Influence of blackout regularization on diverse datasets

Mateusz Garbacz, Michael Schoustra, Ramy Al Sharif, Valerie Pourquie, Thomas Puppels

Introduction

The primary task in Machine Learning (ML) is creating a model that not only performs well on the training samples, but also on samples from the data distribution that have not yet been seen before. Overfitting on training data is a central problem in this task, where a learning model has been attuned too well to the training data at the cost of its generalization performance. There are numerous methods that enforce simplification of the model by, for instance, penalizing high weights [1][2]. By doing this, L1 regularization can turn off some of the connections between (deep) layers in a network. However, currently, one has no direct control over how many connections are disabled, which for L1 regularization is controlled indirectly by a regularization factor. In this project, we introduce a regularization technique based on L1, in which one can define the percentage of connections between layers that are deactivated. As a result, dense networks become sparser, retaining a network of the same depth and number of nodes, but with less parameters to train due to the reduced number of connections in the network. At the same time, the rest of the weights remain unconstrained once the target number of links are deactivated.

Our research will be focused on the influence of blackout regularization on deep learning models.

Methods

Results

Discussion

References

[1] S. Scardapane, D. Comminiello, A. Hussain, and A. Uncini. Group sparse regularization for deep neural networks.
Neuro-computing, 241:81 – 89, 2017.
[2]Y. B. Ian Goodfellow and A. Courville. Deep learning. Book in preparation for MIT Press, 2016.