GDPR
Data Shortage and AI

Qiang Yang
Hong Kong University of Science and Technology

AI and Big Data

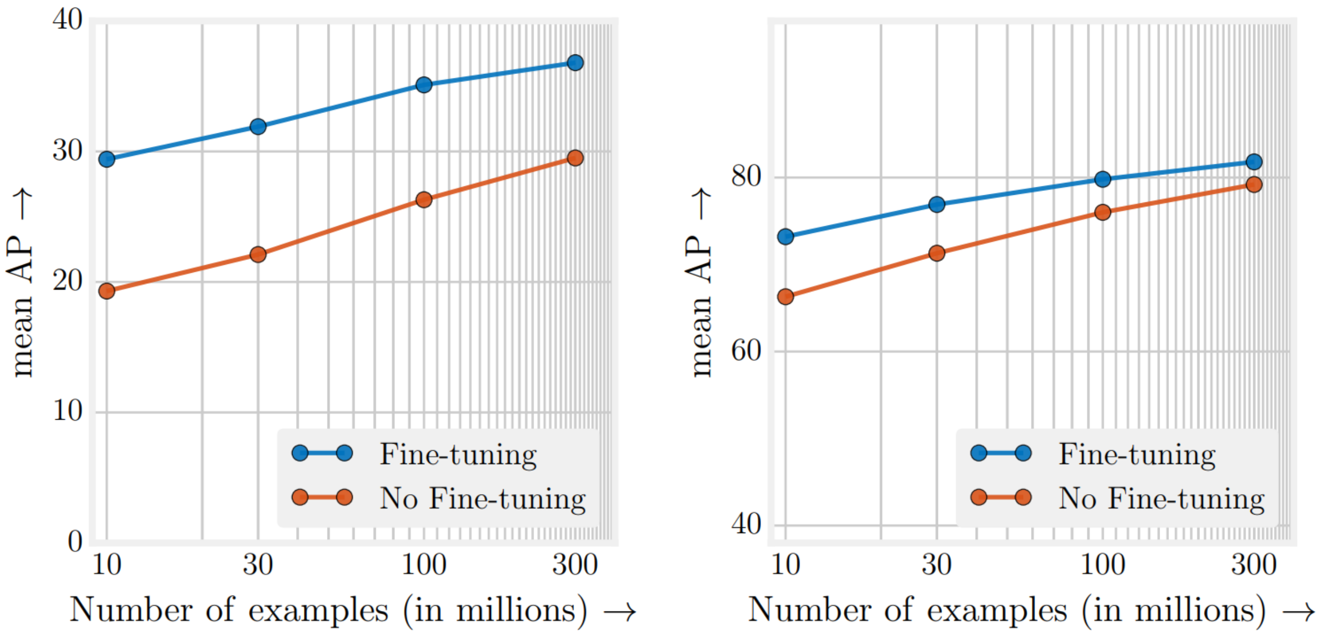


Figure 4. Object detection performance when initial checkpoints are pre-trained on different subsets of JFT-300M from scratch. x-axis is the data size in log-scale, y-axis is the detection performance in mAP@[.5,.95] on COCO minival* (left), and in mAP@.5 on PASCAL VOC 2007 test (right).

“Revisiting Unreasonable Effectiveness of Data in Deep Learning Era.” Google Research, 2017

1. Most Applications Have Only Small Data

- Contract review law firms typically have annotated 10K - 20K of labeled contracts as samples (Bradley Arsenault, Electric Brain 2018)

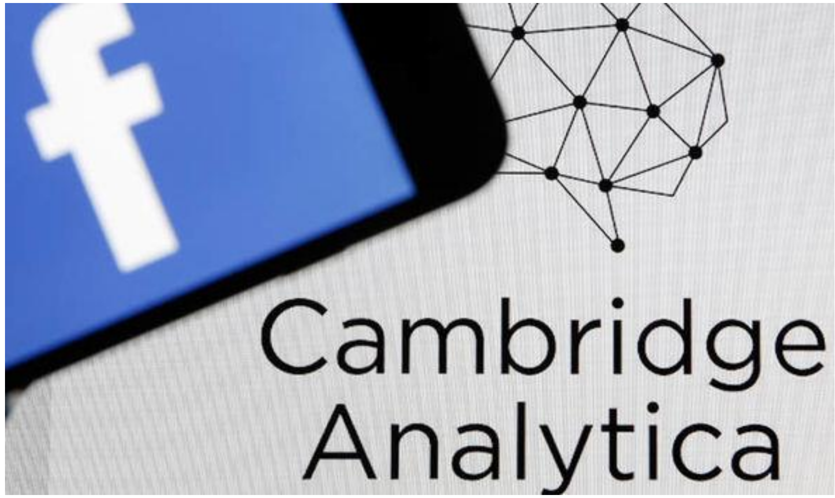
- In finance industry, large loans are few, with only ~ 100 examples as typical samples (4paradigm.com, 2017)

- In medical image recognition, high-quality labeled data are few (A Survey on Deep Learning in Medical Image Analysis, Geert Litjens, et al. 2017 Arxiv.)

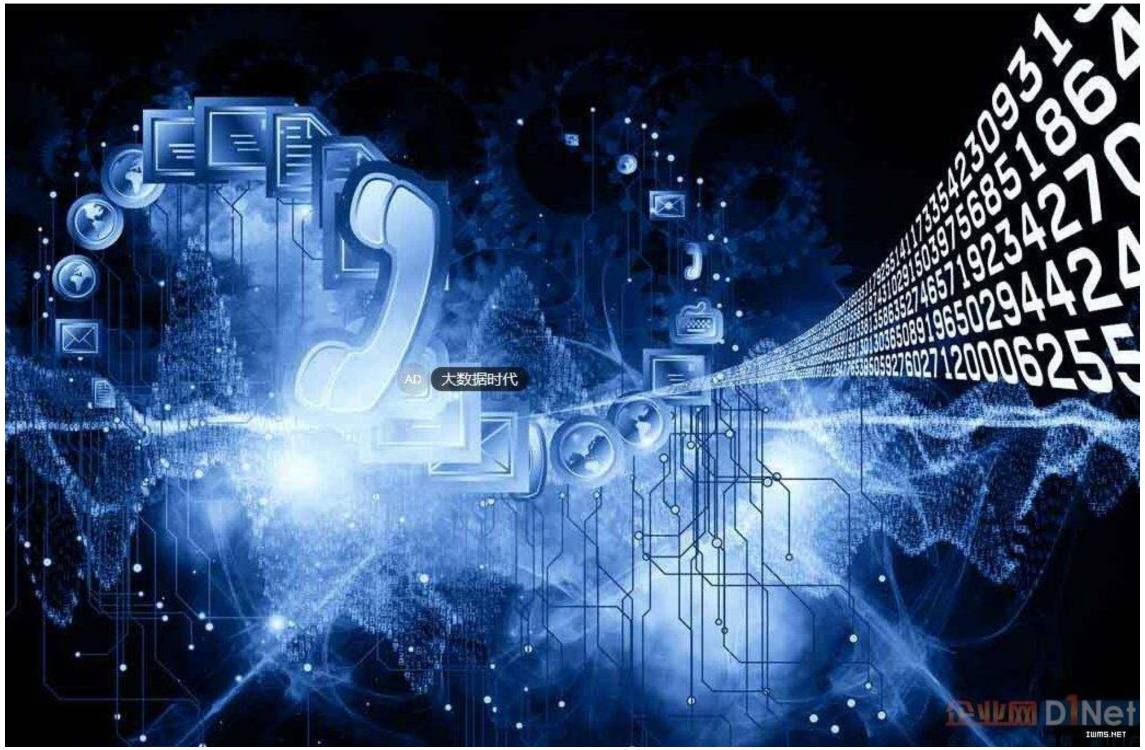
2. Data Sharing Among Parties: Difficult, Impossible or Immoral

- **Medical clinical trial data cannot be shared (by Rogier Stegeman 2018 , on Genemetics)**

- **Our society demands more control on data privacy and security**
  - GDPR, Government Regulations
  - Corporate Security and Confidentiality Concerns
  - Data privacy concerns



Reality: Data often in form of Isolated Islands



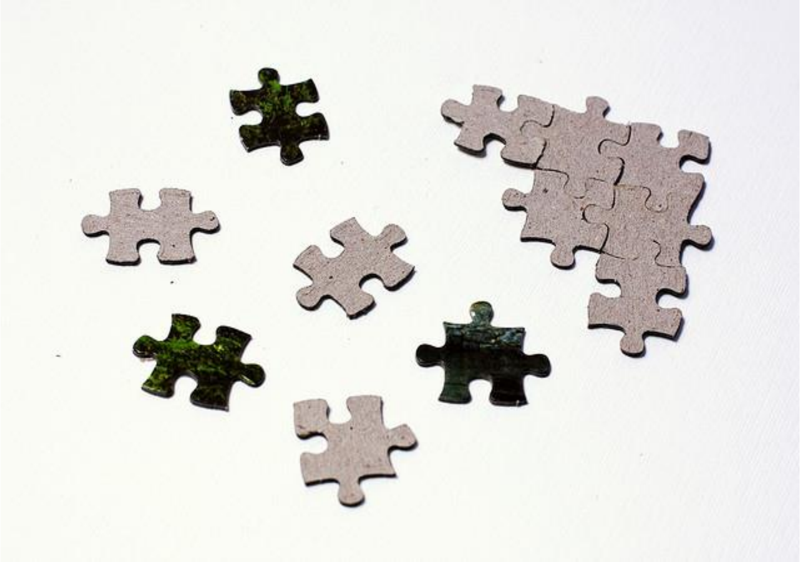
Two Challenges and Two Solutions

- **Small Data**

Transfer Learning from source data and models

- **Fragmented Data**

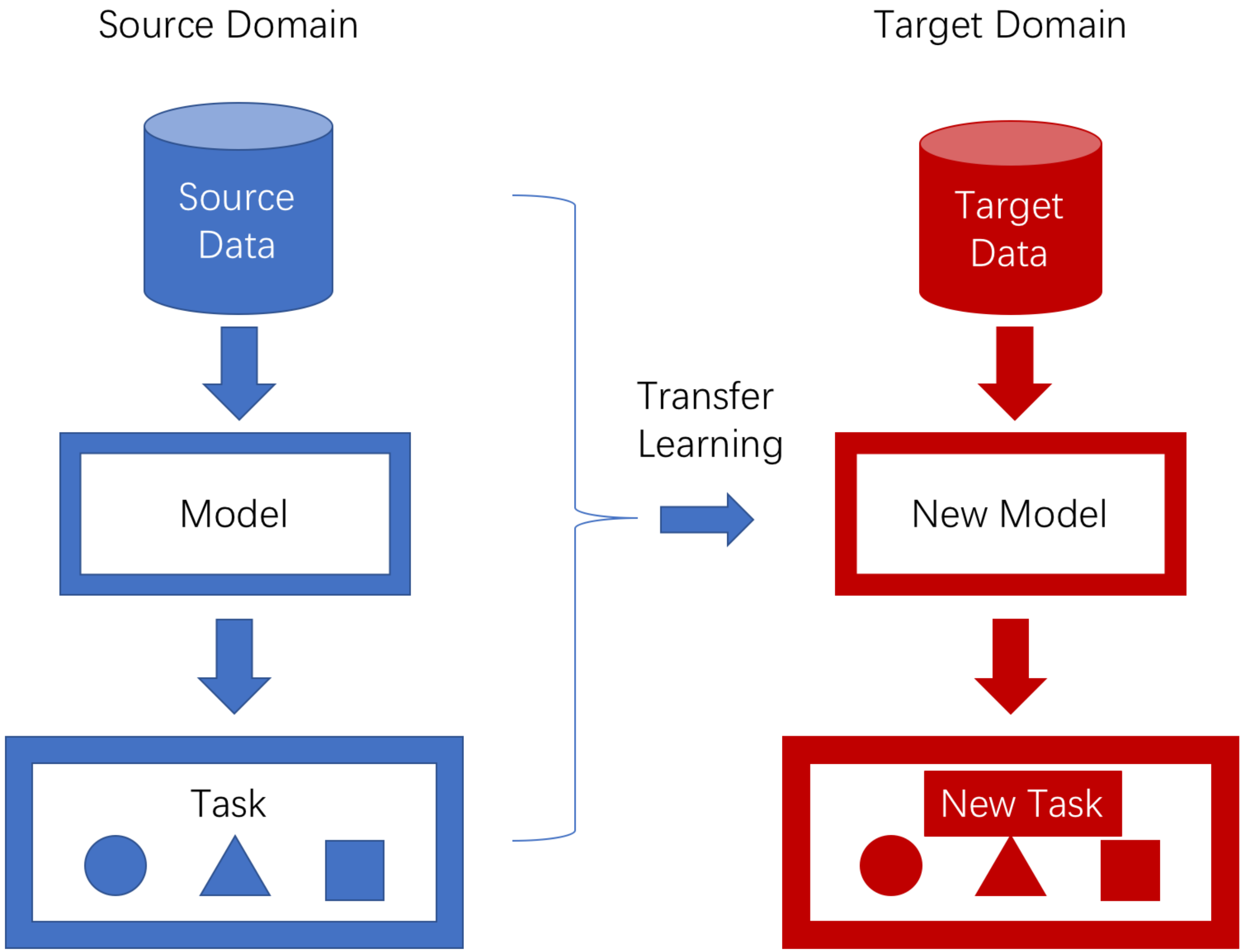
Federated learning with many parties



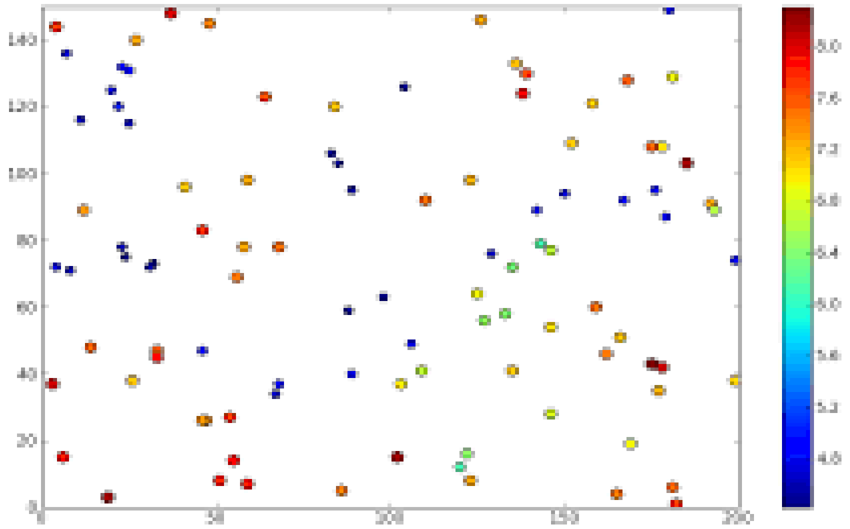
Often, these two problems occur together

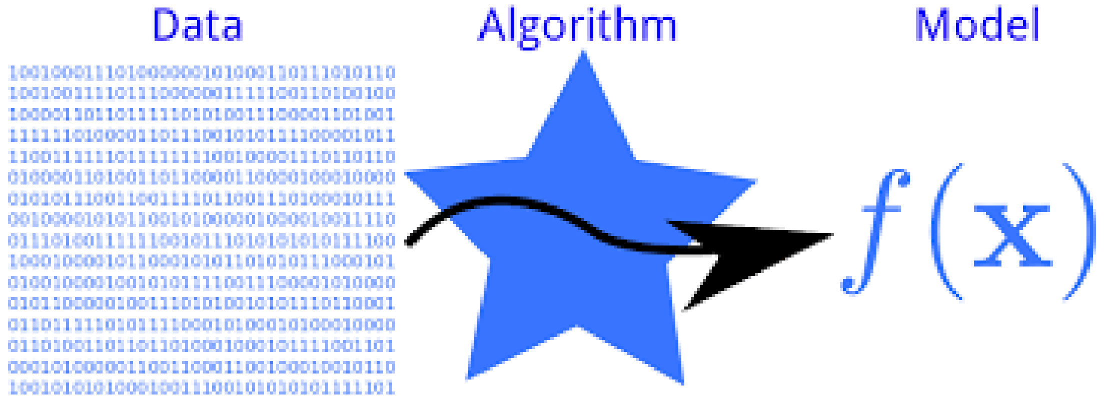
Transfer Learning

Transfer Learning Models

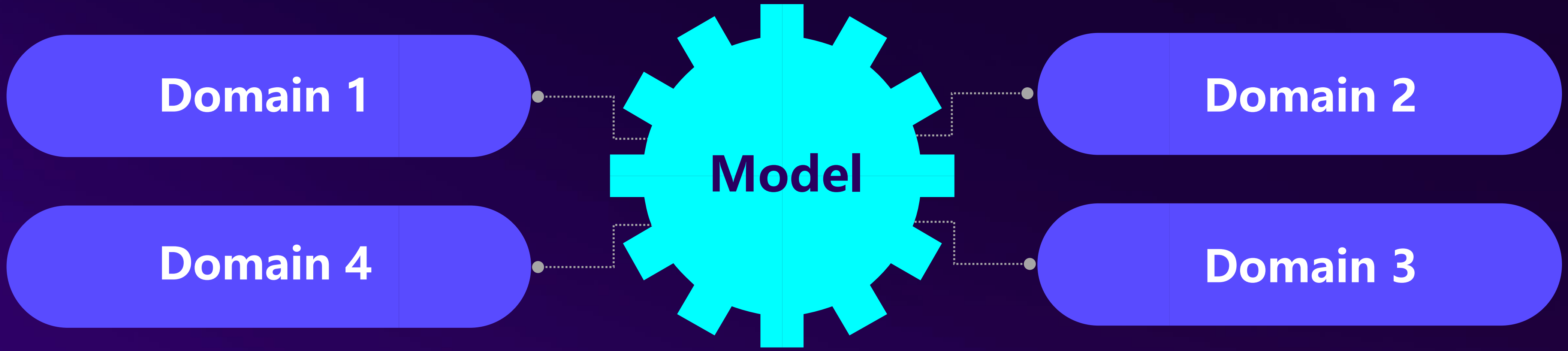


Why Transfer Learning? Small Data





Why Transfer Learning : Reliability



Why Transfer Learning? Personalization

Learning to Transfer

Research issues

- When to transfer
- How to transfer
- What to transfer
- Learning how to learn by transfer learning

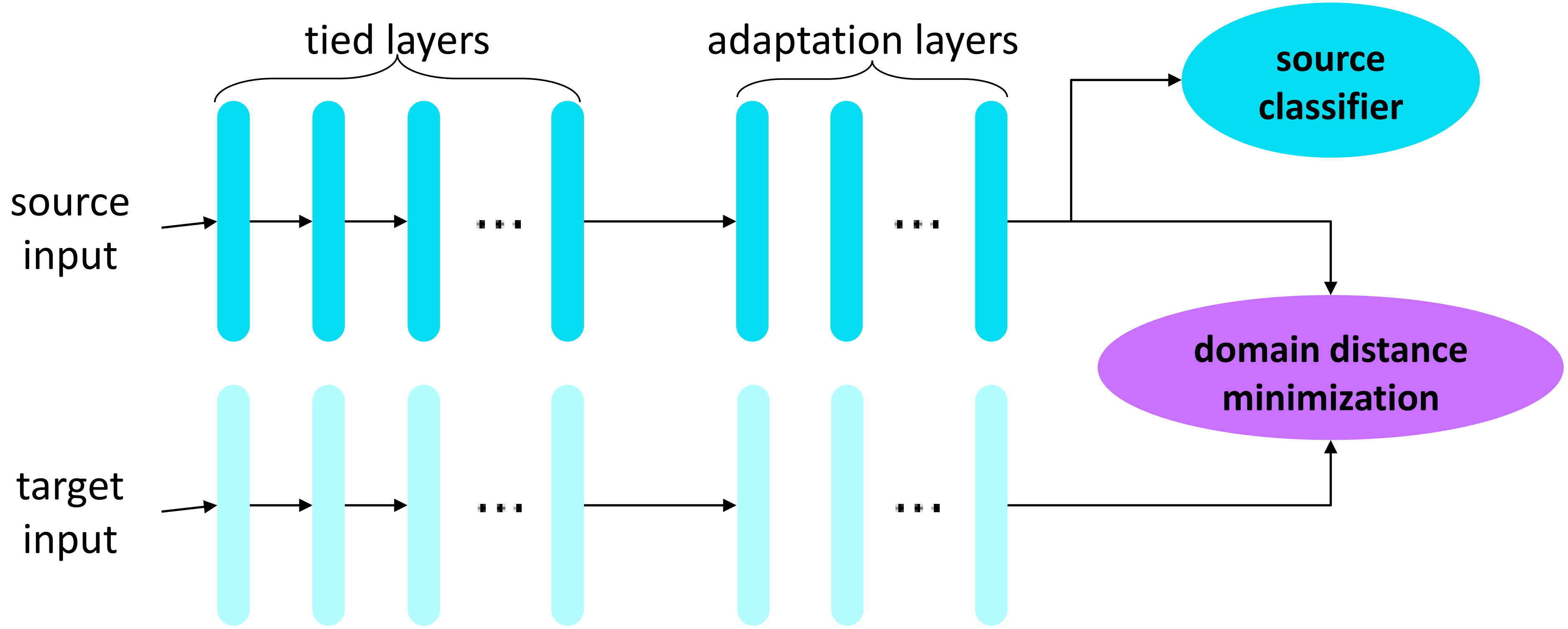
Key to Transfer Learning : Finding the Invariance

Driving in Mainland China

Driving in Hong Kong SAR, China

Transfer Learning in a Deep Model

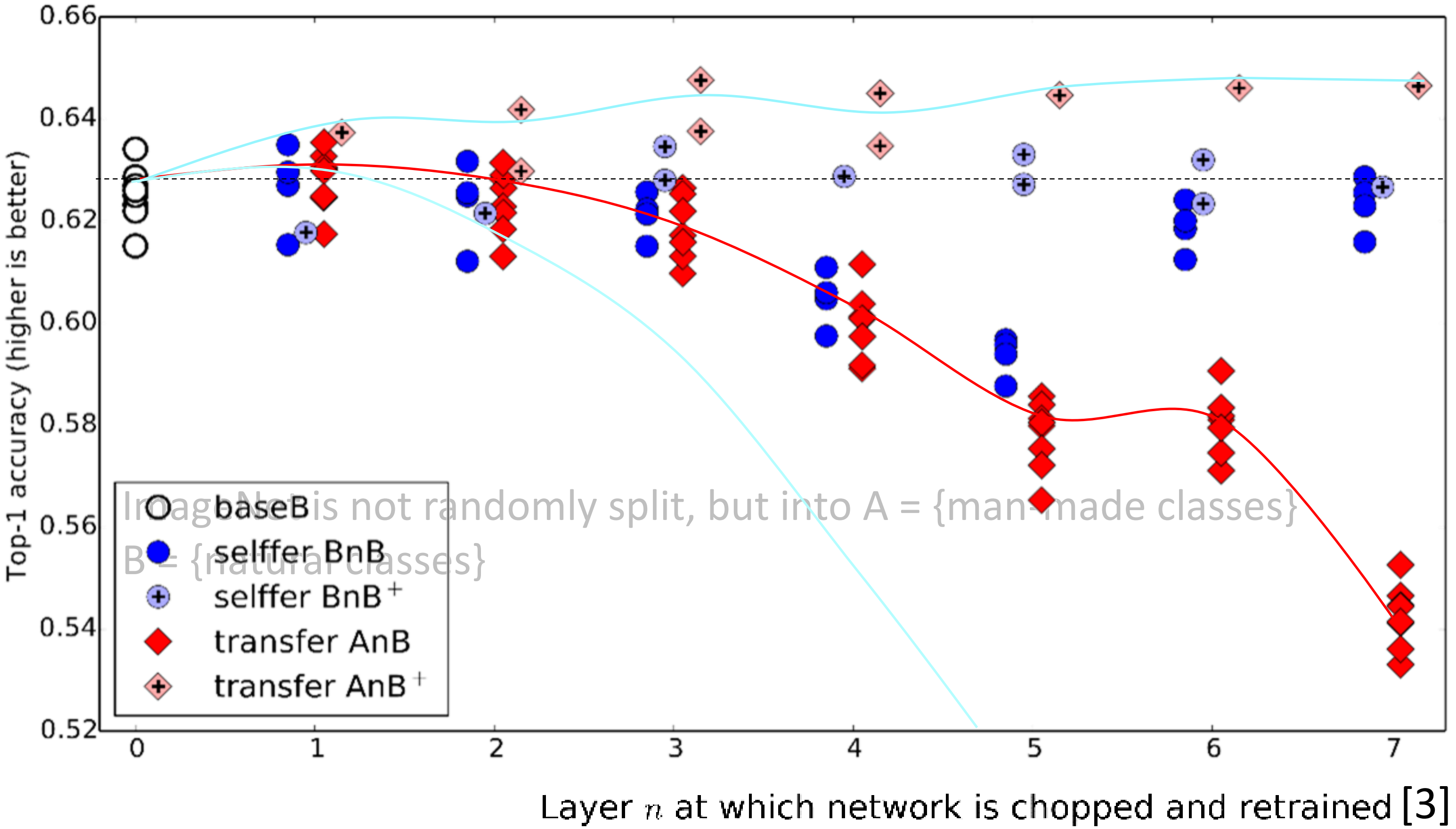
- **Objective** $\mathcal{L} = \mathcal{L}_{\text{source}} + \mathcal{L}_{\text{distance}}$



Learning transferable features with deep adaptation networks. M Long, Y Cao, J Wang, MI Jordan. International Conference on Machine Learning (ICML) 2015

Transfer Learning in a Deep Model

A Quantitative Study



ImageNet is not randomly split, but into A = {man-made classes} B = {natural classes}

[3]

Conclusion: lower layer features are more general and transferrable, and higher layer features are more specific and non-transferrable.

Yosinski, Jason, et al. "How transferable are features in deep neural networks?." NIPS. 2014.

Transfer Learning Setting I :

- Source domain: sufficient labeled data
- Target domain: no labeled data
- Domain Adaptation

Transfer Learning Setting II :

- Source domain: sufficient labeled data
- Target domain: little labeled data
- Supervised Transfer Learning

Transfer Learning Setting I

Source domain: sufficient labeled data

Target domain: no labeled data

Sentiment Analysis

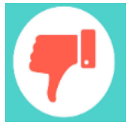
rating 10/10

This movie will blow your mind and break your heart - and make you desperate to go back for more. Brave, brilliant and better than it has any right to be.

shawneofthedead 25 April 2018

Over the past decade, Marvel has earned itself the benefit of the doubt. The studio has consistently delivered smart, funny, brave films that both embrace and transcend their comic-book origins. The 18 blockbuster movies produced since Iron Man first blasted off into the stratosphere in 2008 have not only reinvented superhero films as a genre - they've helped to legitimise it. Indeed, Marvel's two most recent films - Thor: Ragnarok and Black Panther - have received the kind of accolades usually reserved for edgy arthouse flicks.

rating 1/10



I actually laughed out loud at the end

tenaciouspeas 23 May 2018

What a trash heap of a movie. I thought about giving it 2 stars because there were a couple of things that made me chuckle but I left the theater so irritated that I talked myself out of it. I kept singing the "I don't care" song for the last 2 hours of this movie, which seemed to last at least 5 hours long. I'm sure they could have fit at least 2 more bad CGI action fight scenes in there, to make it 6 hours long. I loved the first Avengers. I loved Thor Ragnarok. I hated this movie which can easily be summed up: A really long movie about a boring CGI character titled: Here's Thanos!

44 out of 80 found this helpful. Was this review helpful? Yes No | Report this

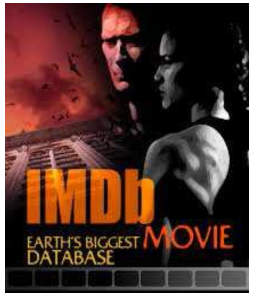
- **Single-Domain Solution**

  depends on sufficient labeled data

- **Cross-domain solution: Transfer Learning**

  Transferring sentiment classification knowledge from one domain to another

Cross-Domain Features: Pivots

Source domain (**Movie**)	Target domain (**Electronics**)
	
Great movie. His characters are **engaging** and **thoughtful**.	This great touchpad feels **glossy** and is **responsive**.
It's a **excellent**, **sobering** drama.	It is very **lightweight**, **excellent** transition from PC.
An **terrible** movie. It is very **plotless** and **insipid**.	It is **blurry** and **fuzzy** in very dark setting. So **terrible** HP.

Domain adaptation with structural correspondence learning, Blitzer et al. EMNLP 2006

Structural Correspondence Learning (SCL)

- Unlabeled step: **pivot predictors**
  - **pseudo-label:** select M pivot features from keywords
  - Each pivot predictor aligns non-pivot features from **source** to **target** domains.

- **Example:**

Movie

review1: Very movie. His characters are **engaging** and **thoughtful**.

Electronics

review2: This touchpad feels **glossy** and is **responsive**.

Binary problem: Does the pivot "**great**" appear in the review?

- **Transformed samples:**

N non-pivot features

	engaging	thoughtful	responsive	glossy	...
review1:	1	1	0	0	...
review2:	0	0	1	1	...

M pseudo labels

great	awful	...
1	0	...
1	0	...

John Blitzer et al. Biographies, bollywood, boomboxes and blenders: Domain adaptation for sentiment classification. **EMNLP 2007**

Movie

	engaging thoughtful responsive glossy	Sobering lightweight	Plotless insipid Blurry fuzzy
	**1**	0	0
	0	**1**	0
	0	0	**1**

Training

$$y = f(\mathbf{x}) = sgn(\mathbf{w}\mathbf{x}^{\mathrm{T}}), \ \mathbf{w} = [1, 1, -1]$$

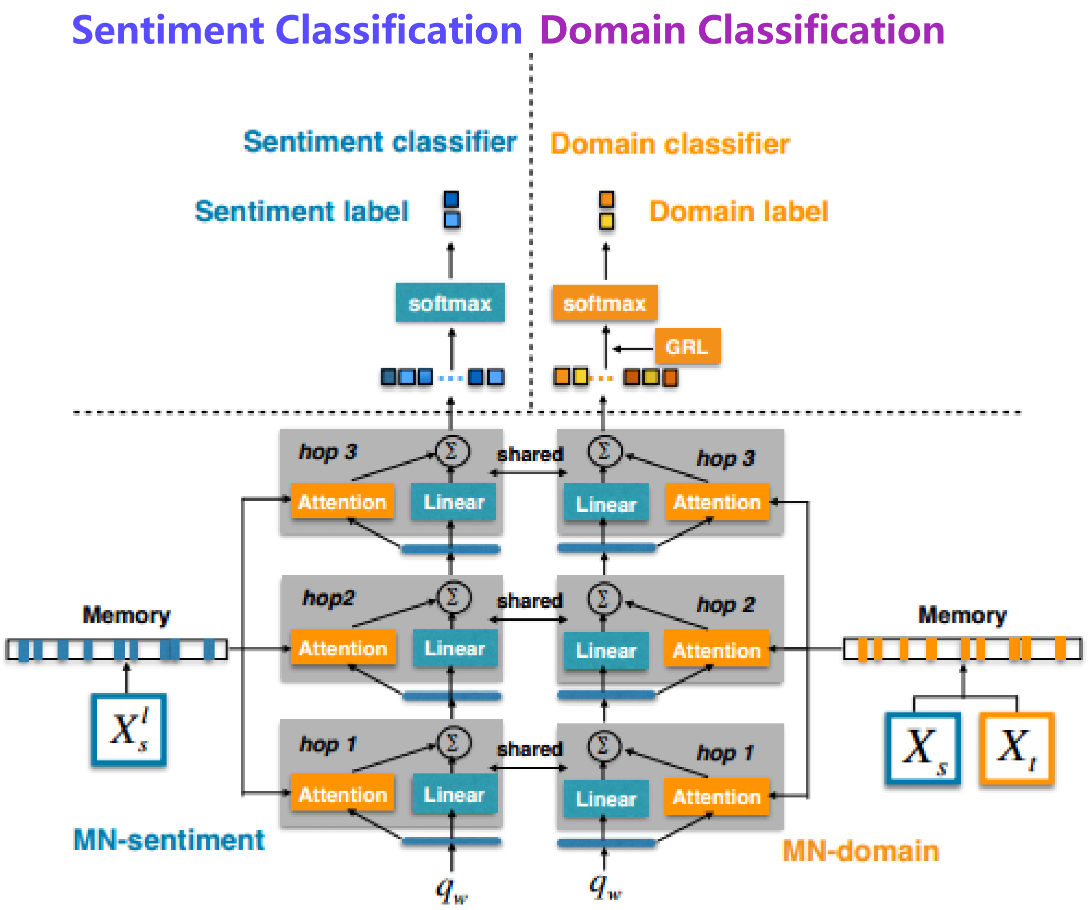
Prediction

Electronics

	engaging thoughtful responsive glossy	Sobering lightweight	Plotless insipid Blurry fuzzy
	**1**	0	0
	0	**1**	0
	0	0	**1**

Sinno Jialin Pan et al. Cross-domain sentiment classification via spectral feature alignment. WWW-10

An Adversarial Approach



Domain Classification Objective:
Maximize domain classification error

Source data X_s
Target data X_t

S

T

Movie
Great movie. His characters are engaging and thoughtful.
Electronics
This great touchpad feels glossy and is responsive.

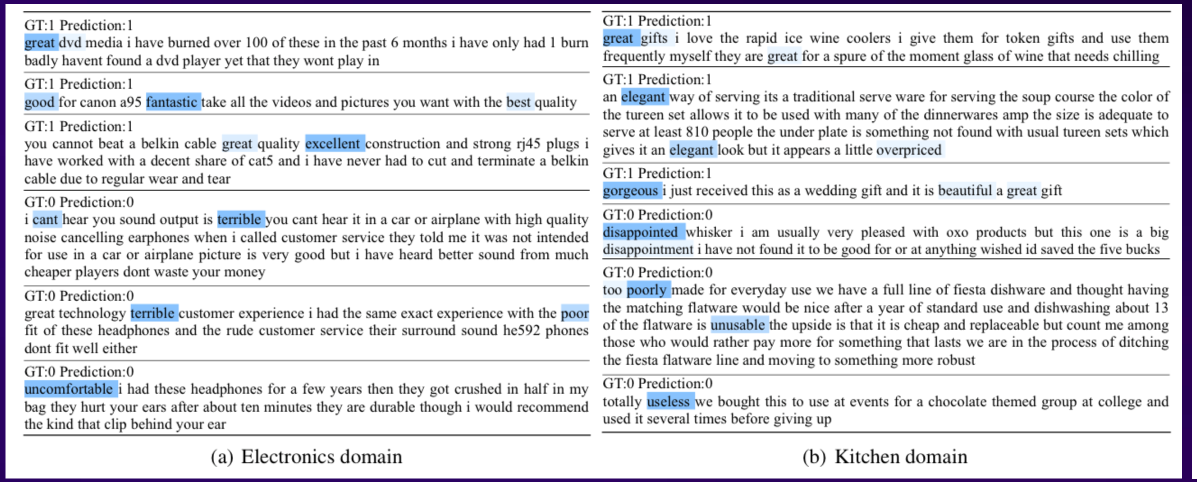
Li, Zheng, Qiang Yang, et al. "End-to-end adversarial memory network for cross-domain sentiment classification." IJCAI 2017.

Comparison with baseline methods

- **Traditional methods:**

  SCL: Structural Correspondence Learning [Blitzer et al., 2006]
  SFA: Spectral Feature Alignment [Pan et al., 2010]

- **AMN model significantly outperforms the traditional methods SFA and SCL on Amazon Reviews Dataset**



GT:1 Prediction:1
great dvd media i have burned over 100 of these in the past 6 months i have only had 1 burn badly havent found a dvd player yet that they wont play in

GT:1 Prediction:1
good for canon a95 fantastic take all the videos and pictures you want with the best quality

GT:1 Prediction:1
you cannot beat a belkin cable great quality excellent construction and strong rj45 plugs i have worked with a decent share of cat5 and i have never had to cut and terminate a belkin cable due to regular wear and tear

GT:0 Prediction:0
i cant hear you sound output is terrible you cant hear it in a car or airplane with high quality noise cancelling earphones when i called customer service they told me it was not intended for use in a car or airplane picture is very good but i have heard better sound from much cheaper players dont waste your money

GT:0 Prediction:0
great technology terrible customer experience i had the same exact experience with the poor fit of these headphones and the rude customer service their surround sound he592 phones dont fit well either

GT:0 Prediction:0
uncomfortable i had these headphones for a few years then they got crushed in half in my bag they hurt your ears after about ten minutes they are durable though i would recommend the kind that clip behind your ear

(a) Electronics domain

GT:1 Prediction:1
great gifts i love the rapid ice wine coolers i give them for token gifts and use them frequently myself they are great for a spure of the moment glass of wine that needs chilling

GT:1 Prediction:1
an elegant way of serving its a traditional serve ware for serving the soup course the color of the tureen set allows it to be used with many of the dinnerwares amp the size is adequate to serve at least 810 people the under plate is something not found with usual tureen sets which gives it an elegant look but it appears a little overpriced

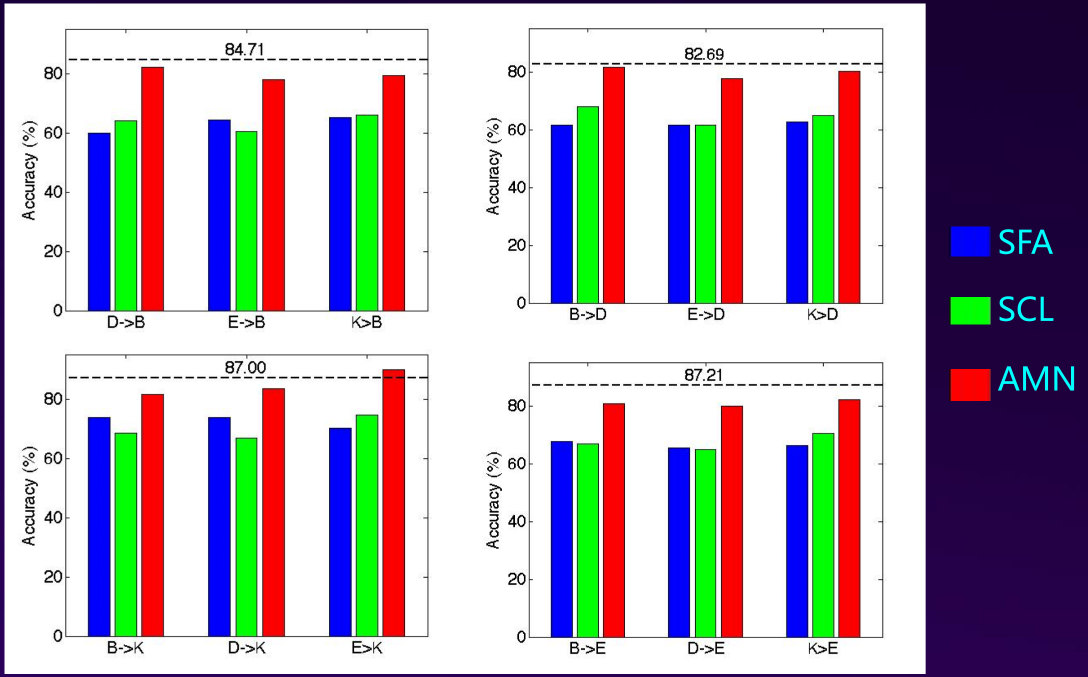
GT:1 Prediction:1
gorgeous i just received this as a wedding gift and it is beautiful a great gift

GT:0 Prediction:0
disappointed whisker i am usually very pleased with oxo products but this one is a big disappointment i have not found it to be good for or at anything wished id saved the five bucks

GT:0 Prediction:0
too poorly made for everyday use we have a full line of fiesta dishware and thought having the matching flatware would be nice after a year of standard use and dishwashing about 13 of the flatware is unusable the upside is that it is cheap and replaceable but count me among those who would rather pay more for something that lasts we are in the process of ditching the fiesta flatware line and moving to something more robust

GT:0 Prediction:0
totally useless we bought this to use at events for a chocolate themed group at college and used it several times before giving up

(b) Kitchen domain



Transfer Learning Setting II : Supervised Transfer Learning

Source domain: sufficient labeled data
Target domain: little labeled data

Transfer Learning in Dialog Systems

Source Domain

	Alice' s 21st Coffee Shopping Dialogue
X_1	Can I have coffee please?
Y_1	What coffee would you like?
X_2	I would like a cup of **Latte**.
Y_2	**Hot Latte** deliver to **No.101 Shandong** Road?
X_3	Yes, exactly!

Alice' s Dialogues

Bob' s Dialogues

…

Transfer

Target Domain

	John' s 3rd Coffee Shopping Dialogue
X_1	Can I have coffee please?
Y_1	What coffee would you like?
X_2	I would like a cup of **Mocha**.
Y_2	?

Candidate Reply Set

Y_{c1}: **Cold Mocha** deliver to **No.1199 Mingsheng** Road?

Y_{c2}: What is your address?

Y_{c3}: Hot Mocha or Iced Mocha?

Kaixiang Mo, Qiang Yang, et al. : Personalizing a Dialogue System With Transfer Reinforcement Learning. AAAI 2018.

Learning Common Dialogue States and a Personalized Q-function

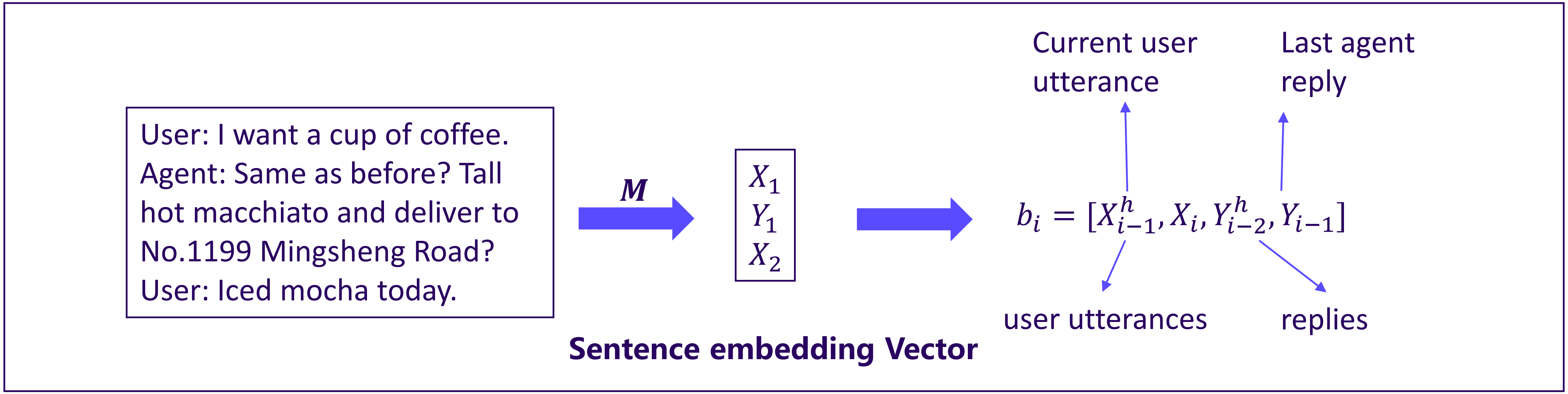
- Common dialogue states are learned in a source domain

  Belief state vector $b_i = f(H_i; \boldsymbol{M})$, where dialogue history $H_i = \{\{X_j, Y_j\}_{j=1}^{i-1}, X_i\}$

- Personalized Q-function

  $$\mathrm{Q}^{\pi_u}(H_i, Y_i|\Theta) = \underbrace{Q_g(H_i, Y_i|\Theta^g)}_{\text{General part}} + \underbrace{Q_p(H_i, Y_i|\Theta_u^p)}_{\text{Personal part}}.$$

  General part: $Q_g(H_i, Y_i|\Theta^g)$, personal part: $Q_p(H_i, Y_i|\Theta_u^p)$



Dialogue Policy Transfer Examples

- Transfer across users

X_1	**Can I have coffee please?**
Y_1	Same as before? Medium Hot Latte and deliver to No.101 Shandong Road?
X_2	Yes, as always. <Success-Feedback>

State=[X_1]
Action=Y_1

→

State=[X_1, Y_1, X_2]
Action=Y_2 <Success>

Transfer

X_1	**Can I have coffee please?**
Y_1	Same as before? Tall Iced Mocha and deliver to No.1199 Mingsheng Road?

State=[X_1]
Action=Y_1

- Transfer across domains

X_1	**I am looking for a restaurant.**
Y_1	What food would you like?

State=[Food=?, Area=?, Price=?]
Action=Request(Food)

Transfer

X_1	**I want to find a hotel.**
Y_1	How many stars should the hotel have?

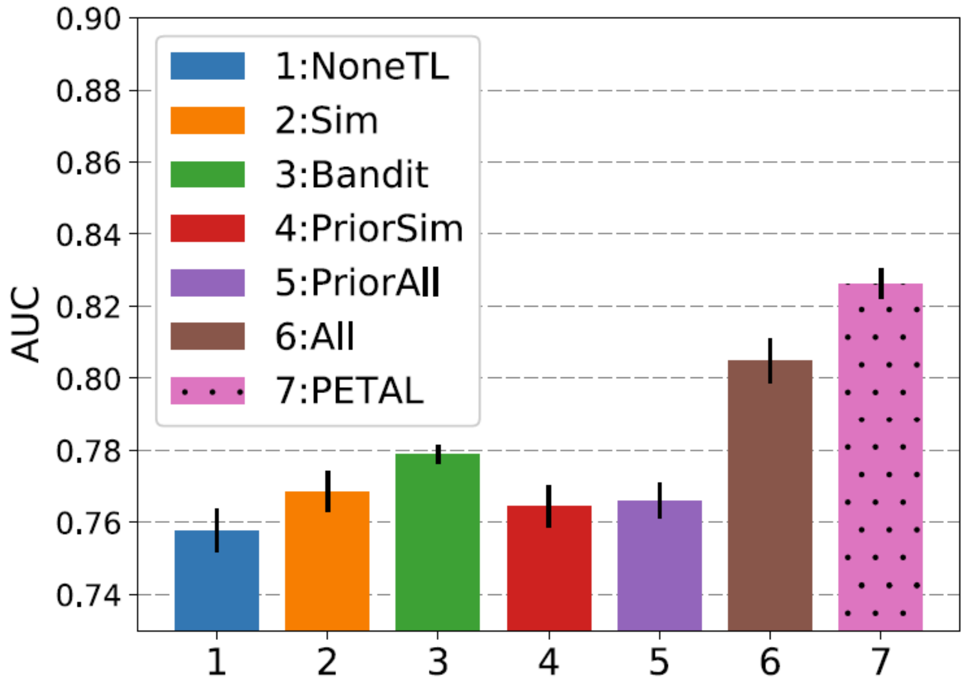
State=[Type=?, Stars=?, Area=?, Park=?]
Action=Request(Stars)

Real-world Experiment

➢ Setting: Coffee ordering

- Collected in O2O company, between real customers and human personal assistants.
- 52 source users and 20 target users, 2000+ multi-turn dialogues.
- Evaluation: AUC

	Source Domain		Target Domain	
	Users	Dialogues	Users	Dialogues
Real Data	52	1859	20	329
Simulation	11	176000	5	100



AUC of Ranking

User utterance	:	I want a cup of coffee.
All	PETAL	Response Candidates
0.86	1.36	* Same as before? Tall hot americano and deliver to Central Conservatory of Music?
0.99	0.92	All right, deliver to No.1199 Beiyuan Road, Chaoyang District, Beijing?
0.72	0.69	What's your address?

Recommendation System

Supervised learning based RecSys

Single Domain RecSys

- Easily get stuck in local optimal and keeps recommending the similar articles.
- Insensitive to fast evolving user interests.

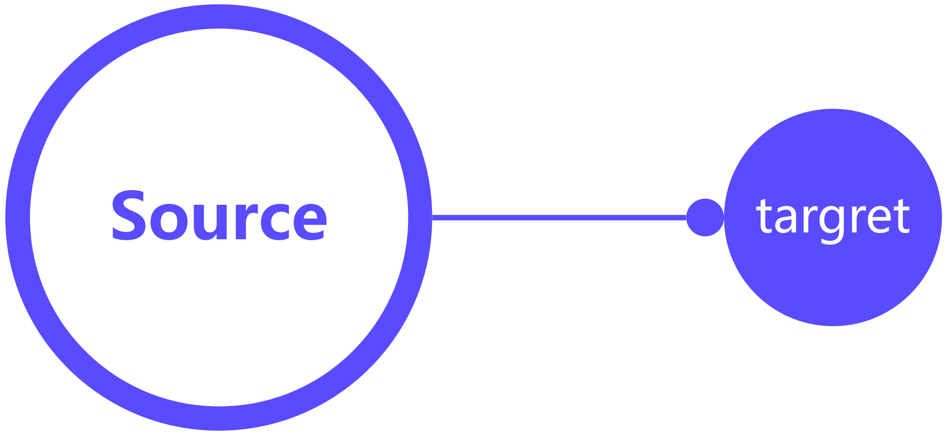
- Performs poor for new user, new article, and new domain.
- Purely explores leading to worse short-term CTR.

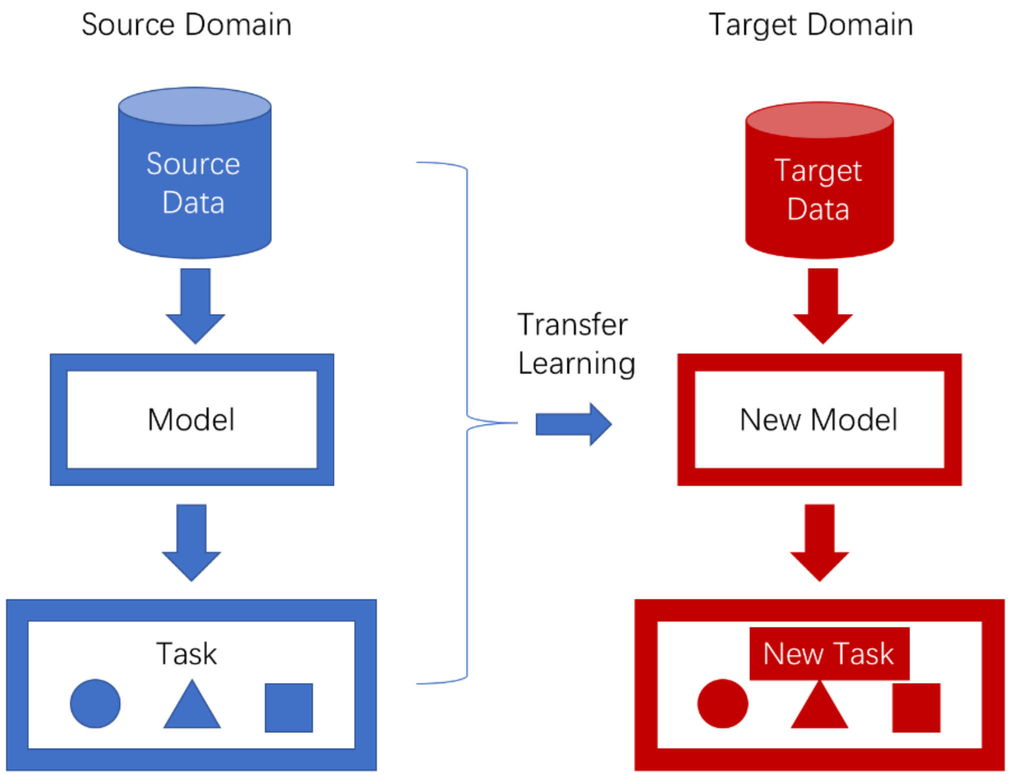
Transferrable Contextual Bandit

Transferable Contextual Bandit for Cross-Domain Recommendation, Bo Liu, Yu Zhang, Qiang Yang et al. AAAI18

Trend in Transfer Learning : Using Huge Pretrained Model

- Source domain: **huge** labeled or unlabeled data
- Target domain: few labeled data
- Objective: transfer model from source domain to target domain **for same or different tasks**

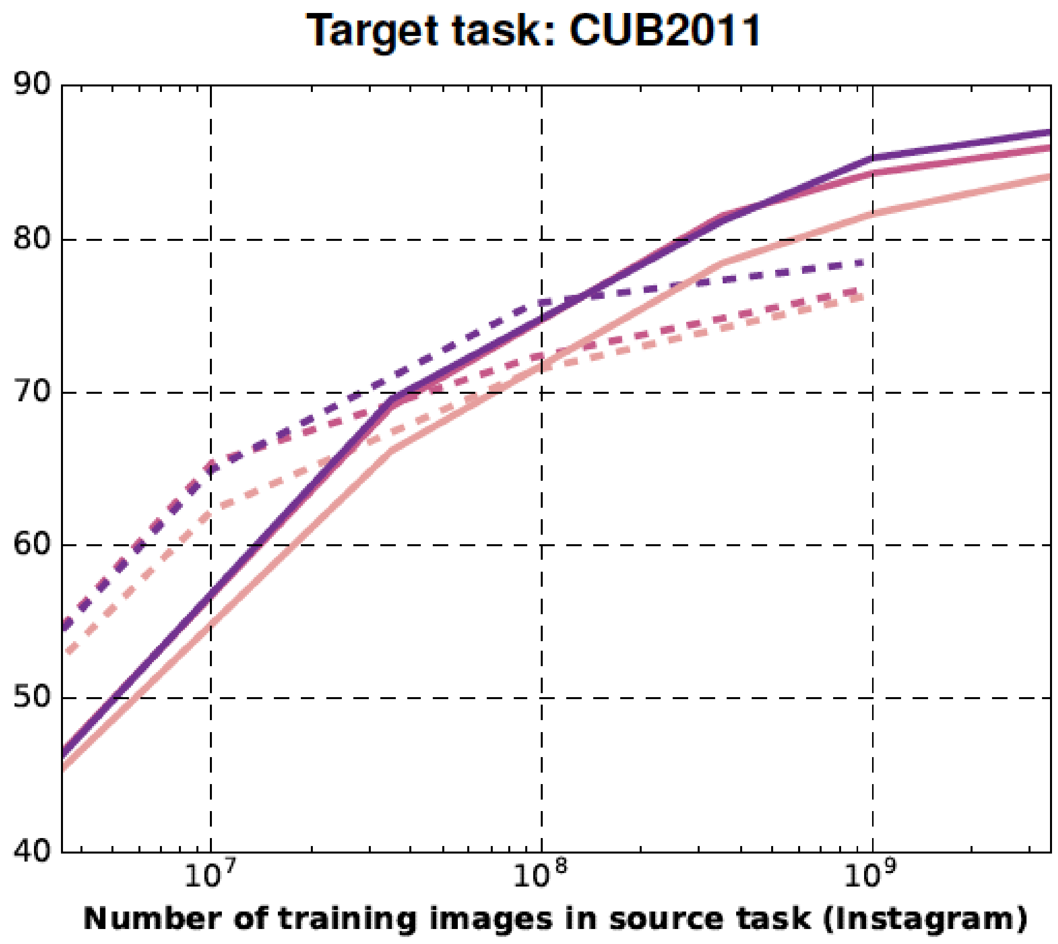




Source-Data Scale Matters in Transfer Learning (image)

- Dhruv Mahajan, et al.: **Exploring the Limits of Weakly Supervised Pretraining**. ECCV (2) 2018

- "Without manual dataset curation or sophisticated data cleaning, models trained on billions of Instagram images using thousands of distinct hashtags as labels exhibit excellent transfer learning performance"



Scale of Source-Data Matters in Transfer Learning (NLP): BERT

Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. CoRR abs/1810.04805 (2018)

" Recent empirical improvements due to transfer learning with language models have demonstrated that rich, unsupervised pre-training is an integral part of many language understanding systems.

Our major contribution is further generalizing these findings to deep bidirectional architectures, allowing the same pre-trained model to successfully tackle a broad set of NLP tasks."

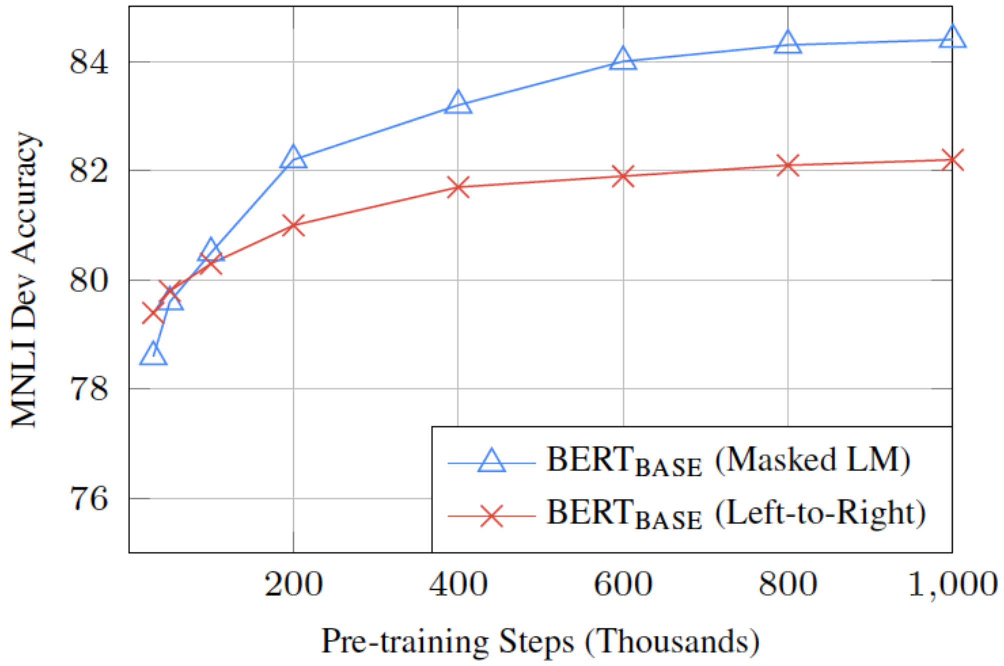
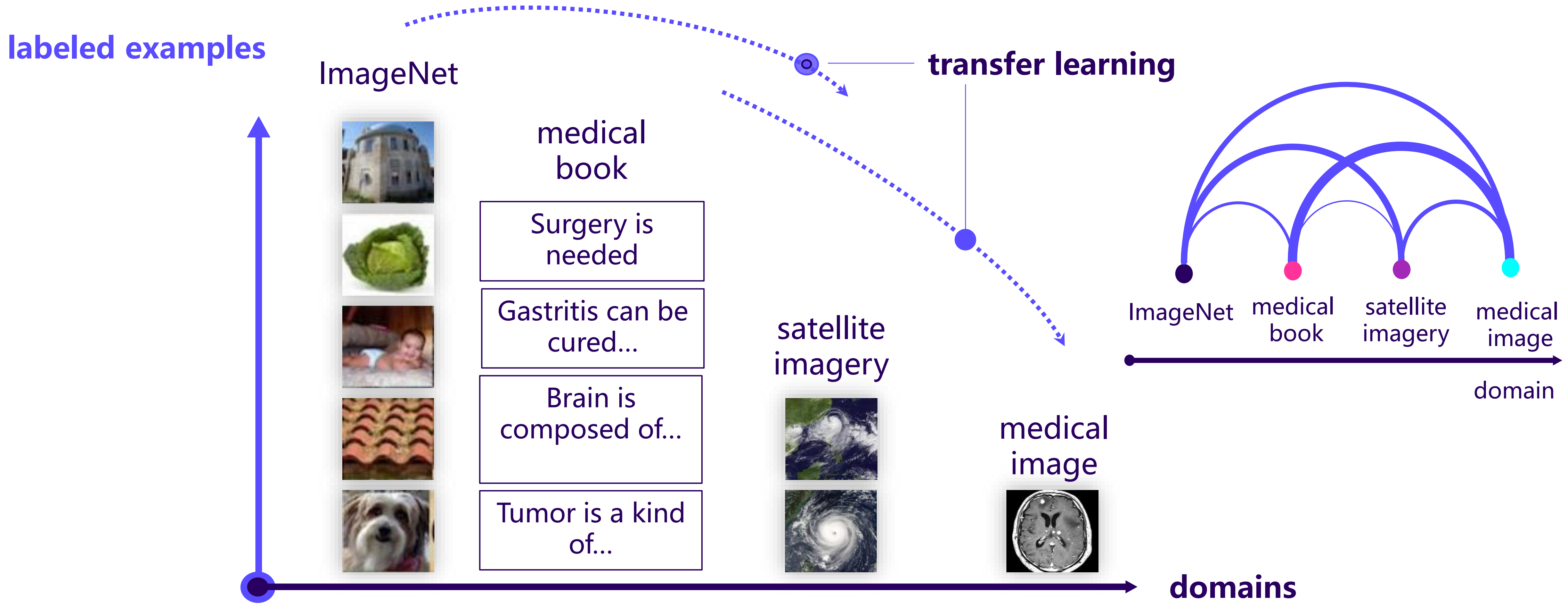


Figure 4: Ablation over number of training steps. This shows the MNLI accuracy after fine-tuning, starting from model parameters that have been pre-trained for k steps. The x-axis is the value of k.

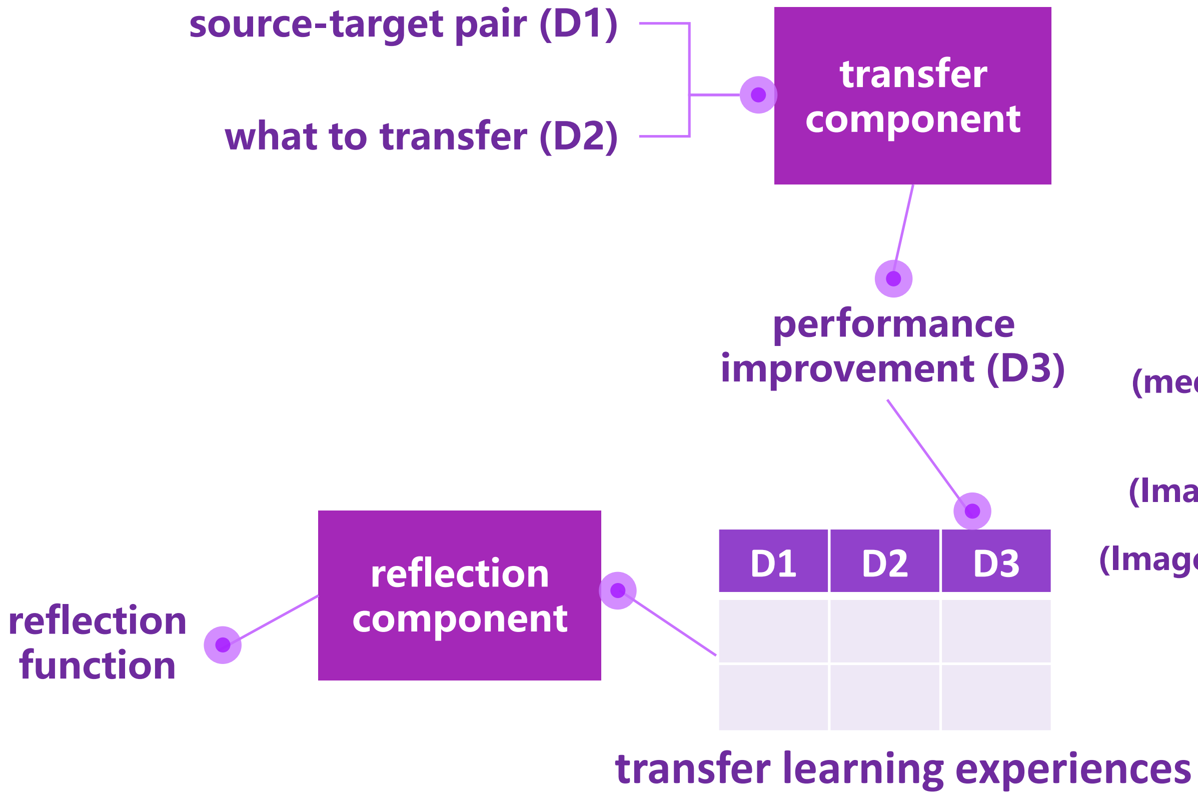
Transfer Learning via Learning to Transfer

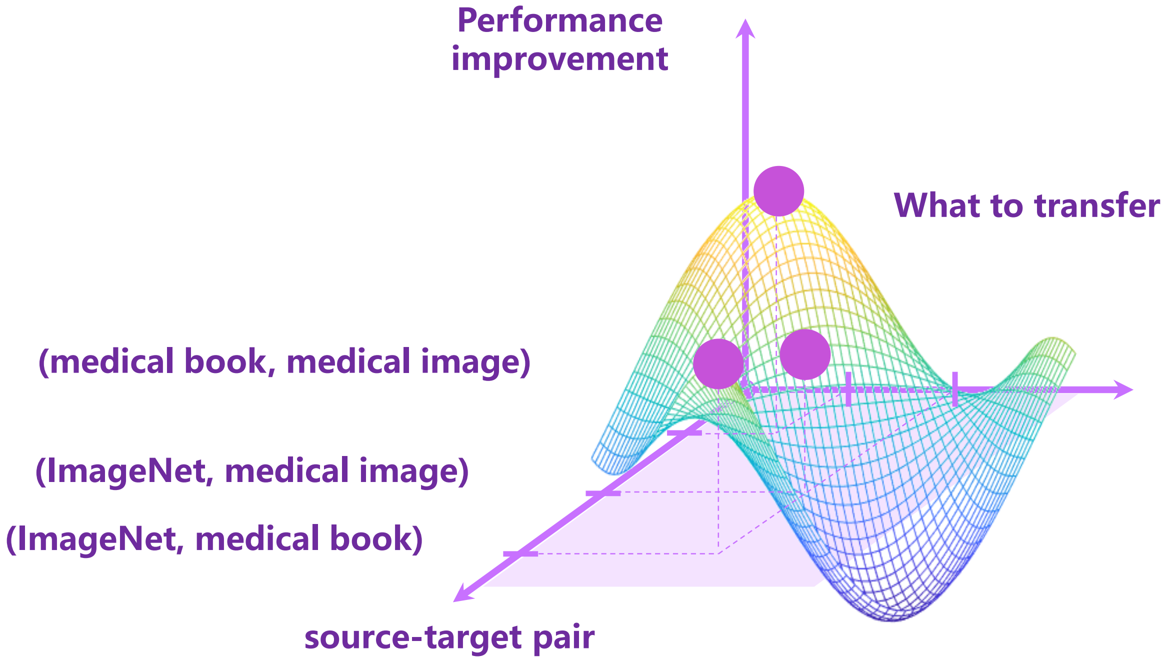


Transfer Learning via Learning to Transfer, Ying Wei, Qiang Yang et al. ICML 2018

Learning-to-Transfer (L2T) Framework

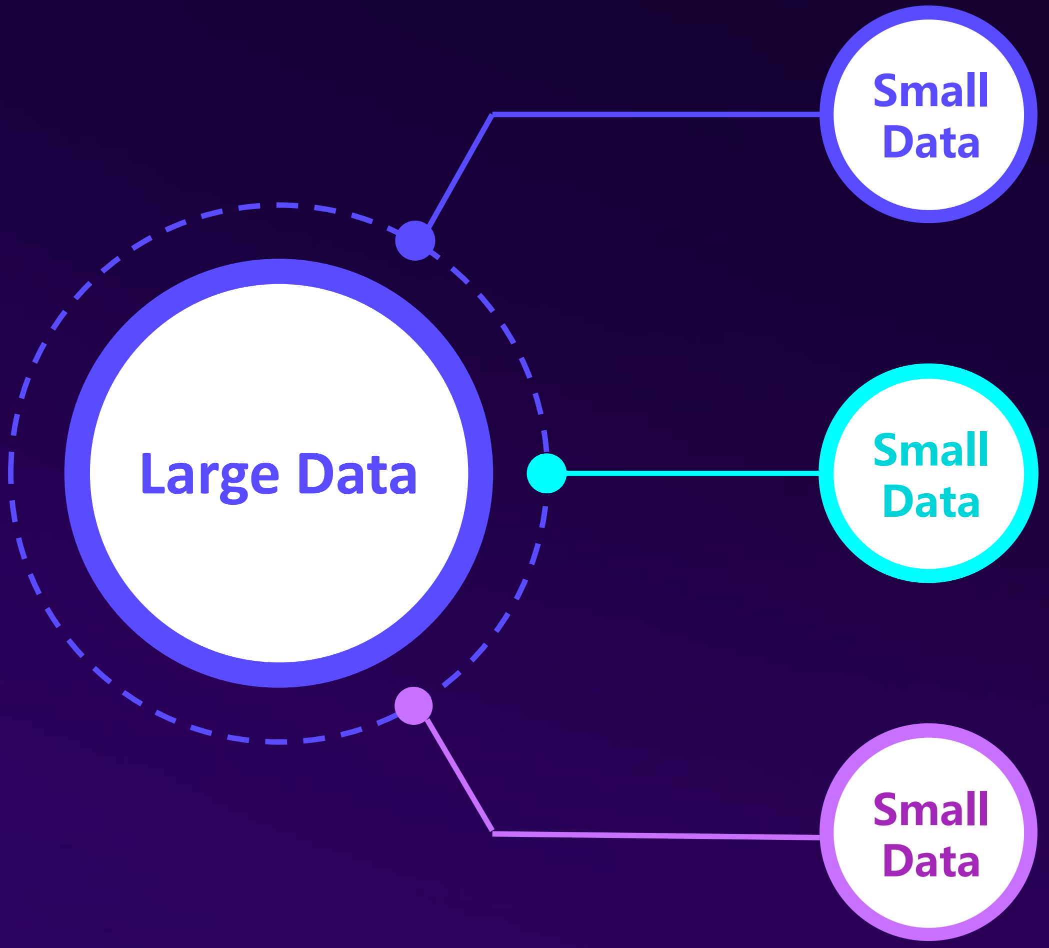
➢ **Training – Learning skills from experiences**



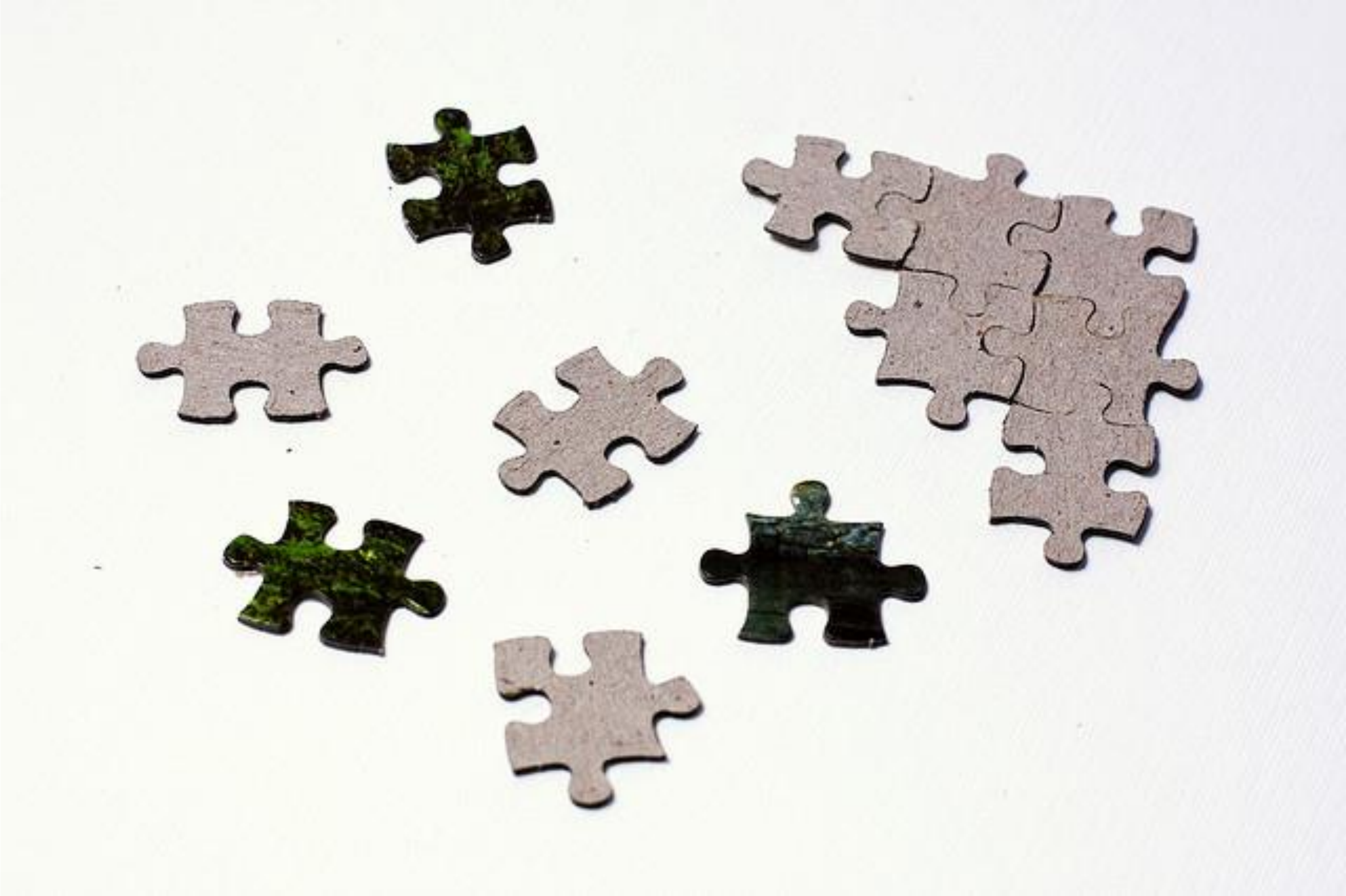


Transfer Learning via Learning to Transfer, Ying Wei, Qiang Yang et al. ICML 2018

Transfer Learning from Large Data to Small Data



Next Problem: Data Are Fragmented



Challenges to AI: Data Privacy and Confidentiality

Facebook's data privacy scandal

2019/1/19

- In 2012, the FTC fined Google $22.5 million over failing to improve privacy practices – a record for such a punishment.
- The Washington Post says that the fine against Facebook is expected to be "much larger."

- More than 50 million people involved
- UK assessed a £500,000 fine to Facebook
- **the worst single-day market value decrease for a public company in the US**, dropping $120 billion, or 19%

The General Data Protection Regulation (GDPR)

- No Autonomous Modeling and Decision
- Interpretability of Model Decisions
- Users'Right for Data to be Forgotten
- Data Privacy By Design
- Explicit Consent for Data Usage

California Consumer Privacy Act (CCPA)

- Takes effect in 2020
- grants consumers the right to know what information is collected and **whom it is shared with**
- Consumers will have the option of barring tech companies from selling their data
- Provides some of the strongest **regulations in the USA.**

China’s Data Cyber Security Law

- Enacted in 2017
- Requires that Internet businesses must not leak or tamper with the personal information
- When conducting data transactions with third parties, they need to ensure that the proposed contract follow legal data protection obligations.
- More to come…

From Report by KPMG 2017

Highlights and interpretation of the Cybersecurity Law

Highlights of the Cybersecurity Law

Comprising 79 articles in seven chapters, the Cybersecurity Law contains a number of cybersecurity requirements, including safeguards for national cyberspace sovereignty, protection of critical information infrastructure and data and protection of individual privacy. The Law also specifies the cybersecurity obligations for all parties. Enterprises and related organisations should prioritise the following highlights of the Cybersecurity Law:

Personal information protection	The Cybersecurity Law clearly states requirements for the collection, use and protection of personal information.
Critical information infrastructure	The Cybersecurity Law frequently mentions the protection of “critical information infrastructure”.
Network operators	“Network operators” are the owners and administrators of networks and network service providers. The Cybersecurity Law clarifies operators’ security responsibilities.
Preservation of sensitive information	The Cybersecurity Law requires personal information/important data collected or generated in China to be stored domestically.
Certification of security products	Critical cyber equipment and special cybersecurity products can only be sold or provided after receiving security certifications.
Legal liabilities	Enterprises and organisations that violate the Cybersecurity Law may be fined up to RMB1,000,000.

Challenges to AI：small data and fragmented data

Low Security in Data Sharing

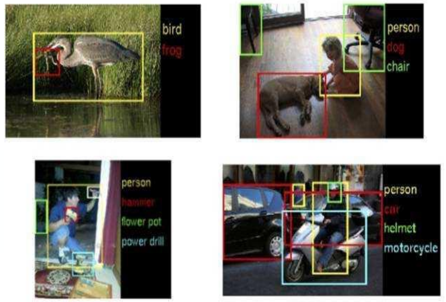
Lack of Labeled Data

Segregated Datasets

Enterprise A

X1

Data silos



Enterprise B

(X2, Y)

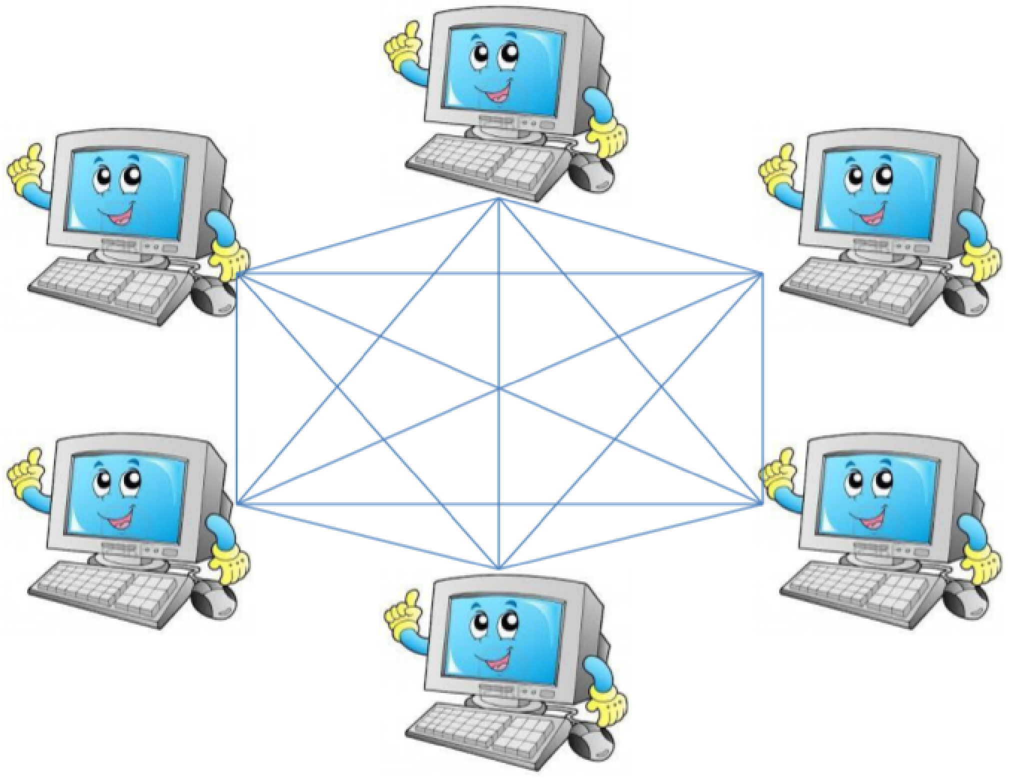
Over 80% of enterprises’ information in data silos!

Privacy-Preserving Technologies

- Secure Multi-party Computation (MPC)
- Homomorphic Encryption (HE)
- Yao’s Garbled Circuit
- Secret sharing
- Differential Privacy (DP)

……

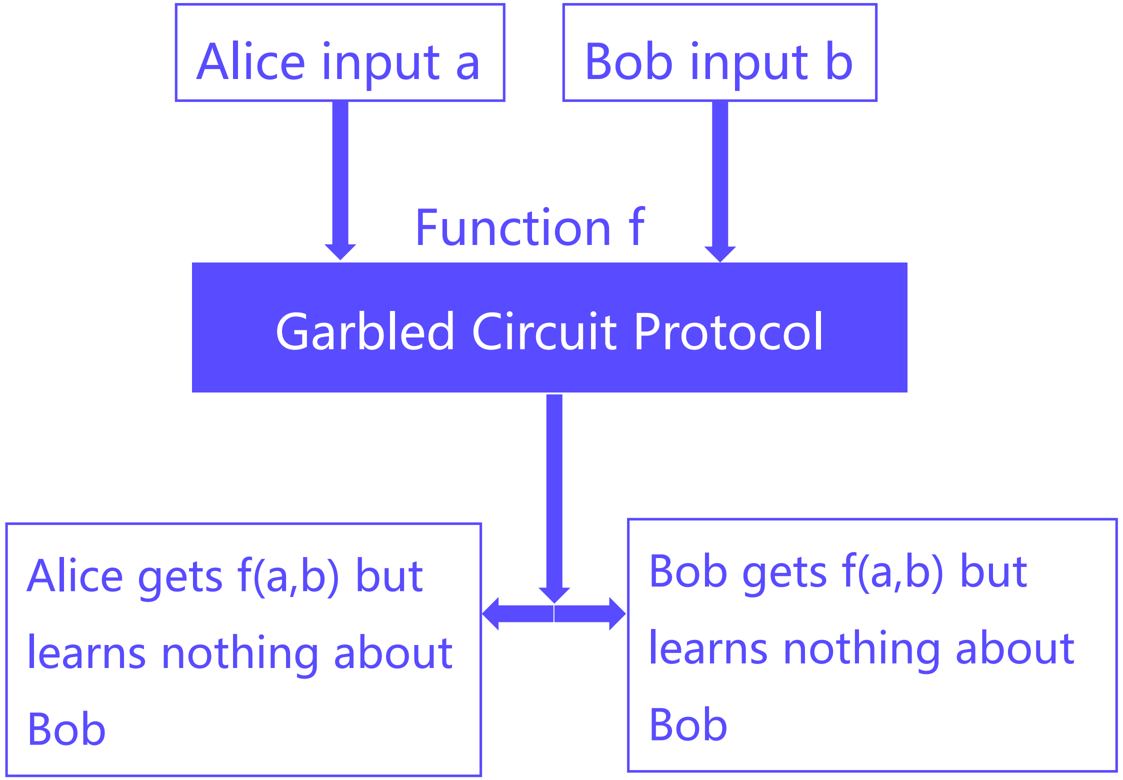
Secure Multi-Party Computation (MPC)



Ran Cohen ,Tel Aviv University, Secure Multiparty Computation: Introduction

- Provides security proof in a well-defined simulation framework
- Guarantees complete zero knowledge
- Requires participants' data to be secretly-shared among non-colluding servers
- Drawbacks:
  - Expensive communication,
  - Though it is possible to build a security model with MPC under lower security requirement in exchange for efficiency

Yao’s Garbled Circuit Protocol (Andrew Yao, 1986)



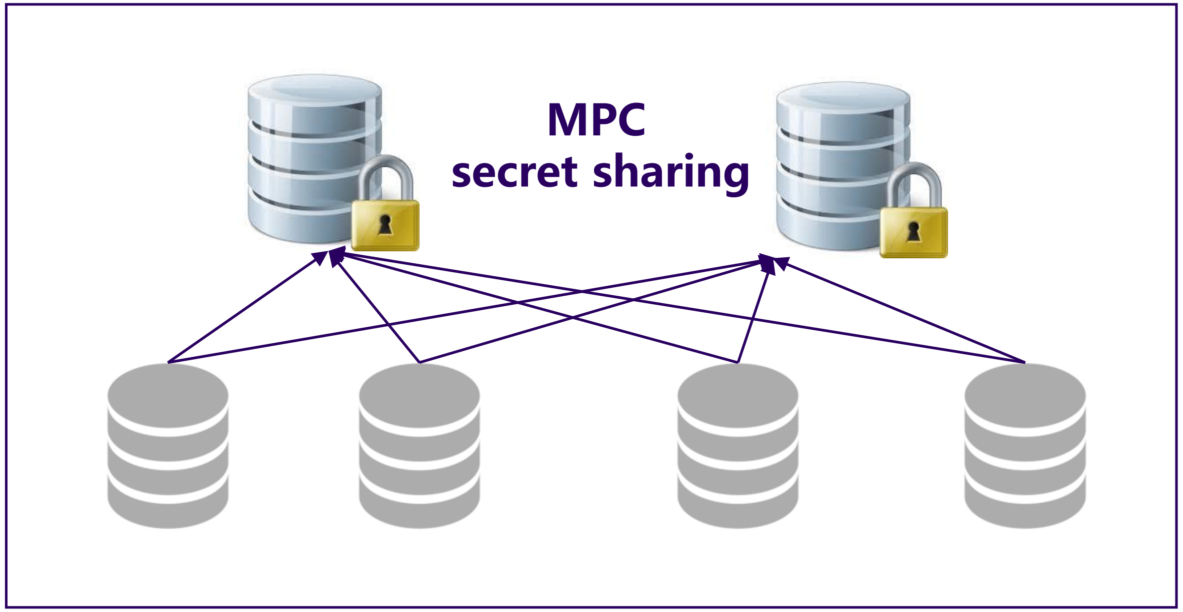
- **Oblivious Transfer**

Steps

- Alice builds a garbled circuits;
- Alice sends her input keys;
- Alice and Bob do Oblivious Transfer;
- Bob gets the output and sends back to Alice;
- Alice and Bob learns nothing about the other value.

SecureML : Privacy-preserving machine learning for linear regression, logistic regression and neural network training

- Combines secret sharing, garbled circuits and oblivious transfer
- Learns via two un-trusted, but non-colluding servers
- Computationally expensive



Mohassel, P., & Zhang, Y. (2017, May). SecureML: A system for scalable privacy-preserving machine learning. In *2017 38th IEEE Symposium on Security and Privacy (SP)* (pp. 19-38). IEEE.

Homomorphic Encryption

- Full Homomorphic Encryption and Partial Homomorphic Encryption.
- **Paillier** partially homomorphic encryption

***Addition*:** $[[u]] + [[v]] = [[u+v]]$
Scalar multiplication: $n[[u]] = [[nu]]$

- For public key ***pk = n***, *the encoded form of* $m \in \{0, \ldots, n-1\}$ is

$$\mathbf{Encode}(m) = r^n (1 + n)^m \bmod n^2$$

r is randomly selected from $\{0, \ldots, n-1\}$.

- For float $q = (s, e)$, encrypt $[[q]] = ([[s]], e)$, here $q = s\beta^e$ is base-β exponential representation.

Rivest, R. L.; Adleman, L.; and Dertouzos, M. L. 1978. On data banks and privacy homomorphisms. Foundations of Secure Computation, Academia Press 169–179.

Applying HE to Machine Learning

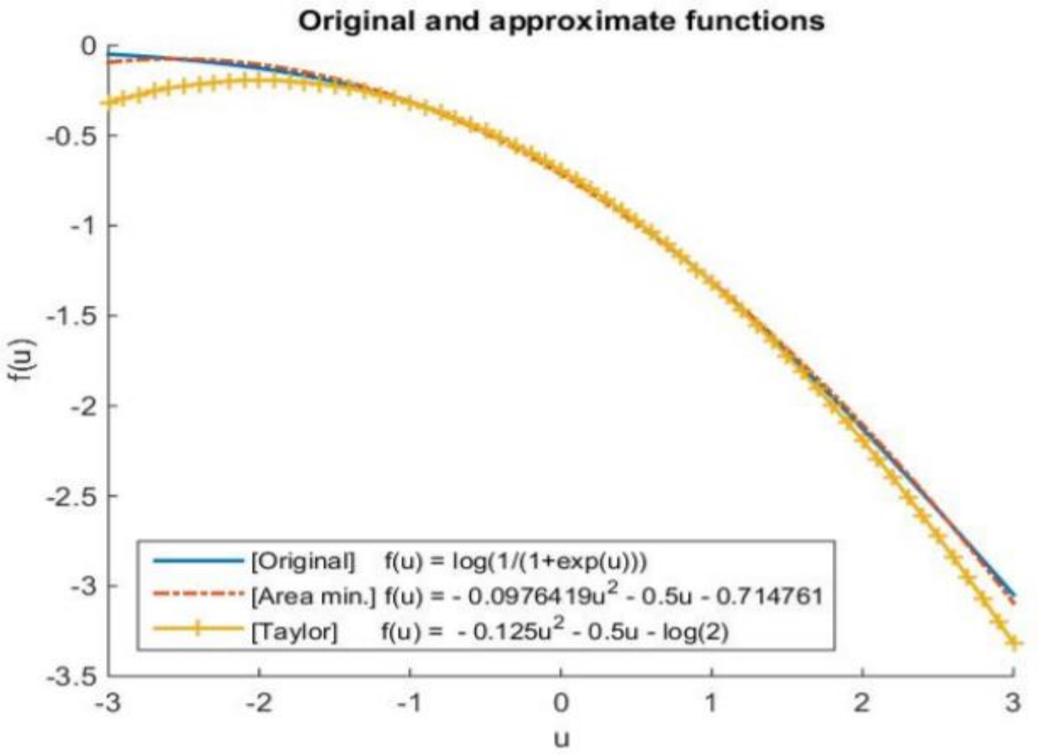
Polynomial approximation for logarithm function

$$\log\left(\frac{1}{1+\exp(u)}\right) \approx \sum_{j=0}^{k} a_j u^j$$

Encrypted computation for each term in the polynomial function

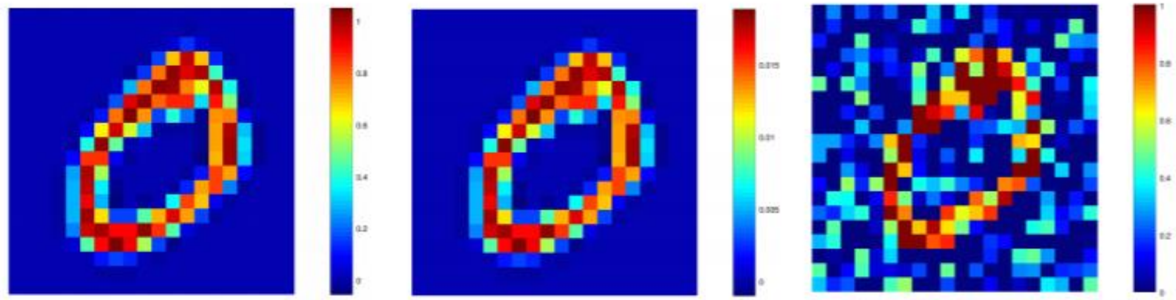
$$loss = \log 2 - \frac{1}{2} y w^T x + \frac{1}{8}(w^T x)^2$$

$$[[loss]] = [[\log 2]] + (-\frac{1}{2}) * [[y w^T x]] + \frac{1}{8}[[(w^T x)^2]]$$



- Kim, M.; Song, Y.; Wang, S.; Xia, Y.; and Jiang, X. 2018. Secure logistic regression based on homomorphic encryption: Design and evaluation. JMIR Med Inform 6(2)
- Y. Aono, T. Hayashi, T. P. Le, L. Wang, Scalable and secure logistic regression via homomorphic encryption, CODASPY16

Is the Gradient Info Safe to Share?



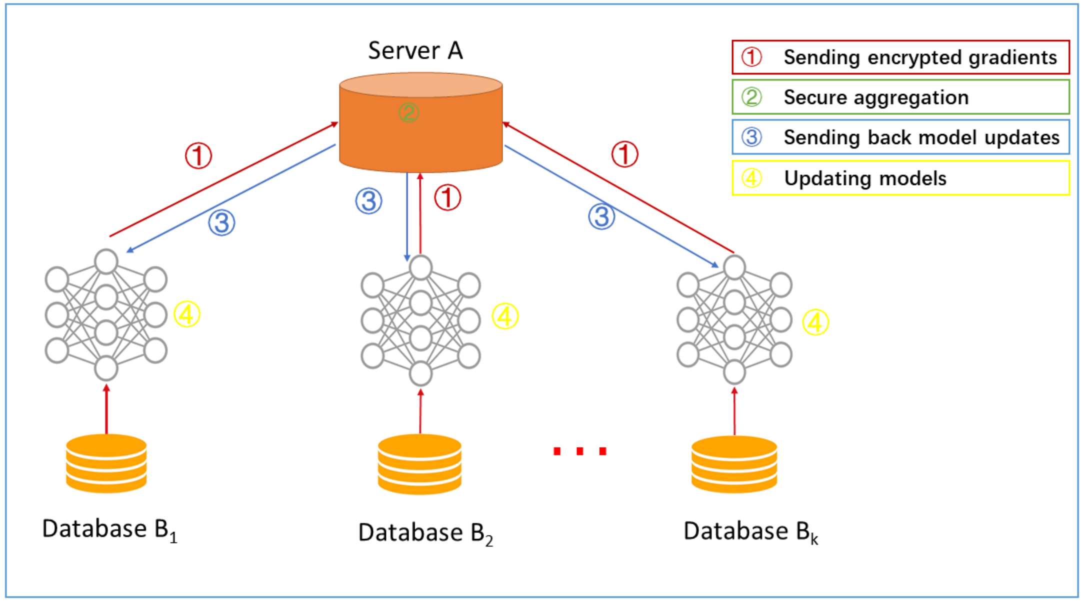
(a) Original 20x20 image of hand-written number 0, seen as a vector over $\mathbb{R}^{400}$ fed to a neural network.

(b) Recovered image using 400/10285 (3.89%) gradients (see Sect.3, Example 2). The difference with the original (a) is only at the value bar.

(c) Recovered image using 400/10285 (3.89%) gradients (see Sect.3, Example 3). There are noises but the truth label 0 can still be seen.

Fig. 3. Original data (a) vs. leakage information (b), (c) from a small part of gradients in a neural network.

Protect gradients with Homomorphic Encryption

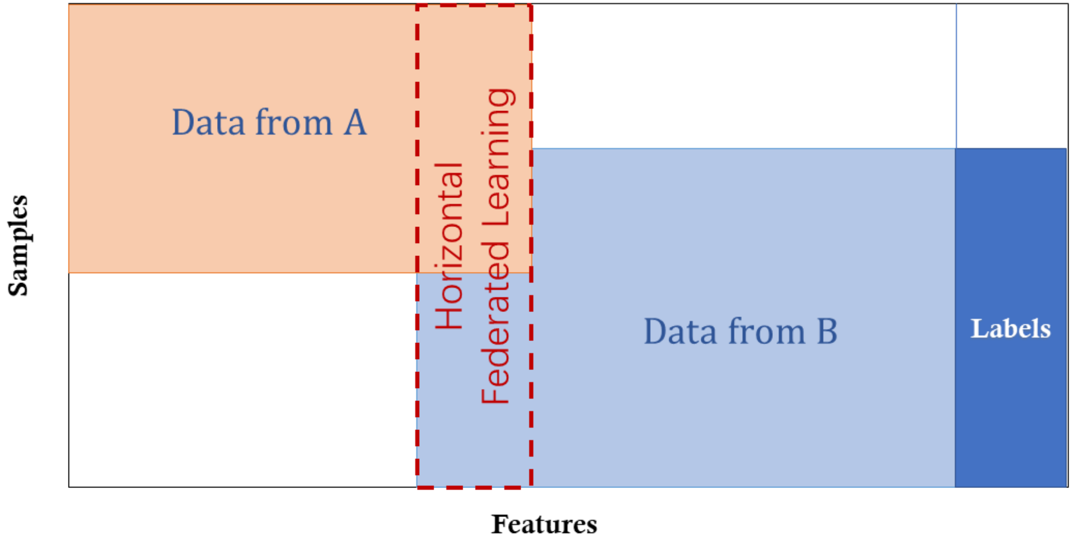


- ➢ Le Trieu Phong, et al. 2018. Privacy-Preserving Deep Learning via Additively Homomorphic Encryption. IEEE Trans. Information Forensics and Security , 13, 5 (2018),1333–1345
- ➢ **Algorithm ensures that no information is leaked to the semi-honest server, provided that the underlying additively homomorphic encryption scheme is secure*.**

*** Q. Yang, Y. Liu, T. Chen, Y. Tong, Federated machine learning: concepts and applications, ACM TIST, ,2018**

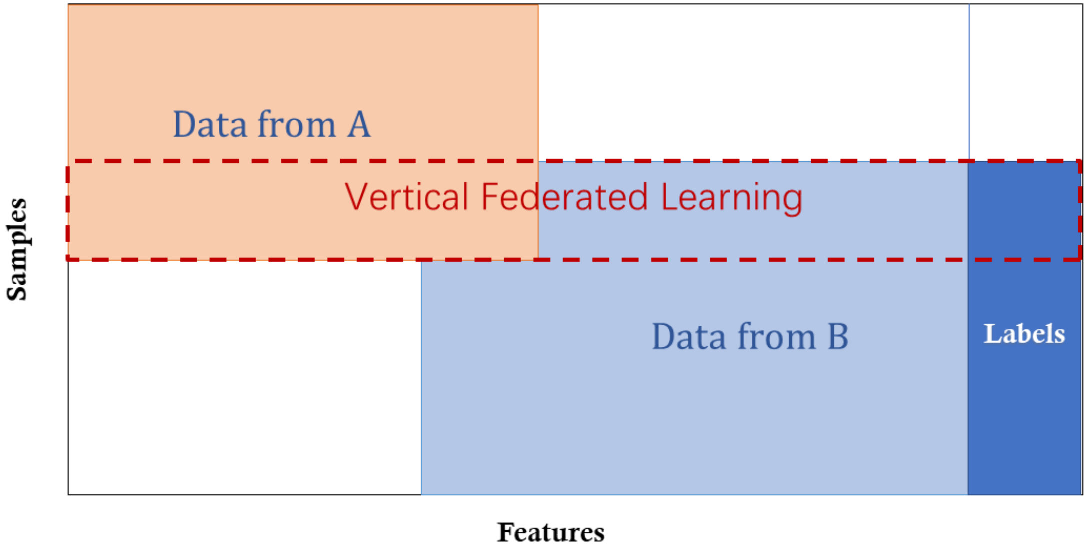
Categorization of Federated Machine Learning

Horizontal FML



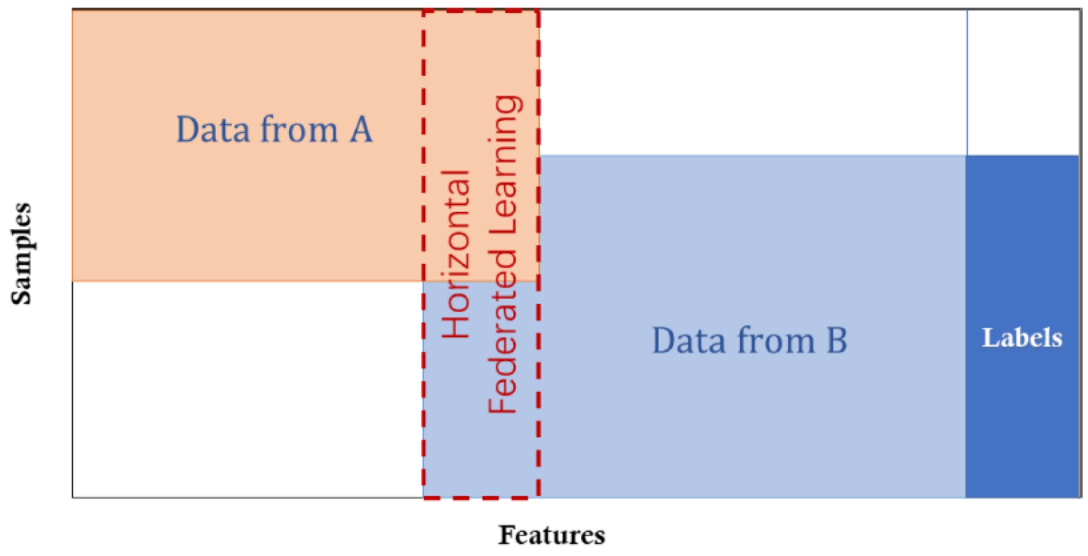
Large overlap of features of the two data sets

Vertical FML



Large overlap of sample IDs (users) of the two data sets

Horizontal Federated Learning: Divide by Users



(a) Horizontal Federated Learning

FEDERATED LEARNING FOR MOBILE KEYBOARD PREDICTION, Andrew Hard, et al., Google, 2018

Step 1: Participants compute training gradients locally

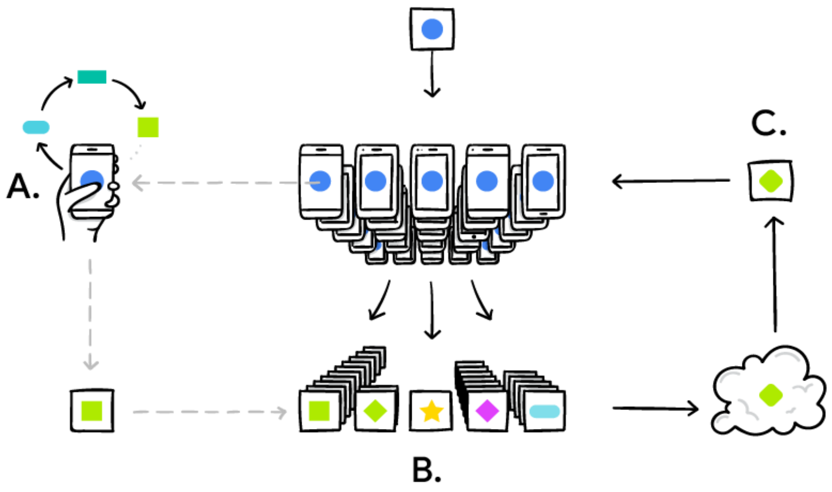
- mask gradients with encryption, differential privacy, or secret sharing techniques
- all participants send their masked results to server

Step 2: The server performs secure aggregation without learning information about any participant

Step 3: The server sends back the aggregated results to participants

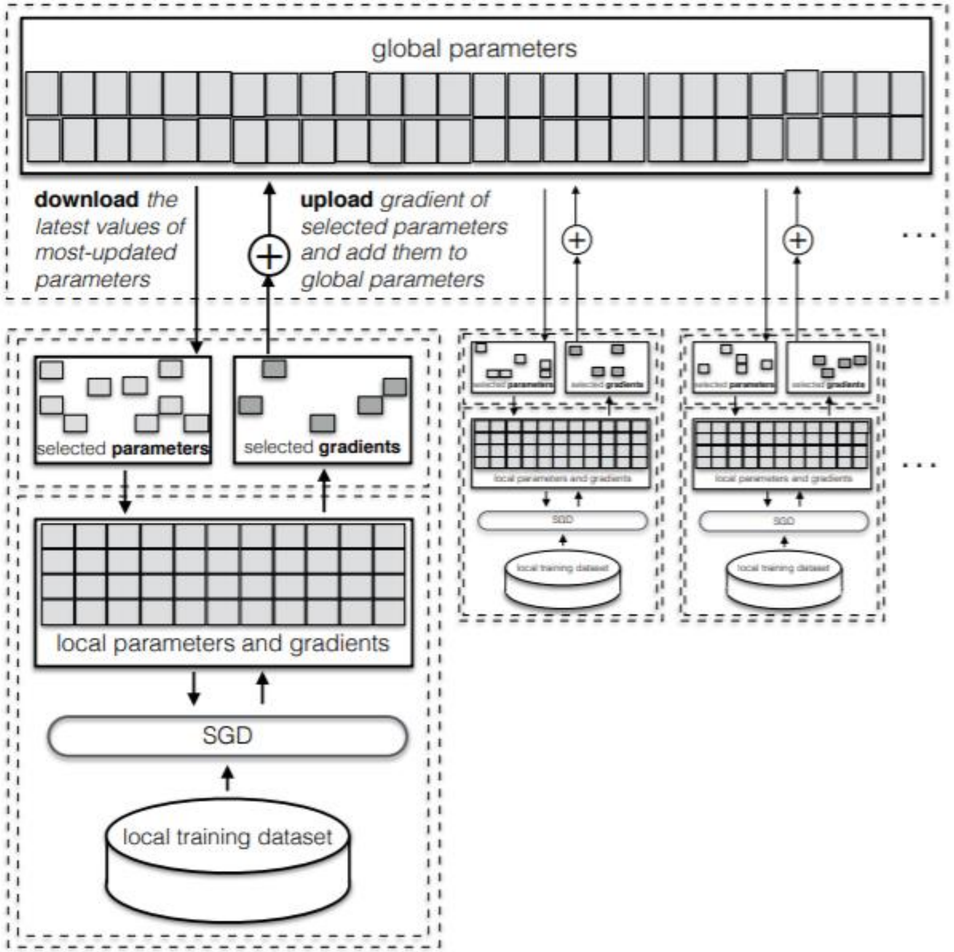
Step 4: Participants update their respective model with the decrypted gradients

Horizontal Federated Learning



H. Brendan McMahan et al, *Communication-Efficient Learning of Deep Networks from Decentralized Data*, Google, 2017

- **Multiple clients, one server**
- **Data is horizontally split across devices, homogeneous features**
- **Local training**
- **Selective clients**



Reza Shokri and Vitaly Shmatikov. 2015. *Privacy-Preserving Deep Learning*. In Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security (CCS ' 15). ACM, New York

Vertical Federated Learning

Objective:

- Party (A) and Party (B) co-build a FML model

Assumptions :

- Only one party has label Y
- Neither party wants to expose their X or Y

Challenges:

- Parties with only X cannot build models
- Parties cannot exchange raw data by law

Expectations :

- Data privacy for both parties
- model is LOSSLESS

(X)

(U, Z)

WeBank 微众银行

(V, Y)

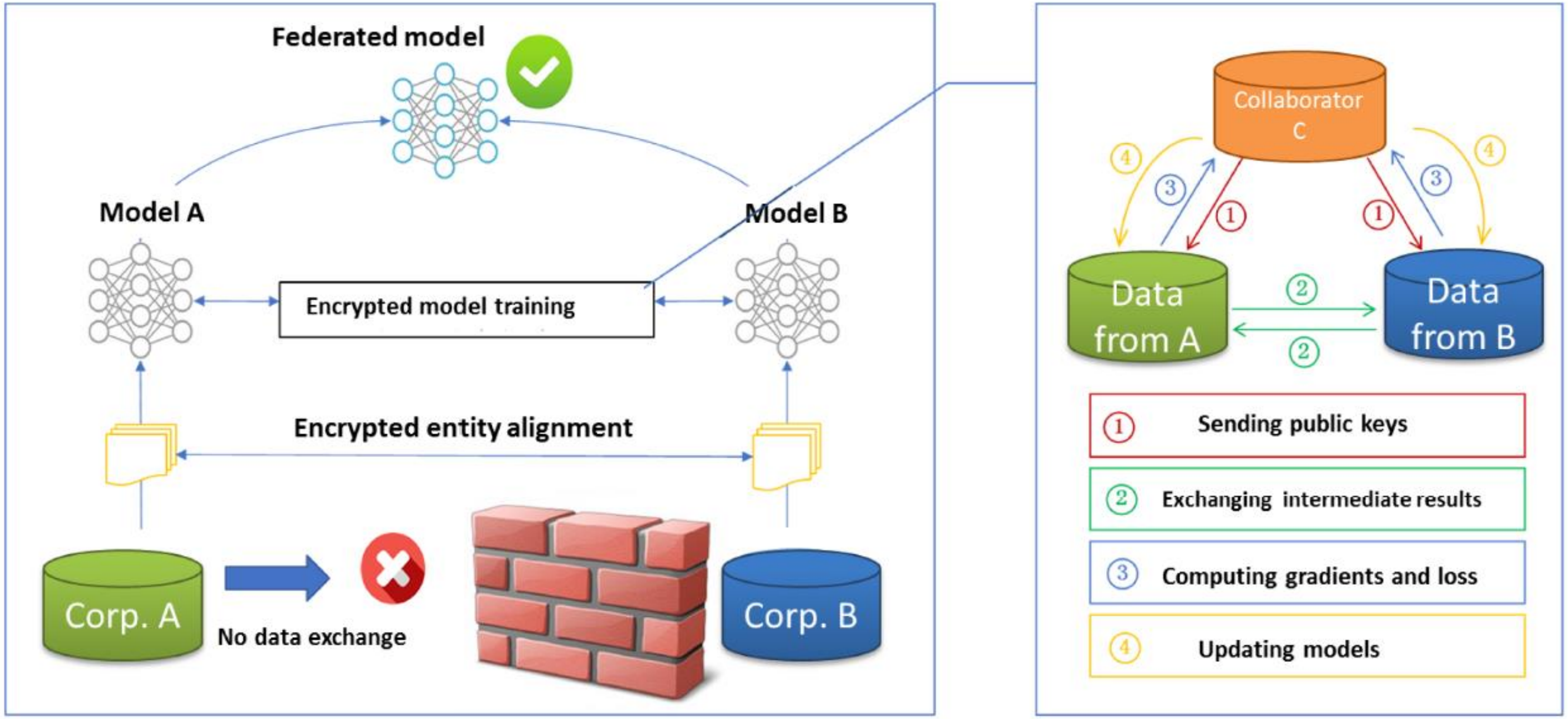
ID	X1	X2	X3
U1	9	80	600
U2	4	50	550
U3	2	35	520
U4	10	100	600
U5	5	75	600
U6	5	75	520
U7	8	80	600

Retail A Data

ID	X4	X5	Y
U1	6000	600	No
U2	5500	500	Yes
U3	7200	500	Yes
U4	6000	600	No
U8	6000	600	No
U9	4520	500	Yes
U10	6000	600	No

Bank B Data

Vertical Federated Learning

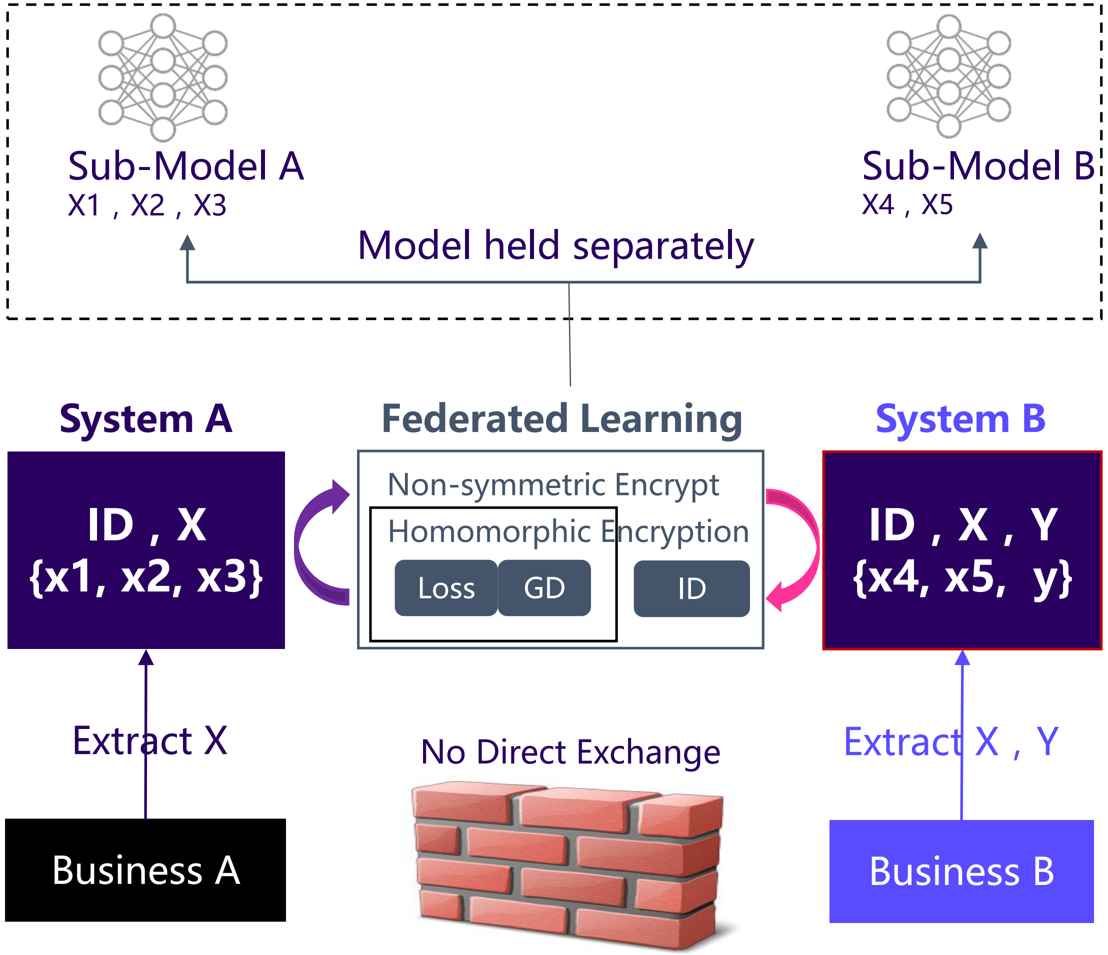


a b

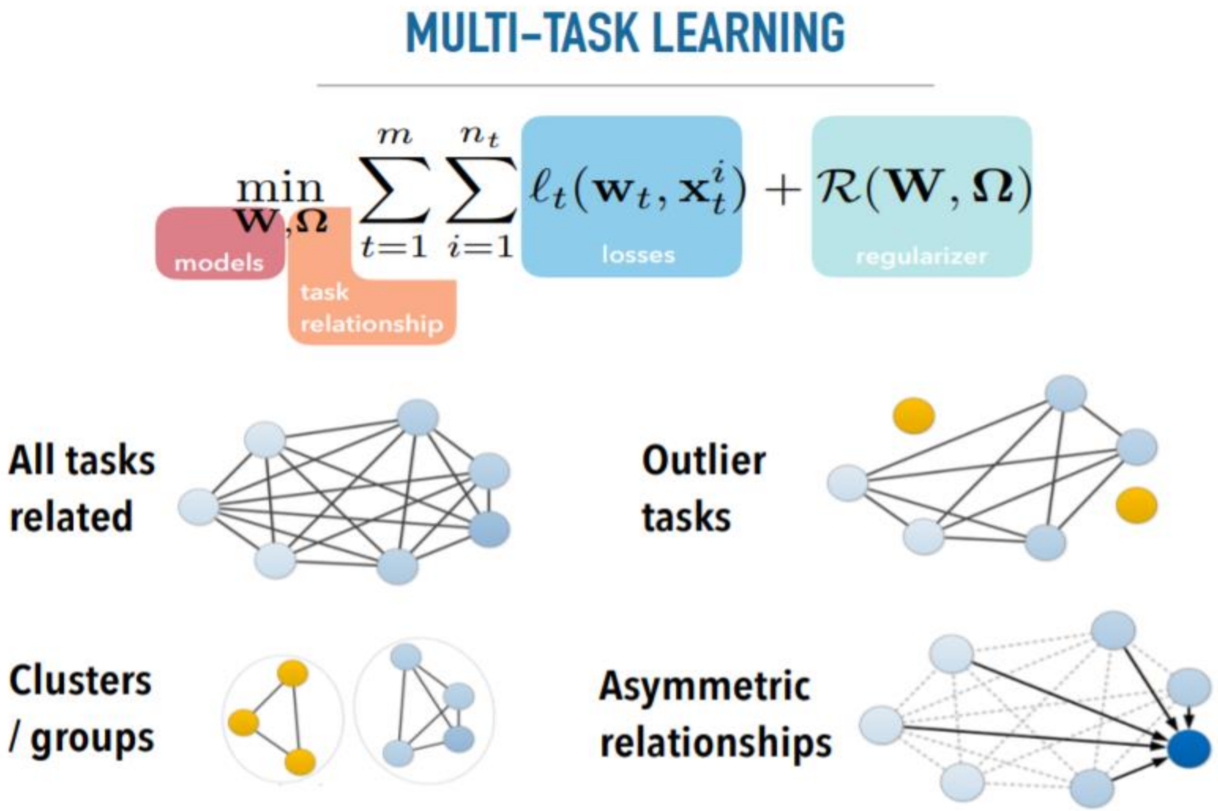
Federated Transfer Learning: Concepts and Applications. Qiang Yang, Yang Liu and Tianjian Chen. ACM TIST 2019.

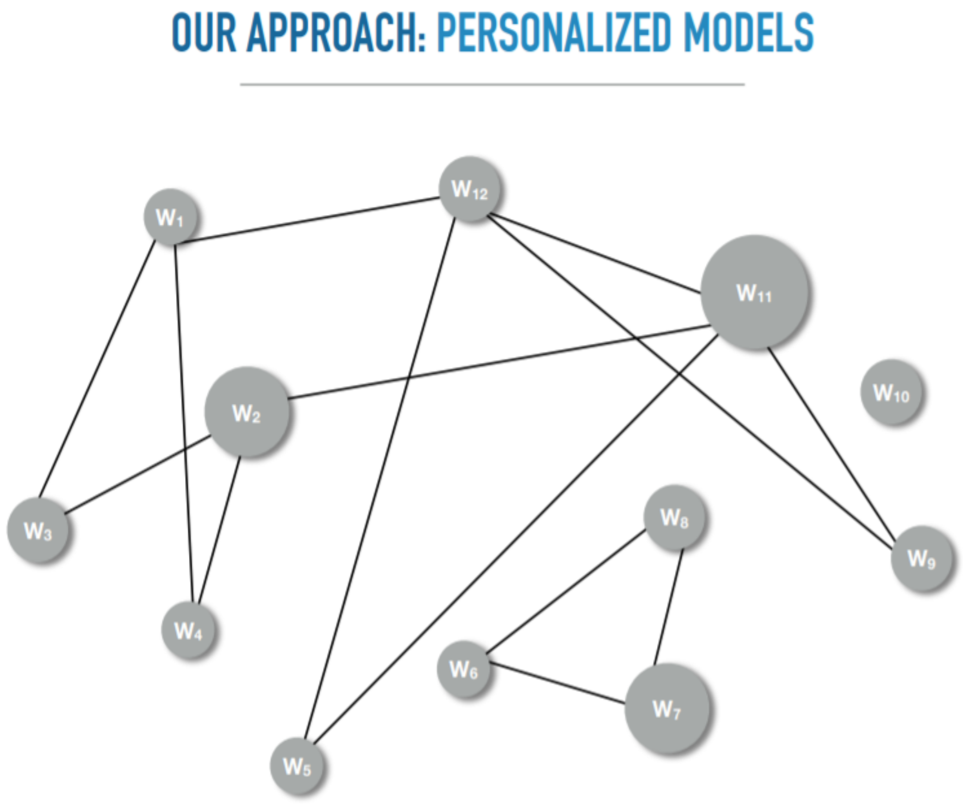
Vertical Federated Transfer Learning：Features

- **Data Protection：**
  - No different sample ID set leaked
  - No (X, Y) leaked
- **Parameter Protection：**
  - Separately held，jointly used
- **Result：**
  - A has Model_A
  - B has Model_B
  - Both models are better than learned separately
- **Property: Lossless**



Federated multi-task learning



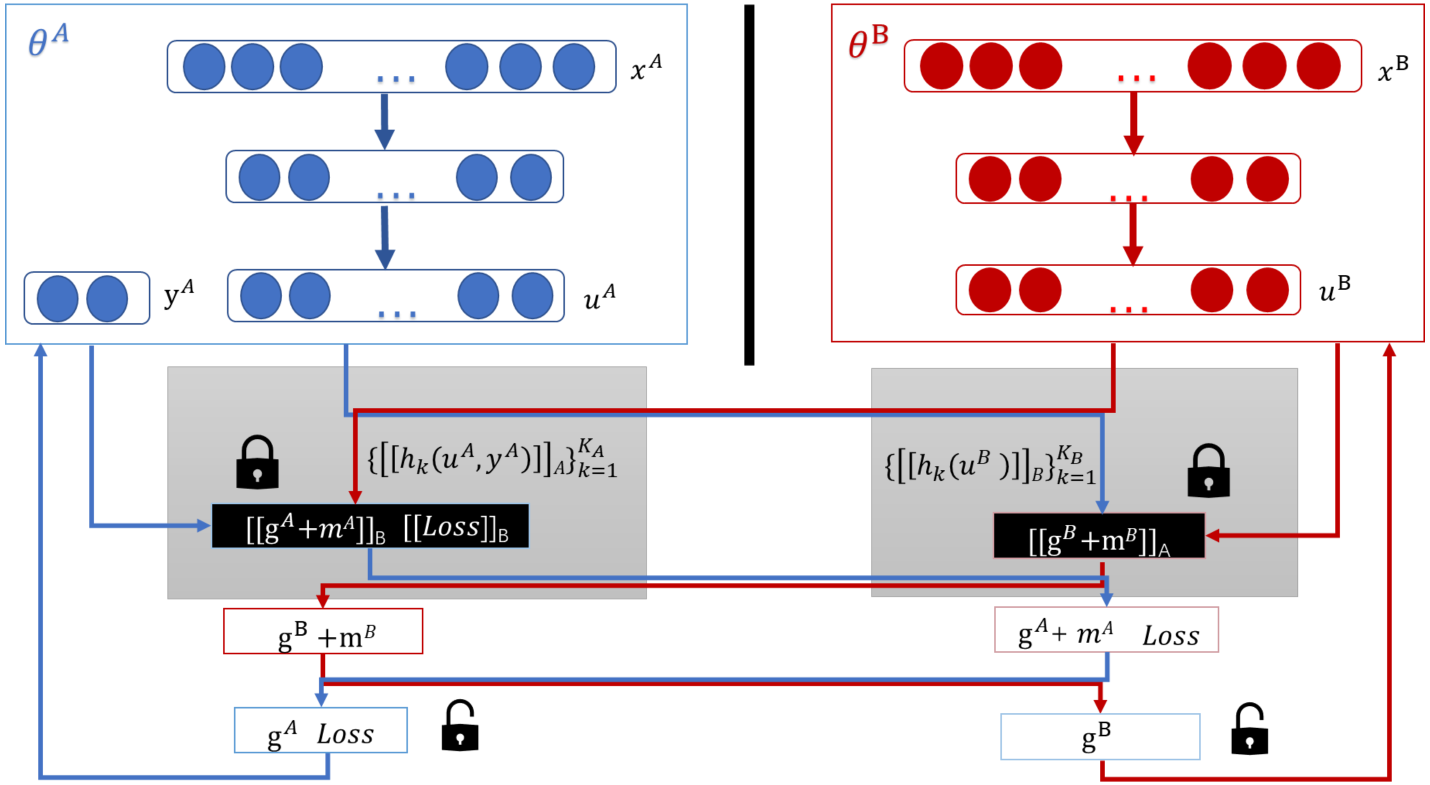


Virginia Smith et al. 2017. Federated Multi-Task Learning. In Advances in Neural Information Processing Systems

How to perform transfer learning without sharing data ?

Federated Transfer Learning

Source Domain Party A **Target Domain Party B**



Step 1
Party A and B send public keys to each other

Step 2
Parties compute, encrypt and exchange intermediate results

Step 3
Parties compute encrypted gradients, add masks and send to each other

Step 4
Parties decrypt gradients and exchange, unmask and update model locally

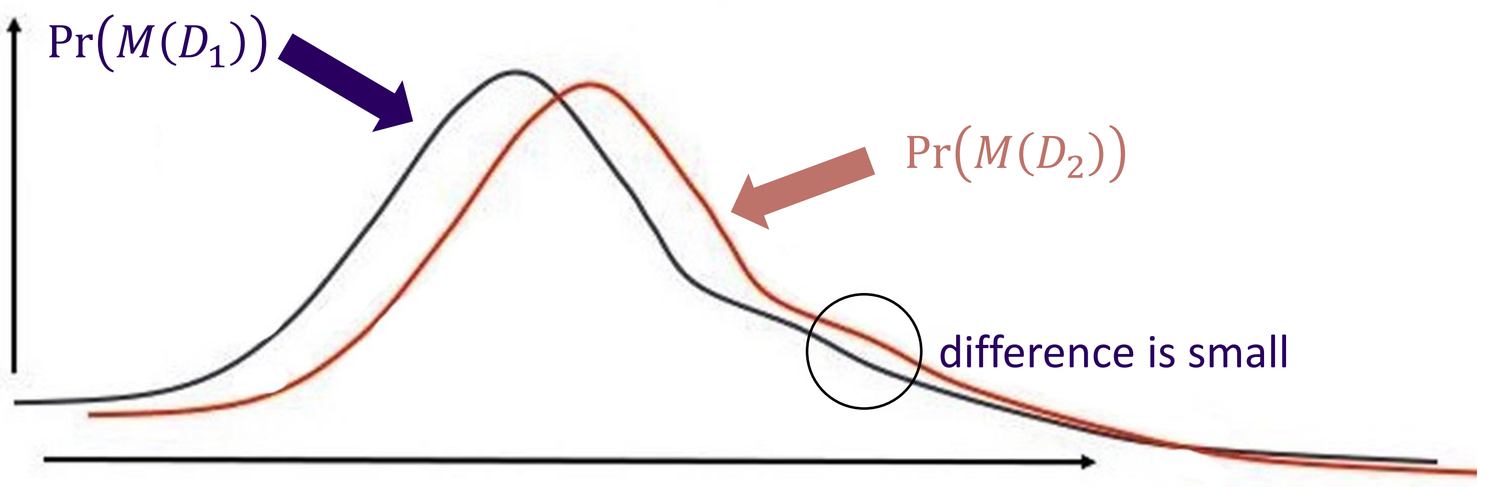
Federated Transfer Learning. Yang Liu, Tianjian Chen, Qiang Yang, https://arxiv.org/pdf/1812.03337.pdf 2018

Differential Privacy: changes in the distribution is too small to be perceived with variations on a single element.

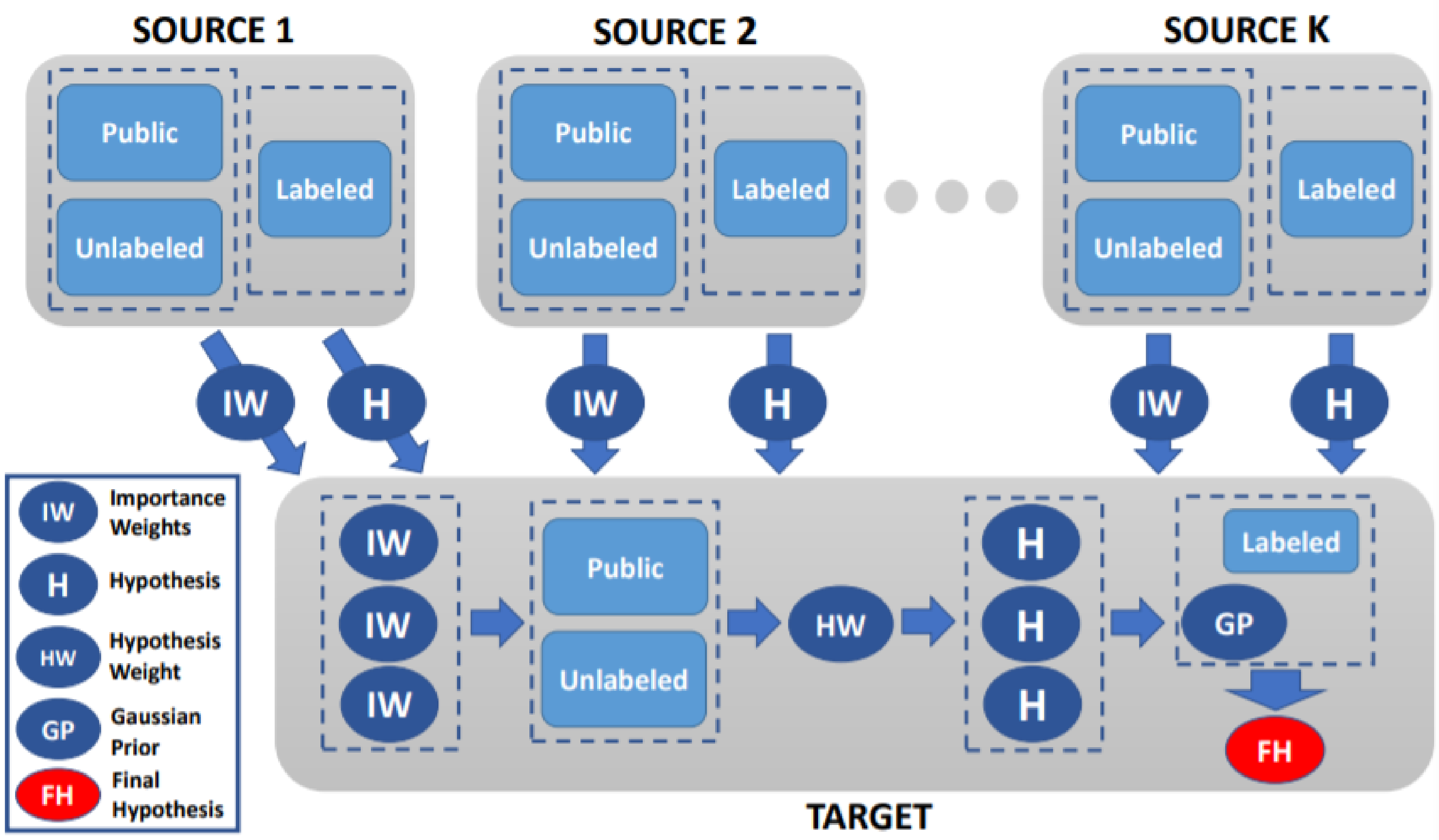
Definition: Differential Privacy (DP) [Dwork et.al. 2006, Dwork 2008]

A randomized mechanism M is ϵ-differentially private, if for all output t of M, and for all databases D_1 and D_2 which differ by at most one element, we have

$$\Pr(M(D_1) = t) = e^{\epsilon}\Pr(M(D_2) = t).$$



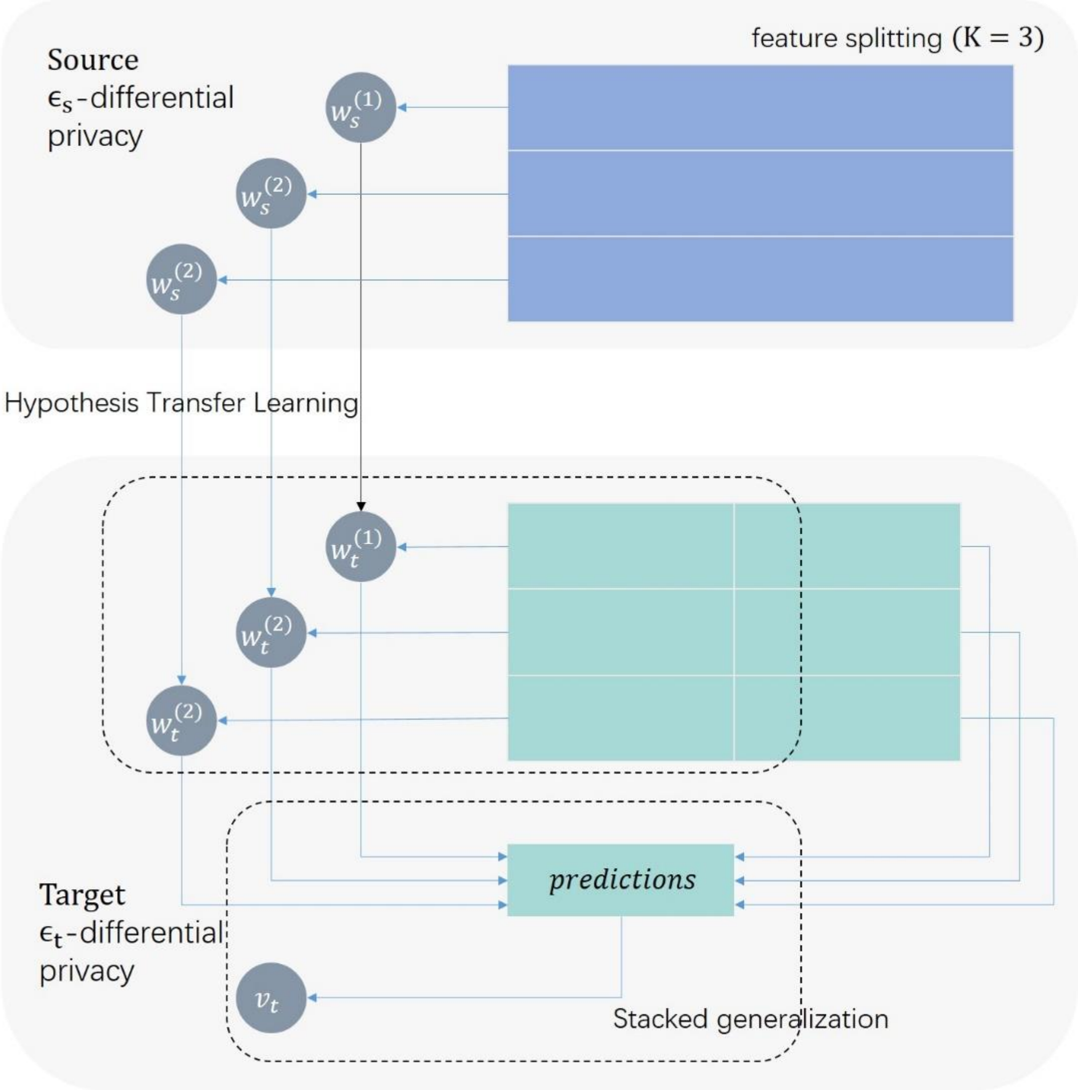
Differential Privacy with Transfer learning (I)



- leveraging the relatively abundant supply of unlabeled samples and an auxiliary public data set;
- derive the relationship between sources and target in a privacy-preserving manner.
- Source model hypothesis is also differential private

Yang Wang , et al., Differentially Private Hypothesis Transfer Learning, 2018

Privacy-preserving Hypothesis Transfer with feature splitting



- PRL in source

$$\mathbf{w}_s^* = \operatorname*{argmin}_{\mathbf{w}} F(\mathbf{w}_s; D, \mathbf{b}, \Delta) + \lambda_s g_s(\mathbf{w}_s)$$

- HTL + PRL in target

$$\mathbf{w}_t^* = \operatorname*{argmin}_{\mathbf{w}} F(\mathbf{w}_t; D, \mathbf{b}, \Delta) + \lambda_t g_t(\mathbf{w}_t) + \frac{1}{2}\|\mathbf{w}_t - \mathbf{w}_s^*\|_2^2$$

Split features randomly in source and make an ensemble of them in the target, Importance assigned with less noise

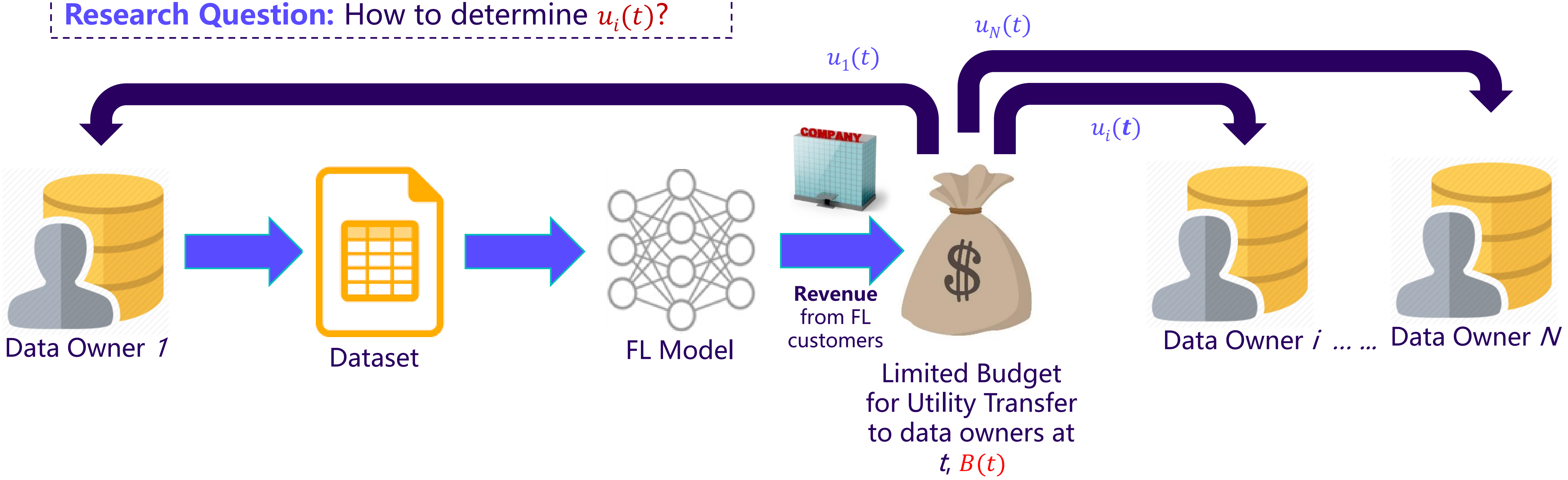
ϵ-differential privacy is guaranteed for both source and target

Privacy-preserving Transfer Learning for Knowledge Sharing. Guo & Yang et.al. Arvix 1811.09491. 2018

Incentivize Parties to Join: Federated Learning Exchange

- **Observation:** The success of a federation depends on data owners to share data with the federation
- **Challenge:** How to motivate continued participation by data owners in a federation?

Research Question: How to determine $u_i(t)$?



Federated Learning Exchange

- **Objective function for the Federated Learning Exchange (FLE) payoff sharing scheme:**

Maximize collective utility while *minimizing* inequality among data-owner regrets &waiting times

Maximize:

Regularization weight term

$$\omega U - \Delta$$

s. t.:

$$\sum_{i=1}^{N} \hat{u}_i(t) \leq B(t), \forall i, t$$

The actual payoff instalment for i at t if the federation does not have enough budget to pay out the full incentive amount for all data owners in one go

Solution:

Owed: $u_i(t) = \omega q_i(t) + Y_i(t) + c_i(t) + Q_i(t)$

Instalment: $\hat{u}_i(t) = \frac{u_i(t)}{\sum_{i=1}^{N} u_i(t)} B(t)$

- The computational time complexity of the algorithm is O(N).
- Once $Y_i(t)$ and $Q_i(t)$ reach 0 after some rounds of pay out with no new cost $c_i(t)$ incurred (i.e. $u_i(t) = \omega q_i(t)$), i will share future payoffs based on the quality of his data contribution.

FL Payoff-Sharing

nature > scientific reports > articles > article

SCIENTIFIC REPORTS

Mitigating Herding in Hierarchical Crowdsourcing Networks

Han Yu, Chunyan Miao, [...] Qiang Yang

Scientific Reports **6**, Article number: 4 (2016)

- **In order to fully commercialize federated learning among different organizations, a fair platform and incentive mechanisms need to be developed**

Individual

Contribution

- $q_i(t)$: Contribution of the dataset towards improving the FL model at t

Data Owner *i*

Cost

- $c_i(t)$: Cost of contributing a dataset at t
- $Y_i(t)$: A "regret queue" to track payoff due for data owner *i* at t
- $Q_i(t)$: A "temporal queue" to track how long data owner *i* has been waiting to receive full payoff from the federation at t

Maximizing collective utility while *minimizing* inequality among data owners' regret and waiting time

Solution:

Owed: $u_i(t) = \omega q_i(t) + Y_i(t) + c_i(t) + Q_i(t)$

Instalment: $\hat{u}_i(t) = \frac{u_i(t)}{\sum_{i=1}^{N} u_i(t)} B(t)$

- The computational time complexity of the algorithm is O(N).
- Once $Y_i(t)$ and $Q_i(t)$ reach 0 after some rounds of pay out with no new cost $c_i(t)$ incurred (i.e. $u_i(t) = \omega q_i(t)$), i will share future payoffs based on the marginal utility of his data

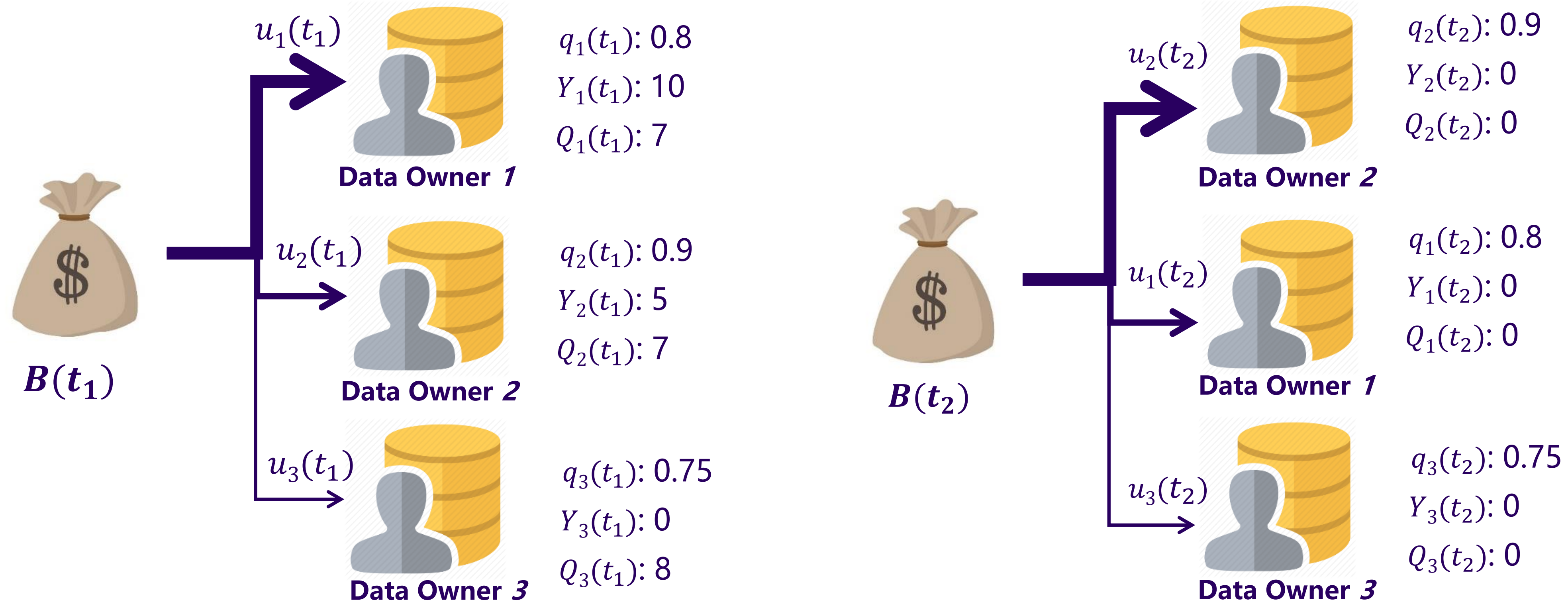
Federated Machine Learning: Advantages

	Coalition games with transferable utility [1]	Labour union games [2,3]	Fair-value games / Shapley games [2,3]	Federated Learning (this work)
Players do not need to engage in complex negotiations	X	√	√	√
Payers can join multiple coalition at the same time, cost for joining a coalition, the value of a dataset does not depreciate after being shared, and players' time spent waiting for cost to be compensated	X	X	X	√
Players' marginal contribution is important	√	X	√	√

1) B. Faltings, G. Radanovic & R. Brachman. *Game theory for data science: Eliciting truthful information*. Morgan & Claypool Publishers, p. 152, 2017.
2) J. Augustine, N. Chen, E. Elkind, A. Fanelli, N. Gravin & D. Shiryaev. Dynamics of profit-sharing games. *Internet Mathematics*, 1:1–22, 2015.
3) S. Gollapudi, K. Kollias, D. Panigrahi & V. Pliatsika. Profit sharing and efficiency in utility games. In *ESA*, pp. 1–16, 2017.

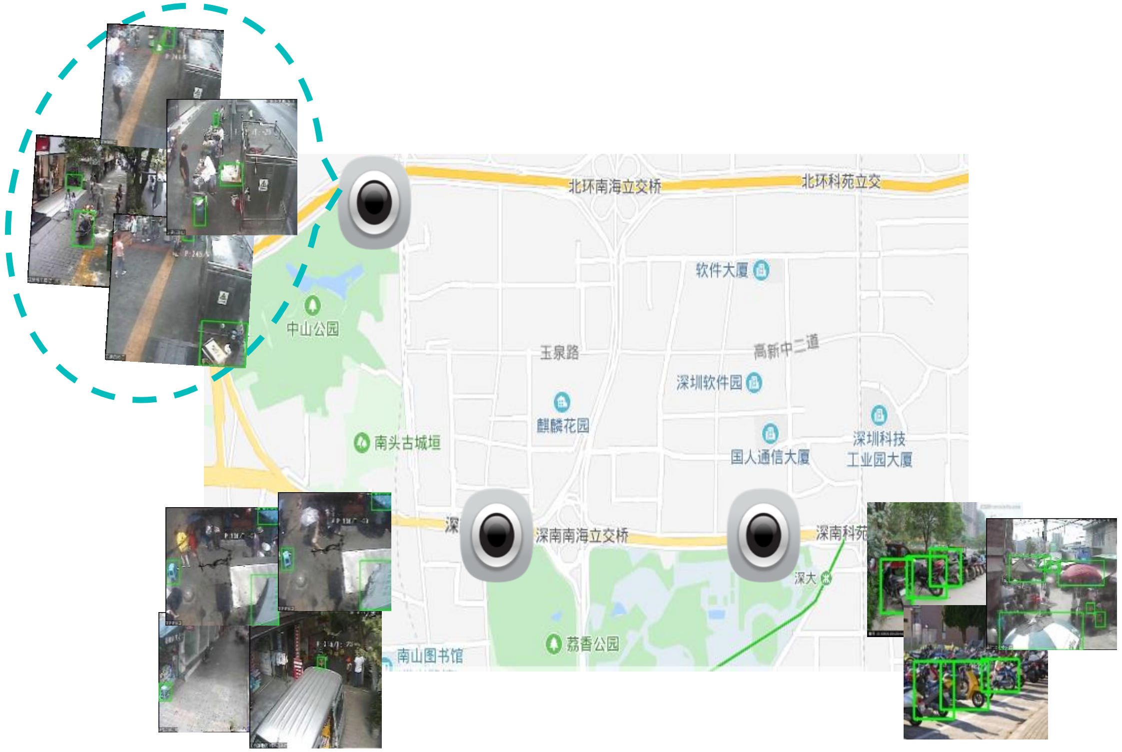
Examples

We set $\omega = 10$ in the following examples



Application: FML Network for Object Detection

- For parking and street vendor violation
- **A partner project of Webank AI and Extreme Vision in Shenzhen, China**



Challenges

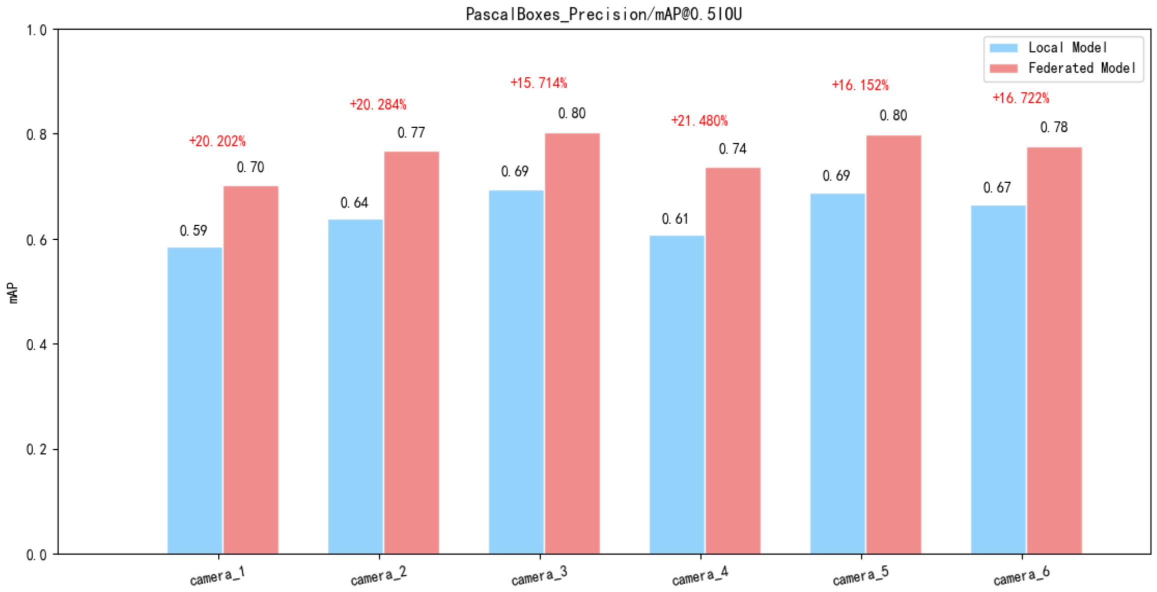
- Difficult detection task with few labels;
- Data are scattered; Expensive to centralize and manage data;
- Delayed feedback and delayed model updates.

A federated learning approach

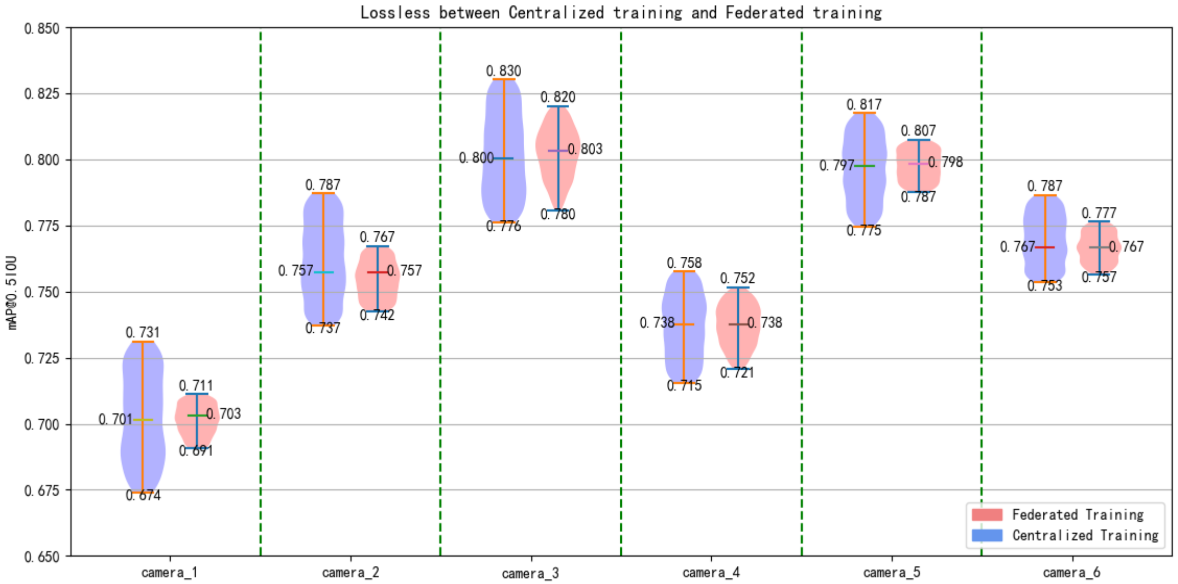
- Online feedback loop and model updates;
- No need to upload and centralize data;
- No sharing data.

7 class : {table, chair, carton, sunshade, basket, gastank, electrombile} with 6 cameras, 1922 images

Federated model improved local model by 15%



lossless performance
(Centralized model vs federated model)

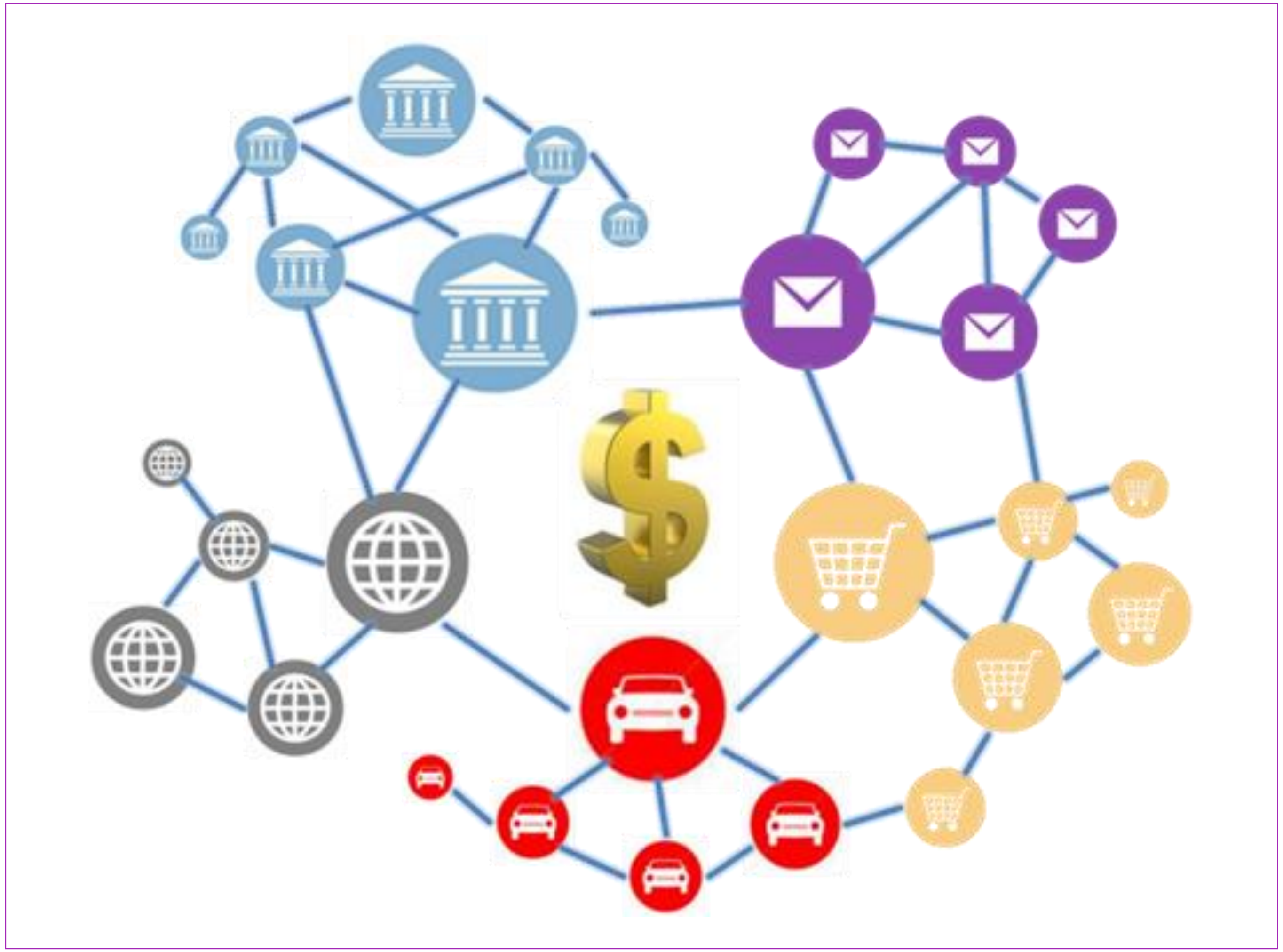


Federated AI Ecosystem

Security & Compliance

Lossless performance

Effective for small data and weak-supervision problem



IEEE Standard P3652.1 – Federated Machine Learning

➢ Title:

- Guide for Architectural Framework and Application of Federated Machine Learning
- Description and definition of federated learningThe types of federated learning and the application scenarios to which each type applied
- Performance evaluation of federated learning
- Associated regulatory requirements

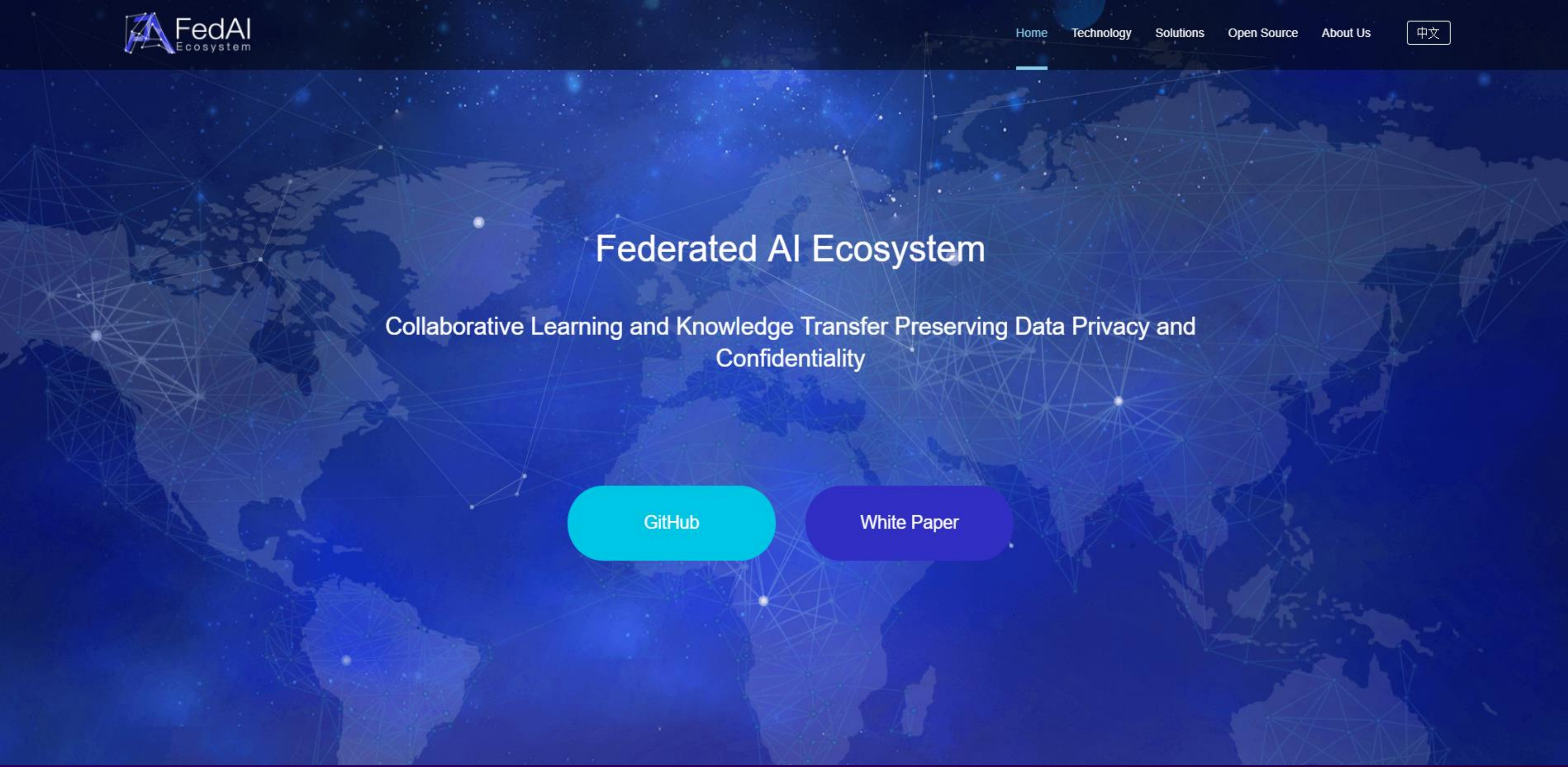
➢ First working group meeting:

- First working group meeting
- Dates: February 21~22, 2019
- Location: Shenzhen, China
- https://sagroups.ieee.org/3652-1/

Open Source in Feb 2019 – Federated AI Technology Enabler (FATE)

➢ **FATE is an open-source project initiated by Webank's AI Department**

- Supports federated learning architectures including horizontal federated learning, vertical federated learning and federated transfer learning
- Implements secure computation protocols based on homomorphic encryption and multi-party computing （MPC）
- Supports the secure computation of various machine learning algorithms, including logistic regression, tree-based algorithms, and deep learning and transfer learning



Find more information at https://www.fedai.org/

Summary

- AI’s Data Challenge: data shortage, regulations, and fragmentation
- Transfer Learning: from pretrained large models to small data
- Federated Machine Learning: secure collaboration in model building
- Federated Transfer Learning, Incentive Mechanisms and Open Source Frameworks
- https://sagroups.ieee.org/3652-1/
- https://www.fedai.org/