

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 №2  
Многослойные полностью связанные сети**

*При поддержке компании Intel*

*Кустикова В.Д.*

Нижний Новгород  
2018

## Содержание

1	Аннотация .....	3
2	Литература .....	3
2.1	Основная литература .....	3

## 1 Аннотация

Цель данной лекции состоит в том, чтобы изучить модель многослойных полностью связанных сетей и метод обратного распространения ошибки для обучения параметров данной модели. Для этого вначале рассматривается детерминистская модель нейрона [1]. Вводится общая структура многослойной полностью связанной нейронной сети. Многослойная нейронная сеть содержит нейроны, которые распределены по слоям. В простейшем случае в сети существует **входной и выходной слою**, и сеть называется **однослойной**. В общем случае сеть может содержать множество промежуточных слоев, называемых **скрытыми слоями**, и является **многослойной**. Если все узлы слоя соединены с узлами следующего слоя, то слой называется **полностью связным**. Если данное условие выполнено для всех слоев сети, то сеть называется **полностью связной** или **полносвязной** (Fully-Connected Neural Network, FCNN). Рассматривается оптимизационная постановка задачи обучения многослойной полносвязной сети. Задача обучения однослойной нейронной сети ставится как задача минимизации **функции ошибки или функция стоимости** (cost function), отражающей разницу выходного сигнала нейронной сети (ожидаемого сигнала) и фактического сигнала, соответствующего текущему входу, по всей обучающей выборке. Данная функция представляет собой меру эффективности обучения. Вводится метод обратного распространения ошибки для обучения нейронных сетей на примере двухслойной полносвязной сети. Метод обратного распространения ошибки определяет стратегию выбора параметров сети с использованием градиентных методов оптимизации [2] в предположении, что целевая функция непрерывна. Градиентные методы на каждом шаге уточняют значения параметров, по которым проводится оптимизация. В лекции приводится детальный вывод математических формул для реализации метода обратного распространения ошибки для двухслойной сети, соответствующей логистической регрессии.

## 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>].