Chapter 1

Highway Layer

A Highway Layer is a type of Neural Network layer that uses a gating mechanism to control the information flow through a layer. Stacking multiple Highway Layers allows for training of very deep networks. Highway Layers work by learning a gating function that chooses which parts of the inputs to pass through and which parts to pass through a transformation function, such as a standard affine layer for example. The basic formulation of a Highway Layer is `T * h(x) + (1 - T) * x`, where T is the learned gating function with values between 0 and 1, $h(x)$ is an arbitrary input transformation and x is the input. Note that all of these must have the same size.