

Федеральное государственное автономное образовательное учреждение
высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»
Физтех-школа Прикладной Математики и Информатики
Кафедра интеллектуальных систем

Направление подготовки / специальность: 03.03.01 Прикладные математика и физика

Направленность (профиль) подготовки: Математическая физика, компьютерные технологии и математическое моделирование в экономике

ПОРОЖДЕНИЕ МОДЕЛЕЙ ЗАДАНОЙ СЛОЖНОСТИ С ИСПОЛЬЗОВАНИЕМ БАЙЕСОВСКИХ ГИПЕРСЕТЕЙ

(бакалаврская работа)

Студент:

Гребенькова Ольга Сергеевна

(подпись студента)

Научный руководитель:

Бахтеев Олег Юрьевич,

(подпись научного руководителя)

Консультант (при наличии):

(подпись консультанта)

Москва 2021

Аннотация

В работе исследуется задача построения модели глубокого обучения. Предлагается способ контроля ее сложности. Под сложностью модели понимается минимальная длина описания, минимальный объем информации, который требуется для передачи информации о модели и о выборке. Сложность рассматривается как параметр, который может быть задан на этапе получения модели на основе имеющихся вычислительных и других эксплуатационных требований. Чтобы контролировать сложность модели, вводятся вероятностные предположения о распределении параметров модели глубокого обучения. В работе предложены три формы регуляризации, которые позволяют управлять распределением параметров модели. Предлагается метод оптимизации параметров модели, основанный на представлении модели глубокого обучения в виде гиперсети с использованием байесовского подхода. Под гиперсетью понимается модель, которая генерирует параметры оптимальной модели. Предлагается подход, максимизирующий нижнюю вариационную оценку байесовской обоснованности модели. Вариационная оценка рассматривается как условная величина, зависящая от требуемой сложности модели. Для анализа качества предлагаемого алгоритма проведены эксперименты на выборках MNIST и CIFAR.

Contents

1	Introduction	4
2	Problem statement	6
3	Hypernetworks for the model complexity control	8
4	Experiments	10
4.1	Experimental settings	10
4.2	MNIST experiment results	11
4.3	CIFAR experiment results	13
5	Conclusion	25

References

- [1] Graves, A. 2011. Practical Variational Inference for Neural Networks. *Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011*. Granada. 2348–2356.
- [2] Ha, D., A.M. Dai, and Q.V. Le. 2017. HyperNetworks. *The International Conference on Learning Representations (ICLR) 2017*. Toulon.
- [3] Saxena, S., and J. Verbeek. 2016. Convolutional Neural Fabrics. *Advances in Neural Information Processing Systems 29: 30th Annual Conference on Neural Information Processing Systems 2016*. Barcelona. 4053–4061
- [4] Xie, S., H. Zheng, C.Liu., and L. Lin. 2019. SNAS: Stochastic Neural Architecture Search. *The International Conference on Learning Representations (ICLR) 2019*. New Orleans.
- [5] Wu, B., X. Dai, P. Zhang, Y. Wang, F. Sun, Y. Wu, Y. Tian, P. Vajda, Y. Jia., and K. Keutzer. 2019. FBNet: Hardware-Aware Efficient ConvNet Design via Differentiable Neural Architecture Search. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach. 1:10726–10734
- [6] Kuznetsov, M., A. Tokmakova, and V. Strijov. 2016. Analytic and Stochastic Methods of Structure Parameter Estimation. *Informatica*. 27:607–624.
- [7] Bakhteev, O. Yu., and V. V. Strijov. 2018. Deep Learning Model Selection of Sub-optimal Complexity. *Autom Remote Control*. 79:1474–1488.
- [8] Lorraine, J., and D. Duvenaud. 2018. Stochastic Hyperparameter Optimization through Hypernetworks. *CoRR*. Available at: <http://arxiv.org/abs/1802.09419>.
- [9] LeCun, Y., C. Cortes, and C. Burges. 1998. The MNIST dataset of handwritten digits. Available at: <http://yann.lecun.com/exdb/mnist/index.html>.
- [10] Krizhevsky A., N. Vinod and G. Hinton. The CIFAR-10 (Canadian Institute for Advanced Research) dataset. Available at <http://www.cs.toronto.edu/~kriz/cifar.html>
- [11] Bishop, C. 2006. Pattern Recognition and Machine Learning. *Springer*.
- [12] Song, H., J. Pool, J. Tran and W. Dally. 2015. Learning both Weights and Connections for Efficient Neural Network. *Advances in Neural Information Processing Systems*. 28.