

Нижегородский государственный университет им. Н.И. Лобачевского
Институт информационных технологий, математики и механики
Кафедра Математического обеспечения и суперкомпьютерных технологий

**Образовательный курс
«Введение в глубокое обучение
с использованием Intel® neon™ Framework»**

**Лекция №8
Обзор возможностей инструмента
Intel® nGraph™**

При поддержке компании Intel

Кустикова В.Д.

Нижний Новгород
2018

Содержание

1	Аннотация	3
2	Литература	3
2.1	Основная литература	3
2.2	Ресурсы сети Интернет	3

1 Аннотация

Цель настоящей лекции состоит в том, чтобы познакомить слушателя с инструментом Intel® nGraph™ (nGraph) [4 – 7]. Инструмент позволяет сделать работу аналитиков данных, разработчиков алгоритмов, программных инженеров и программистов портативной, адаптивной и развертываемой для самого современного аппаратного обеспечения, доступного сегодня для решения задач машинного обучения [4].

В основе концепции nGraph лежит идея представления глубокой нейросетевой модели в виде модели *графа потоков данных* (data-flow graph). На основании указанного представления формируется состав библиотеки nGraph: мосты с инструментами глубокого обучения (framework bridges), преобразователи операций (transformer ops) и элементы для построения графа (graph shaping). В лекции более подробно рассматриваются концептуальные вопросы устройства nGraph.

Выделяется две схемы работы с nGraph: сборка стороннего инструмента глубокого обучения для работы с nGraph и последующее обучение глубокой модели, конвертация обученной модели в формат Open Neural Network Exchange Format (ONNX) [8] для запуска на аппаратной инфраструктуре, поддерживаемой nGraph. В лекции рассматривается первый вариант использования nGraph на примере инструмента Intel® neon™ Framework. Приводится процедура сборки и установки nGraph, а также Neon для работы с nGraph [9]. Рассматривается пример обучения и тестирования полносвязной сети для решения задачи распознавания рукописных цифр. Отмечаются отличия от рассмотренных в курсе примеров.

2 Литература

2.1 Основная литература

1. Хайкин С. Нейронные сети. Полный курс. – М.: Издательский дом «Вильямс». – 2006. – 1104 с.
2. Осовский С. Нейронные сети для обработки информации. – М.: Финансы и статистика. – 2002. – 344 с.
3. Goodfellow I., Bengio Y., Courville A. Deep Learning. – MIT Press. – 2016. – [http://www.deeplearningbook.org].

2.2 Ресурсы сети Интернет

4. Документация Intel® nGraph™ [http://ngraph.nervanasys.com/docs/latest].
5. Репозиторий исходных кодов Intel® nGraph™ [https://github.com/NervanaSystems/ngraph].
6. Статья на странице Intel AI Academy [https://ai.intel.com/ngraph-a-new-open-source-compiler-for-deep-learning-systems].
7. Intel® nGraph™. An Intermediate Representation, Compiler, and Executor for Deep Learning [https://arxiv.org/pdf/1801.08058.pdf].
8. Open Neural Network Exchange Format (ONNX) [http://onnx.ai].
9. A Deep Learning framework powered by Intel® nGraph [https://github.com/NervanaSystems/ngraph-neon].