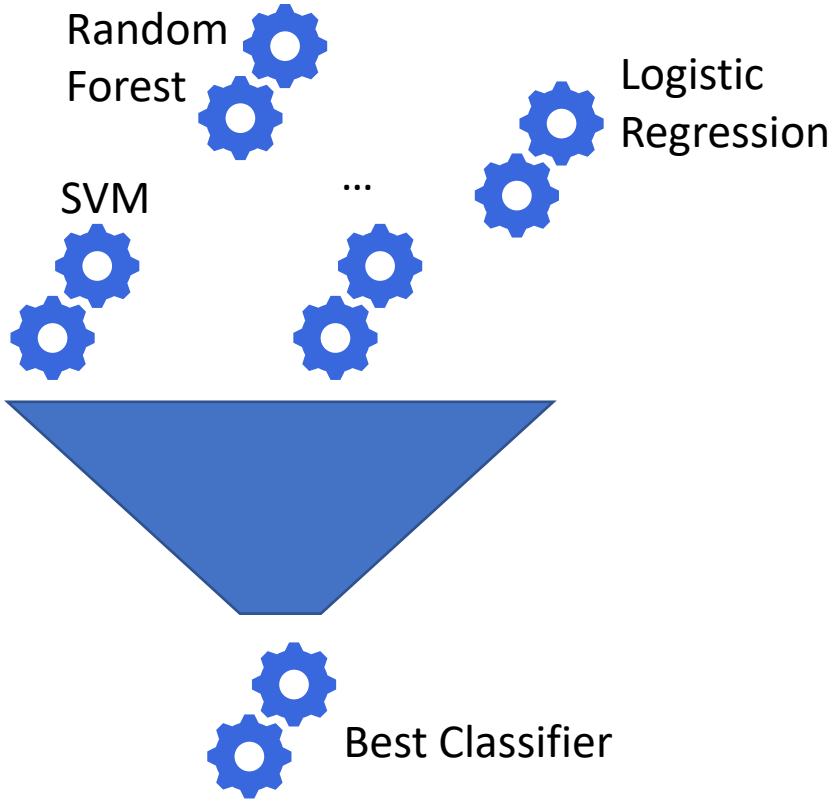
Ensemble Learning 101

Ensemble Learning

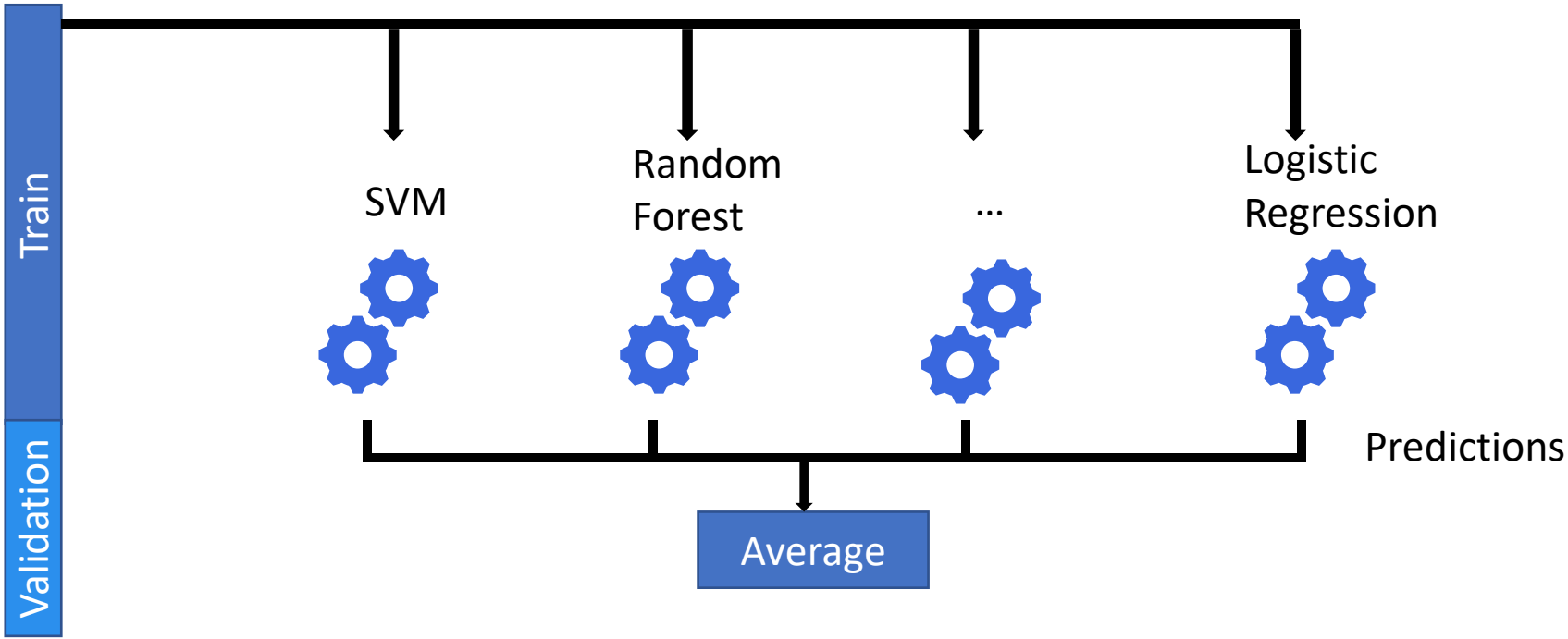
Classical Model Selection



Ensemble Learning

Stacking

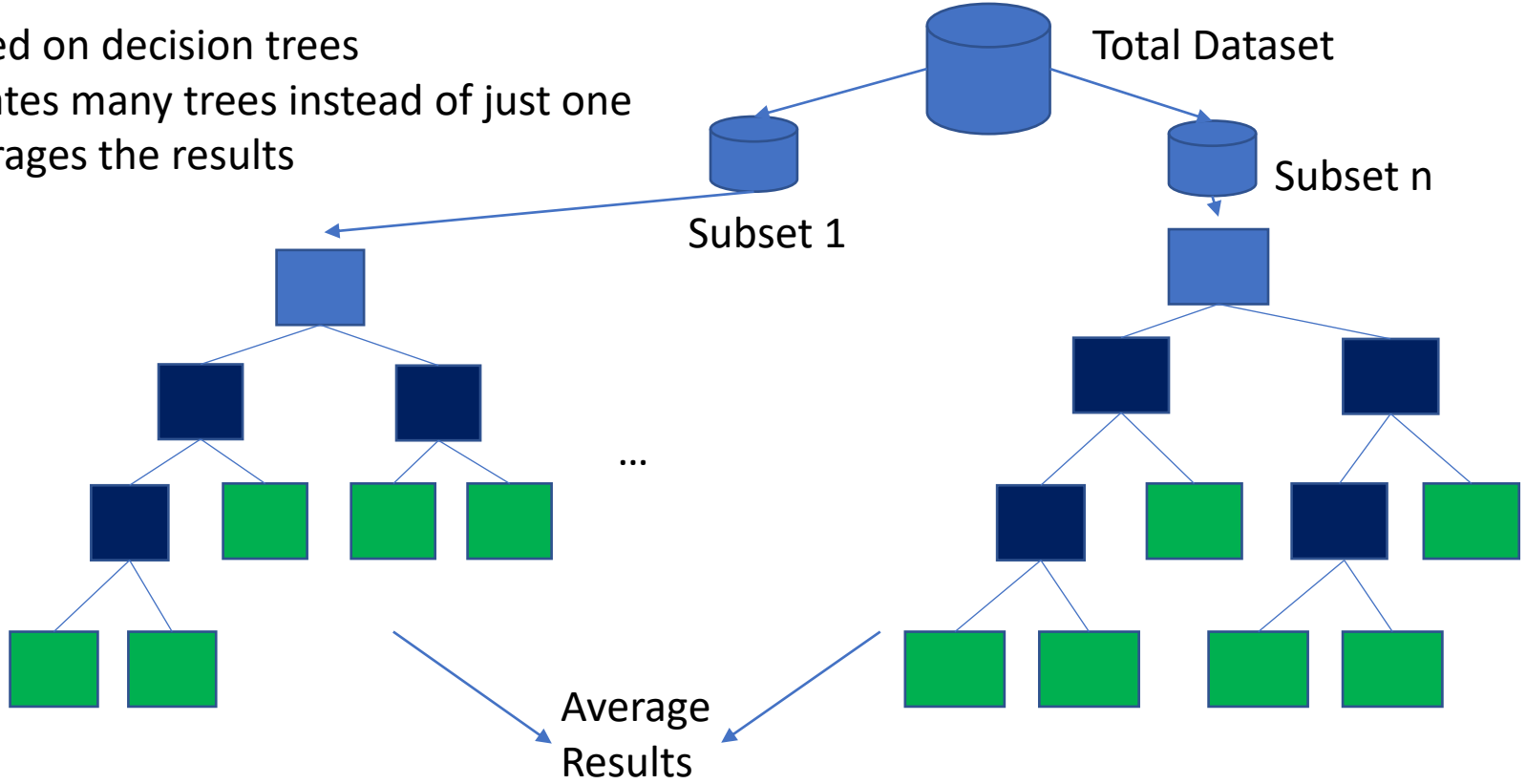
- Combines several models
- Target: improve metric compared to single algorithm



Ensemble Learning

Bagging

- Based on decision trees
- Creates many trees instead of just one
- Averages the results



Ensemble Learning

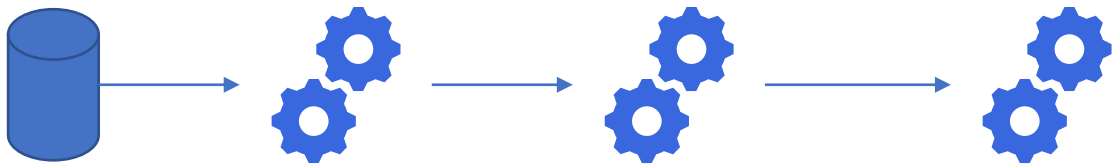
Bagging

- Uses different version of the same model type,
- Models created in parallel
- Example: Random Forest
- Bagging can reduce variance
- Each model differs from the others

Ensemble Learning

Boosting

- Takes previous models into account for next models
- Models created in sequence
- Examples: xgboost, catboost, lightbgm



Ensemble Learning

Advantages / Disadvantages

+

- Better performance than individual method
- Reduces overfitting

-

- Black-box
- High computational effort