

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 №5  
Перенос обучения глубоких нейронных сетей**

*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

Перенос обучения является одним из широко применяемых подходов при решении практических задач. Цель настоящей лекции состоит в том, чтобы ввести понятие *переноса обучения* глубоких нейросетевых моделей [4] и продемонстрировать на практическом примере эффективность применения данного подхода.

Идея подхода состоит в том, чтобы заимствовать обученную модель для решения поставленной задачи. Если имеется некоторая *исходная задача*, то для ее решения необходимо собрать тренировочные данные и выполнить их разметку. Аналогичная процедура может быть выполнена для какой-то другой *целевой задачи*. Если эти задачи связаны некоторым образом, то имеет смысл использовать обученную модель решения исходной задачи для решения целевой. Цель переноса обучения состоит в том, чтобы накопить знания, необходимые для решения исходной задачи, и использовать их для решения близкой по смыслу целевой задачи [5].

Согласно общему подходу вводится четыре типа переноса обучения: прямое использование модели, использование структуры глубокой модели, использование модели в качестве фиксированного метода извлечения признаков, тонкая настройка параметров исходной модели для решения целевой задачи. Первая схема предполагает, что целевая задача является подзадачей исходной. В этом случае допустимо прямое использование обученной модели. Очевидно, что такой подход не всегда возможно реализовать. Вторая схема предполагает, что модель, построенная для решения исходной задачи, обучается на данных, подготовленных для решения целевой задачи. При этом веса модели инициализируются случайным образом. Идея третьей схемы состоит в том, чтобы удалить из глубокой модели классификатор (последние полностью связанные слои) и рассматривать начальную часть сети как метод извлечения признаков. При этом взамен старого классификатора можно поместить новый классификатор (например, другой набор полностью связанных слоев) и обучить его на признаках, построенных с использованием начальной части сети. Четвертый подход подобен третьему, разница лишь в том, что после замены классификатора вся модель обучается как единая система. Основной вопрос состоит в том, как определить, когда и какой тип переноса знаний необходимо использовать. В лекции рассматриваются основные факторы, которые оказывают значительное влияние на принимаемое решение [6].

Демонстрация практического применения переноса обучения выполняется на задаче классификации пола человека по фотографии. Приводятся результаты экспериментов, которые доказывают перспективность использования данного подхода на практике.

## 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>].
4. Pan S.J., Yang Q. A Survey on Transfer Learning // IEEE Transactions on Knowledge and Data Engineering. – 2010. – Vol. 22, Issue 10. – P.1345-1359.

### 2.2 Ресурсы сети Интернет

5. Transfer Learning – Machine Learning's Next Frontier» – [<http://ruder.io/transfer-learning>].
6. Transfer Learning [<http://cs231n.github.io/transfer-learning>].