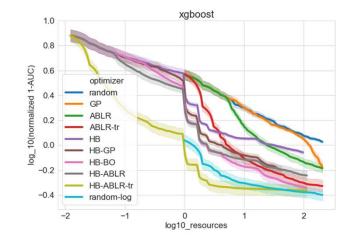# A Simple Transfer-Learning Extension of Hyperband

Lazar Valkov, Rodolphe Jenatton, Fela Winkelmolen, Cédric Archambeau

- Setting: Hyperparameter Optimisation
- Hyperband (HB):
    - Incrementally allocates more resources to the best-performing candidates initially taken from a pool of randomly sampled candidates.
    - Evaluates different number of initial candidates n_i for r_i
- We enhance HB with model-based sampling, using ABLR (Peronne *et al.*)

$$\boldsymbol{\Phi_z}(\boldsymbol{X_t}, \boldsymbol{r_i})$$

$$P(\boldsymbol{y}_t \mid w_t, z, \beta_t) = \mathcal{N}\big(\boldsymbol{\Phi_z}(\boldsymbol{X_t}, \boldsymbol{r_i})\boldsymbol{w_t}, \beta_t^{-1} I_{N_t}\big) \, P(\boldsymbol{w_t}\mid\alpha_t) = \mathcal{N}(\boldsymbol{0}, \alpha_t^{-1}I_D)$$

- Benefits:
    - Makes use of all data produced by a HB run
    - Can use data from past HB runs to learn better basis function
    - We don't use heuristics for low number of data points, nor to encourage exploration

amazon