Defined JSON schema
- Created S7 class to hold amino acid sequence + rich set of annotations
- Read/write/validate functions
- Functions to import data from popular databases (e.g. UniProt) and convert to A3
- Custom visualization function



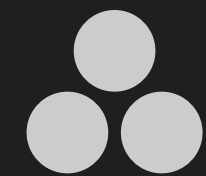
# S7 used as template to port across languages



[cran.r-project.org/package=rtemis.a3](https://cran.r-project.org/package=rtemis.a3)



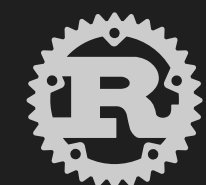
[pypi.org/project/rtemis-a3](https://pypi.org/project/rtemis-a3)



[juliahub.com/ui/Packages/General/RtemisA3](https://juliahub.com/ui/Packages/General/RtemisA3)



[npmjs.com/package/@rtemis/a3](https://npmjs.com/package/@rtemis/a3) & [jsr.io/@rtemis/a3](https://jsr.io/@rtemis/a3)



[crates.io/crates/rtemis-a3](https://crates.io/crates/rtemis-a3)



MAPT



- Primary structure
  - Disease-Associated Variant
  - N-terminal Repeat
  - Microtubule Binding Domain (R1, R3, R4)
  - Microtubule Binding Domain (R2)
- Regions**
- KXGS
  - Canonical KFERQ
  - Potential KFERQ
  - Phosphodegrom
- PTMs**
- Phosphorylation
  - Acetylation
  - Methylation
  - Glycation
  - Ubiquitination
  - O-Glycosylation
  - SUMOylation





**rtemislive**



# rtemislive Mk I

- Originally dataset-specific shiny apps with ability to visualize datasets, fit models++
- → `rtemislive()`: function that outputs custom shiny app for any dataset with an extensive set of modules
- Deployed at UCSF using shinyproxy
- (rtemislive name predated shinylive)
- shinylive failed with the massive rtemislive app



# rtemislive Mk II

- Rewrite using Next.js, DuckDB, {ECharts, Sigma.js, MapLibre}, WebR, Pyodide, AI-SDK, assistant-ui
- Data loading, cleaning, preprocessing, feature engineering
- Visualization using ECharts, Sigma.js, MapLibre
- Hypothesis testing using WebR
- Supervised Learning, Decomposition, Clustering using rtemis.server to connect to local R session and run rtemis pipelines





# Local, secure, in-browser analytics

- Support for tabular, hierarchical, A3 data: read, edit, write
- GUI & Console (WebR, Pyodide, DuckDB)
- High performance, interactive visualization (ECharts, Sigma.js, MapLibre)
- Hypothesis testing through WebR
- Supervised Learning, Decomposition, Clustering through rtemis.server (WebSocket) to local R session running rtemis
- Context-aware agentic Chat with custom tools, structured output, direct context injection, skills library, MCP server support can provide assistance at any step





Live

Data

View



Chat



# rtemislive

Integrated environment for data analysis,  
machine learning, and visualization.  
Catalyzing human × AI collaboration.

## DATA ?



### Load

From a file on your computer



### Enter manually

Type or paste data directly

## WORKSPACES ?

No saved workspaces yet.  
Use Live menu (top left) to save  
and manage workspaces.

## WALKTHROUGHS



### Tabular data

CSV, Parquet, JSON arrays



### Hierarchical data

Treemap and sunburst charts



### A3 data

Annotated protein sequen...

## EXAMPLE DATA



### Load example

Built-in sample datasets



This app runs entirely in your browser.  
Your datasets never leave your computer.

Toggle cloud services

Command palette /



Idle



Load data to begin.



© 2026 EDG



Live

Data

View

View

Check

Summarize

Plot

Test

Model

Chat

Data

Metadata

Select

Edit

Showing

100

of 132 columns

| #  | PIDN    | Dx          | Age_at_Scan | Age_at_First_Visit | Age_at_Onset | Sex         | Race          | Handedness  | Education_Level | APOE4   | CDR_Date            | CDR_Sum |
|----|---------|-------------|-------------|--------------------|--------------|-------------|---------------|-------------|-----------------|---------|---------------------|---------|
|    | numeric | categorical | numeric     | numeric            | numeric      | categorical | categorical   | categorical | numeric         | numeric | datetime            | numeric |
| 1  | 10002   | E0AD        | 66          | 64                 | 63           | Female      | Caucasian     | Right       | 18              | 1       | 2010-08-18 00:00:00 |         |
| 2  | 12552   | E0AD        | 59          | 59                 | 57           | Female      | Caucasian     | Right       | 14              | 2       | 2011-03-16 00:00:00 |         |
| 3  | 12709   | PCA         | 55          | 55                 | 50           | Female      | Not_caucasian | Right       | 16              | 1       | 2011-04-29 00:00:00 |         |
| 4  | 10931   | PCA         | 62          | 61                 | 58           | Female      | Caucasian     | Right       | 13              | 0       | 2011-06-08 00:00:00 |         |
| 5  | 8687    | E0AD        | 63          | 60                 | 50           | Female      | Caucasian     | Left        | 16              | 1       | 2011-04-22 00:00:00 |         |
| 6  | 11725   | LOAD        | 75          | 74                 | 72           | Male        | Not_caucasian | Right       | 12              | 1       | 2011-05-20 00:00:00 |         |
| 7  | 4670    | E0AD        | 69          | 63                 | 60           | Female      | Not_caucasian | Right       | 0               | 0       | 2011-08-09 00:00:00 |         |
| 8  | 12585   | E0AD        | 56          | 56                 | 52           | Male        | Caucasian     | Left        | 14              | 0       | 2011-12-22 00:00:00 |         |
| 9  | 14063   | PCA         | 79          | 79                 | 77           | Female      | Caucasian     | Right       | 8               | 1       | 2012-01-25 00:00:00 |         |
| 10 | 13520   | LOAD        | 80          | 80                 | 77           | Female      | Not_caucasian | Right       | 16              | 0       | 2012-06-28 00:00:00 |         |
| 11 | 14435   | E0AD        | 63          | 63                 | 56           | Female      | Caucasian     | Right       | 14              | 2       | 2012-05-11 00:00:00 |         |
| 12 | 7646    | PCA         | 72          | 68                 | 61           | Female      | Caucasian     | Left        | 17              | 0       | 2012-10-17 00:00:00 |         |
| 13 | 14810   | E0AD        | 56          | 56                 | 49           | Female      | Caucasian     | Left        | 18              | 0       | 2012-07-18 00:00:00 |         |
| 14 | 15017   | E0AD        | 51          | 51                 | 48           | Male        | Caucasian     | Right       | 16              | 1       | 2012-10-19 00:00:00 |         |
| 15 | 15623   | PCA         | 71          | 71                 | 69           | Male        | Caucasian     | Right       | 16              | 0       | 2013-01-24 00:00:00 |         |
| 16 | 14788   | 1vPPA       | 58          | 58                 | 56           | Female      | Caucasian     | Right       | 20              | 0       | 2013-02-01 00:00:00 |         |
| 17 | 13811   | 1vPPA       | 71          | 70                 | 65           | Male        | Caucasian     | Right       | 16              | 0       | 2013-03-08 00:00:00 |         |
| 18 | 15590   | 1vPPA       | 68          | 68                 | 65           | Female      | Caucasian     | Right       | 16              | 2       | 2013-05-07 00:00:00 |         |
| 19 | 16001   | E0AD        | 58          | 58                 | 54           | Female      | Caucasian     | Right       | 12              | 0       | 2013-04-17 00:00:00 |         |
| 20 | 15610   | PCA         | 58          | 58                 | 51           | Female      | Not_caucasian | Left        | 17              | 0       | 2013-06-05 00:00:00 |         |
| 21 | 15295   | 1vPPA       | 66          | 66                 | 62           | Female      | Caucasian     | Right       | 18              | 1       | 2013-06-24 00:00:00 |         |
| 22 | 14342   | E0AD        | 71          | 69                 | 60           | Female      | Not_caucasian | Right       | 12              | 0       | 2013-03-13 00:00:00 |         |

1-100 of 190 rows

Loaded ND\_GMD.csv

190 rows, 132 columns · streamed from disk

ND\_GMD.csv · 190 × 132

Ready

© 2026 EDG



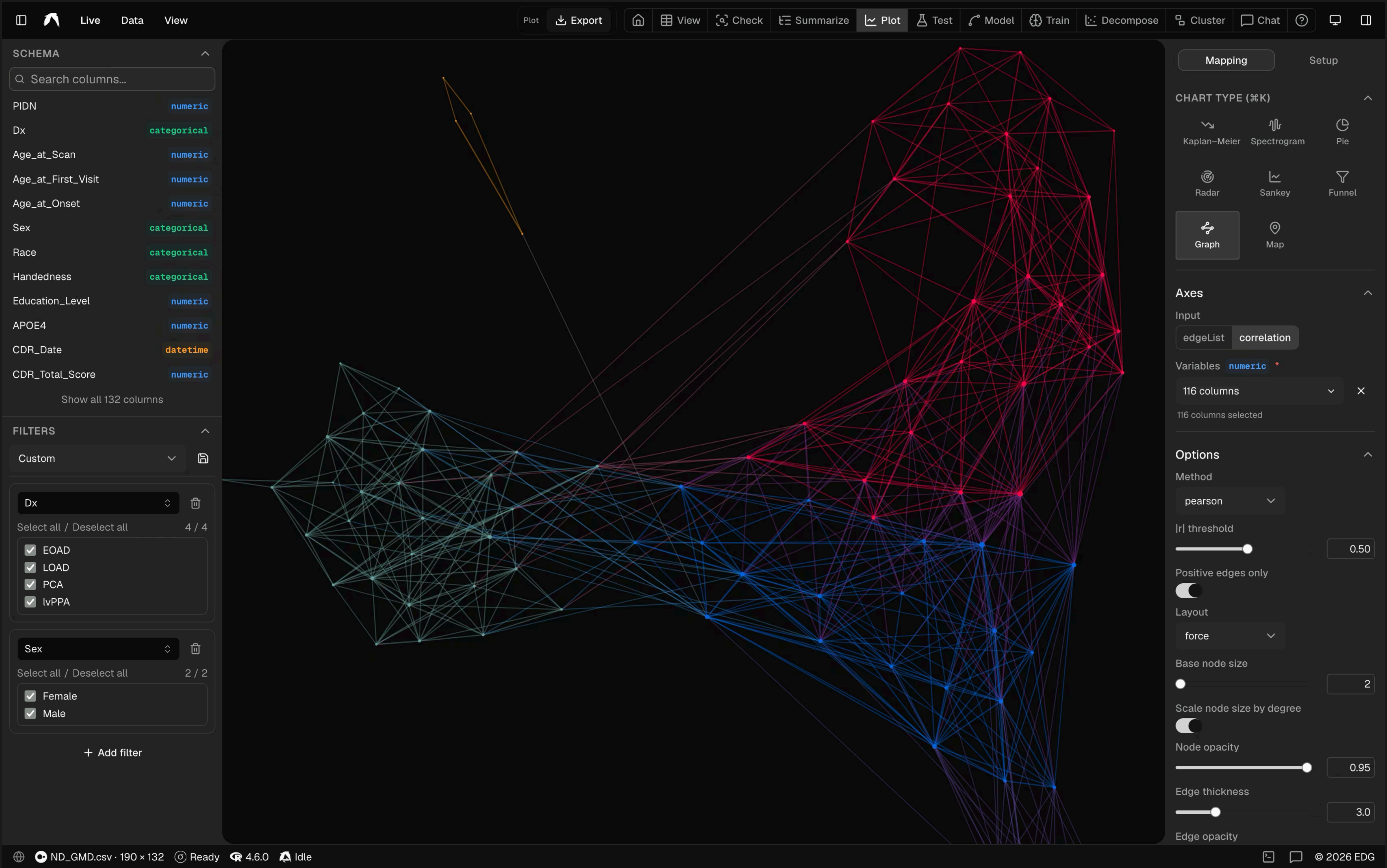




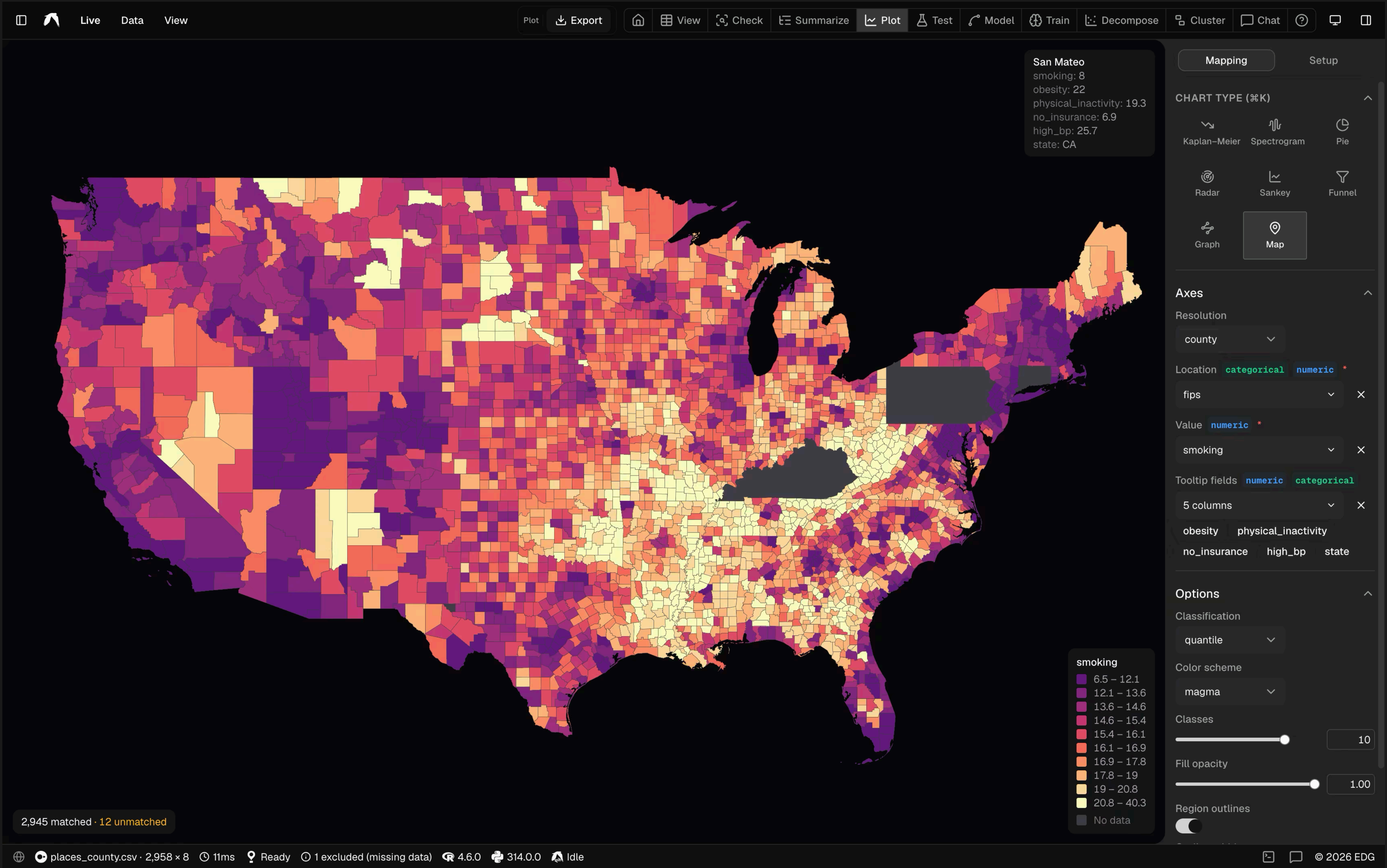












Live

Data

View

Home

View

Check

Summarize

Plot

Test

Model

Train

Decompose

Cluster

Chat

Help

Setup

### Choose a hypothesis test

Organized by the question being asked. Within a question, the variants answer it under different assumptions.

Search tests by name or keyword...

Value

Is a feature of the distribution equal to a hypothesized value?

LOCATION

t-test

Parametric

One-sample

Compare a mean to a hypothesized value ( $\mu_0$ ). Assumes approximately normal data.

Wilcoxon

Rank-based

One-sample (signed rank)

Rank-based non-parametric alternative to the one-sample t-test (signed rank). Tests a median against a hypothesized value. No...

Comparison

Is a feature (or the whole distribution) equal across groups?

LOCATION

ANOVA

Parametric

One-way analysis of variance — compare means across 2+ groups under the assumption of normal residuals with equal...

t-test

Parametric

Two-sample

Compare two group means, or matched pairs (paired option). Assumes approximately normal data.

Wilcoxon

Rank-based

Two-sample (rank sum)

Rank-based non-parametric alternative to the t-test, for two groups or matched pairs (paired option). Tests whether distributions /...

Welch ANOVA

Parametric

Welch's ANOVA — like one-way ANOVA but does not assume equal variances. Recommended when group variances differ.

SCALE

Bartlett

Parametric

Tests whether  $\geq 2$  groups have equal variances. Sensitive to non-normality.

F-test (variances)

Parametric

F-test comparing two group variances. Highly sensitive to non-normality.

Fligner–Killeen

Rank-based

Rank-based test for homogeneity of variances across  $\geq 2$  groups. Robust to non-

ND\_GMD.csv · 190 × 132

Ready

Idle

Pick a test.

© 2026 EDG

2026 Gennatas E.D. | rtemis.org



🏠

📊

🔍

📝

📈

🧪

🔄

🧠

🔗

💬

❓

🖥️

LiveDataView

🧪

SetupResults

WA

AA

WILCOXON RANK SUM TEST WITH CONTINUITY CORRECTION

GMD\_Insula\_R by Sex

two-sample

p = 0.0347 \*

W = 5300.00

ESTIMATE

difference in location0.01

95% CONFIDENCE INTERVAL

[8.1e-04, 0.03]

H<sub>1</sub>: true location shift ≠ 0.00

▼ R code

```
wilcox.test(
  GMD_Insula_R ~ Sex,
  alternative = "two.sided",
  conf.int = TRUE,
  conf.level = 0.95,
  exact = FALSE,
  correct = TRUE
)
```

BoxplotDensityHistogramStrip

GMD\_Insula\_R

FemaleMale

Sex

0.70

0.65

0.60

0.55

0.50

0.45

0.40

Ask AI about this

🌐

📄 ND\_GMD.csv · 190 × 132

🟢 Ready

🔧 4.6.0

🛑 Idle

💬

© 2026 EDG



2026 Gennatas E.D. | rtemis.org

LiveDataView

HomeViewCheckSummarizePlotTestModelTrainDecomposeClusterChatHelp

SetupResults

LA

Ask AI about this

True vs predictedCoefficientsResiduals vs fittedQ-Q residualsScale-locationResiduals vs leverage

MULTIVARIABLE LINEAR REGRESSION · GAUSSIAN (IDENTITY)  
**Age\_at\_Scan**  
Age\_at\_Scan ~ GMD\_Hippocampus\_R + GMD\_Insula\_R + Dx + Sex

| Term              | Estimate | SE    | t     | Pr(> t )   |
|-------------------|----------|-------|-------|------------|
| (Intercept)       | 97.54    | 6.29  | 15.51 | <0.001 *** |
| GMD_Hippocampus_R | -21.41   | 7.95  | -2.69 | 0.0078 **  |
| GMD_Insula_R      | -44.91   | 13.31 | -3.37 | <0.001 *** |
| DxLOAD            | 12.78    | 1.62  | 7.91  | <0.001 *** |
| DxPCA             | 1.49     | 1.30  | 1.14  | 0.254      |
| DxlvPPA           | 2.17     | 1.12  | 1.95  | 0.0528     |
| SexMale           | 0.73     | 0.94  | 0.78  | 0.439      |

R<sup>2</sup>†0.50Adj. R<sup>2</sup>†0.48Dispersion39.42

Null14363.81Resid.7213.85AIC1246.18

dev.(189 df)dev.(183 df)

n190

† Deviance R<sup>2</sup> (1 - resid./null deviance); not in summary.glm. Equals the lm R<sup>2</sup> for gaussian.

▼ R code

glm(  
  Age\_at\_Scan ~ GMD\_Hippocampus\_R + GMD\_Insula\_R + Dx + Sex,  
  family = gaussian(link = "identity"),  
  data = df  
)

True vs predicted

True  
Predicted

ND\_GMD.csv · 190 × 132

Ready

4.6.0

Idle

© 2026 EDG



🏠

📊

🔍

📝

📈

🧪

🔄

🧠

🔗

📁

💬

❓

🖥️

🗑️

LiveDataView

📄

🧠

Config

Results

Timeline

+

AG

AR

RANGER · REGRESSION

Age\_at\_Scan\_Ranger

✓ Done in 0.60s

🧠 Ask AI

📄 Save

job-28a037287b4611f1

Random Forest was used for regression. Mean R-squared was 0.91 on the training set and 0.44 on the test set across 10 independent folds.

Metrics

Mean (sd) across 10 independent folds

| TRAINING       |                 | TEST           |                 |
|----------------|-----------------|----------------|-----------------|
| MAE            | 2.137 (0.030)   | MAE            | 5.269 (0.574)   |
| MSE            | 6.997 (0.230)   | MSE            | 42.154 (10.378) |
| RMSE           | 2.645 (0.043)   | RMSE           | 6.449 (0.795)   |
| R <sup>2</sup> | 0.907 (2.9e-03) | R <sup>2</sup> | 0.440 (0.099)   |

True vs. predicted

● Training ● Test

The plot shows Predicted vs. Actual values. The x-axis (Actual) and y-axis (Predicted) both range from 43.4 to 86.6. A dashed diagonal line represents perfect prediction. Training data points (blue) are tightly clustered along this line, indicating high performance on the training set. Test data points (green) are more scattered and generally fall below the diagonal line, indicating poorer performance on unseen data. A green trend line for the test data is flatter than the diagonal, illustrating the 'regression to the mean' phenomenon.

Fireworks AI Qwen3.7 Plus ↺

+ New 📄 ↗️ 🗑️

🧠 Reasoning

- **Key Issue:** Explicitly state "Overfitting".
- **Follow-up Recommendations:** List actionable steps (tuning, simpler models, feature engineering).

Refining the "Follow-up" section:

- The test R<sup>2</sup> of 0.44 isn't *terrible* for biological age prediction (brain age gap studies often get 0.5-0.8), but the gap is the problem.
- The "regression to the mean" in the plot (green line is flatter than the diagonal) is a specific phenomenon. The model predicts values closer to the average age for extreme cases. This is typical of models trained with MSE loss that are uncertain.

Let's write it out.

Based on the provided metrics and the true-vs-predicted plot, here is a description of the model's results and recommendations for follow-up.

Model Results Description

1. Significant Overfitting

The most prominent feature of these results is a classic case of **overfitting**. The model performs exceptionally well on the data it was trained on but struggles significantly to generalize to unseen data.

- **Training Performance:** The model explains ~91% of the variance ( $R^2 = 0.907$ ) with a very low error (MAE  $\approx$  2.1 years). The standard deviations across folds are tiny, indicating the model consistently memorized the training patterns.
- **Test Performance:** The performance drops sharply on the test set. The  $R^2$  falls to 0.440, meaning the model only explains about 44% of the variance in new data. The error (MAE) more than doubles to  $\approx$  5.3 years.

2. Visual Analysis of the Plot

- **Training Data (Blue):** The blue points are tightly clustered along the diagonal dashed line (the line of perfect prediction). The blue trend line is almost identical to the diagonal, confirming the high  $R^2$ .

Send a message...

+ 🌐 W 📄 📎 🔄

⌵ {} ⬆️

🌐 🔄 ND\_GMD.csv · 190 × 132 🟢 Ready 🔄 4.6.0 🛑 Idle

🗑️ 💬 © 2026 EDG



Live

Data

View

Home

View

Check

Summarize

Plot

Test

Model

Train

Decompose

Cluster

Chat

Help

Fullscreen

+

AG

AR

Config

Results

Timeline

Age\_at\_Scan\_Ranger

job-28a037287b4611f1

valid

SuperConfig v1

schema.rtemis.org/supervised/v1/schema.json

```
{
  "$schema": "https://schema.rtemis.org/supervised/v1/schema.json",
  "algorithm": "Ranger",
  "hyperparameters": {
    "num_trees": 500,
    "importance": "impurity",
    "write_forest": true,
    "probability": false,
    "replace": true,
    "num_random_splits": 1,
    "alpha": 0.5,
    "minprop": 0.1,
    "poisson_tau": 1,
    "scale_permutation_importance": false,
    "local_importance": false,
    "regularization_factor": 1,
    "regularization_usedepth": false,
    "keep_inbag": false,
    "holdout": false,
    "quantreg": false,
    "oob_error": true,
    "save_memory": false,
    "verbose": true,
    "node_stats": false,
    "na_action": "na.learn",
    "ifw": false
  },
  "outer_resampling_config": {
    "n_resamples": 10,
    "type": "KFold",
    "train_p": 0.75,
    "strat_n_bins": 4,
    "verbosity": 1
  }
}
```

ND\_GMD.csv · 190 × 132

Ready

4.6.0

Idle

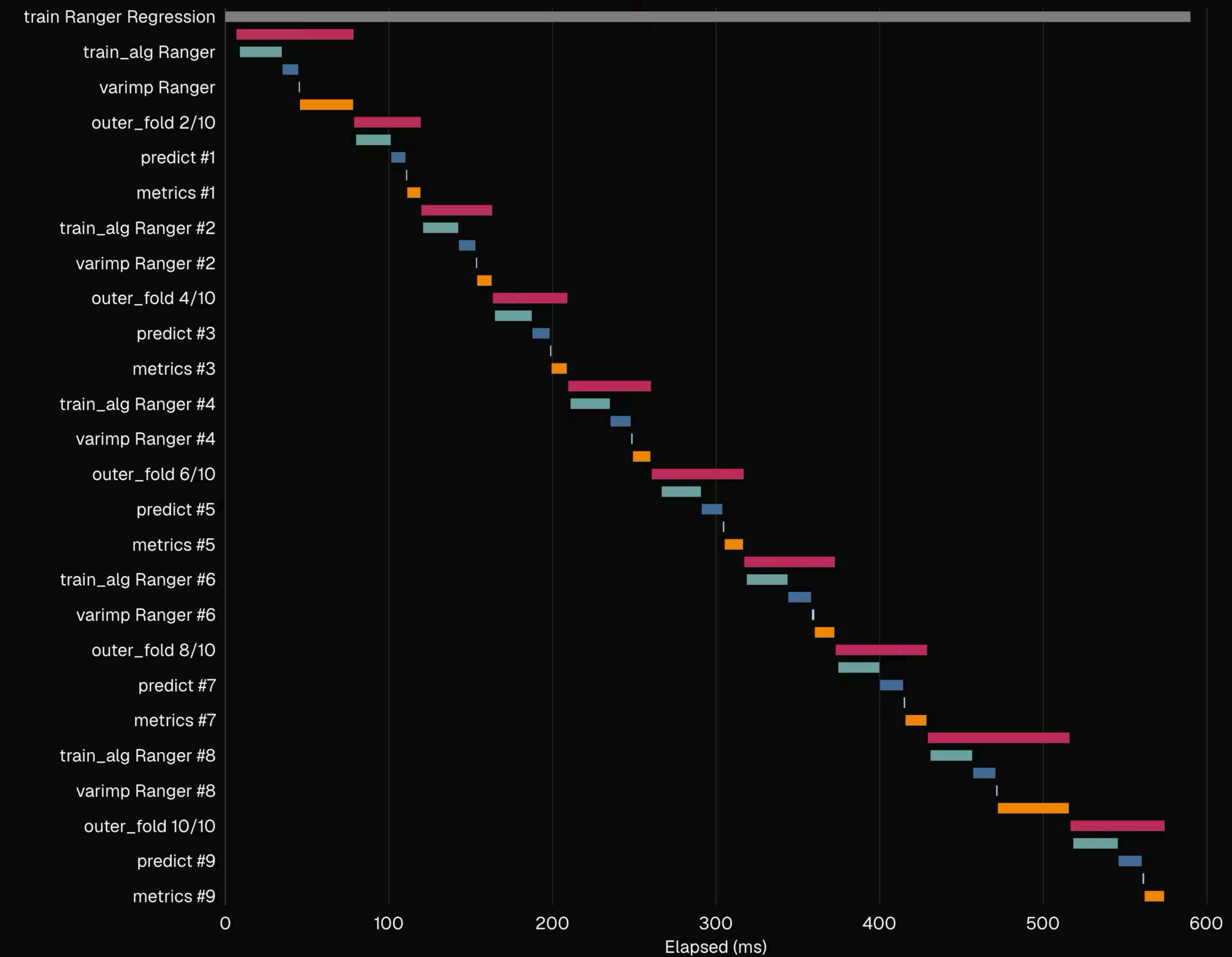
© 2026 EDG





Row height  15  Export

train outer\_fold train\_alg predict varimp metrics



Live

Data

View

View

Check

Summarize

Plot

Test

Model

Train

Decompose

Cluster

Chat

SCHEMA

Q Search columns...

PIDN

numeric

Dx

categorical

Age\_at\_Scan

numeric

Age\_at\_First\_Visit

numeric

Age\_at\_Onset

numeric

Sex

categorical

Race

categorical

Handedness

categorical

Education\_Level

numeric

APOE4

numeric

CDR\_Date

datetime

CDR\_Total\_Score

numeric

Show all 132 columns

FILTERS

Custom

Dx

Select all / Deselect all

4 / 4

☒ EOAD

ND\_GMD.csv

190 rows × 132 columns

126 numeric

5 categorical

1 datetime

Numeric

(126)

Q Search columns...

| STAT     | PIDN     | Age_at_Scan | Age_at_First_Visit | Age_at_Onset | Education_Level | APOE4  | CDR_Total_Score | CDR_Sum_of_Boxe... |
|----------|----------|-------------|--------------------|--------------|-----------------|--------|-----------------|--------------------|
| N Unique | 206      | 37          | 39                 | 41           | 11              | 3      | 4               | 21                 |
| Mean     | 1.425e+4 | 63.03       | 62.39              | 58.15        | 16.39           | 0.5967 | 0.8016          | 4.368              |
| SD       | 4186     | 8.718       | 8.481              | 8.458        | 2.745           | 0.6564 | 0.4178          | 2.325              |
| Median   | 1.478e+4 | 63.00       | 62.00              | 57.00        | 16.00           | 1.000  | 0.5000          | 4.000              |
| Min      | 747.0    | 45.00       | 45.00              | 40.00        | 0.000           | 0.000  | 0.000           | 0.000              |
| Max      | 2.156e+4 | 85.00       | 85.00              | 79.00        | 20.00           | 2.000  | 2.000           | 11.00              |
| Missing  | -        | -           | -                  | -            | 2               | 9      | 1               | 1                  |

Showing first 50 of 126 columns — search to find others.

Categorical

(5)

| STAT     | Dx | Sex | Race | Handedness | QC_Status |
|----------|----|-----|------|------------|-----------|
| N Unique | 4  | 2   | 2    | 2          | 2         |
| Missing  | -  | -   | -    | -          | -         |

PREPROCESSING

COLUMN NAMES

Title Case

lowercase

TYPES

Convert columns with ≤ unique values to categorical

-

5

+

Apply

MISSING DATA

Remove columns ≥ % missing

-

50

+

Apply

Remove rows ≥ % missing

-

50

+

Apply

Keep complete cases only

CLEANUP

Remove constant columns

Remove duplicate rows

RTemis

..:rtemis v.1.2.8 wasm32-unknown-emsripten (32-bit) (1 core available)

Defaults

Theme: whitegrid

Font: Helvetica

Palette: rtms

Resources

Docs: https://docs.rtemis.org/r/ml

Learn R: https://pdsr.rtemis.org

Cite: > citation(rtemis)

PSA: Do not throw data at algorithms. Compute responsibly!

Mounted data.parquet (190 × 132) — load with:

dat ← read("data.parquet")

Edit it, then write back to the app with:

sync(dat)

>

ND\_GMD.csv · 190 × 132

Ready

4.6.0

314.0.0

Idle

© 2026 EDG

SCHEMA

Q

Search columns...

PIDN

numeric

Dx

categorical

Age\_at\_Scan

numeric

Age\_at\_First\_Visit

numeric

Age\_at\_Onset

numeric

Sex

categorical

Race

categorical

Handedness

categorical

Education\_Level

numeric

APOE4

numeric

CDR\_Date

datetime

CDR\_Total\_Score

numeric

Show all 132 columns

FILTERS

Custom

▼

📄

Dx

⬇

🗑

Select all / Deselect all4 / 4

☒ EOAD

ND\_GMD.csv

190 rows × 132 columns

126 numeric

5 categorical

1 datetime

Numeric (126)

Q

Search columns...

| STAT     | PIDN     | Age_at_Scan | Age_at_First_Visit | Age_at_Onset | Education_Level | APOE4  | CDR_Total_Score | CDR_Sum_of_Boxe... |
|----------|----------|-------------|--------------------|--------------|-----------------|--------|-----------------|--------------------|
| N Unique | 206      | 37          | 39                 | 41           | 11              | 3      | 4               | 21                 |
| Mean     | 1.425e+4 | 63.03       | 62.39              | 58.15        | 16.39           | 0.5967 | 0.8016          | 4.368              |
| SD       | 4186     | 8.718       | 8.481              | 8.458        | 2.745           | 0.6564 | 0.4178          | 2.325              |
| Median   | 1.478e+4 | 63.00       | 62.00              | 57.00        | 16.00           | 1.000  | 0.5000          | 4.000              |
| Min      | 747.0    | 45.00       | 45.00              | 40.00        | 0.000           | 0.000  | 0.000           | 0.000              |
| Max      | 2.156e+4 | 85.00       | 85.00              | 79.00        | 20.00           | 2.000  | 2.000           | 11.00              |
| Missing  | -        | -           | -                  | -            | 2               | 9      | 1               | 1                  |

Showing first 50 of 126 columns — search to find others.

Categorical (5)

| STAT     | Dx | Sex | Race | Handedness | QC_Status |
|----------|----|-----|------|------------|-----------|
| N Unique | 4  | 2   | 2    | 2          | 2         |
| Missing  | -  | -   | -    | -          | -         |

PREPROCESSING

COLUMN NAMES

Title Case

lowercase

TYPES

Convert columns with ≤ unique values to categorical

—

5

+

Apply

MISSING DATA

Remove columns ≥ % missing

—

50

+

Apply

Remove rows ≥ % missing

—

50

+

Apply

Keep complete cases only

CLEANUP

Remove constant columns

Remove duplicate rows

RTemis

..:rtemis v.1.2.8 🐢 wasm32-unknown-emscripten (32-bit) (1 core available)

Defaults

Theme: whitegrid

Font: Helvetica

Palette: rtms

Resources

Docs: https://docs.rtemis.org/r/ml

Learn R: https://pdsr.rtemis.org

Cite: > citation(rtemis)

PSA: Do not throw data at algorithms. Compute responsibly!

Mounted data.parquet (190 × 132) — load with:

dat ← read("data.parquet")

Edit it, then write back to the app with:

sync(dat)

> ?resample

← resample rtemis

×

+

🔗 ↗

×

Resample data

DESCRIPTION

Create resamples of your data, e.g. for model building or validation. "KFold" creates stratified folds, , "StratSub" creates stratified subsamples, "Bootstrap" gives the standard bootstrap, i.e. random sampling with replacement, while "StratBoot" uses StratSub and then randomly duplicates some of the training cases to reach the original length of the input, or the length defined by target\_length .

USAGE

resample(x, config = setup\_Resampler(), verbosity = 1L)

ARGUMENTS

R

Python

SQL

🌐

ND\_GMD.csv · 190 × 132

🔄 Ready

📦 4.6.0

🧠 314.0.0

🛑 Idle

📄

💬

© 2026 EDG









Live

Data

View

View

Check

Summarize

Plot

Test

Model

Train

Decompose

Cluster

Chat

Overview

Data Input

Chart Types

TABULAR DATA

Explore

Feature Engineering

Test

Model

rtemis Server

Train

Decompose

Cluster

A3 DATA

{A3}

 A3 Editor

CHAT

Local LLM

Cloud providers

Console

Workspaces

Session Log

Local-only mode

Tech Stack

Chart Types

WASM

Data input shapes and chart options.

Supported File Formats

Tabular:

 CSV, Parquet, JSON (array of objects), Arrow IPC

Hierarchical:

 JSON, YAML, TOML (with nested children)

A3:

 Amino Acid Annotation JSON (via @rtemis/a3)

Data Shapes

per point

 One value per row. X and Y are read in parallel — row i of X pairs with row i of Y.

per group

 The full column of raw values, split by group. Used to compute statistics like quartiles or density.

per category

 One scalar value per category level. Multiple rows per category are aggregated.

tree

 A nested hierarchy of nodes with name, value, and children arrays.

protein

 An amino acid sequence with annotation layers (sites, regions, PTMs, processing annotations, variants).

Tabular Charts

Box

tabular

grouping

Displays the distribution of a numeric variable — median, quartiles, and outliers. Compare distributions across groups or multiple variables side by side.

DATA INPUTS

Value

numeric

per group

 Raw values used to compute quartiles, median, and whiskers. Multiple columns get separate boxes.

Group

opt

categorical

per group

 Splits data into side-by-side boxes per group level.

OPTIONS

Group layout

 auto | x-axis | legend

Color boxes

 single | group

Box fill opacity

 0–1

Points

 outliers | all | none

Point size

 1–12

Points

 Show none, outliers only, or all points a...

Histogram

tabular

grouping

Bins a numeric variable into equal-width intervals and shows the count per bin. Visualize the shape, center, spread, and modality of a distribution.

DATA INPUTS

X

numeric

per group

 Raw values binned into intervals to compute counts.

ND\_GMD.csv · 190 × 132 → 190 × 131

Ready

4.6.0

314.0.0

Idle

© 2026 EDG





LiveDataView

PlotExportViewCheckSummarizePlotTestModelTrainDecomposeClusterChat

SCHEMA

Search columns...

PIDNnumeric

Dxcategorical

Age\_at\_Scannumeric

Age\_at\_First\_Visitnumeric

Age\_at\_Onsetnumeric

Sexcategorical

Racecategorical

Handednesscategorical

Education\_Levelnumeric

APOE4numeric

CDR\_Datdatetime

CDR\_Total\_Scorenumeric

Show all 132 columns

FILTERS

Custom

Dx

Select all / Deselect all4 / 4

Ollamagemma...+ New



# Hello.

What shall we explore today?

Send a message...

+🌐W📄PMC🔗⌵⌨️⬆️

```
> library(rtemis); suppressPackageStartupMessages(library(data.table))
```



```
..:rtemis v.1.1.2.8 🐛wasm32-unknown-emsripten (32-bit) (1 core available)
```

Defaults

Theme: whitegrid

Font: Helvetica

Palette: rtms

Resources

Docs: https://docs.rtemis.org/r/ml

Learn R: https://pdsr.rtemis.org

Cite: > citation(rtemis)

PSA: Do not throw data at algorithms. Compute responsibly!

Mounted data.parquet (190 × 131) – load with:

dat ← read("data.parquet")

Edit it, then write back to the app with:

sync(dat)

```
> 
```

+🔍↗️✕

R

Python

SQL

🌐🔄ND\_GMD.csv · 190 × 132 → 190 × 131🕒Ready🔧4.6.0📦314.0.0🐦Idle

🖼️💬© 2026 EDG



2026 Gennatas E.D. | rtemis.org

# Ecosystem overview







# Summary



- Schema-driven, multi-language, low-code, no-code, validated, observable, auditable AI/ML/Visualization









Thank you



[www.rtemis.org](http://www.rtemis.org)  
[stathis@rtemis.org](mailto:stathis@rtemis.org)