THE UNIVERSITY OF CHICAGO

Department of Statistics

MASTER'S THESIS PRESENTATION

KENNETH RUITER

Department of Statistics
The University of Chicago

Experimenting with Encoded Classifiers

THURSDAY, September 19, 2019, at 11:00 AM
Jones 304, 5747 S. Ellis Avenue

ABSTRACT

Using artificially created data sets to train a neural network can save a considerable amount of time and money. If this data is a 'clean' version of the actual test data, the noise from the latter must somehow be filtered out during the classification process. In this report we introduce the notion of encoded classifiers, and experiment with some examples. The MNIST data set is used as the clean data, and modifications are used as the noisy data. We consider Gaussian noise, inverting the colors, a mix of the two and finally a fully, randomly colored version of the data set. The neural network consists of two input layers, one for the clean data to train the classifier, and one for the noisy data to train the encoder. In an intermediate hidden layer, the L^2 loss is computed. A linear combination of this loss and the cross entropy loss for the classifier is minimized. The L^2 loss is computed at the first hidden layer for the basic network. An extended version is considered where the loss is computed at the second hidden layer. Finally we experiment with freezing the weights that correspond with the classifier input layer, when passing data that trains the encoder, so that the classifier and encoder are trained disjointly.

The Gaussian Noise Data set shows a 98.68% accuracy when passing the data through the classifier input layer, using the basic network after 20 epochs. When passing the data through the intended input layer an accuracy of 98.96% is reached, however this shows that the data is not noisy enough, so the other data sets are considered. The Inverted Color Data set, Combined Data set and Colored data set reach a test accuracy of 97.78%, 89.49% and 48.53% respectively. For a better performance, the extended network is used on the same three data sets, resulting in accuracies of 98.44%, 98.13% and 95.42% respectively. Freezing the weights generally gives similar, but slightly worse results. The more interesting cases are discussed. Using the basic network and the

---

For information about building access for persons with disabilities, please contact Keisha Prowoznik at 773.702-0541 or send an email to kprowoznik@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following web site: https://lists.uchicago.edu/web/subscribe/statseminars.

inverted color data set, a test accuracy of 94.87% is reached. For the extended network and the inverted color, combined and colored data sets respectively, the test accuracies are 98.76%, 97.87% and 95.40%. An advantage that freezing the weights has, is that there is one less parameter that needs to be optimized, which can save a considerable amount of time for larger networks or data sets.