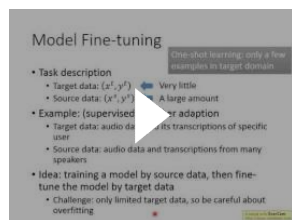


# Transfer Learning

17 February 2022 11:57

## ML Lecture 19: Transfer Learning



- Data **no directly related** to the task considered
- Speech recognition, image recognition, text analyze(specific domain)
- Overview

## Transfer Learning - Overview

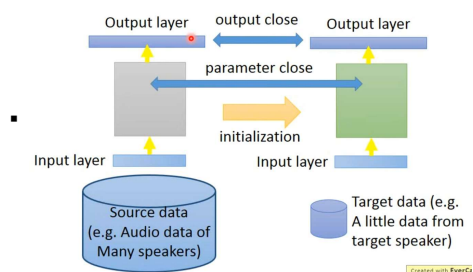
		Source Data (not directly related to the task)	
		labelled	unlabelled
Target Data	labelled	<b>Fine-tuning</b> <b>Multitask Learning</b>	<b>Self-taught learning</b> Rajat Raina , Alexis Battle , Honglak Lee , Benjamin Packer , Andrew Y. Ng, Self-taught learning: transfer learning from unlabeled data, ICML, 2007
	unlabelled	<b>Domain-adversarial training</b> <b>Zero-shot learning</b>	<b>Self-taught Clustering</b> Wenyuan Dai, Qiang Yang, Gui-Rong Xue, Yong Yu, "Self-taught clustering", ICML 2008

Source data 具有泛性, target data 为specific domain, 才能有好的结果, 反之不可。

### Model Fine-tuning

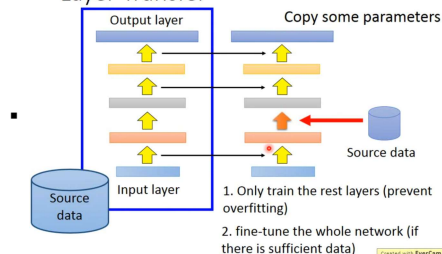
#### Conservative Training

#### Conservative Training



#### Layer Transfer

#### Layer Transfer



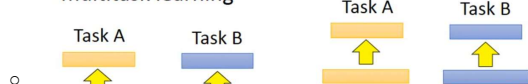
- Low level feature被使用
- Which layer should be chosen, based on the specific case
- 选择cover不同的层, 结果会不同

### Multitask Learning

- 同时关心target domain和source domain的效果。
- Examples: Translation, multilingual speech recognition

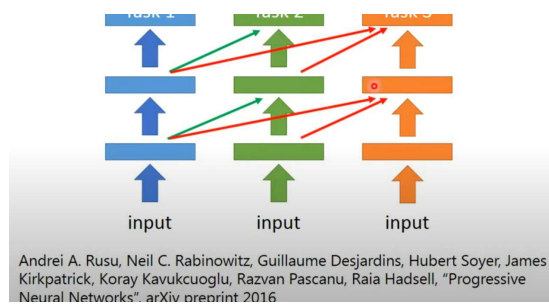
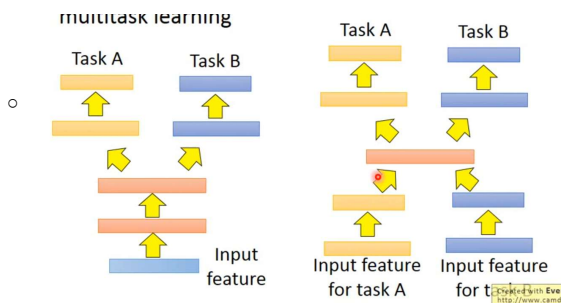
## Multitask Learning

- The multi-layer structure makes NN suitable for multitask learning



## Progressive Neural Networks



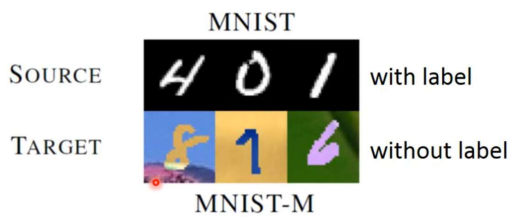


Andrei A. Rusu, Neil C. Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, Raia Hadsell, "Progressive Neural Networks", arXiv preprint 2016

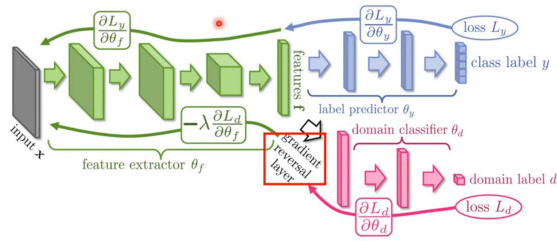
## Domain-adversarial training

### Task description

- Source data:  $(x^s, y^s) \rightarrow$  Training data
  - Target data:  $(x^t) \rightarrow$  Testing data
- mismatch



### Domain-adversarial training



Yaroslav Ganin, Victor Lempitsky, Unsupervised Domain Adaptation by Backpropagation, ICML, 2015

Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, Domain-Adversarial Training of Neural Networks, JMLR, 2016

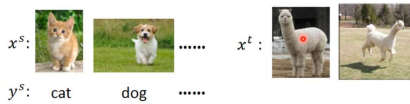
## Zero-shot learning

Target data 没有出现在Source data中

### Zero-shot Learning

<http://evchk.wikia.com/wiki/%E8%8D%B9%E6%B3%A5%E9%A6%A6>

- Source data:  $(x^s, y^s) \rightarrow$  Training data
  - Target data:  $(x^t) \rightarrow$  Testing data
- Different tasks

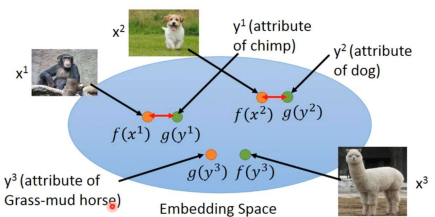


In speech recognition, we can not have all possible words in the source (training) data.

How we solve this problem in speech recognition?

### Zero-shot Learning

- Attribute embedding



### Zero-shot Learning

$$f^*, g^* = \arg \min_{f, g} \sum_n \|f(x^n) - g(y^n)\|_2 \quad \text{Problem?}$$

$$f^*, g^* = \arg \min_{f, g} \sum_n \max \left( 0, k - f(x^n) \cdot g(y^n) \right) + \max_{m \neq n} f(x^n) \cdot g(y^m)$$

Margin you defined

$$\text{Zero loss: } k - f(x^n) \cdot g(y^n) + \max_{m \neq n} f(x^n) \cdot g(y^m) < 0$$

$$f(x^n) \cdot g(y^n) - \max_{m \neq n} f(x^n) \cdot g(y^m) > k$$

$f(x^n)$  and  $g(y^n)$  as close  $f(x^n)$  and  $g(y^m)$  not as close

## Self-taught learning

### Self-taught learning

- Learning to extract better representation from the source data (unsupervised approach)
- Extracting better representation for target data

Domain	Unlabeled data	Labeled data	Classes	Raw features
Image classification	10 images of outdoor scenes	Caltech101 image classification dataset	101	Intensities in 14x14 pixel patch
Handwritten character recognition	Handwritten digits ("0"-"9")	Handwritten English characters ("a"-"z")	26	Intensities in 28x28 pixel character/digit image
Font character recognition	Handwritten English characters ("a"-"z")	Font characters ("a"/"A"/"z"/"Z")	26	Intensities in 28x28 pixel character image
Song genre classification	Song snippets from 10 genres	Song snippets from 7 different genres	7	Log-frequency spectrogram over 50ms time windows
Webpage classification	100,000 news articles (Reuters newswire)	Categorized webpages (from DMOZ hierarchy)	2	Bag-of-words with 500 word vocabulary
UseNet article classification	100,000 news articles (Reuters newswire)	Categorized UseNet posts (from "SRAA" dataset)	2	Bag-of-words with 377 word vocabulary