

Supervised ML **tasks** contain:

- data + target feature
- performance measure
- resampling technique (e.g., 10-fold CV) and train-test splits

 **Data**
currently > 20.000

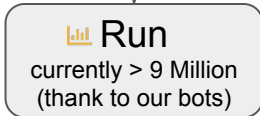
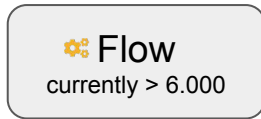
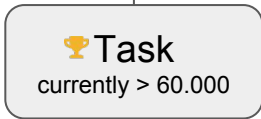
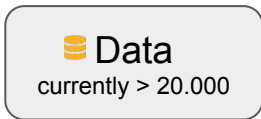
 **Task**
currently > 60.000

 **Flow**
currently > 6.000

A **flow** can be a ML algorithm or a complex ML pipeline that solves a **task**.

Supervised ML **tasks** contain:

- data + target feature
- performance measure
- resampling technique (e.g., 10-fold CV) and train-test splits

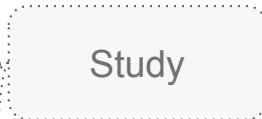
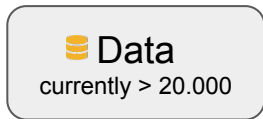


A **flow** can be a ML algorithm or a complex ML pipeline that solves a **task**.

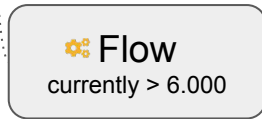
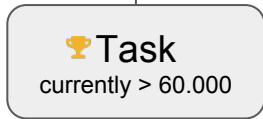
The result of **flows** if applied to **tasks**.
Includes: hyperparameters, predictions, performance evaluation, and runtime.

Supervised ML **tasks** contain:

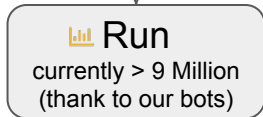
- data + target feature
- performance measure
- resampling technique (e.g., 10-fold CV) and train-test splits



A **study** links multiple data sets, tasks, flows, and runs.
Aim: better reproducibility and sharing in publications.



A **flow** can be a ML algorithm or a complex ML pipeline that solves a **task**.



The result of **flows** if applied to **tasks**.
Includes: hyperparameters, predictions, performance evaluation, and runtime.

Supervised ML **tasks** contain:

- data + target feature
- performance measure
- resampling technique (e.g., 10-fold CV) and train-test splits