Data Abstraction

Donghee Hong
2020312657
hdh12345@g.skku.edu

April 7th, 2023

1 Dataset Description

The *BOHB Sampled Configuration & Results* dataset, which I made, consists of the sampled configuration's data and its evaluation results while performing the hyperparameter optimization method, BOHB. The dataset's type is a table with 575 items, each comprising thirteen attributes. One attribute, **sample_id**, is a unique key attribute, and others are just values.

2 Abstraction Results

- Nominal Attributes

Name	Semantics	Values
sample_id	Key of Sample data	{bracket}-{round}-{trial}
sample_type	Where the configuration was sampled	random, samples, BO
config_optimizer	The type of optimizer in the sampled configuration	adam, sgd, rms
config_scheduler_p	Whether the learning rate scheduler is enabled or not	True, False
config_activation	The type of activation function in the sampled configuration	relu, lrelu, tanh

- Ordinal Attributes

Name	Semantics	Values	Direction
config_batch_size	The value of batch size in the sampled configuration	8, 16, 32	Sequential
config_hidden_size	The value of hidden size in the sampled configuration	16, 32, 64	Sequential

- Quantitative Attributes

Name	Type	Semantics	Range	Direction
budget	Ratio	Budget (e.g., epoch) used to evaluate the configuration	[1, 300]	Sequential
config_momentum	Ratio	The value of momentum in the sampled configuration	[0, 1]	Sequential
config_learning_rate	Ratio	The learning rate value in the sampled configuration	[1e-4, 1e-1]	Sequential
config_weight_decay	Ratio	The value of weight decay in the sampled configuration	[0, 1e-3]	Sequential
sample_loss	Ratio	Loss value for the sampled configuration	[0, inf)	Sequential
sample_acc	Ratio	Accuracy value for the sampled configuration	[0, 100]	Sequential