

1 Giới thiệu về Keras

Keras là một thư viện 'high level' với phần backend là TensorFlow, CNTK, hoặc Theano

2 Optimizer

2.1 Adam

Adam là một thuật toán tối ưu hóa nó có thể được sử dụng thay thế cho stochastic gradient descent để cập nhật trọng số của network thông qua việc training dữ liệu.

Adam được đề xuất bởi Diederik Kingma từ OpenAI và Jimmy Ba từ Đại học Toronto và năm 2015 trên báo ICLR với tiêu đề "Adam: A Method for Stochastic Optimization".

Adam hoạt động khác với stochastic gradient descent.

Stochastic gradient descent duy trì một [learning rate](#) độc lập cho việc cập nhật tất cả các trọng số và con số này sẽ không thay đổi trong suốt quá trình training.

[Learning rate](#) được duy trì cho từng trọng số mạng (tham số) và được điều chỉnh riêng trong quá trình học.