We propose a method to stabilize training of Recurrent Neural Networks (RNNs). The RNN is one of the most successful models to handle the time-series data in many applications such as speech recognition or machine translation. However, training of RNNs requires trial and error, and expertise since training of RNNs is difficult due to the gradient exploding problem. In this study, we focus on the Gated Recurrent Unit (GRU), which is one of the modern RNN models. We reveal the parameter point at which training of GRUs is disrupted by the gradient exploding problem and propose an algorithm to prevent the gradient from exploding. Our method can reduce time for trial and error, and does not require in-depth expertise to tune the hyper-parameters for training of GRU.

Reference