

CCL23-Eval 任务3系统报告：苏州大学CFSP系统

刘亚慧, 李正华, 张民
苏州大学 计算机科学与技术学院, 江苏 苏州
yahuiliu.nlp@foxmail.com,
{zhli13,minzhang}@suda.edu.cn

摘要

本文介绍了我们在第二十二届中国计算语言学大会汉语框架语义解析评测中提交的参赛系统。框架语义解析是自然语言处理领域中重要的任务，其目标是从句子中提取框架语义结构。本次评测任务针对汉语框架语义的三个子任务（框架识别、论元范围识别和论元角色识别）使用不同的端到端框架进行解析，并利用数据增强和投票方法进一步提高预测的精度，最终，在A榜测试集上取得第二名，B榜测试集上取得第三名。

关键词： 汉语框架语义解析；数据增强；投票；端到端

System Report for CCL23-Eval Task3: SUDA CFSP System

Yahui Liu, Zhenghua Li, Min Zhang
School of Computer Science and Technology, Soochow University, Suzhou, China
yahuiliu.nlp@foxmail.com,
{zhli13,minzhang}@suda.edu.cn

Abstract

This paper introduces the model we submitted in the shared task of Chinese Frame Semantic Parsing (CFSP) at the Twenty-second China National Conference on Computational Linguistics. Frame Semantic Parsing (FSP) is an important task in Natural language Processing, aiming to extract the frame semantic structure from the sentence. This work uses end-to-end frameworks to parse the three sub-tasks of CFSP (Frame Identification, Argument Identification, and Role Identification), and employs Data Augmentation and Voting Technique to further improve the accuracy of prediction. In the end, we achieved second place in the testA set and third place in the testB set.

Keywords: Chinese Frame Semantic Parsing , Data Enhancement , Voting Technique, End-to-End

1 任务介绍

汉语框架语义解析（Chinese Frame Semantic Parsing, CFSP）是一种浅层的语义解析任务(宋衡et al., 2022)，使用汉语框架网（Chinese FrameNet, CFN）(You and Liu, 2005)资源作

©2023 中国计算语言学大会

根据《Creative Commons Attribution 4.0 International License》许可出版

项目资助：国家自然科学基金（62176173）、江苏高校优势学科建设工程资助项目

框架名称:	等同
框架定义:	表示两个实体具有相等、相同、同等看待等的关系。
框架元素:	实体集 具有同等关系的两个或多个实体
	实体1 与实体2具有等同关系的实体
	实体2 与实体1具有等同关系的实体
	施动者 判断实体集具有同等关系的人
	方式 修饰用来概括无法归入其他更具体的框架元素的任何语义成分, 包括认知的修饰(如很可能、大概、神秘地), 辅助描述(安静地, 大声地), 和与事件相比较的一般描述(同样的方式)
	时间 实体之间具有等同关系的时间

Table 1: 汉语框架语义网中关于名为“等同”的框架示例

为基础, 旨在识别出句子中给定目标词所触发的框架及其对应的框架元素(王晓晖 et al., 2022)。该任务对阅读理解(Guo et al., 2020b; Guo et al., 2020a)、文本摘要(Guan et al., 2021)、关系抽取(Zhao et al., 2020)等下游任务具有非常重要的意义。

CFN是由山西大学以Charles J. Fillmore提出的框架语义学(Frame Semantics)为理论基础, 以英文FrameNet(Fillmore et al., 2003)为参照, 以汉语的真实语料为依据构建而成。CFN是一种结构化的知识表示, 它建立了词汇和概念之间的框架关系。每个框架包含了与特定概念相关的语义信息。表1给出了CFN中名为“等同”的框架, 该框架表达的概念为两个实体具有相等、相同、同等看待等的关系, 对应的词元有“是”、“为”等(对应的词元就是句子中要标注的目标词)。框架中的框架元素用于捕获句子中与框架相关的语义信息。不同的框架元素被赋予不同的含义, 如“实体1”表示与实体2具有等同关系的实体, “时间”表示实体之间具有等同关系的时间。

本次评测将汉语框架语义解析分为三个子任务: 框架识别、论元范围识别和论元角色识别。框架识别任务意为根据目标词在句子中的含义从所有框架中找到其对应的框架。框架元素也称为论元, 在句子中论元可能是一个词, 也可能是一个片段, 论元范围识别是指在句子中识别出框架元素的位置, 并确定论元的边界。而论元角色识别则是为识别出的框架元素分配相应的框架元素名称(又称为语义角色标签)。论元范围识别和论元角色识别两个任务合起来称为论元识别任务。以图1(a)为例, 在句子“餐饮业是天津市在海外投资的重点之一”中, 给定目标词“是”, 该目标词对应的框架为表1中的“等同”。在句子中有两个与框架相关的框架元素, 即“餐饮业”和“天津市在海外投资的重点之一”, 它们分别对应于语义角色标签“实体1”和“实体2”。

在本次评测任务中, 对于框架识别任务, 我们借鉴Zhou et al. (2022)的方法, 将语义框架解析转换为一个基于词的图解析任务, 在端到端的框架中将目标词的框架和对应元素一起识别出来, 并在解码时融合高阶信息, 利用目标词与论元之间的边来帮助框架的识别。但该框架下论元识别任务的性能没有达到预期, 我们借鉴Zhang et al. (2021)的方法, 通过建模论元片段的内部结构, 将基于片段的论元识别任务转换为一个树解析任务来增强识别论元范围和论元角色的能力。我们的代码发布在<https://github.com/yahui19960717/CFN-FINETUNE.git>。

2 相关工作

评测任务SemEval2007 Task-19(Baker et al., 2007)首次提出了框架语义结构抽取任务, 包括目标词识别、框架识别、论元识别等任务。目前已有的框架语义解析任务可以分为两类: 基于特征工程的方法和基于神经网络的方法。

对于框架识别任务, 早期研究采用传统的机器学习方法, 人工构建特征, 使用最大熵(李济洪 et al., 2011; 李国臣 et al., 2013)、支持向量机(Johansson and Nugues, 2007)等模型进行框架的识别。对于论元识别任务, 早期也采用人工构建的特征, 使用条件随机场模型(李济洪 et al., 2010)和最大熵模型(王蔚林, 2010)来实现论元的识别。另外, 屠寒非 et al. (2016)使用主动学习

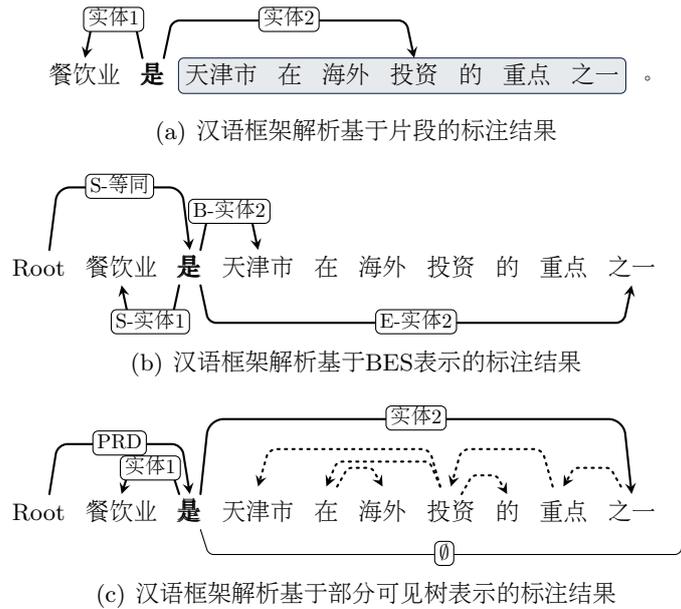


Figure 1: 汉语框架语义解析示例

方法来提升论元的识别能力。但是传统算法在构建特征时费时费力，且特征较为稀疏。

随着深度学习的发展，许多工作采用神经网络进行框架语义解析。对于框架识别任务，赵红燕et al. (2016)提出使用DNN架构来学习目标词的上下文特征进行框架的识别；张力文et al. (2017)在分布式表征的基础上，提出基于距离和基于词相似度矩阵进行CFN上的框架识别；Su et al. (2021)通过利用框架关系和定义来帮助框架的识别。但上述工作都不是针对于目标词的向量表示，为了约束模型的学习，You et al. (2022)基于GCN和门机制利用句法信息帮助框架的识别。对于论元识别任务，李济洪et al. (2010)用BIO策略将CFN转化词层面的线性序列标注，采用条件随机场模型来实现论元的识别。在与框架语义解析任务相似的语义角色标注任务上，有些研究尝试将基于片段的输入转为端到端的模型(Tan et al., 2018; Ouchi et al., 2018)。王晓晖et al. (2022)在此基础上，引入self-attention机制来提升CFN上论元识别的性能。

考虑到框架元素与框架有密不可分的关系，本文借鉴SRL任务上Zhou et al. (2022)的方法，在一个端到端的任务中将框架元素和框架一起识别出来，并在解码的时候利用目标词和框架元素之间的关系帮助框架的识别。对于论元识别任务，之前的工作都没有考虑论元片段的内部结构。在语义角色标注 (Semantic Role Labeling, SRL) 任务上，Zhang et al. (2021)显示对论元片段进行建模可提升论元识别的性能。SRL任务和框架语义解析任务非常相似，它们都涉及到对句子中语义信息的解析和标注。本文借鉴Zhang et al. (2021)的方法，通过建模论元片段的内部关系来帮助框架语义解析中的论元范围识别和论元角色识别任务。

3 方法介绍

针对任务一框架识别任务，本文采用BES的图表示结构将基于片段的语义框架解析转换成基于词的图解析任务。BES是一种将句子转换为一个序列的策略，被用来标记句子中的论元。当论元为多个词时，B表示论元片段开始，E表示论元片段的结束，当论元为单个词时则使用S进行标记。基于Zhou et al. (2022)，本文在端到端框架中把句子中目标词对应的框架和论元放在一起识别，在解码时考虑多条边的关系，利用目标词的论元信息来帮助其框架的识别。具体而言，就是在一个句子的开始添加一个“Root”头节点，然后将其与句子中的目标词相连，和Zhou et al. (2022)不同，本文将Zhou et al. (2022)中目标词与“Root”之间的边标签“PRD”改成了“S-r”（注意，目标词一般都是单个词），用于表示不同的框架。对于目标词和论元之间边的标签，如果论元是单个词，则使用“S-r”，如果论元由多个词组成，那么只将目标词与对应论元中的开始词和结束词相连接，并分别使用“B-r”和“E-r”作为边标签，其中“r”表示目标词框架名称或原始的论元标签，如图1(b)所示。因为对每条边的标签预测是独立的，所有可能会出现非法图的情况，例如单个词的论元被标注成了“B-r”或“E-r”，则在解码的时候使用一个受限的

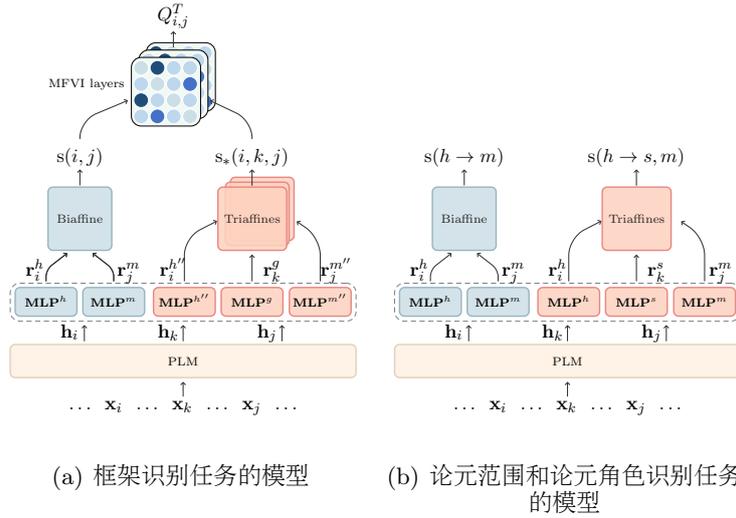


Figure 2: 模型框架

维特比来解决这个问题。该方法的具体实现可参考Zhou et al. (2022)。

如图1(a), 句子中目标词的论元可能是单个词, 也可能由多个词组成, 显式地建模论元内部结构可能对框架解析有帮助。针对论元识别的任务二和任务三 (论元范围识别和论元角色识别任务), 借鉴Zhang et al. (2021), 本文将基于片段的框架语义解析转换成一个树解析任务, 即将扁平的论元建模成树结构。具体而言, 在句子的开始添加一个“Root”头节点, 将其与句子中的目标词相连, 并用“PRD”作为它们的边标签。对于单个词的论元直接分配语义角色标签, 对于多个词的论元, 将所有潜在的论元子树作为目标词的后代, 即列举出论元内所有可能的子树, 并为其使用Eisner算法(Eisner, 2000)找出最高得分的子树, 例如图1(c), 虚线箭头所组成的树是从若干个潜在子树中找到的最高得分的子树, 将该论元片段的语义角色标签分配到目标词和子树的词头“之一”之间的边上, 最后利用抽取子树的后代来识别论元的范围, 并将目标词和词头之间的标签作为论元片段的标签。

4 模型

本节介绍了这次评测所使用的模型结构。具体而言, 我们借鉴Zhou et al. (2022), 利用框架元素与目标词之间的边来帮助目标词的框架识别任务。在给定目标词的情况下, 本文借鉴Zhang et al. (2021)的方法, 无需得知框架类别, 直接预测目标词的论元范围和角色。这两个模型都分为编码器和评分模块两部分, 它们的编码器相同, 但评分模块不同。

4.1 编码器

如图2所示, 编码器都是使用预训练模型BERT(Devlin et al., 2018)编码输入句子。给定一个句子 $\mathbf{x} = x_0, x_1, \dots, x_n$, 对于每一个词 x_i , 使用BERT最后四层输出的加权之和作为最终输入 \mathbf{h}_i :

$$\mathbf{h}_0, \mathbf{h}_1, \dots, \mathbf{h}_n = \text{BERT}(x_0, x_1, \dots, x_n) \quad (1)$$

4.2 评分模块

4.2.1 框架识别任务

边预测 : 对于每一条边 $i \rightarrow j$, 模型需要计算一个得分 logit_{ij} , 这个得分分为一阶得分和二阶得分两部分。模型用两个多层感知机 (Multilayer Perceptron, MLP) 分别获得每个词作为目标词的表示 \mathbf{r}_i^h 和作为修饰词的表示 \mathbf{r}_j^m , 然后使用一个Biaffine层(Cai et al., 2018)来获得一阶得

分, 其中 $\mathbf{W} \in \mathbb{R}^{(d+1) \times d}$:

$$\begin{aligned} \mathbf{r}_i^h; \mathbf{r}_i^m &= \text{MLP}^h(\mathbf{h}_i); \text{MLP}^m(\mathbf{h}_i) \\ s(i, j) &= \begin{bmatrix} \mathbf{r}_j^m \\ 1 \end{bmatrix}^\top \mathbf{W} \mathbf{r}_i^h \end{aligned} \quad (2)$$

对于二阶得分, 模型首先用三个MLPs分别获得每个词作为头节点 $\mathbf{r}_i^{h''}$ 、依存节点 $\mathbf{r}_i^{m''}$ 和孙子节点 \mathbf{r}_i^g 的表示。然后使用TriAffine(Zhang et al., 2020)层来获得子树三种类型的得分, 图2中 $s_*(i, k, j)$ 对应于二阶的得分, 其中 $*$ $\in \{sib, cop, grd\}$:

$$\mathbf{r}_i^{h''}; \mathbf{r}_i^{m''}; \mathbf{r}_i^g = \text{MLP}^{h''/m''/g}(\mathbf{h}_i) \quad (3)$$

$$s_{sib}(i, j, k) = \text{TriAFF}_{sib}(\mathbf{r}_i^{h''}, \mathbf{r}_j^{m''}, \mathbf{r}_k^{m''}) \quad (4)$$

$$s_{cop}(i, j, k) = \text{TriAFF}_{cop}(\mathbf{r}_i^{h''}, \mathbf{r}_j^{m''}, \mathbf{r}_k^{h''}) \quad (5)$$

$$s_{grd}(i, j, k) = \text{TriAFF}_{grd}(\mathbf{r}_i^{h''}, \mathbf{r}_j^{m''}, \mathbf{r}_k^g) \quad (6)$$

最后使用平均场变分推断(Mean Field Variational Inference, MFVI)将一阶得分 logit_{ij}^T 和二阶得分 Q_{ij}^T 进行整合:

$$\begin{aligned} \mathcal{M}_{ij}^{t-1} &= \sum_{k \neq i, j} Q_{ik}^{t-1} s_{sib}(i, j, k) \\ &\quad + Q_{kj}^{t-1} s_{cop}(i, j, k) \\ &\quad + Q_{jk}^{t-1} s_{grd}(i, j, k) \\ \text{logit}_{ij}^t &= s(i, j) + \mathcal{M}_{ij}^{t-1} \\ Q_{ij}^t &= \sigma(\text{logit}_{ij}^t) \end{aligned} \quad (7)$$

其中, $t \in [1, T]$ 是迭代的次数, \mathcal{M}_{ij} 是来自二阶得分的中间变量, 初始值 Q_{ij}^0 是通过将Sigmoid函数应用于 $s(i, j)$ 得到的结果。 T 次迭代后最终得到预测得分 logit_{ij}^T 和概率 Q_{ij}^T 。

标签预测 : 和边预测的一阶得分类似, 使用了两个额外的MLPs和一组Biaffines来获得。

$$\begin{aligned} \mathbf{r}_i^{h'}; \mathbf{r}_i^{m'} &= \text{MLP}^{h'}(\mathbf{h}_i); \text{MLP}^{m'}(\mathbf{h}_i) \\ s(i, j, \ell) &= \begin{bmatrix} \mathbf{r}_j^{m'} \\ 1 \end{bmatrix}^\top \mathbf{W}_\ell^{\text{label}} \begin{bmatrix} \mathbf{r}_i^{h'} \\ 1 \end{bmatrix} \\ p(\ell|i, j) &= \frac{\exp(s(i, j, \ell))}{\sum_{\ell' \in \mathcal{L}} \exp(s(i, j, \ell'))} \end{aligned} \quad (8)$$

其中, $s(i, j, \ell)$ 是边 (i, j) 的标签得分, $p(\ell|i, j)$ 是在所有标签上使用softmax获得的概率。

训练目标 : 模型的损失来自边和标签的预测。给定句子 X 和其正确边的图 G , C 表示 X 的全连接图, 边和标签的损失计算如下:

$$\begin{aligned} L_e(\theta) &= - \sum_{(i, j) \in G} \log Q_{ij}^T - \sum_{(i, j) \in C \setminus G} \log(1 - Q_{ij}^T) \\ L_l(\theta) &= - \sum_{(i, j) \in G} \log p(\hat{\ell}|i, j) \end{aligned} \quad (9)$$

其中, θ 是模型的参数, $C \setminus G$ 表示不正确边的集合, $\hat{\ell}$ 表示正确边的标签。最终模型的损失是这两部分的加权之和, 设置 $\lambda = 0.06$:

$$L(\theta) = \lambda L_l(\theta) + (1 - \lambda) L_e(\theta) \quad (10)$$

4.2.2 论元范围识别和论元角色识别任务

边预测：该模型的边预测也用到了二阶信息。将一棵树 t 分解为一个只有边的骨架树 y 和对应排好序的标签序列 l ，对于在骨架树中的每个头节点 \rightarrow 修饰节点对 $h \rightarrow m$ ，使用两个MLPs和一个Biaffine层来计算一阶得分，和公式2类似。然后利用三个MLPs计算每个词作为头、依存和兄弟的表示，然后用一个TriAffine层计算二阶子树的得分：

$$\begin{aligned} \mathbf{r}_i^{h/m/s} &= \text{MLP}^{h/m/s}(\mathbf{h}_i) \\ s(h \rightarrow s, m) &= \text{TriAFF}(\mathbf{r}_h^h, \mathbf{r}_m^m, \mathbf{r}_s^s) \end{aligned} \quad (11)$$

其中 s 和 m 是 h 的修饰词， s 在 m 和 h 之间。

一阶和二阶信息融合之后，我们可以获得骨架树 y 的得分，并通过一个片段约束的Tree CRF来计算它的概率，其中 $Z(\mathbf{x})$ 是归一化项：

$$\begin{aligned} s(\mathbf{x}, \mathbf{y}) &= \sum_{h \rightarrow m} s(h \rightarrow m) + \sum_{h \rightarrow s, m} s(h \rightarrow s, m) \\ P(\mathbf{y} | \mathbf{x}) &= \frac{\exp(s(\mathbf{x}, \mathbf{y}))}{Z(\mathbf{x}) \equiv \sum_{\mathbf{y}' \in Y(\mathbf{x})} \exp(s(\mathbf{x}, \mathbf{y}'))} \end{aligned} \quad (12)$$

标签预测：模型使用额外的两个MLPs和多个Biaffines来获得词对 $h \rightarrow m$ 之间的标签得分，每个树骨架对应序列的概率如下计算：

$$\begin{aligned} \mathbf{r}_i^{h'}; \mathbf{r}_i^{m'} &= \text{MLP}^{h'}(\mathbf{h}_i); \text{MLP}^{m'}(\mathbf{h}_i) \\ s(i, j, \ell) &= \begin{bmatrix} \mathbf{r}_j^{m'} \\ 1 \end{bmatrix}^\top \mathbf{W}_\ell^{\text{label}} \begin{bmatrix} \mathbf{r}_i^{h'} \\ 1 \end{bmatrix} \\ P(l | \mathbf{x}, \mathbf{y}) &= \prod_{h \xrightarrow{l} m \in t} P(l | \mathbf{x}, h \rightarrow m) \end{aligned} \quad (13)$$

训练目标：最终每个带标签的树的概率为骨架树 y 和其标签序列 l 的概率之积：

$$P(t | \mathbf{x}) = P(\mathbf{y} | \mathbf{x}) \cdot P(l | \mathbf{x}, \mathbf{y}) \quad (14)$$

在训练过程中，我们最大化每个目标词所有转换树的概率作为目标，计算如下：

$$\mathcal{L} = - \sum_p \log P(T_p | \mathbf{x}) \quad (15)$$

其中 $P(T_p | \mathbf{x})$ 可以具体化为：

$$\begin{aligned} P(T_p | \mathbf{x}) &= \sum_{t \in T_p} \underbrace{P(\mathbf{y} | \mathbf{x}) \cdot P(l | \mathbf{x}, \mathbf{y})}_{P(t | \mathbf{x})} \\ &= \frac{1}{Z(\mathbf{x})} \sum_{t \in T_p} \underbrace{\exp(s(\mathbf{x}, \mathbf{y})) \cdot P(l | \mathbf{x}, \mathbf{y})}_{\exp(s(\mathbf{x}, t))} \end{aligned} \quad (16)$$

4.3 数据增强和投票技术

数据增强是一种在机器学习和深度学习中广泛使用的技术，旨在通过对原始训练数据进行一系列随机变换或扩展来增加数据的多样性和数量(Feng et al., 2021)。本次评测中，我们通过从CoNLL09(Hajič et al., 2009)数据集中抽取一定量的数据，然后用训练好的模型进行预测，将预测好的数据扩充到训练数据集，来增加样本的多样性和数量，以增强模型的鲁棒性和泛化能力，提升模型的性能。

在机器学习和深度学习模型中，投票技术常用于集成多个模型的预测结果(Gandhi and Pandey, 2015)，来提高整体性能和准确率。本文通过使用不同种子获得多个模型的预测结果。在后处理阶段，根据多数票来决定最终的预测结果，以提高模型的性能。

#sent	Train	Dev	TestA	TestB	Frame
CFN	10000	2000	4000	4000	695

Table 2: CFN数据统计

5 实验

5.1 实验设置

官方提供的数据如表2所示，本文使用了官方提供的训练集和验证集训练模型，另外，还从CoNLL09中随机抽取与训练集不重合的10000个句子、20000个句子用于数据增强。

两个模型中的编码器都使用了bert-chinese-based¹，在框架识别任务中，迭代训练60轮，最初的学习率为5e-5,超参 $\lambda=0.06$ 。论元识别任务中的参数可以参照(Zhang et al., 2021)。

5.2 评测指标

本次评测使用主办方提供的评测指标，框架识别任务采用正确率ACC，论元范围识别和论元角色识别任务采用准确率P、召回率R和F1值，计算公式如下：

$$ACC = \frac{\text{正确识别的个数}}{\text{总数}} \quad (17)$$

$$P = \frac{(\text{gold} \cap \text{pred})}{\text{Count}(\text{pred})} \quad (18)$$

$$R = \frac{\text{Count}(\text{gold} \cap \text{pred})}{\text{Count}(\text{gold})} \quad (19)$$

$$F1 = \frac{2 * P * R}{P + R} \quad (20)$$

其中，*gold*和*pred*分别表示真实结果和预测结果，对于论元范围识别任务，*Count*(*)表示结果中的token数，而对于论元角色识别任务*Count*(*)表示计算集合元素的数量。

5.3 评测结果和分析

由于参赛队伍较多，表3只列出了本次评测任务中前五名的得分，结果保留到小数点后两位，队伍名称使用的是其参赛单位的简写。可以看到，本文的方法在两个测试集中分别取得第二和第三的成绩。在TestA中，和第一名仅相差0.75的分数，task3的结果甚至超过了第一名，但Task2上的召回率太低。在TestB中，和第二名相比，分数低了1.5，但是比第四名的分数高出2.7。

表4展示了不同方法在TestB上的结果。名称中带有“基础模型”的模型意为只使用官方提供的数据进行训练，其中“基础模型-图解析”是指将语义框架解析转换成基于词的图解析任务模型，可以看出它在论元识别任务上的总得分在全部结果中最低。“基础模型-树解析”表示论元片段内部结构的模型，可以看到它可以提升模型在论元识别上的性能。“基础模型”是指使用“基础模型-图解析”中框架识别任务和“基础模型-树解析”中论元识别任务的结果。在“基础模型-结合”基础上，我们使用了投票和数据增强方式来提升模型性能。

“基础模型-投票-3”是指基于三个不同种子(seed=1, 33, 777)获得的模型结果进行投票得到的，可以看出总分提升了2.5，说明投票策略对模型性能的提升有帮助。“数据增强-10000”是指将CoNLL09数据中随机抽取的10000条数据通过基础模型预测之后，把它们混入训练集中进行数据增强的方法。模型性能提升不太明显，反而Task2和Task3的结果都比基础模型还要低。“数据增强-20000”则是使用CoNLL09中20000个句子进行数据增强，比只使用10000个句子的分数高0.52。说明数据增强的效果和数据的数量有关。“模型增强-20000-3”是指在CoNLL09的20000条数据进行数据增强的基础上，使用不同的种子

¹<https://huggingface.co/bert-base-chinese>

Team	Task1		Task2		Task3			总分
	ACC	P	R	F1	P	R	F1	
	TestA							
SCU	73.05	89.62	86.82	88.20	60.68	61.24	60.96	72.76
SUDA	72.50	88.72	83.74	86.16	68.60	54.97	61.03	72.01
BLCU	72.90	90.09	86.61	88.31	58.77	58.98	58.88	71.91
HIT	70.05	89.41	87.67	88.53	57.47	58.40	57.93	70.75
UIR-ISC	69.70	90.20	87.90	89.04	56.61	57.07	56.84	70.36
	TestB							
SCU	73.87	90.83	82.88	86.68	59.58	57.41	58.47	71.55
BLCU	72.34	90.37	84.28	87.22	57.90	55.83	56.84	70.60
SUDA	70.42	89.44	82.08	85.60	64.14	49.57	55.92	69.18
HIT	64.36	90.22	84.89	87.47	52.37	52.28	52.36	66.48
UIR-ISC	65.14	90.41	85.60	87.94	51.31	50.39	50.85	66.26

Table 3: 在测试集A和测试集B上，参赛队伍中前五名的得分

模型名称	Task1 (ACC)	Task2 (F1)	Task3 (F1)	总分
基础模型-图解析	67.62	69.87	46.95	60.11
基础模型-树解析	-	83.54	51.66	-
基础模型	67.62	83.54	51.66	66.01
基础模型-投票-3	69.78	84.07	55.83	68.58
数据增强-10000	68.62	79.78	51.48	66.33
数据增强-20000	68.51	84.66	52.24	66.85
模型增强-20000-3	69.15	84.49	54.47	67.88
模型增强-投票-4	70.42	85.60	55.92	69.18

Table 4: 不同方法在测试集B上的结果

(seed=1, 33, 777) 获得3个模型后进行投票，但是结果没有基础模型投票的性能佳。考虑到“模型增强-20000-3”中任务2的结果高一些，所以本文使用基础模型和数据增强模型进行混合，经过多次交叉验证，我们发现，任务1在“基础模型” (seed=1, 33)、 “数据增强-10000” (seed=1) 和“数据增强-20000” (seed=1) 四个模型上进行投票的结果最好，任务2在“数据增强-20000” (seed=1, 777) 和“基础模型” (seed=1, 33) 四个模型上进行投票的结果最好。而任务3在三个模型“基础模型” (seed=777)、 “数据增强-10000” (seed=777) 和“数据增强-20000” (seed=777) 进行投票的结果最好，即“模型增强-投票-4”的结果。总体而言，数据增强和投票方法对提升模型的预测性能都有帮助，而投票方法的帮助相对较大。

6 结语

在本次CFSP2023评测任务中，我们将基于片段的语义框架解析转换成一个基于词的图解析任务，在端到端框架下用于识别句子中目标词对应的框架，解码的时候利用目标词和对应论元之间的关系来帮助目标词的框架识别任务。就论元识别任务，我们将基于片段的框架语义解析转换成树解析任务，建模论元内部的结构，以便更好地识别论元的边界。另外使用了数据增强和投票方法来提高模型的预测性能。最终取得A榜第二名，B榜第三名的成绩。

但我们的模型仍然有很大的不足，例如任务2的召回率与其他队伍相比较低。另外文中的投票技术目前只是以后处理的方式进行应用，后续可以探究不同模型的结构，对投票技术进行改进，或探索多任务学习框架，将不同模型融合在一起，以更好地支持对框架解析中三个任务的预测。

参考文献

- Collin F Baker, Michael Ellsworth, and Katrin Erk. 2007. Semeval-2007 task 19: Frame semantic structure extraction. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 99–104.
- Jiaxun Cai, Shexia He, Zuchao Li, and Hai Zhao. 2018. A full end-to-end semantic role labeler, syntactic-agnostic over syntactic-aware? In *Proceedings of the 27th International Conference on*

- Computational Linguistics*, pages 2753–2765, Santa Fe, New Mexico, USA, August. Association for Computational Linguistics.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Jason Eisner. 2000. Bilexical grammars and their cubic-time parsing algorithms. *Advances in probabilistic and other parsing technologies*, pages 29–61.
- Steven Y Feng, Varun Gangal, Jason Wei, Sarath Chandar, Soroush Vosoughi, Teruko Mitamura, and Eduard Hovy. 2021. A survey of data augmentation approaches for nlp. *arXiv preprint arXiv:2105.03075*.
- Charles J Fillmore, Christopher R Johnson, and Miriam RL Petruck. 2003. Background to framenet. *International journal of lexicography*, 16(3):235–250.
- Isha Gandhi and Mrinal Pandey. 2015. Hybrid ensemble of classifiers using voting. In *2015 international conference on green computing and Internet of Things (ICGCIoT)*, pages 399–404. IEEE.
- Yong Guan, Shaoru Guo, Ru Li, Xiaoli Li, and Hu Zhang. 2021. Integrating semantic scenario and word relations for abstractive sentence summarization. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 2522–2529.
- Shaoru Guo, Yong Guan, Ru Li, Xiaoli Li, and Hongye Tan. 2020a. Incorporating syntax and frame semantics in neural network for machine reading comprehension. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 2635–2641.
- Shaoru Guo, Ru Li, Hongye Tan, Xiaoli Li, Yong Guan, Hongyan Zhao, and Yueping Zhang. 2020b. A frame-based sentence representation for machine reading comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 891–896.
- Jan Hajič, Massimiliano Ciaramita, Richard Johansson, Daisuke Kawahara, Maria Antònia Martí, Lluís Màrquez, Adam Meyers, Joakim Nivre, Sebastian Padó, Jan Štěpánek, Pavel Straňák, Mihai Surdeanu, Nianwen Xue, and Yi Zhang. 2009. The CoNLL-2009 shared task: Syntactic and semantic dependencies in multiple languages. In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL 2009): Shared Task*, pages 1–18, Boulder, Colorado, June. Association for Computational Linguistics.
- Richard Johansson and Pierre Nugues. 2007. LTH: Semantic structure extraction using nonprojective dependency trees. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 227–230, Prague, Czech Republic, June. Association for Computational Linguistics.
- Hiroki Ouchi, Hiroyuki Shindo, and Yuji Matsumoto. 2018. A span selection model for semantic role labeling. *arXiv preprint arXiv:1810.02245*.
- Xuefeng Su, Ru Li, Xiaoli Li, Jeff Z Pan, Hu Zhang, Qinghua Chai, and Xiaoqi Han. 2021. A knowledge-guided framework for frame identification. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5230–5240.
- Zhixing Tan, Mingxuan Wang, Jun Xie, Yidong Chen, and Xiaodong Shi. 2018. Deep semantic role labeling with self-attention. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32.
- Liping You and Kaiying Liu. 2005. Building chinese framenet database. In *2005 international conference on natural language processing and knowledge engineering*, pages 301–306. IEEE.
- Yanan You, Ru Li, Xuefeng Su, Zhichao Yan, Minshuai Sun, and Chao Wang. 2022. Chinese frame disambiguation method based on GCN and gate mechanism. In *Proceedings of CCL 2022*, pages 201–210, Nanchang, China, October. Chinese Information Processing Society of China.
- Yu Zhang, Zhenghua Li, and Min Zhang. 2020. Efficient second-order treecrf for neural dependency parsing. *arXiv preprint arXiv:2005.00975*.
- Yu Zhang, Qingrong Xia, Shilin Zhou, Yong Jiang, Guohong Fu, and Min Zhang. 2021. Semantic role labeling as dependency parsing: Exploring latent tree structures inside arguments. *arXiv preprint arXiv:2110.06865*.

- Hongyan Zhao, Ru Li, Xiaoli Li, and Hongye Tan. 2020. CFSRE: context-aware based on frame-semantics for distantly supervised relation extraction. *Knowledge-Based Systems*, 210:106480.
- Shilin Zhou, Qingrong Xia, Zhenghua Li, Yu Zhang, Yu Hong, and Min Zhang. 2022. Fast and accurate end-to-end span-based semantic role labeling as word-based graph parsing. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 4160–4171.
- 宋衡, 曹存根, 王亚, and 王石. 2022. 一种细粒度的汉语语义角色标注数据集的构建方法. *中文信息学报*, 36(12):52–66,73.
- 屠寒非, 李茹, 王智强, and 周铁峰. 2016. 一种基于主动学习的框架元素标注. *中文信息学报*, 30(4):44–55.
- 张力文, 王瑞波, 李茹, and 张晟. 2017. 基于词分布式表征的汉语框架排歧模型. *中文信息学报*, 31(6):50–57.
- 李国臣, 张立凡, 李茹, 刘海静, and 石佼. 2013. 基于词元语义特征的汉语框架排歧研究. *中文信息学报*, 27(4):44–51.
- 李济洪, 王瑞波, 王蔚林, and 李国臣. 2010. 汉语框架语义角色的自动标注. *软件学报*, 21(4):597–611.
- 李济洪, 高亚慧, 王瑞波, and 李国臣. 2011. 汉语框架自动识别中的歧义消解. Ph.D. thesis.
- 王晓晖, 李茹, 王智强, 柴清华, and 韩孝奇. 2022. 基于self-attention 的句法感知汉语框架语义角色标注. *中文信息学报*, 36(10):38–44.
- 王蔚林. 2010. 基于最大熵模型的汉语框架语义角色自动标注. Ph.D. thesis, 太原: 山西大学硕士学位论文.
- 赵红燕, 李茹, 张晟, and 张力文. 2016. 基于dnn 的汉语框架识别研究. *中文信息学报*, 30(6):75–83.