Lecture 1: Introduction

Yang Feng[1], Ye Tian[2]

[1]Department of Biostatistics, School of Global Public Health, New York University
[2]Department of Statistics, Columbia University

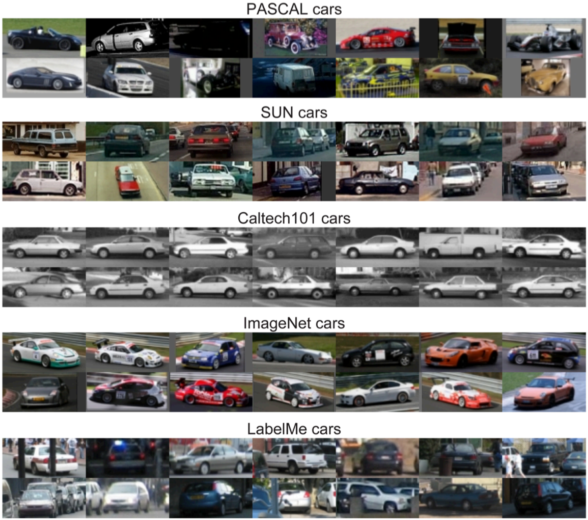
May 20, 2024

The slides were initially used in a full-day short course "A (Selective) Introduction to the Statistical Foundations of Transfer Learning" at the 37th New England Statistics Symposium (NESS 2024), University of Connecticut. All the slides are publicly available at http://www.columbia.edu/~yt2661/STL.html and free to use in the educational context. The source should be mentioned.

Overview I

§1.1: Why, when, and where

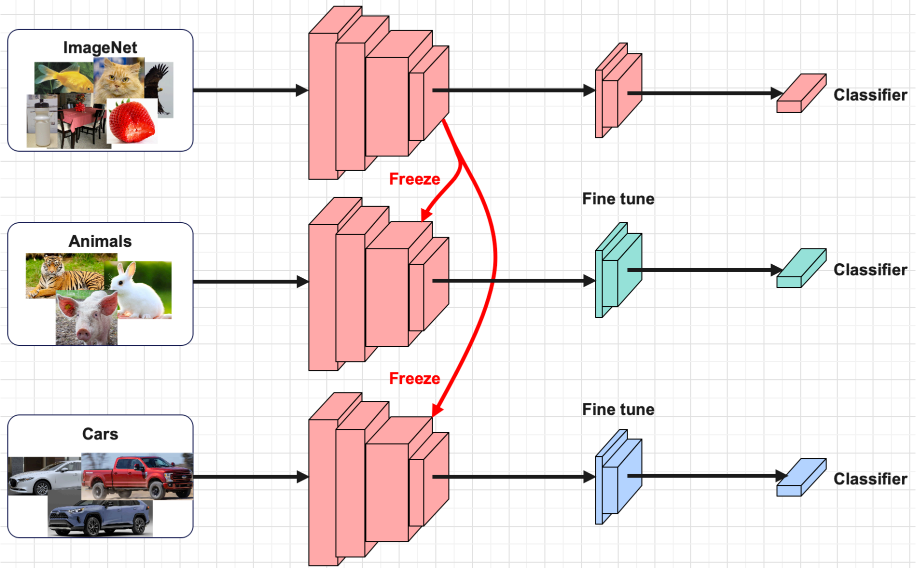
Transfer learning in computer vision



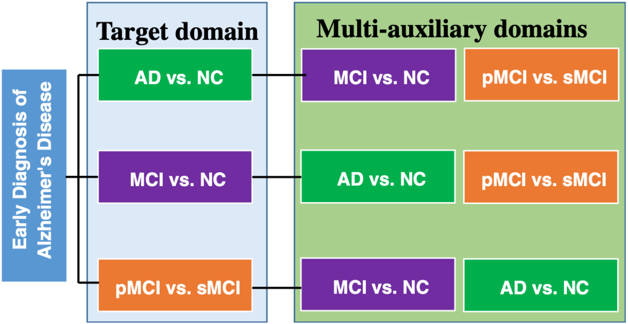
Car images across datasets

Torralba, A., & Efros, A. A. (2011, June). Unbiased look at dataset bias. In CVPR 2011 (pp. 1521-1528). IEEE.

Transfer learning in computer vision



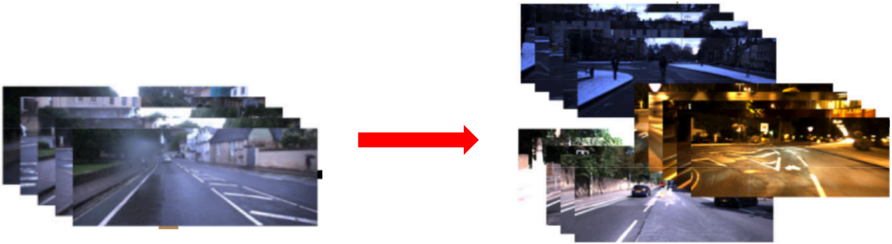
Transfer learning in disease diagnosis



- AD: Alzheimer's Disease
- MCI: mild cognitive impairment
  - pMCI: progressive MCI
  - sMCI: stable MCI
- NC: normal controls

Cheng, B., Liu, M., Shen, D., Li, Z., Zhang, D., & Alzheimer's Disease Neuroimaging Initiative. (2017). Multi-domain transfer learning for early diagnosis of Alzheimer's disease. Neuroinformatics, 15, 115-132.

Transfer learning in autonomous driving

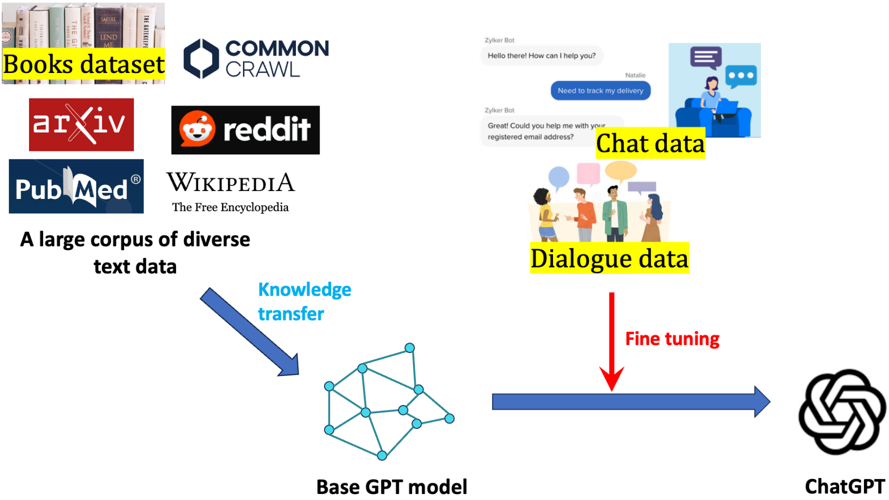


Training scenario **Real application scenario**

Image source:
https://nbviewer.org/github/vistalab-technion/cs236605-tutorials/blob/master/tutorial6/tutorial6-TL_DA.ipyn

Akhauri, S., Zheng, L. Y., & Lin, M. C. (2020, October). Enhanced transfer learning for autonomous driving with systematic accident simulation. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 5986-5993). IEEE.

Transfer learning in ChatGPT



Why, when, and where

- **Why:** Improve the performance on the target dataset of limited (or zero) sample size
- **When:** There is some similarity between the target and source to transfer
- **Where:** Various applications in many areas
  - Computer vision
  - Healthcare
  - Natural Language Processing
  - Finance
  - Autonomous Driving
  - ......

See Pan and Yang (2009); Weiss et al. (2016); Zhuang et al. (2020) for more comprehensive surveys of transfer learning.

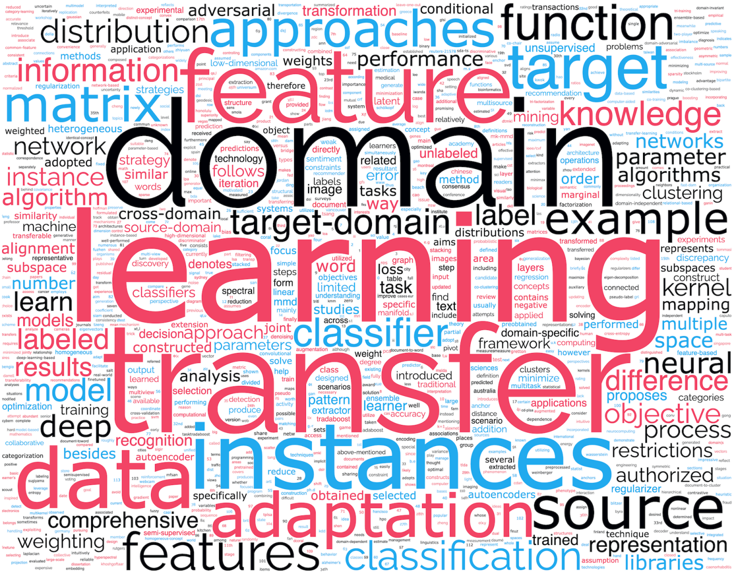
[1] Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.

[2] Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. Journal of Big data, 3, 1-40.

[3] Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. Proceedings of the IEEE, 109(1), 43-76.

§1.2: Concepts

Terminologies do not matter



The word cloud is generated based on "Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. Proceedings of the IEEE, 109(1), 43-76."

Concepts: domain and task

Most definitions are quoted from Redko et al. (2019, 2020).

- **Domain:** We may call $(\mathcal{S}^{(0)}, \mathbb{P}^{(0)})$ as the **target domain** and $(\mathcal{S}^{(1)}, \mathbb{P}^{(1)})$ as the **source domain**, where $\mathbb{P}^{(0)}$ and $\mathbb{P}^{(1)}$ are the distributions defined on corresponding spaces $\mathcal{S}^{(0)}$ and $\mathcal{S}^{(1)}$. [1]
- **Task:** There is a learning **task** associated with each domain, which can be classification, clustering, regression, on $\mathcal{S}^{(k)}$. For supervised learning problems, $\mathcal{S}^{(k)}$ consists a feature space $\mathcal{X}$ and an output space $\mathcal{Y}$. The task usually is to learn a predictor from $\mathcal{X}$ to $\mathcal{Y}$. We can denote the task for target and source as $t^{(0)}$ and $t^{(1)}$
- **Transfer learning:** How to improve the performance on target domain for $t^{(0)}$ by the knowledge of the source domain for $t^{(1)}$.

[1] There can be more than one sources.

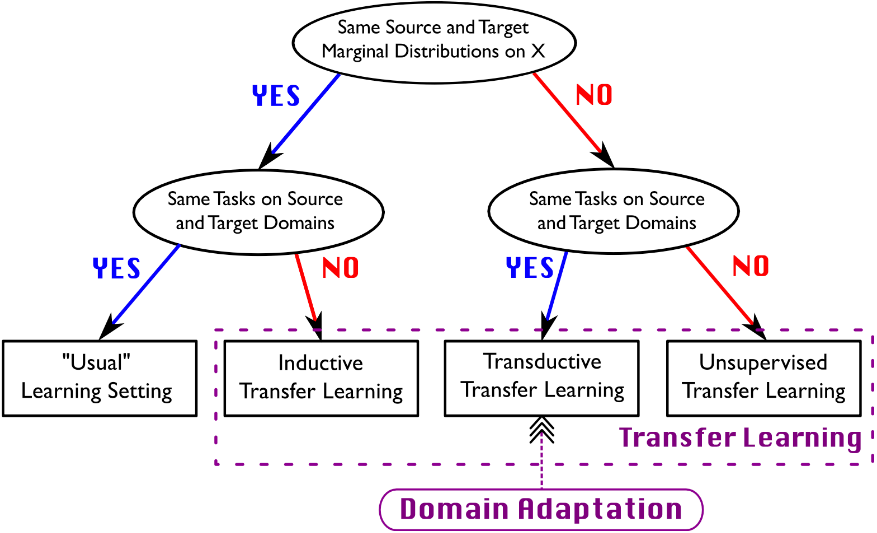
[1] Redko, I., Morvant, E., Habrard, A., Sebban, M., & Bennani, Y. (2019). Advances in domain adaptation theory. Elsevier.

[2] Redko, I., Morvant, E., Habrard, A., Sebban, M., & Bennani, Y. (2020). A survey on domain adaptation theory: learning bounds and theoretical guarantees. arXiv preprint arXiv:2004.11829.

Categories of TL by goals

- **Transfer learning:** Using the information from the **source domain** to improve the performance on **target domain**.
- **Multi-task learning:** Each domain is called a **task** and we want to simultaneously learn well on **all domains** by borrowing information from each other.

Categories of TL by tasks

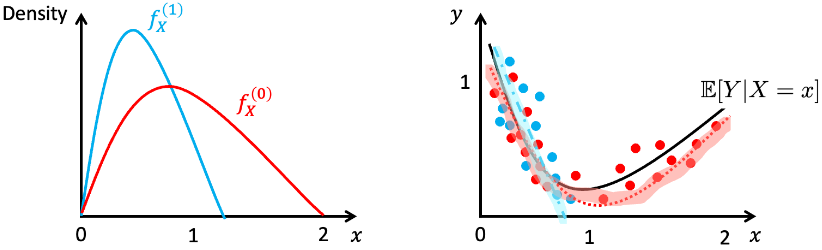


Picture source: Redko, I., Morvant, E., Habrard, A., Sebban, M., & Bennani, Y. (2020). A survey on domain adaptation theory: learning bounds and theoretical guarantees. arXiv preprint arXiv:2004.11829.

Categories of TL by shift types

Consider the supervised learning setting, and we have the same feature space $\mathcal{X}$ and output space $\mathcal{Y}$ for target and source [2]. $\mathbb{P}^{(k)}$ is the joint distribution of (X, Y) on target $(k = 0)$ or source $(k = 1)$ domains.

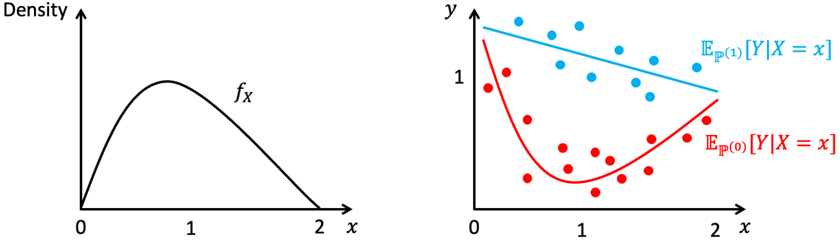
- **Covariate shift:** $\mathbb{P}_X^{(0)} \neq \mathbb{P}_X^{(1)}$, $\mathbb{P}_{Y|X}^{(0)} = \mathbb{P}_{Y|X}^{(1)}$ (in the sense of $\mathbb{P}_X^{(0)}$-a.e.)
- **Posterior drift:** $\mathbb{P}_X^{(0)} = \mathbb{P}_X^{(1)}$, $\mathbb{P}_{Y|X}^{(0)} \neq \mathbb{P}_{Y|X}^{(1)}$
- **Concept drift:** $\mathbb{P}_X^{(0)} \neq \mathbb{P}_X^{(1)}$, $\mathbb{P}_{Y|X}^{(0)} \neq \mathbb{P}_{Y|X}^{(1)}$



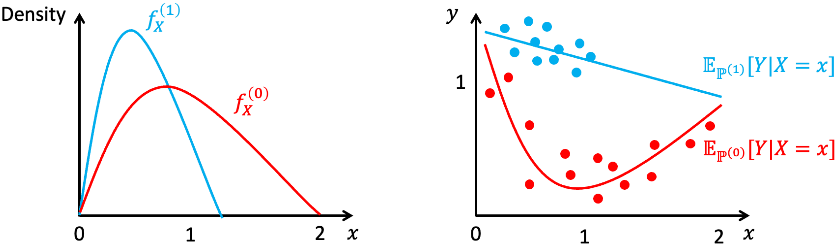
"Covariate shift"

[2] We will also discuss distribution shift where the space $\mathcal{X} \times \mathcal{Y}$ can be different for target and source, in Lecture 3.

Categories of TL by shift types



"Posterior drift"



"Concept drift"

§1.3: Overview of the short course

Overview of the short course

A (Selective) Introduction to the Statistical Foundations of Transfer Learning

Lecture 2: Generic Analysis of Domain Adaptation by Divergence Notions

- No target data
- Few target data + many source data
- Posterior drift
- Hardness results
- Other similarity notions beyond divergence

Lecture 3: Covariate Shift

- Adaptivity to covariate shift for free?
- The reweighting method
- Approaches of density ratio estimation
- Other methods

Lecture 4: Posterior Drift and Biased Regularization

- Biased regularization
- Extension to high-dimensional regressions
- Other applications of biased regularization
- Statistical inference
- Unsupervised multi-task learning
- Representation learning

Next time!

Lecture 5: Selected Topics (coming in future)

- Adversarial contamination and outlier tasks
- Federated learning
- Differential privacy
- ......

A few things

- We will mainly focus on the intuition, methodology, and simple theoretical justification. But compared to other tutorials on transfer learning and domain adaptation online, this short course is <u>much more theoretical</u> (that's why it concerns the "statistical foundations" of TL)
- This is a big area, and different people may view it from different perspectives. That's why we added a big "selective" to the name of the course.
- You may not be able to catch all the intuitions for now. Neither do us. But after this short course, you should be able to read most of papers in this field, no matter they are theoretical, methodological, or applied.

References I

Akhauri, S., Zheng, L. Y., and Lin, M. C. (2020). Enhanced transfer learning for autonomous driving with systematic accident simulation. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5986–5993. IEEE.

Cheng, B., Liu, M., Shen, D., Li, Z., Zhang, D., and Initiative., A. D. N. (2017). Multi-domain transfer learning for early diagnosis of alzheimer's disease. *Neuroinformatics*, 15:115–132.

Pan, S. J. and Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359.

Redko, I., Morvant, E., Habrard, A., Sebban, M., and Bennani, Y. (2019). *Advances in domain adaptation theory*. Elsevier.

Redko, I., Morvant, E., Habrard, A., Sebban, M., and Bennani, Y. (2020). A survey on domain adaptation theory: learning bounds and theoretical guarantees. *arXiv preprint arXiv:2004.11829*.

Torralba, A. and Efros, A. A. (2011). Unbiased look at dataset bias. In *CVPR 2011*, pages 1521–1528. IEEE.

References II

Weiss, K., Khoshgoftaar, T. M., and Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, 3(1):1–40.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., and He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1):43–76.

Special thanks to

Haotian Lin (Penn State University, Statistics)
Binghe Zhu (Columbia University, Statistics)
Gan Yuan (Columbia University, Statistics)
Haiyan Zheng (University of Bath, Math)

for helpful discussions and providing references