Temas Selectos de Aprendizaje Automático

Sesion II

PhD. Edwyn Javier Aldana Bobadilla

Ensambling Methods

Bagging

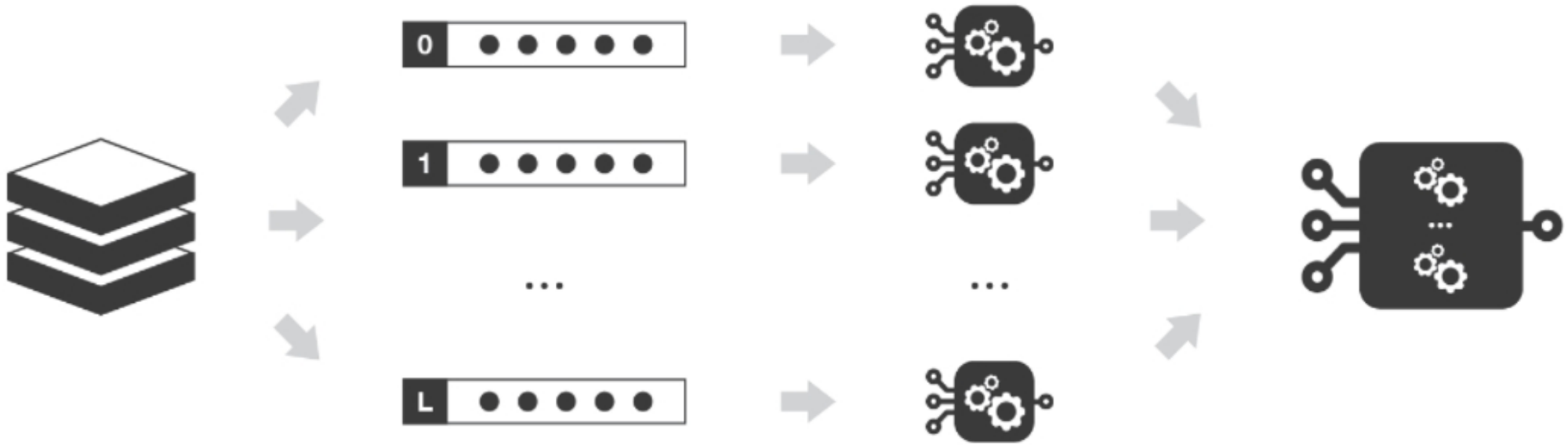
- This approach often considers homogeneous weak learners.
- They learn independently from each other in parallel
- The solutions are combined following some kind of deterministic averaging process

Ensambling Methods

Bagging

- The most common method in this approach is "Bootstrap"
- This consists in generating samples of size B (bootstrap samples) from initial dataset of size N.
- Each sample is evaluated independently by an estimator or learner
- Based on individual estimations, variance and confidence intervals are computed.

Ensambling Methods



Source: https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205

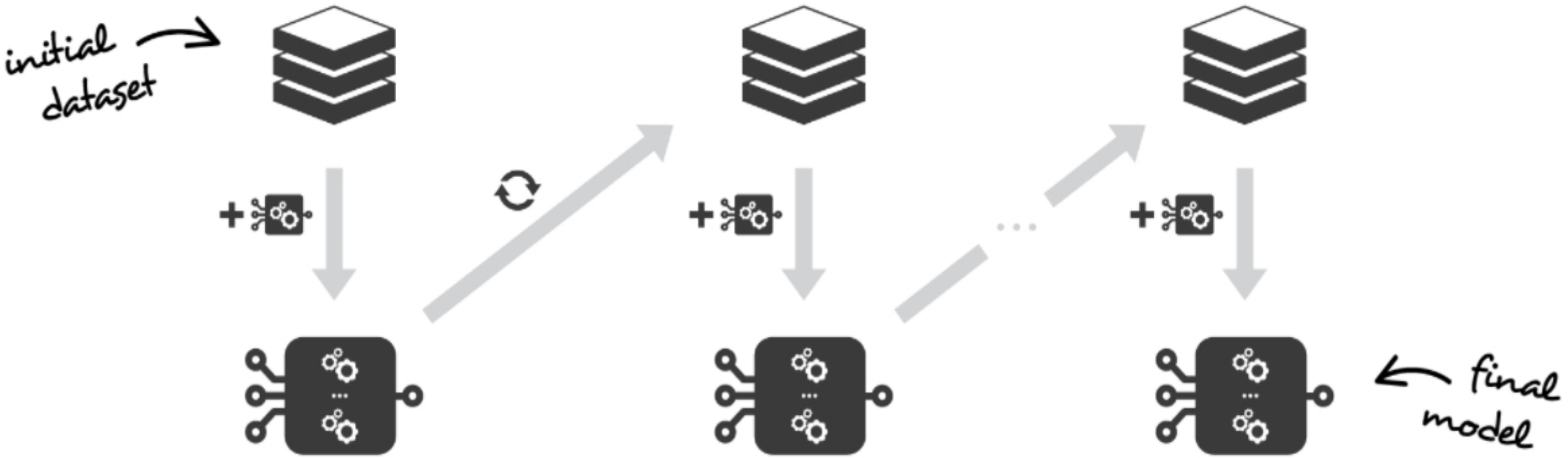
Ensambling Methods

Boosting

- This approach often considers homogeneous weak learners.
- They learn sequentially in a very adaptive way
- A model depends on the previous ones.
- Each model in the sequence is fitted giving more importance to observations in the dataset that were badly handled by the previous models in the sequence.
- Each new model focus its efforts on the most difficult observations
- Solutions are combined following a deteministic strategy

Ensambling Methods

Boosting



Source: https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205

Exercise 1:

1. Dowload the script S2_AdvancedApproaches.ipynb
2. Download the dataset titanic.csv
3. Review different approaches of ensambling methods