Datasets

* Boston Housing Price prediction
* MNIST Handwritten digits
* MNIST fashion dataset
* CIFAR-10
* CIFAR-100

Algorithms

* BGD
* BGD+HM
* BGD+M
* BGD+M+HM
* BGD+N+M
* BGD+M+N+HM
* RMSprop
* RMSprop+HM
* Adam
* Adam+HM
* Adamax
* Adamax+HM

Algorithm parameters

* learning rate = 0.01 (for SGD, SGD+M, SGD+M+N), and 0.001 (for RMSprop, Adam, Adamax)
* momentum = 0.6 (for SGD, SGD+M, SGD+M+N, RMSprop)

Models

* DNN (2 hidden layers with 64 neurons)
* ResNet
* VGG

Parameters

* Training loss
* Training accuracy
* Test accuracy
* Training time