

Nizhny Novgorod State University
Institute of Information Technologies, Mathematics and Mechanics
Department of Computer Software and Supercomputer Technologies

Educational course
«Introduction to deep learning
using the Intel® neon™ Framework»

Lecture №3
Introduction to the Intel® neon™ Framework

Supported by Intel

Kustikova Valentina

Nizhny Novgorod
2018

Content

1	Abstract	3
2	Literature	3
2.1	Books.....	3
2.2	References	3

1 Abstract

The goal of this lecture is to introduce the listener to the capabilities of the Intel® neon™ Framework. The lecture describes the Intel® neon™ Framework: the purpose of developing the framework, the license, the necessary resources to get start [4 – 7], technical information about software and hardware requirements. The general procedure of installing the Intel® neon™ Framework for Linux is provided, which includes a sequence of command lines. A scheme for developing a deep model using the Intel® neon™ Framework is considered. The development of the model assumes the generating a backend, loading data, specifying a deep model, training and evaluating a deep model. The backend determines where the calculations will be performed: CPU, CPU + MKL, GPU (Pascal, Maxwell or Kepler). Loading data can be performed using standard loaders provided by library developers, or implemented by the programmer. The development of own data loaders involves the development of an iterator for working with data in the HDF5 format. To specify the architecture of a deep model, different types of layers, activation functions, and initializers are introduced. To specify the parameters of the training method, a cost function, a metric, and an optimization method are chosen. Along with this, the framework provides opportunities for organizing your own schedule of updating the learning rate parameter. Training and testing of the model are performed using the built-in functions of the Intel® neon™ Framework. In the lecture, the implementation of each stage is discussed in more detail, classes and basic methods are given. The description of the stage is accompanied by an example of the program development. At the end of the lecture, a complete example for solving the problem of classifying a person's sex from a photograph based on the IMDB-WIKI dataset [8] using a multilayered fully-connected network is considered [1 – 3].

2 Literature

2.1 Books

1. Haykin S. Neural Networks: A Comprehensive Foundation. – Prentice Hall PTR Upper Saddle River, NJ, USA. – 1998.
2. Osofsky S. Neural networks for information processing. – 2002.
3. Goodfellow I., Bengio Y., Courville A. Deep Learning. – MIT Press. – 2016. – [<http://www.deeplearningbook.org>].

2.2 References

4. Intel® neon™ Framework [<https://github.com/NervanaSystems/neon>].
5. Intel® neon™ Framework Documentation (API + tutorials) [<http://neon.nervanasys.com/docs/latest>].
6. Intel AI Academy [<https://software.intel.com/ru-ru/ai-academy/frameworks/neon>].
7. A set of trained models for Intel® neon™ Framework [<https://github.com/NervanaSystems/ModelZoo>].
8. IMDB-WIKI dataset [<https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki>].