

# Intro to tidymodels



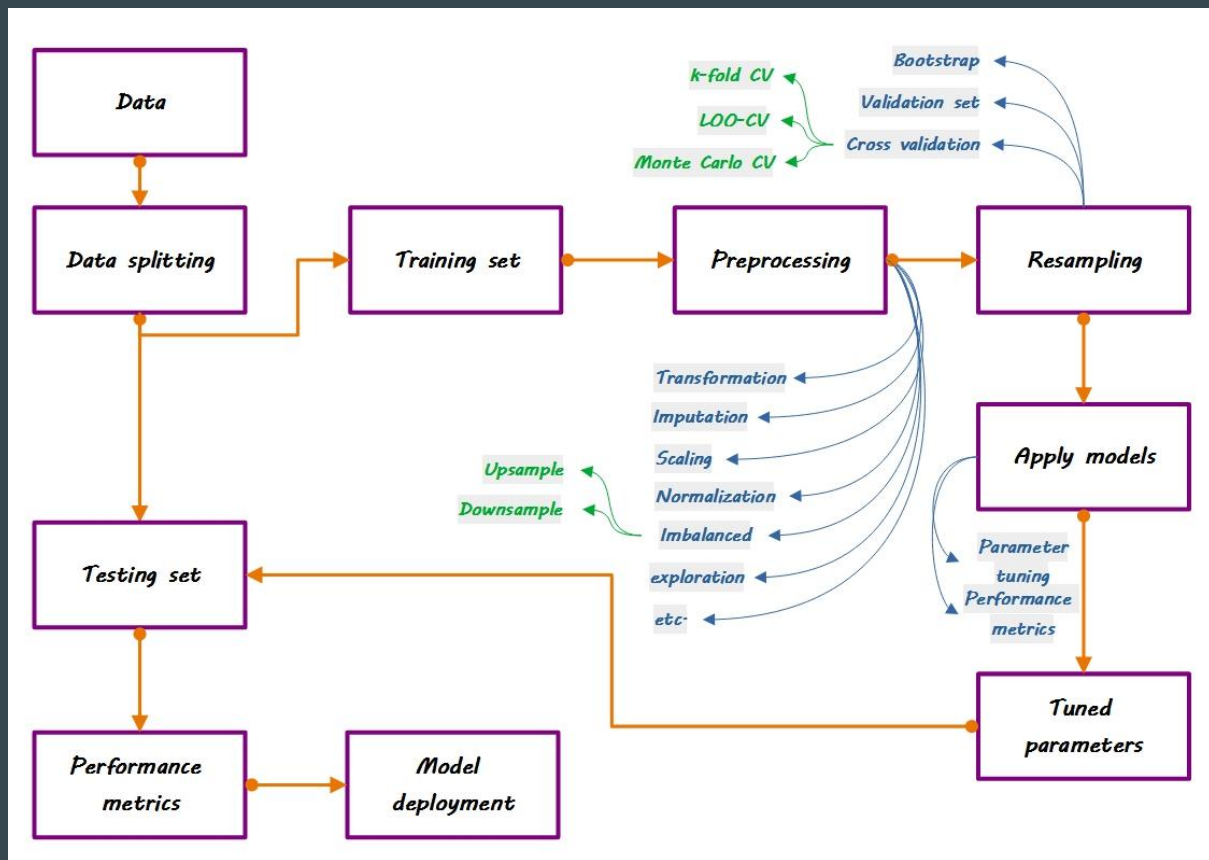
Tengku Muhammad Hanis  
Sept 28, 2021

# What is tidymodels?

- Collection of packages equivalent to tidyverse, but for machine learning
- Successor of caret package
- Lead by Max Kuhn and his team at R Studio
- Why tidymodels:
  - Overcome limitations of caret
  - Developed by R Studio team
  - Developed by Max Kuhn - experienced biostatistician, previously work with Pfizer (> 10 years)
  - Probably going to be a standard approach in R for ML (personal opinion)

- Why not tidymodels?
  - Still in development (lacking documentation in certain areas)
  - Complicated especially for beginners

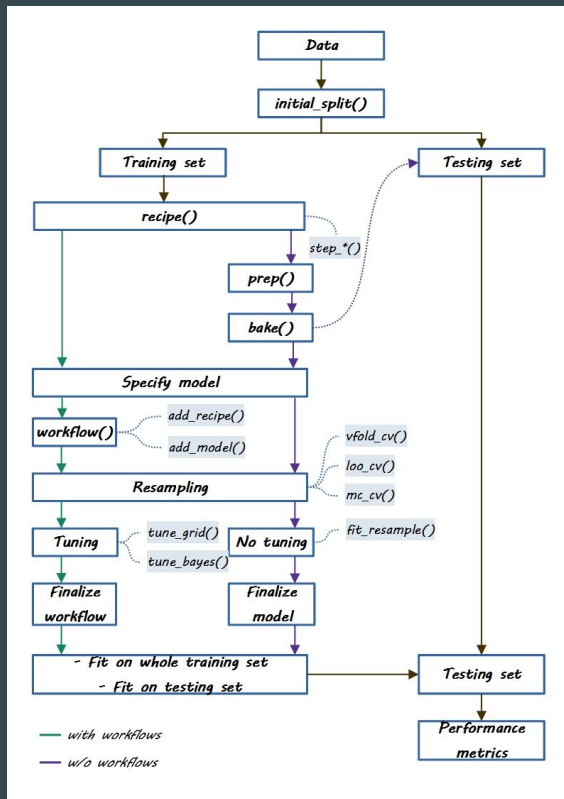
# General flow of machine learning



- Recommended preprocessing steps\*:
  - a. Imputation
  - b. Handle factor levels
  - c. Individual transformation for skewness and other issues
  - d. Discretize (last resort)
  - e. Create Dummy variable
  - f. Create interactions
  - g. Normalization steps (center, scale, etc)
  - h. Multivariate transformation (PCA, spatial sign, etc)

\*<https://recipes.tidymodels.org/articles/Ordering.html>

# General flow of tidymodels



# Packages in tidymodels

- Core packages:
  - Recipes: preprocessing
  - rsamples: resample process
  - parsnip: apply various ML models
  - tune, dials: parameters tuning
  - yardstick: performance metrics
  - workflows: bind all the process in workflow
  - brooms: extract information from ML model

- Specialized packages\*:
  - corr: for working with correlation matrices
  - spatialsample: resampling process for spatial data
  - finetune: more advanced tuning parameter approach
  - themis: upsample and downsample for class imbalance
  - discrim: discriminant analysis model
  - plsmod: partial least square model
  - rule: rule based models (cubist, RuleFit, c5.0, etc)
  - baguette: ensemble ML (bagging @ bootstrap aggregating method)
  - stacks: ensemble ML (stacking method)
  - poissonreg: various type of poisson regression
  - etc.

\*<https://www.tidymodels.org/packages/#specialized-packages>



# Good references

- <https://www.tidymodels.org/>
- <https://www.tmwr.org/> (online book)
- <https://www.youtube.com/c/JuliaSilge> (Julia Silge's youtube)

# Questions?



<https://tengkuhanis.netlify.app/>