Preparation

- **Crossmatching**
  - >150 catalogs
  - ~8 M sources
  - ~100 (sub) types
  - based on positional and magnitude difference

- **Training set for General Variability Detection**
  - 2 types: 60 K ‘variables’ and 63 K ‘others’

- **Training set for General classification**
  - 40 type groups: 60 K variables

Execution

- **Classifier framework H2O** ([www.h2o.ai](http://www.h2o.ai))

- **General Variability Detection**
  - XGBoost: binary classifier
  - Completeness: 99%
  - Contamination in variables: 0.4%

- **General classification**
  - About 100 one-vs-rest classifiers for each variable type using XGBoost (XGB) and Distributed Random Forest (DRF).
  - Several multi-class classifiers using XGB and DRF.
  - Several DRF meta-classifiers combining the one-vs-rest or multi-class classifier results.

Verification & Validation

Systematic semi-automated procedure for each class (incl. literature comparison)

**Examples from Gaia Data Release 2**: HRD-Gaia Collaboration, L. Eyer et al. (2018)

Time series: Holl et al. (2021)

Results

- To be published in Gaia DR3 (mid 2022)
- About 11 million variables
- 24 variability types + galaxies