Born Again Neural Networks

Radek Bartyzal

Let's talk ML in Prague

Date TBA

Prior work

Ensembles

Diverse models with similar validation performances can be often be combined to achieve predictive power superior to each of the constituent models. [3]

Born again trees

Learn a single tree that is able to recover the performance of a multiple-tree predictor. [4]

Knowledge distillation = model compression

Transfer knowledge acquired by a learned teacher model to a new simpler student model. [5]

Knolewdge distillation

Teacher

- high-capacity model
- good performance

Student

- more compact model
- not as good performance as the teacher but better than if it was trained without it

By transferring knowledge, one hopes to benefit from the student's compactness while suffering only minimal degradation in performance.

Born Again Networks

- not compressing models
- students are parameterized identically to their parents.
-

Surprisingly, these born again networks (BANs), tend to outperform their teacher models. Our experiments with born again dense networks demonstrate state-of-the-art performance on the CIFAR-100 dataset reaching a validation error of 15.5% with a single model and 14.9% with our best ensemble. Additionally, we investigate knowledge transfer to architectures that are different, but with capacity comparable to their teachers. In these experiments, we show that similar advantages can be achieved by transferring knowledge between dense networks and residual networks of similar capacity.

Sources

- Tommaso Furlanello et al. "Born Again Neural Networks." Workshop on Meta-Learning (MetaLearn 2017) at NIPS. Accessible from: http://metalearning.ml/papers/metalearn17_furlanello.pdf
- Breiman, Leo. "Statistical modeling: The two cultures (with comments and a rejoinder by the author)." Statistical science 16.3 (2001): 199-231. Accessible from: https://projecteuclid.org/download/pdf_1/euclid.ss/1009213726%20
- Hansen, Lars Kai, and Peter Salamon. "Neural network ensembles." IEEE transactions on pattern analysis and machine intelligence 12.10 (1990): 993-1001.
- Breiman, Leo, and Nong Shang. "Born again trees." Accessible from: https://www.stat.berkeley.edu/~breiman/BAtrees.pdf ps (1996).