GAN, Cross-domain

2017/11/22

周杰

Unsupervised Domain Adaptation by Backpropagation

ICML15

Yaroslav Ganin, Victor Lempitisky

Outline

- **Introduction**
- **Deep Domain Adaptation**
- **Experiments**

Outline

- **Introduction**
- **Deep Domain Adaptation**
- **Experiments**

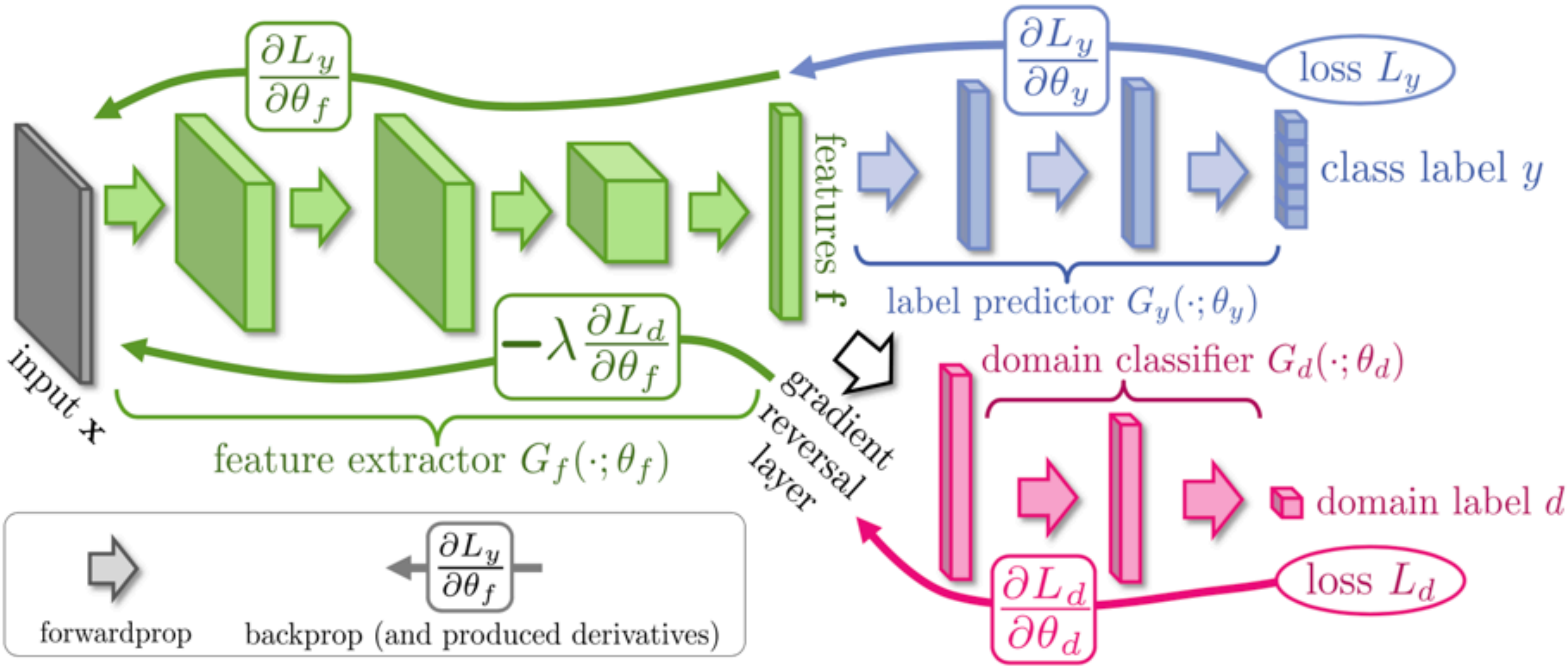
Introduction

- Domain adaptation in deep architectures
- Training
  - large amount of labeled data from the source domain
  - large amount of unlabeled data from the target domain
  - no labeled target-domain data is necessary
- Features
  - Discriminative for the main learning task on the source domain
  - Invariant with respect to the shift between the domains

Outline

- Introduction
- **Deep Domain Adaptation**
- Experiments

Deep Domain Adaptation



Deep Domain Adaptation

- **Minimize the loss of the label classifier**
- **Maximize the loss of the domain classifier**
    - **Have the similar distributions in the source and the target domains**

$$E(\theta_f, \theta_y, \theta_d) = \sum_{\substack{i=1..N \\ d_i=0}} L_y\left(G_y(G_f(\mathbf{x}_i;\theta_f);\theta_y), y_i\right) -$$

$$\lambda \sum_{i=1..N} L_d\left(G_d(G_f(\mathbf{x}_i;\theta_f);\theta_d), y_i\right) =$$

$$= \sum_{\substack{i=1..N \\ d_i=0}} L_y^i(\theta_f, \theta_y) - \lambda \sum_{i=1..N} L_d^i(\theta_f, \theta_d) \qquad (1)$$

Deep Domain Adaptation

- **Seek the parameters θ_f of the feature mapping that maximize the loss of the domain classifier**
- **Seek the parameters θ_d of the domain classifier that minimize the loss of the domain classifier**
- **Seek to minimize the loss of the label predictor**

$$(\hat{\theta}_f, \hat{\theta}_y) = \arg\min_{\theta_f, \theta_y} E(\theta_f, \theta_y, \hat{\theta}_d)$$

$$\hat{\theta}_d = \arg\max_{\theta_d} E(\hat{\theta}_f, \hat{\theta}_y, \theta_d)\,.$$

Deep Domain Adaptation

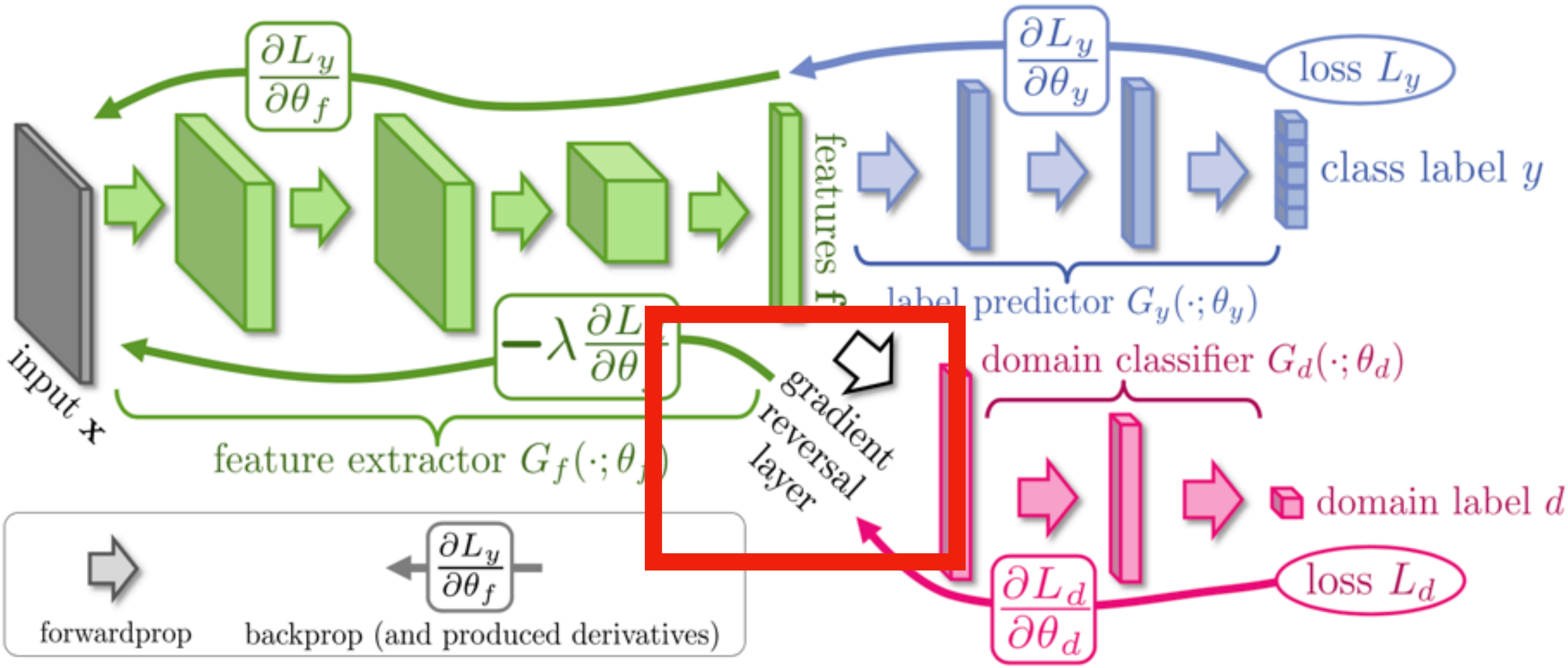
$$\theta_f \longleftarrow \theta_f - \mu\left(\frac{\partial L_y^i}{\partial \theta_f} - \lambda \frac{\partial L_d^i}{\partial \theta_f}\right) \quad (4)$$

$$\theta_y \longleftarrow \theta_y - \mu \frac{\partial L_y^i}{\partial \theta_y} \quad (5)$$

$$\theta_d \longleftarrow \theta_d - \mu \frac{\partial L_d^i}{\partial \theta_d} \quad (6)$$

- **Difference**
  - **-λ factor in (4)**
  - **Direct implementation of (4)-(6) as SDG is not possible**

Gradient reversal layer(GRL)



Gradient reversal layer(GRL)

Definition

$$R_\lambda(\mathbf{x}) = \mathbf{x}$$

$$\frac{dR_\lambda}{d\mathbf{x}} = -\lambda \mathbf{I}$$

$$\frac{\partial L_d}{\partial \theta_f} \rightarrow -\lambda \frac{\partial L_d}{\partial \theta_f}$$

Objective function

$$\tilde{E}(\theta_f, \theta_y, \theta_d) = \sum_{\substack{i=1..N \\ d_i=0}} L_y\left(G_y(G_f(\mathbf{x}_i;\theta_f);\theta_y), y_i\right) + \sum_{i=1..N} L_d\left(G_d(R_\lambda(G_f(\mathbf{x}_i;\theta_f));\theta_d), y_i\right) \qquad (9)$$

Outline

- Introduction
- Deep Domain Adaptation
- **Experiments**

Experiments

Classification accuracies for digit image classification.

Method (Source / Target)	MNIST / MNIST-M	Syn Numbers / SVHN	SVHN / MNIST	Syn Signs / GTSRB
Source only	.5225	.8674	.5490	.7900
SA (Fernando et al., 2013)	.5690 (4.1%)	.8644 (−5.5%)	.5932 (9.9%)	.8165 (12.7%)
Proposed approach	**.7666** (52.9%)	**.9109** (79.7%)	**.7385** (42.6%)	**.8865** (46.4%)
Train on target	.9596	.9220	.9942	.9980

Experiments

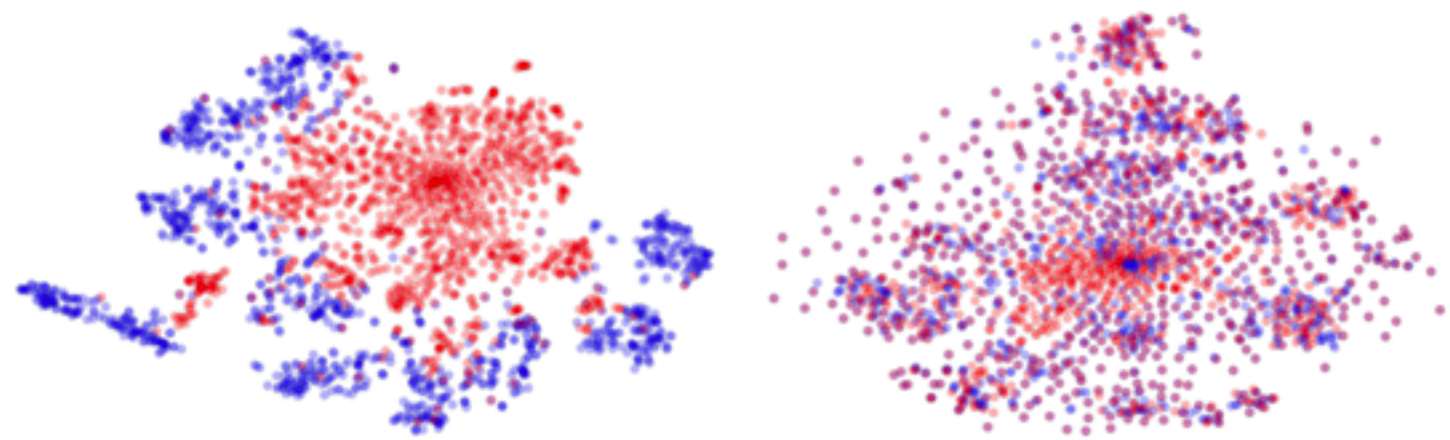
Accuracy evaluation of different DA approaches on the standard OFFICE dataset

Method	Source / Target	Amazon / Webcam	DSLR / Webcam	Webcam / DSLR
GFK(PLS, PCA) (Gong et al., 2012)		.214	.691	.650
SA* (Fernando et al., 2013)		.450	.648	.699
DLID (S. Chopra & Gopalan, 2013)		.519	.782	.899
DDC (Tzeng et al., 2014)		.605	.948	.985
DAN (Long & Wang, 2015)		.645	.952	.986
Source only		.642	.961	.978
Proposed Approach		**.730**	**.964**	**.992**

Experiments

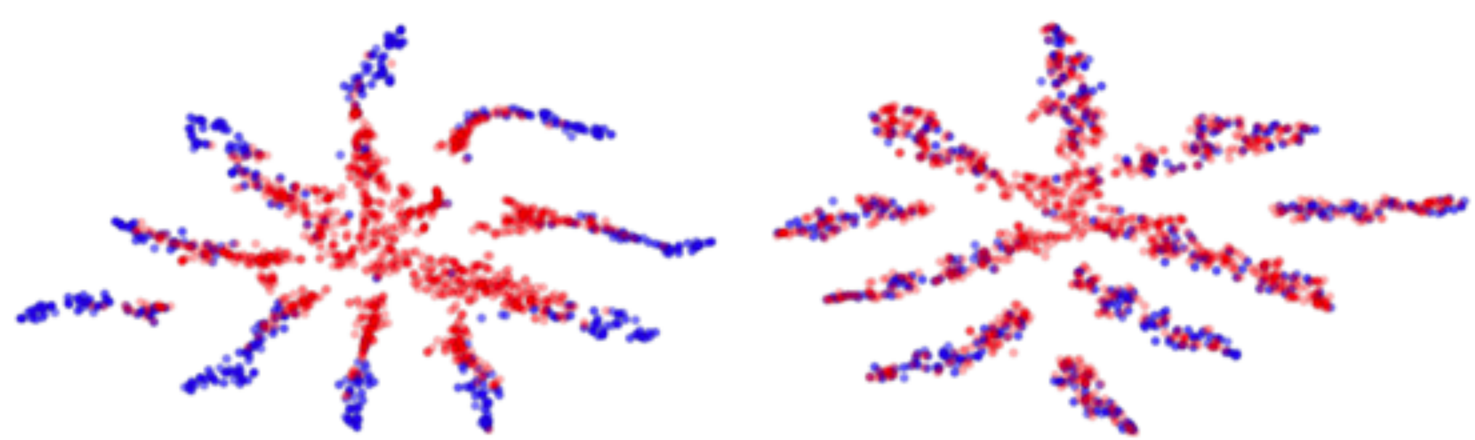
Visualize feature distributions

MNIST → MNIST-M: top feature extractor layer



(a) Non-adapted (b) Adapted

SYN NUMBERS → SVHN: last hidden layer of the label predictor



(a) Non-adapted (b) Adapted

End-to-End Adversarial Memory Network for Cross-domain Sentiment classification

IJCAI17
Zheng Li, Yu Zhang, Ying Wei, Yuixiang Wu, Qiang Yang

Outline

- **Introduction**
- **Adversarial Memory Network**
- **Experiment**

Outline

- **Introduction**
- **Adversarial Memory Network**
- **Experiment**

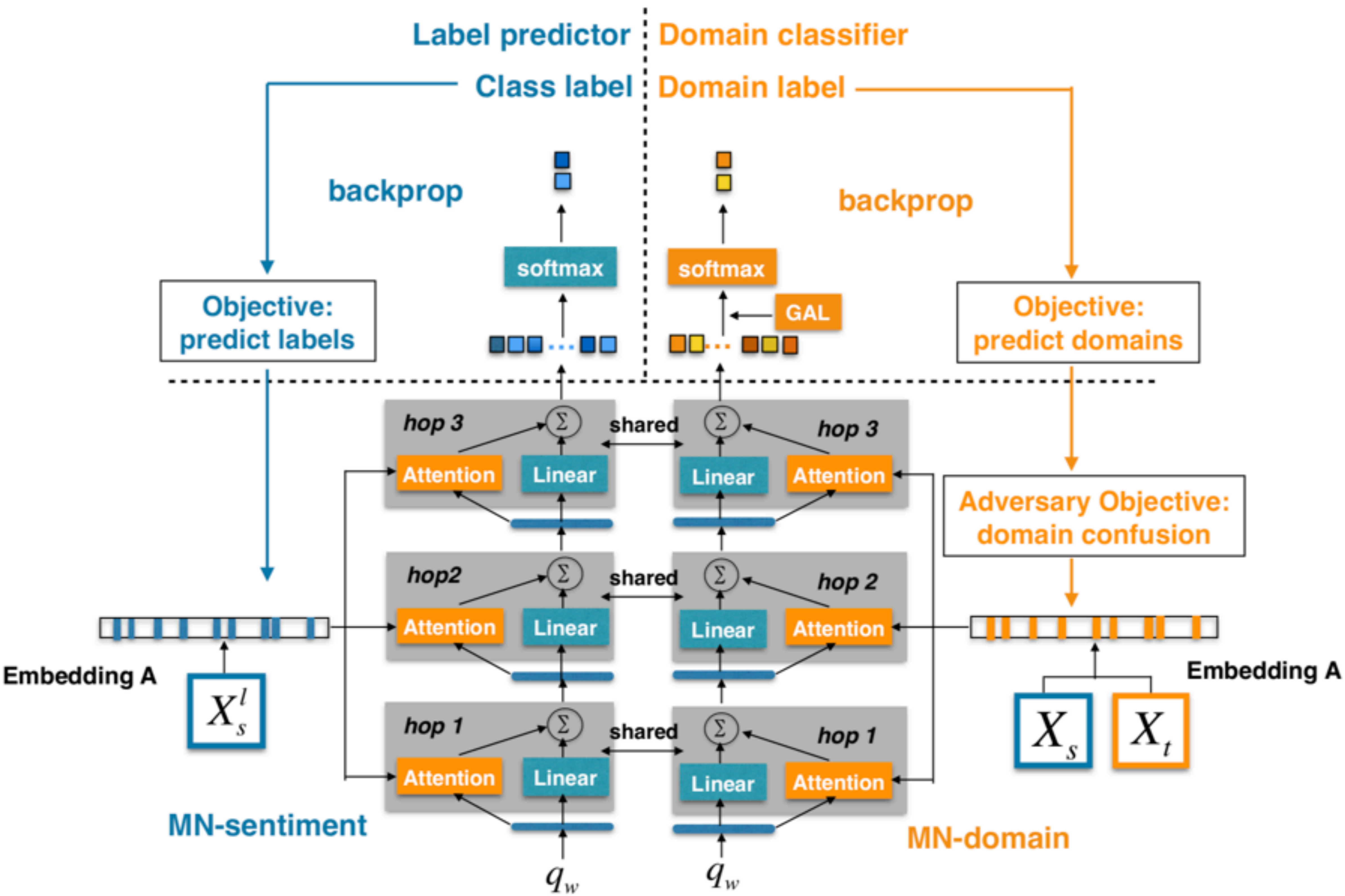
Introduction

- **Cross-domain sentiment classification**
- **Adversarial Memory Network**
- **AMN model:**
  - **capture the pivots using the attention mechanism**
  - **Offer a direct visualization of the pivots**

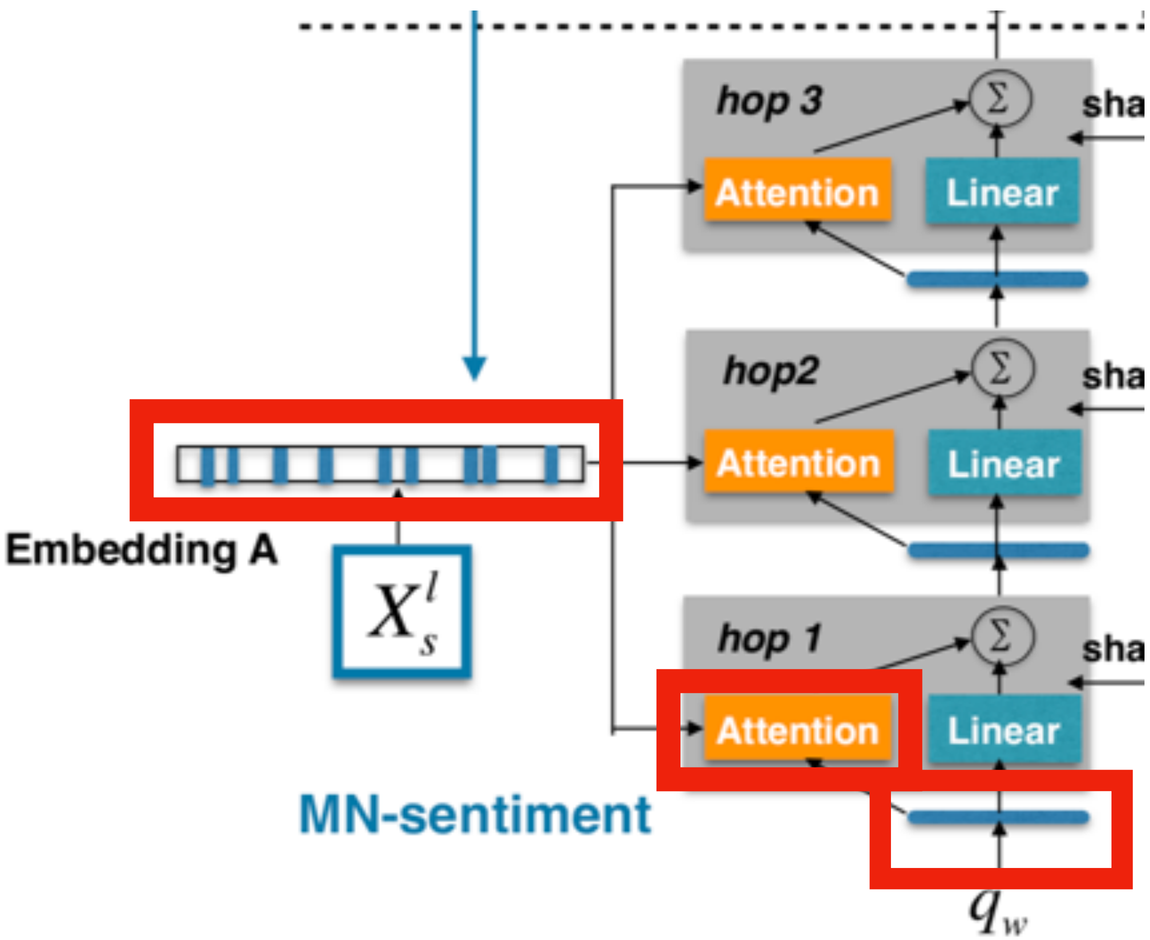
Outline

- Introduction
- **Adversarial Memory Network**
- Experiment

Adversarial Memory Network



Word Attention

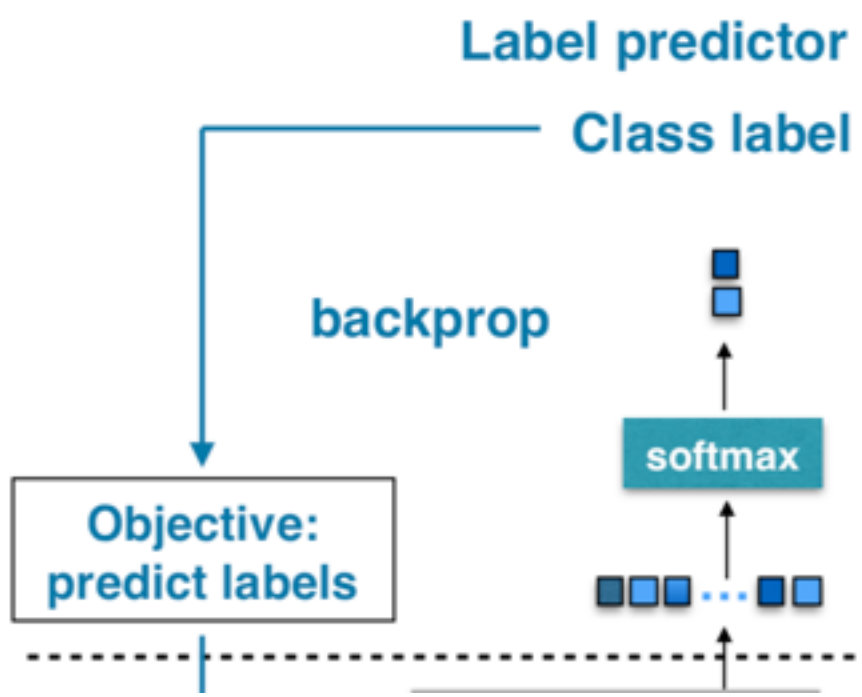


$$v = \sum_{i=1}^{m} a_i m_i.$$

$$a_i = \frac{\exp\left(h_i^T q_w\right)}{\sum_{j=1}^{n} \exp\left(h_j^T q_w\right)},$$

$$h_i = \tanh(W_s m_i + b_s)$$

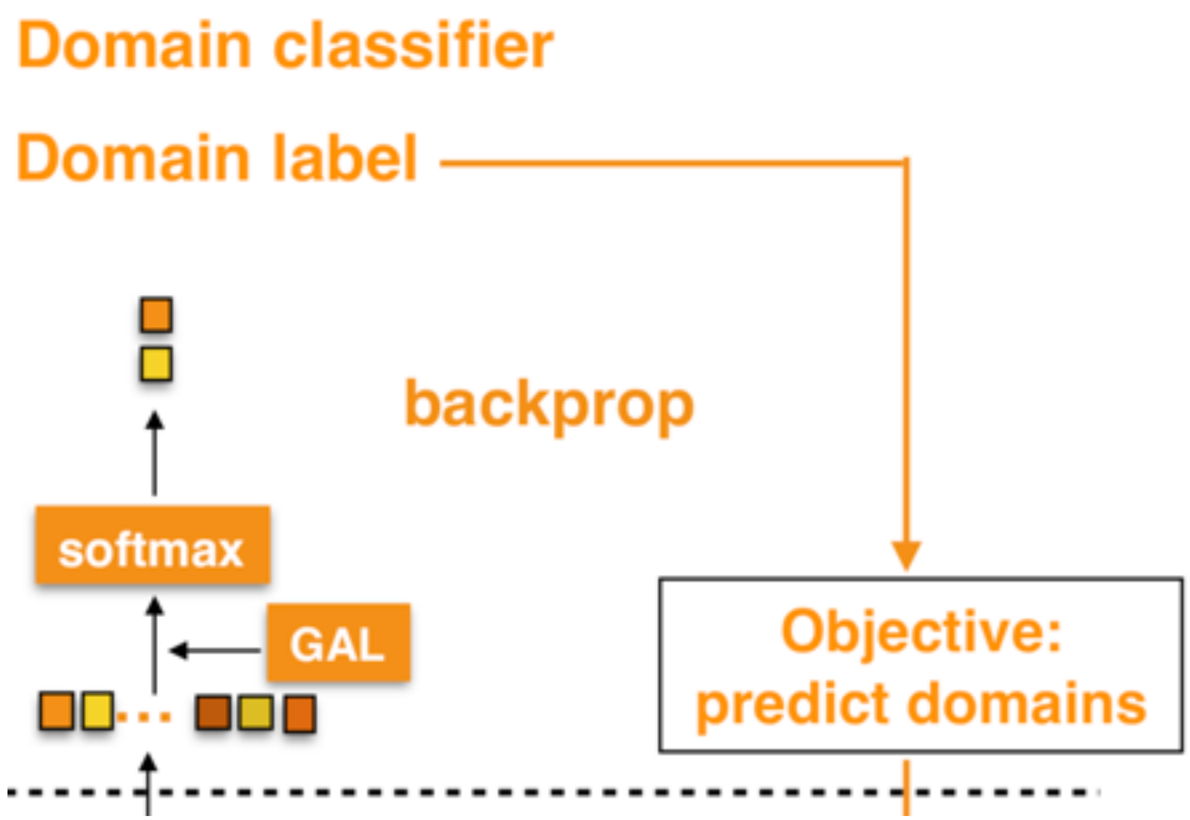
Sentiment Classifier



$$y = \mathrm{softmax}\left(W_s v_s + b_s\right).$$

$$L^{sen} = -\frac{1}{N_s^l} \sum_{i=1}^{N_s^l} \left(\hat{y}_i \ln y_i + (1 - \hat{y}_i) \ln (1 - y_i)\right),$$

Domain Classifier



$$Q_\lambda(v_d) = \hat{v}_d$$

$$d = \mathrm{softmax}\left(W_d \hat{v}_d + b_d\right)$$

$$L^{dom} = -\frac{1}{N_s + N_t} \sum_{i=1}^{N_s+N_t} \hat{d}_i \ln d_i + \left(1 - \hat{d}_i\right) \ln\left(1 - d_i\right)$$

Outline

- Introduction
- Adversarial Memory Network
- **Experiment**

Experiment

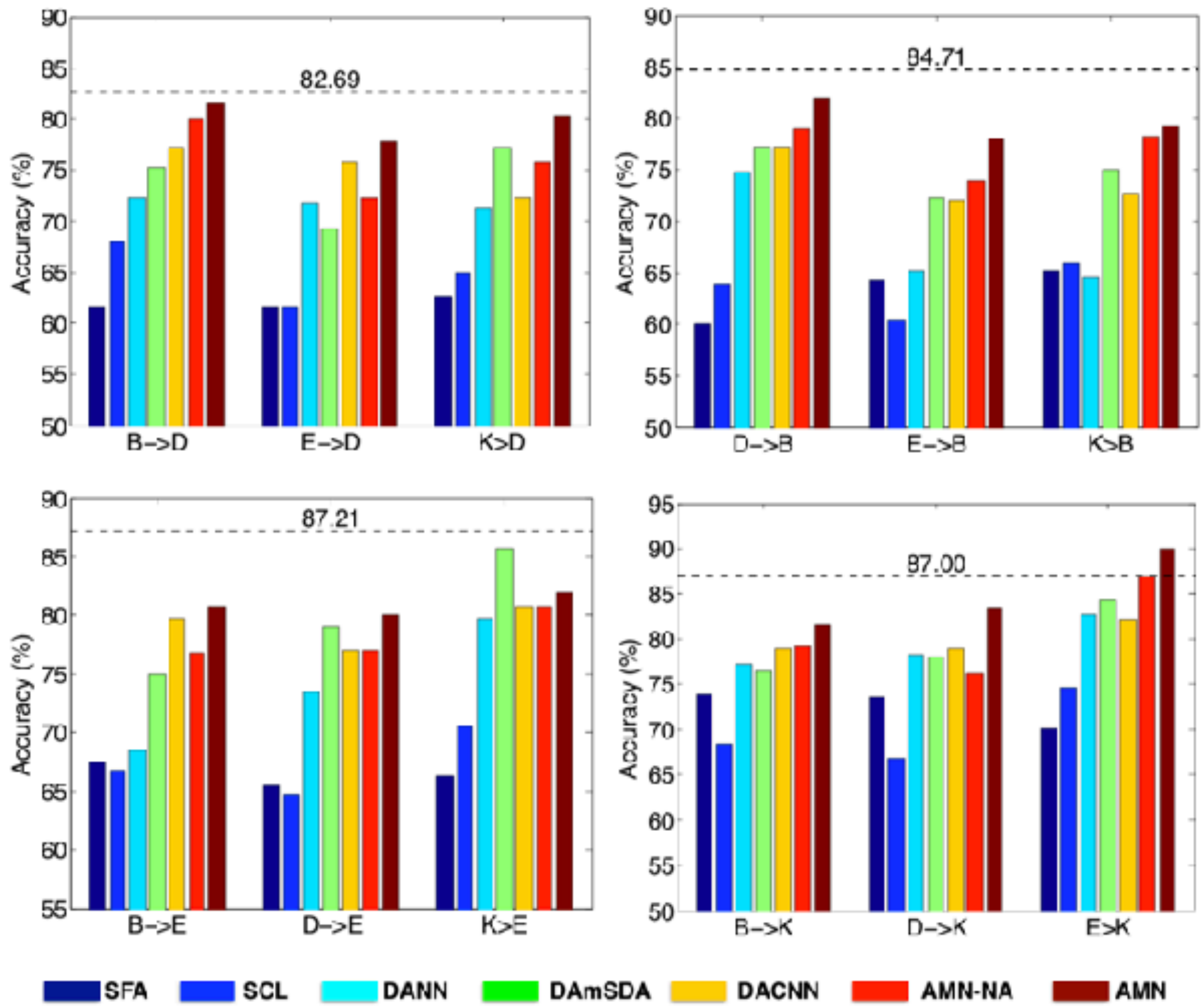


Figure 2: Average results for cross-domain sentiment classification on the Amazon reviews dataset.

Experiment

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

Figure 3: Samples from the Amazon reviews dataset in the E→K task. Deeper color implies larger attention weights. Label 1 denotes positive sentiment and label 0 denotes negative sentiment.

Thanks