Train In Data

Bayesian Optimization

Bayesian Optimization

Bayesian Optimization is a sequential approach for global optimization of black-box objective functions, that are costly to evaluate.

Bayesian Optimization

Mathematically, we want to find the global maximizer (or minimizer) of an unknown (black-box) objective function *f*:

$$\mathbf{x}^\star = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$$

where x are the hyperparameters.

The objective function *f*

- *f* is continuous
- *f* is difficult to evaluate ➔ too much time or money
- *f* lacks known structure, like concavity or linearity ➔ *f* is black-box
- *f* has no derivative ➔ we can't evaluate a gradient
- *f can be evaluated at arbitrary points of* x *(the hyperparameters)*
  - We can make point-wise observations of *f*

Bayesian Optimization

f is unknown

$$\mathbf{x}^\star = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$$

1. In Bayesian optimization we treat *f* as a random function and place a **prior** over it. (the prior is a function that captures the belief -distribution, behaviour- of *f)*

2. Then, we evaluate *f* at certain points.

3. With the new data, the prior (*f* original belief) is updated to a new the **posterior** distribution.

4. The posterior distribution is used to construct an acquisition function to determine the next query point.

Estimating the prior

- Gaussian processes
- Tree-parzen estimator
- Random Forests

Acquisition Function

- Expected Improvement (EI)
- Gaussian process upper confidence bound (UCB)

Coming up next:

- Bayes Inference
- Conditional probability
- Bayes Rule
- Sequential Model based Global Optimization (SMBO)
- Gaussian processes
- Acquisition functions
- Specific algorithms for hyperparameter tuning.

Train In Data

THANK YOU

www.trainindata.com