

REGULARIZATION PROPERTIES OF DROPOUT

Anna Shalova

Eindhoven University of Technology, Netherlands

a.shalova@tue.nl

Generalization is a crucial aspect of training algorithms in machine learning. Dropout training has been empirically shown to improve generalization of different models including neural networks and generalized linear models. In this work, we give a theoretical explanation of this phenomenon. We introduce a time-continuous analog of dropout gradient descent called Ornstein-Uhlenbeck dropout and study its behavior in the small noise limit. We obtain an effective limit model, in which the regularization term induced by dropout is explicit.

Joint work with Mark Peletier (Eindhoven University of Technology, Netherlands) and André Schlichting (WWU Münster, Germany).