Learning to Selectively Transfer: Reinforced Transfer Learning for Deep Text Matching

Chen Qu[1], Feng Ji[2], Minghui Qiu[2 §], Liu Yang[1], Zhiyu Min[3], Haiqing Chen[2], Jun Huang[2], W. Bruce Croft[1]

[1]Center for Intelligent Information Retrieval, University of Massachusetts Amherst

[2]Alibaba Group

[3]Carnegie Mellon University

UMassAmherst
College of Information & Computer Sciences
Center for Intelligent Information Retrieval

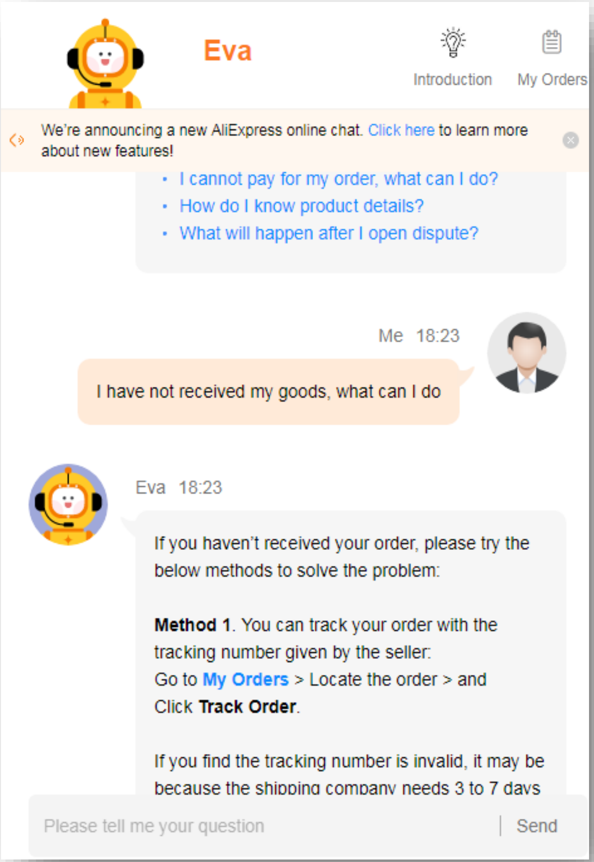
Alibaba Group 阿里巴巴集团

CARNEGIE MELLON UNIVERSITY PITTSBURGH PENNSYLVANIA 1900

§ Corresponding Author

Motivation

AliMe: a retrieval-based chatbot

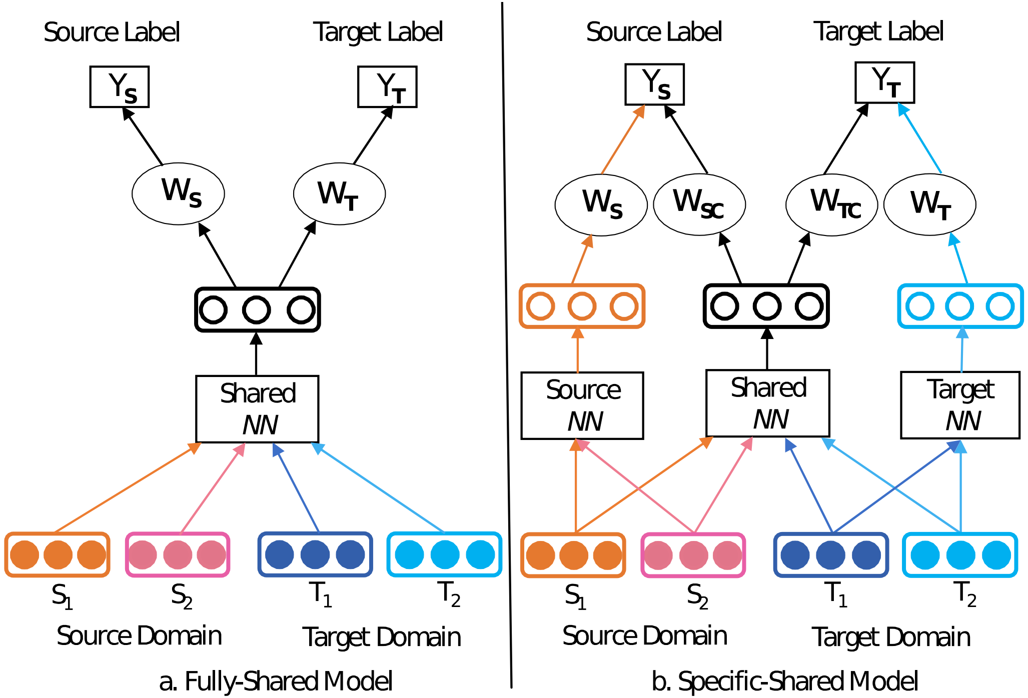


Motivation (Cont’d)

- Text matching is important in a **retrieval-based QA** systems.
- Use transfer learning to handle domains with **insufficient labeled data**.
- Avoid **negative transfer** in transfer learning.

Domain	Sentence 1	Sentence 2
Source (Open)	Which answers does Quora show first for each question?	How does Quora decide the ***order*** of the answers to a question?
	What ***order*** should the Matrix movies be watched in	Is there any particular ***order*** in which I should watch the Madea movies
	How can I ***order*** a cake from Walmart online?	How do I ***order*** a cake from Walmart?
Target (E-comm)	How long is my ***order*** arriving? Will I have the refund?	I have escalated an ***order*** and have not been updated in over a week
	How can i get an ***order*** receipt or invoice?	How do I get an invoice to pay?
	Why my ***order*** have been closed?	I need to understand why my ***orders*** have been cancelled

Neural Transfer Learning Models



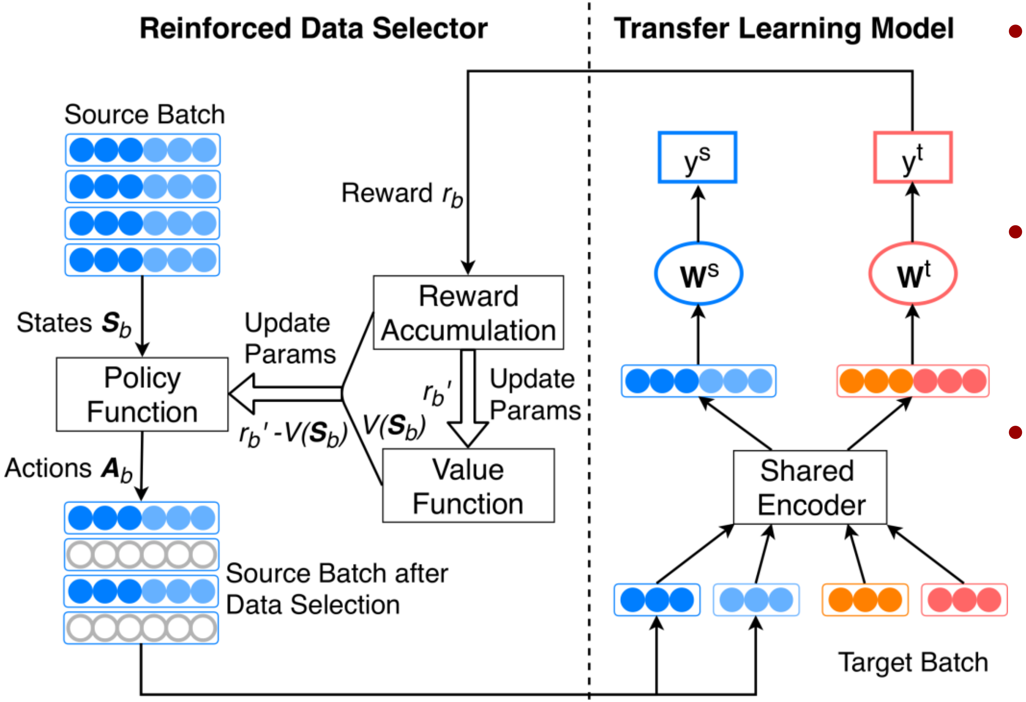
- Proven to be **effective** for text matching in QA.
- Most of existing data selection methods **don't fit well** with neural TL models.
- Data selection methods for transfer learning **need to be revisited under the DNN based TL setting**.

Figure is from Yu et al. "Modelling Domain Relationships for Transfer Learning on Retrieval-based Question Answering Systems in E-commerce". WSDM 2018.

Task Definition

- Text Matching
  - (sentence 1, sentence 2) → **semantically related** or not.
  - Paraphrase Identification (**PI**) / Natural Language Inference (**NLI**).
- Transfer Learning
  - The **same task** but **different domains**.
  - Both **labeled**; The source data is much larger than the target data.
- Data Selection
  - Setting: **DNN based transfer learning**.
  - Intervenes the TL model before each source batch.

Our Approach – RTL



- **Base model**:
  - Decomposable Attention Model (**DAM**[1]) for text matching.
- **Transfer Learning Model**:
  - To leverage a large amount of source domain data to help the target domain.
- **Reinforced Data Selector**:
  - Handle **source domain data selection**
  - Serves as an **agent** that **interacts** with the **environment** constructed by the TL model.

[1]Parikh, Ankur P. et al. "A Decomposable Attention Model for Natural Language Inference." EMNLP 2016.

Reinforced Data Selector

- **Model the problem as a Markov Decision Process**
  - **State**:
    - A hidden **representation** generated by the shared encoder.
    - The training/testing losses and probabilities on source/target model.
  - **Action**: **binary**, **drop or keep** the current training instance.
    - Action **sampled** according to a learned policy.
    - Policy approximated with a **policy network** with two fully-connected layers.
  - **Reward**:
    - The prediction **accuracy** on the **target validation** data.
    - Estimate the total return by a **value network**.
  - **Episode**:
    - Each epoch is an episode and each batch is a step to take actions.

Experiments

- Dataset:
  - **PI**: Quora Question Pairs (**open** domain) ➔ AnalytiCup (**E-commerce**)
  - **NLI**: MultiNLI (**open** domain) ➔ SciTail (**science**)

Task	Domain	Data	Train	Validation	Test
PI	Source	Quora QP	404,287/149,263	N/A	N/A
	Target	AnalytiCup	6,668/1,731	3,334/830	3,330/820
NLI	Source	MultiNLI	261,799/130,899	N/A	N/A
	Target	SciTail	23,596/8,602	1,304/657	2,126/842

- Baselines: base model, transfer baseline, Ruder and Plank (data selection with Bayesian optimization for TL)

Experiments (Cont'd)

- Experimental results:
  - RTL shows statistically significant improvements on both PI and NLI

Methods	PI		NLI	
	Acc	AUC	Acc	AUC
Base Model [22]	0.8393	0.8548	0.7300	0.7663
Transfer Learning Model	0.8488	0.8706	0.7453	0.8044
Ruder and Plank [27]	0.8458	0.8680	0.7521	0.8062
RTL	**0.8616**‡	**0.8829**	**0.7672**‡	**0.8163**

Ablation Analysis

- Reward functions and policy optimization method:

Methods		PI		NLI	
Reward	RL	Acc	AUC	Acc	AUC
AUC	REINFORCE	0.8557	0.8818	0.7486	0.8070
AUC	Actor-Critic	0.8545	0.8793	0.7613	0.8067
Acc	REINFORCE	0.8428	0.8788	0.7587	0.8121
Acc	Actor-Critic	**0.8616**	**0.8829**	**0.7672**	**0.8163**

- Acc is a better reward function
- Actor-critic is better than vanilla PG

- State features:

Features	PI		NLI	
	Acc	AUC	Acc	AUC
Transfer Learning Model	0.8488	0.8706	0.7521	0.8044
(1)	0.8539	0.8813	0.7594	0.8135
(2) (3) (4) (5)	0.8529	0.8778	0.7507	0.7916
(1) (2) (3) (4) (5)	**0.8616**	**0.8829**	**0.7672**	**0.8163**

- State representation with all the features gives the best performance.

Performance Interpretation

- **Wasserstein distance**
- **Term distributions**

Name	Domains in Comparison	PI	NLI
D_{origin}	Target ↔ Source	5.250E-06	3.256E-06
D_{select}	Target ↔ Source (Selected)	4.963E-06	3.190E-06
D_{drop}	Target ↔ Source (Dropped)	5.320E-06	3.290E-06
D_{rand}	Target ↔ Source (Random)	5.232E-06	3.243E-06

- $D_{rand} \approx D_{origin}$
- $D_{select} < D_{origin}$
- $D_{drop} > D_{origin}$

Our method can select source domain data whose Wasserstein distance is **close to the target domain data**.

Conclusions and Future Work

- A **reinforced data selection method** for **DNN based transfer learning**
  - Different settings of **states**, **rewards**, and **policy optimization**
  - Extensive experiments on **PI** and **NLI** demonstrated our effectiveness.
  - Used **Wasserstein distance** to interpret the model performance.

- Future work:
  - Explore more effective state representations
  - Adapt our method to other tasks.

Thank You!

Come to see our poster!