

Dual Problem Reading Notes

伍瀚杰
华南理工大学

1. Dual Learning for Machine Translation

1.1. 论文题目和作者

He, Di, Xia, Yingce, Qin, Tao, Wang, Liwei, Yu, Nenghai, Liu, Tie-Yan, and Ma, Wei-Ying. Dual learning for machine translation. In Advances In Neural Information Processing Systems, pp. 820–828, 2016a.

1.2. 问题提出

目前在机器翻译的领域，神经机器翻译 (NMT) 快速发展，取得了很好的效果，但是为了得到一个好的 NMT 的模型，需要大量的人工标记的训练集即双语的语句对。一般采集的数据都是未标记的数据，通过人工去标记这些未标记的数据成本十分高。本文介绍一种新的学习的思想 (对偶学习) 去解决标记数据量不足的问题，并且将未标记的数据有效地利用起来。(利用部分训练的模型，利用对偶结构形成的闭环，来增加反馈)

1.3. 现有的解决方案

- Label Propagation (标签传播): 以图像分类为例，如果只有 1 万张标注数据，但我想获得 100 万张标注数据，我可以从网上抓取到很多图像，如果一个未标注的图像 A 和标注的图像 B 很相似 (利用已经训练好的图像分类模型)，那么就认为图像 A 具有和 B 相同的类别标签。这样就可以给很多未标注的数据加上标签，增加训练的数据量。
- Transductive learning (转导学习): 一种半监督学习的方法。
- Multi-task Learning (多任务学习): 每个任务都有自己的标注数据，那么多个任务在训练时可以共享这些数据，从而每个任务训练自己的模型时

都能看到更多的数据。

- Transfer Learning (迁移学习): 这也是最近比较火的一种训练方法。比如我有一个标注数据较少的目标任务 (target task)，但我的另一个 source task 的标注数据量比较充足，那么我们可以通过模型的迁移或将数据通过变换借鉴过来，以辅助 target task 的学习。

1.4. 作者提出的解决方案

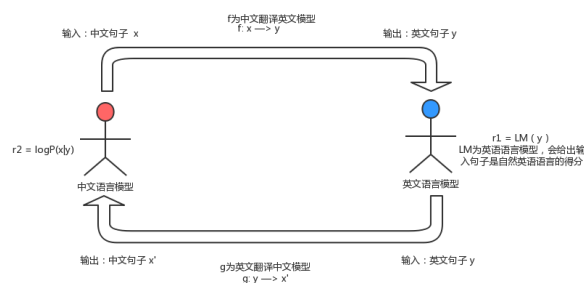


图 1. The Framework

作者提出的 dual learning 的主要框架为: 首先, 拿到一个无标注的中文句子 x , 我们并不知道的正确英文翻译 (no ground truth)。我们通过 primal model (f), 翻译成一个英文句子 y 。因为我们没有进行标注, 因此无从判断句子 y 的正误。但英语语言模型可以判断作为一个英语句子 y , 是否为一个通顺的、语法正确的句子。因此, 英语语言模型可以给出一个 partial feedback, 反馈句子 y 的质量如何。随后, 我们通过对偶模型 g , 将英文句子再翻译为中文句子 x' 。中文语言模型收到这个句子后, 它可以比较 x 与 x' 的相似度。如果 f 和 g 的表现很好, 那么 x 与 x' 应该非常相近。如果反之, 可能模型就需要改进。这个过程实际上非常

像强化学习的过程。在训练过程中，没有人告诉机器某个状态下正确的 action 是什么，只能通过「试错-反馈」的过程来反复尝试。以围棋为例子，可能需要走上百步才能知道输赢，但通过最终的反馈，就能训练提高这个模型的优劣。对于我们机器翻译在 x 的状态下，我们无从知道正确的 action 是什么，因此只能通过已有的 policy (f) 来 take action 得到 y ，再用另一个 policy (g) 得到 x' ，从而通过比较 x 和 x' 获得反馈。这实际上也是一个不断试错的过程，而且像强化学习一样，是具有延迟的反馈，最开始采取第一个 action 时， y 只能获得部分反馈，只有到流程结束，才能获得更有效的完整反馈，比如说 x 和 x' 的相似性。

Algorithm 1 The dual-learning algorithm

```

1: Input: Monolingual corpora  $D_A$  and  $D_B$ , initial translation models  $\Theta_{AB}$  and  $\Theta_{BA}$ , language models  $LM_A$  and  $LM_B$ ,  $\alpha$ , beam search size  $K$ , learning rates  $\gamma_{1,t}, \gamma_{2,t}$ .
2: repeat
3:    $t = t + 1$ .
4:   Sample sentence  $s_A$  and  $s_B$  from  $D_A$  and  $D_B$  respectively.
5:   Set  $s = s_A$ .  $\triangleright$  Model update for the game beginning from  $A$ .
6:   Generate  $K$  sentences  $s_{mid,1}, \dots, s_{mid,K}$  using beam search according to translation model  $P(\cdot|s; \Theta_{AB})$ .
7:   for  $k = 1, \dots, K$  do
8:     Set the language-model reward for the  $k$ th sampled sentence as  $r_{1,k} = LM_B(s_{mid,k})$ .
9:     Set the communication reward for the  $k$ th sampled sentence as  $r_{2,k} = \log P(s|s_{mid,k}; \Theta_{BA})$ .
10:    Set the total reward of the  $k$ th sample as  $r_k = \alpha r_{1,k} + (1 - \alpha)r_{2,k}$ .
11:  end for
12:  Compute the stochastic gradient of  $\Theta_{AB}$ :

$$\nabla_{\Theta_{AB}} \hat{E}[r] = \frac{1}{K} \sum_{k=1}^K [r_k \nabla_{\Theta_{AB}} \log P(s_{mid,k}|s; \Theta_{AB})].$$

13:  Compute the stochastic gradient of  $\Theta_{BA}$ :

$$\nabla_{\Theta_{BA}} \hat{E}[r] = \frac{1}{K} \sum_{k=1}^K [(1 - \alpha) \nabla_{\Theta_{BA}} \log P(s|s_{mid,k}; \Theta_{BA})].$$

14:  Model updates:

$$\Theta_{AB} \leftarrow \Theta_{AB} + \gamma_{1,t} \nabla_{\Theta_{AB}} \hat{E}[r], \Theta_{BA} \leftarrow \Theta_{BA} + \gamma_{2,t} \nabla_{\Theta_{BA}} \hat{E}[r].$$

15:  Set  $s = s_B$ .  $\triangleright$  Model update for the game beginning from  $B$ .
16:  Go through line 6 to line 14 symmetrically.
17: until convergence

```

图 2. Algorithm Description

使用 policy gradient theorem 进行梯度下降计算

1.5. 实验过程

- 模型: dual-NMT (本文的)、standard NMT、pseudo-NMT
- 数据集: WMT'14
- GPU: K40m GPU
- 评价标准: BLEU, 越高越好

Table 2: Reconstruction performance of En \leftrightarrow Fr task

	En \rightarrow Fr \rightarrow En (L)	Fr \rightarrow En \rightarrow Fr (L)	En \rightarrow Fr \rightarrow En (S)	Fr \rightarrow En \rightarrow Fr (S)
NMT	39.92	45.05	28.28	32.63
pseudo-NMT	38.15	45.41	30.07	34.54
dual-NMT	51.84	54.65	48.94	50.38

图 3. Experimental Result

1.6. 作者方案的效果和局限

- 使用 dual-NMT 模型一起训练中翻法和法翻中两个模型，模型的 BLEU 得分。比使用标准的和目前最好的 NMT 模型单独训练两个模型的效果要好。
- 局限文中没描述，我暂时也没想到。

1.7. 我的思考

对偶学习的基本思想是两个对偶的任务能形成一个闭环反馈系统，使我们得以从未标注的数据上获得反馈信息，进而利用该反馈信息提高对偶任务中的两个机器学习模型。在将该思想扩展到三个任务或多个任务时，前提是只要它们能形成一个闭环反馈系统。例如，从中文翻译到英文，然后从英文翻译到日文，再从日文翻译到中文。另外一个例子是从图片转化成文字，然后从文字转成语音，再从语音转化成图片。利用中间结果的反馈和转了一圈回来后的输出和原输入比较的反馈来充分利用数据。

2. Dual Supervise Learning

2.1. 论文题目和作者

Yingce Xia, Tao Qin, Wei Chen, Jiang Bian, Nenghai Yu and Tie-Yan Liu, Dual Supervised Learning, In 34th International Conference on Machine Learning, (ICML-2017)

2.2. 问题提出

在传统的监督学习中，两个任务的训练过程都是分开的。我们根据之前那篇论文可知，对偶结构在无监督学习中能起到提高数据的利用性，从而提高模型的正确率。考虑在监督学习中引入对偶结构。假设 x, y 为两种语言的数据空间。两个模型的联合概率 $P(x, y)$ 应该是一样的，但是将两个模型分开训练时，没有保证让两个分开训练的模型他们的联合概率是一样的。所

以在这篇论文中，利用对偶结构，并且显式地定义一个正则项 $\min |P(x)P(y|x; f) - P(y)P(x|y; g)|$ 将两个概率值的 gap 最小化，从而达到保证他们联合概率非常逼近的效果。并且这个正则项和其他的正则项不同，一般像 SVM 模型的正则项都是模型参数有关，但本文的正则项是和数据有关的。

2.3. 现有解决方案

- Co-training
- Multi-task
- Transfer Learning

2.4. 作者提出的解决方案

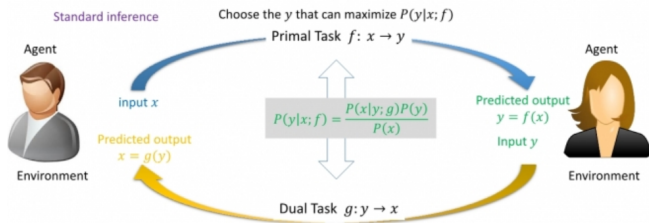


图 4. The Framework

最主要的创新就是用联合概率函数的公式，把两种语言的数据用进去，进行优化。同时作者利用 Rademacher complexity 证明了利用对偶结构后会比普通的监督学习的方法收敛得更快，更容易学习。

Algorithm 1 Dual Supervise Learning Algorithm

Input: Marginal distributions $\hat{P}(x_i)$ and $\hat{P}(y_i)$ for any $i \in [n]$; Lagrange parameters λ_{xy} and λ_{yx} ; optimizers Opt_1 and Opt_2 ;

repeat

Get a minibatch of m pairs $\{(x_j, y_j)\}_{j=1}^m$;

Calculate the gradients as follows:

$$\begin{aligned} G_f &= \nabla_{\theta_{xy}} (1/m) \sum_{j=1}^m [\ell_1(f(x_j; \theta_{xy}), y_j) \\ &\quad + \lambda_{xy} \ell_{\text{duality}}(x_j, y_j; \theta_{xy}, \theta_{yx})]; \\ G_g &= \nabla_{\theta_{yx}} (1/m) \sum_{j=1}^m [\ell_2(g(y_j; \theta_{yx}), x_j) \\ &\quad + \lambda_{yx} \ell_{\text{duality}}(x_j, y_j; \theta_{xy}, \theta_{yx})]; \end{aligned} \quad (4)$$

Update the parameters of f and g :

$\theta_{xy} \leftarrow Opt_1(\theta_{xy}, G_f)$, $\theta_{yx} \leftarrow Opt_2(\theta_{yx}, G_g)$.

until models converged

图 5. Algorithm Description

2.5. 实验过程

将这种模型用在三种应用场景（文字翻译、图像分类-图像生成、情感分析-根据情感词生成句子）都得到了比目前的模型所能得到的最好结果还要好的结果。

2.6. 方案优势和局限

- 体现了对偶学习不仅在无监督学习可行，在监督学习中同样能提高模型训练的效果。

2.7. 我的思考

结合这两篇文章的 future work, dual learning 未来工作主要是有几个：

- 将 dual learning 这种模型结构应用到更多的 AI 任务中，像最新的 cvpr 那篇应用 dual 到 VQA 和 VQG 的。同时在同一种类型的任务，改变训练过程或者优化过程。
- duality 的理论研究
- dual learning 监督学习和 dual learning 无监督学习结合的训练方式结合
- dual inference 结合在训练预测的时候同时结合对偶模型。

监督学习的 dual learning 的代码目前是有情感分析和图片分类和生成的，上次找到那个是无监督学习的文本翻译的 dual learning 代码，我想先跑下他的代码，然后再在那个基础上加，因为无监督学习加的方法相对来说比较直观，就是对训练一个语言模型。监督学习这里应该是从三变量的联合概率入手，仿照两变量的把他们串起来。