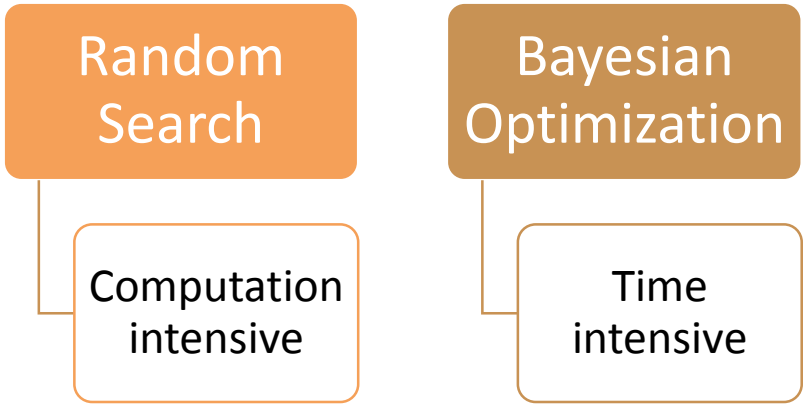
Train In Data

Multi-Fidelity Optimization

Search strategies



- Both strategies aim to find the hyperparameters, x, that minimize f(x)
- To find the "global" minimum we require a good coverage of f(x)
- As x dimension increases, the number of evaluations needed to cover f(x) increases exponentially
  - More computing resource or more time
- Impracticable

Time budget

- In practice, one constraint for optimization techniques is the **time budget**.

- The aim of the optimization algorithm is to select the hyperparameters that can achieve **(near)-**optimal performance within a defined time budget (or cost).

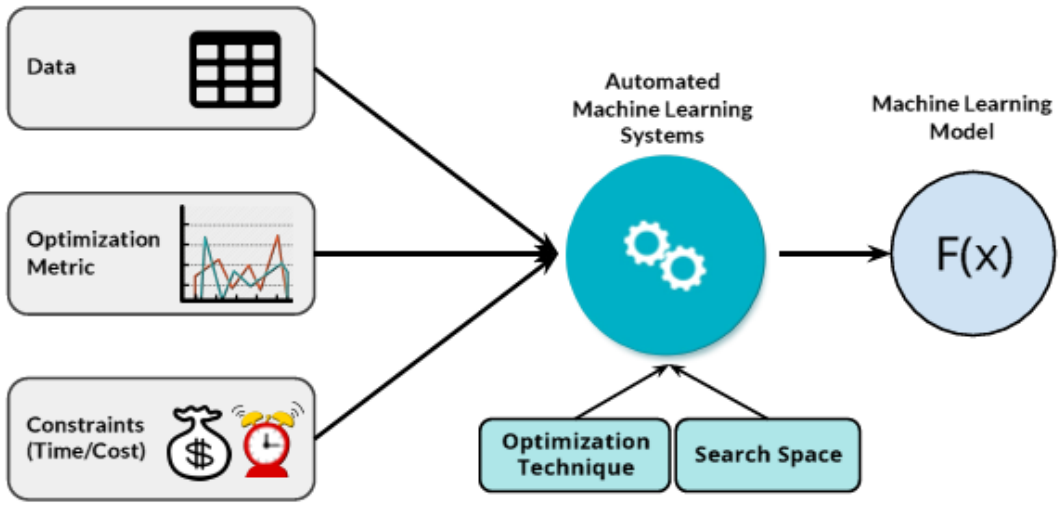


Image taken from Elshawi, et. Al., "Automated Machine Learning:State of the Arti and Open Challenges", ArXiv, 2019

Black-box strategies

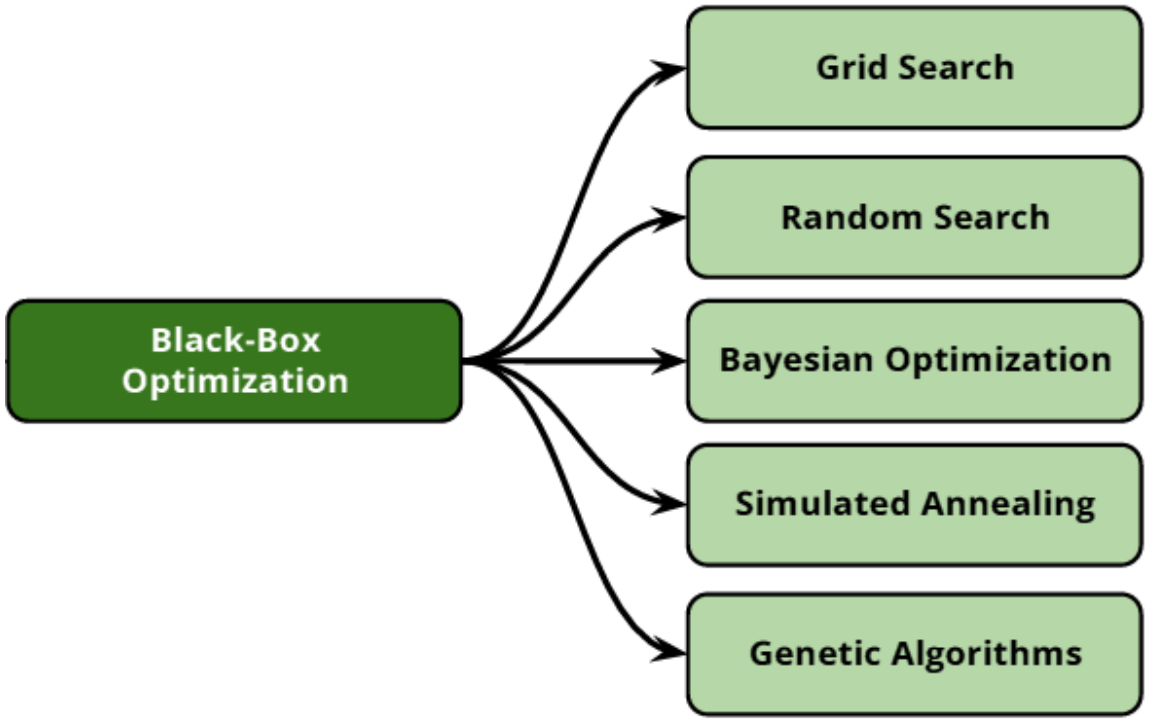


Image taken from Elshawi, et. Al., "Automated Machine Learning:State of the Arti and Open Challenges", ArXiv, 2019

We can only know f(x) after we have fully evaluated it at certain values of the hyperparameters.

- F(x) is often expensive to obtain
- Often, no getting around the fact that a thorough search of f(x) will require sampling a large number of xs.
- **What if, we could use a cheaper version of f(x)?**

Multi-fidelity

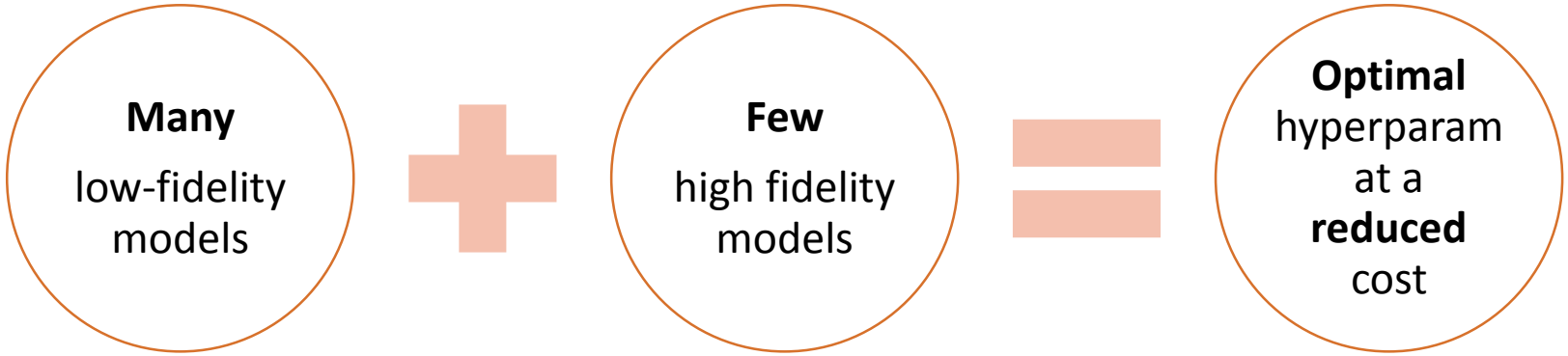
- Multi- fidelity optimization focuses on decreasing the evaluation cost by combining **large number** of cheap **low- fidelity** evaluations and a **small number** of expensive **high- fidelity** evaluations.

- The **high- fidelity** evaluation outputs **precise performance** at a higher cost

- The **low-fidelity** evaluations output a **(good) approximation** of the best f(x) at a much cheaper cost

Multi-fidelity optimization

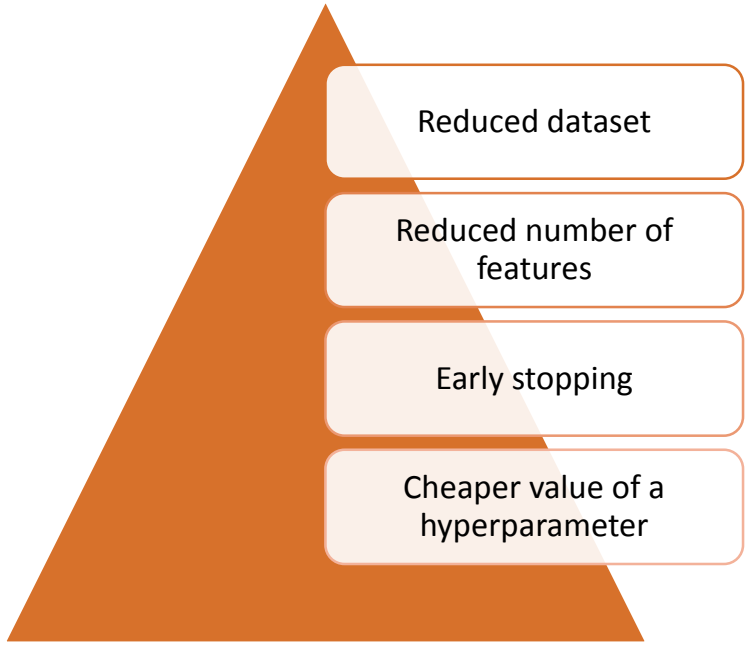
Multi- fidelity optimization uses many low-fidelity models to **reduce the total evaluation cost**, find a subset of potentially good solutions, and then, examine the best solutions with high-fidelity model(s).



What are low- and high-fidelity models?

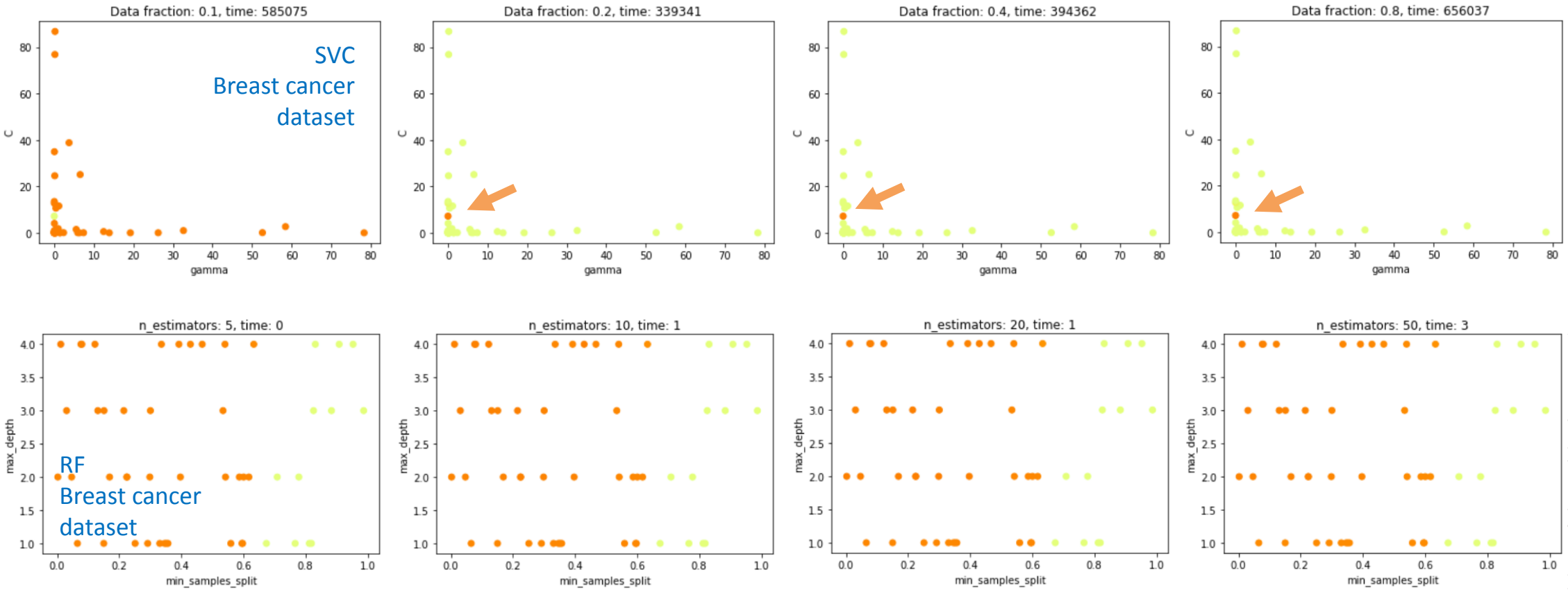
Low-fidelity models

Low-fidelity models are cheaper representations of f(x)



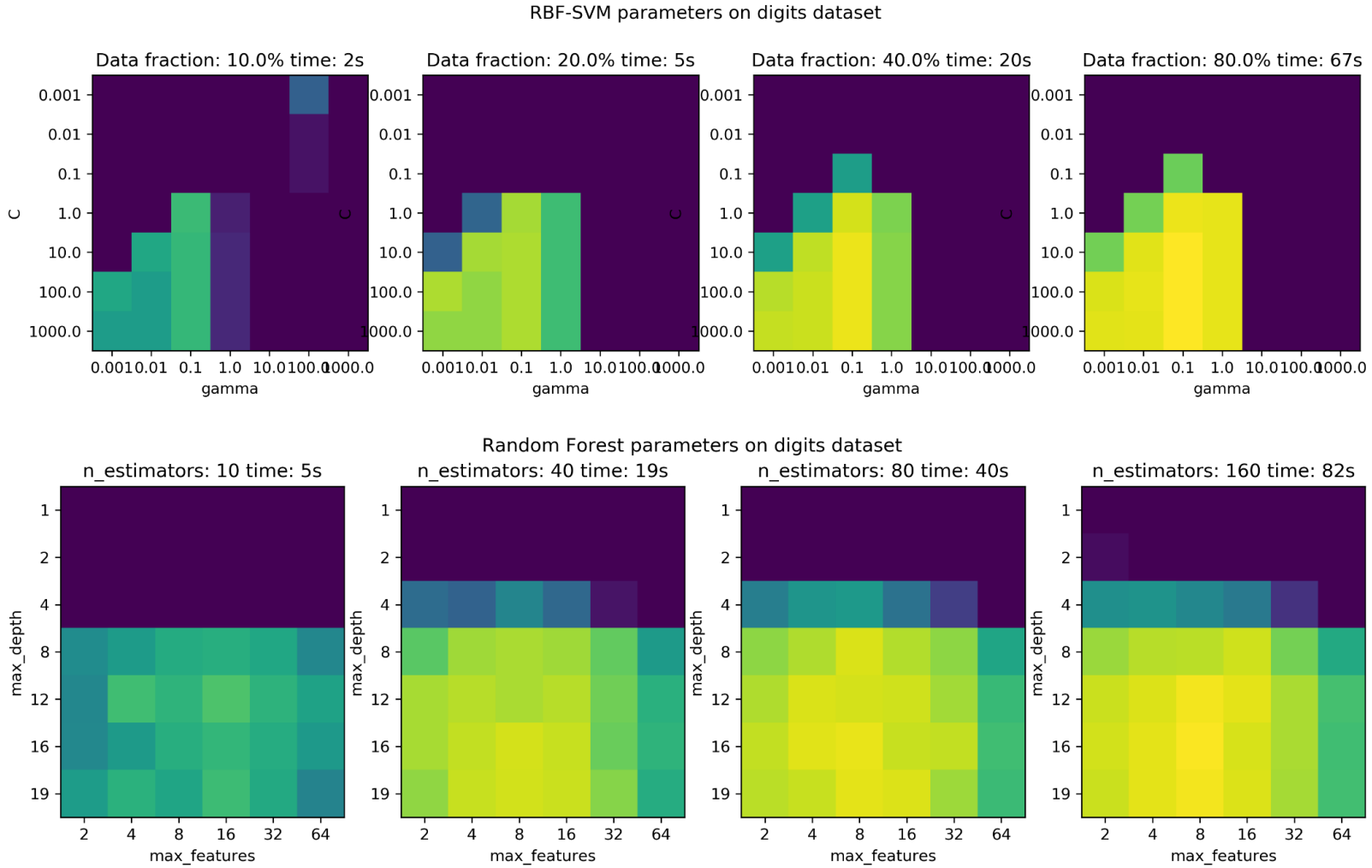
- Low-fidelity models are created to reduce computational cost when a large number of expensive simulations

Low-fidelity models approximate f(x)



With low-fidelity models we can already detect where the optimal hyperparameters are at a fraction of the time.

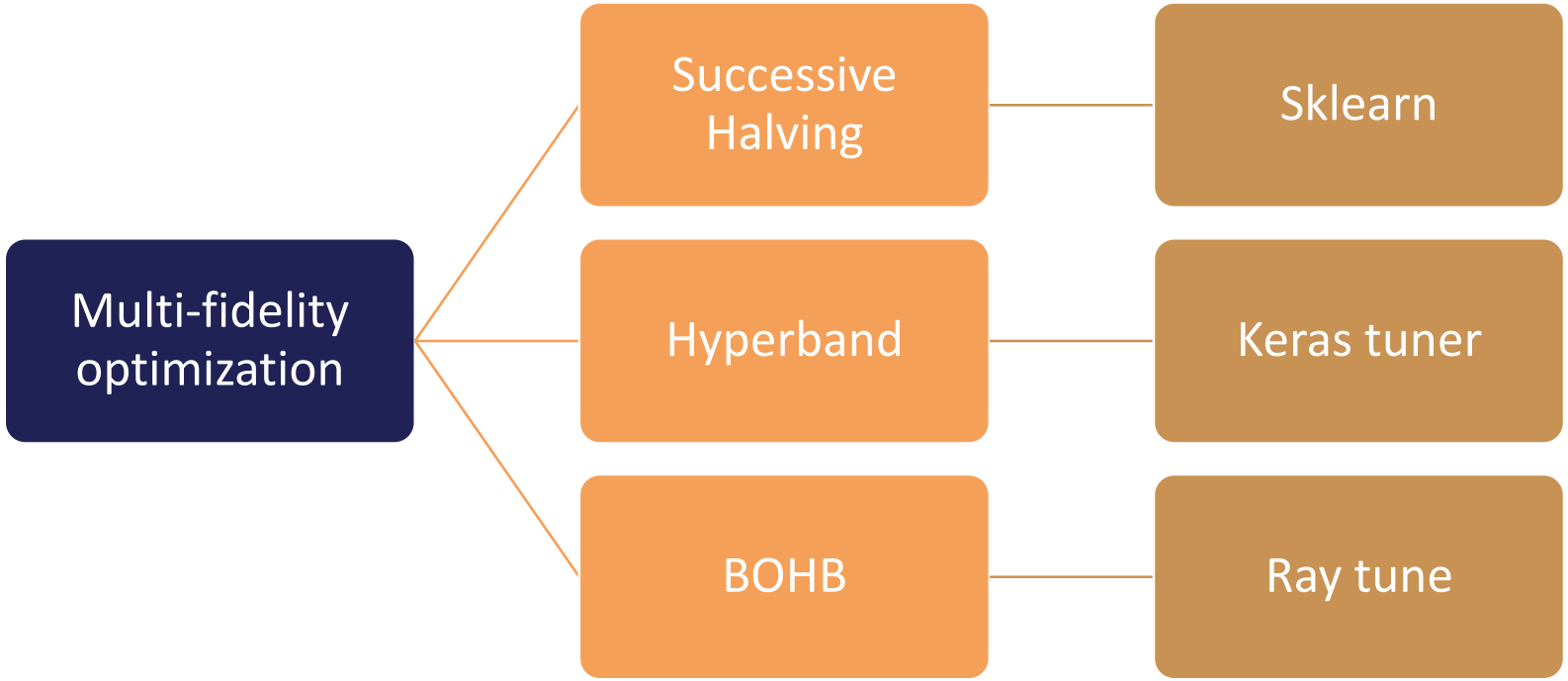
Low-fidelity models approximate f(x)



With low-fidelity models we can already detect where the optimal hyperparameters are at a fraction of the time.

Image taken from A. Mueller class: https://amueller.github.io/COMS4995-s19/slides/aml-13-parameter-tuning-automl/#17

Multi-fidelity optimization models



Train In Data

THANK YOU

www.trainindata.com