

基于 Tri-Training 的事件关系分类方法研究*

丁思远, 洪宇, 朱珊珊, 姚建民, 朱巧明

(苏州大学 江苏省计算机信息处理重点实验室, 江苏 苏州 215006)

摘要: 事件关系分类是一项研究事件之间存在何种逻辑关系的自然语言处理技术。该文针对事件关系分类任务中训练语料不足的问题, 提出了基于Tri-Training的事件关系分类方法。该方法首先根据已标注的语料训练三个不同的分类器, 以多数投票的方式从未标注集中抽取置信度较高的样本对训练集进行扩充, 然后利用新的训练集重新训练分类器, 如此反复迭代, 不断完善分类模型, 最终达到提升事件关系分类性能的目的。实验结果表明, 以 F_1 值为评价标准, 基于Tri-Training的事件关系分类方法在四大类事件关系上的分类性能为64.36%。

关键词: 事件关系; 框架语义; 半监督学习; Tri-Training

Research of Event Relation Classification based on Tri-Training

DING Siyuan, HONG Yu, ZHU Shanshan, YAO Jianmin, ZHU Qiaoming

(Provincial Key Laboratory of Computer Information Processing Technology Soochow University, Suzhou, Jiangsu 215006, China)

Abstract: As one of the natural language processing technologies, event relation detection aims at exploring logical relationship between pairwise events. To solve the problem of lacking enough training data in event relation detection tasks, we proposed a novel approach based on Tri-Training to augment the training corpus. The approach firstly used labeled training data to learn three different classifiers, and then exploited majority voting to expand training corpus with higher confidence samples, kept optimizing the model and eventually reached the purpose of improving the performance of event relation classification. The experiments results show that when compared to other methods, the Tri-Training method achieves 64.3% F_1 -score over four general semantic relations.

Key words: Event relation; Frame; Semi-supervised learning; Tri-Training

1 引言

事件关系是指事件与事件之间相互依存和关联, 且具有客观性和规律性的逻辑形式。事件关系检测以事件为主体元素, 通过分析事件文本的结构信息及语义特征, 挖掘事件之间深层的逻辑关系, 进而辅助事件的衍生、发展以及信息的推理与预测。事件关系检测包含事件关系识别和事件关系判定两个部分。本文只研究事件关系判定任务, 将事件关系判定看作分类问题。

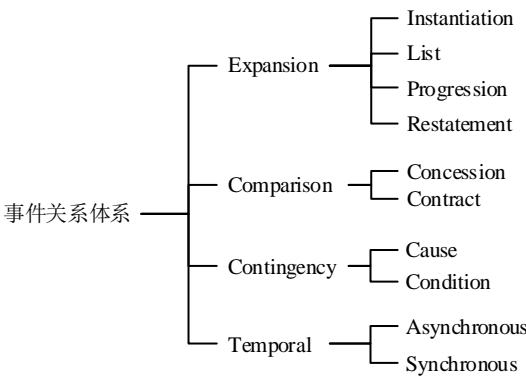


图1 事件关系体系结构图

* 收稿日期:

定稿日期:

基金项目: 国家自然科学基金 (61003152, 61272259, 61272260)

目前，针对事件关系分类的研究相对较少，仍缺乏完善的事件关系体系定义。而与事件关系较为类似的篇章关系研究已然获得更多关注，相关任务定义和研究方法也得到一定程度的验证和实践。基于此，杨等^[1]通过分析事件关系检测与篇章关系检测的异同点，建立一套基于篇章关系分析的事件关系体系。本文延续杨等^[1]提出的事件关系检测研究方法，以篇章关系为桥梁，借助篇章关系的任务定义，辅助指导事件关系分类。事件关系类型分成两层（如图 1 所示）：第一层包含四种主要关系：Expansion（扩展关系）、Contingency（偶然关系）、Comparison（对比关系）和时序关系（Temporal）；第二层关系针对第一层的事件关系进行细分，共包含十种关系。本文主要研究第一层的四种主要事件关系。图 2 为包含四种事件关系类型的文本实例，其中 E1-E4 为四个相关的事件片段（EventArg）。

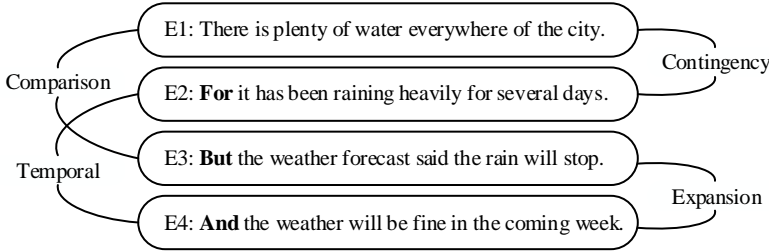


图 2 四种事件关系实例图

此外，事件关系研究尚缺乏权威的标注语言学资源，为解决该问题，杨等^[1]标注了部分事件关系样例，然而该样例的数量仍不足以解决事件关系研究中语料不足的问题。针对该问题，本文提出一种基于 Tri-Training 的事件关系分类方法，该方法旨在利用少量的已标注事件关系样本训练三个不同的分类器，通过分类器之间的协同合作，从大规模未标注数据资源中挖掘错误率较小的未标注数据，进而扩充训练样本，以提高分类器的分类性能。具体实现过程中，本文首先从大规模语言学资源 GIGAWORD 纽约时报新闻语料（LDC2003T05）中挖掘包含 Golden 连接词¹的“事件对”，借助框架语义知识库（FrameNet）^[2]过滤不具备事件特征的样本，然后将 Golden 连接词所对应的事件关系作为未标注事件对的先验关系类别，以此构造未标记事件关系数据集。其次，基于 Tri-Training 方法在大规模人工标注事件关系数据集上训练三个不同的分类器，在协同训练过程中，每个分类器新增的训练样本均由其余两个分类器协作投票提供，以此对少量人工标注的事件关系数据集进行扩充，最终达到提高事件关系分类性能的目的。图 3 为事件关系分类任务框架，输入为待测事件关系片段对 EventArg1 和 EventArg2，输出则为具体的事件关系类型。

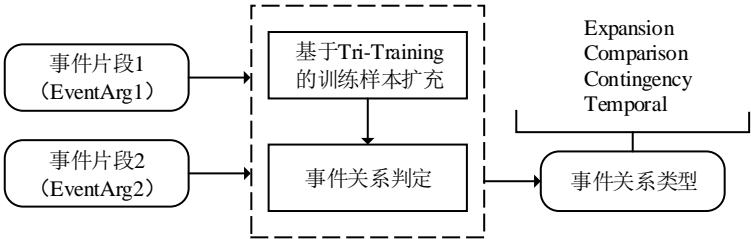


图 3 事件关系分类任务框架

本文的组织结构如下：第 2 节介绍相关工作；第 3 节介绍框架语义知识库及未标注样本数据的获取；第 4 节给出基于 Tri-Training 的事件关系样本扩展方法；第 5 节给出实验结果与分析；最后总结全文并展望未来工作。

¹ Golden 连接词：指向某一事件关系概率大于 96% 的连接词，如连接词“now”指向时序关系的概率为 100%。

2 相关工作

由于事件关系检测的研究刚刚起步,仍缺少权威的事件关系体系,相关研究工作主要集中在某种特定事件关系类型的判定^{[3][4][5]},主流的研究方法大致可以分为模板匹配法和元素分析法。

• 模版匹配法

事件关系检测的主要方法之一是借助事件特征的模式匹配法。例如,利用事件触发词(Trigger)的关系模式匹配,根据人工定义的模板,对文本中符合模板的事件关系进行抽取。Chklovski^[4]等首先定义六种时序关系:“*similarity*”(“相似”关系),“*strength*”(“加强”关系),“*antonymy*”(“相反”关系),“*enablement*”(“支持”关系),“*happens*”(“发生”关系)和“*before*”(“前”关系),再利用人工收集的 LSP (Lexical-Syntactic Pattern, 即词-句匹配模板)抽取包含这六种时序关系的“事件对”,并将抽取的结果形成称为“*VerbOcean*”的知识库。人工定义的事件关系模板往往受数量限制,造成关系检测的低召回率问题。Pantel^[5]提出利用 Espresso 算法进行自动模板的构建,该算法根据给定的少量关系实例,通过机器学习方法对现有模板进行迭代扩展,在一定程度上提高了模板匹配方法的召回率。

• 元素分析法

以事件元素为线索的研究大都继承了 Harris^[6]的分布假设。Harris 分布假设指出,处在同一上下文环境中的词语往往具有相同或相似的含义。Lin^[7]提出了一种结合 Harris 分布假设和建立依存树思想的无监督方法,称为 DIRT 算法。该算法将所有事件构造成依存树形式,树中的每条路径表示一个事件,路径的节点表示事件中的词语,若两条路径的词语完全相同,则这两条路径所表示的事件相同或者相似。

• 其它

杨等^[1]在 2014 年的工作中,通过分析事件关系检测与篇章关系检的异同点,建立了一套基于篇章关系分析的事件关系检测体系,同时提出了基于跨场景推理的事件关系检测方法,该方法通过构建事件场景向量,计算待测事件对与已挖掘到的包含连接词的事件对之间的事件场景向量相似度,将最相似的事件对所指派的事件关系作为该待测事件对的最终关系类别。但是这种通过计算相似度的方法,不仅要求语料规模足够大。而且,仅仅根据 Golden 连接词的映射获取事件关系的方法很难保证准确性。

针对事件关系分类中语料不足的问题,本文提出一种基于 Tri-Training 事件关系分类方法。该方法根据已有的少量人工标注的事件关系数据集,利用 Tri-Training 方法,从挖掘到的大量未标注数据集中选择高置信度的样本扩充训练语料,继而提高事件关系分类性能。

3 框架语义知识库及未标注样本数据获取

3.1 框架语义知识库

框架语义知识库(FrameNet)是由美国加州大学伯克利分校构建的基于框架语义学(Frame Semantics)的词汇资源,知识库通过框架(即词语背后隐藏的概念结构和语义等信息)描述单词的释义,对词语意义和句法结构研究提供一种理论框架。该语言学资源共标注 1028 个框架语义,其中能够触发事件属性的框架语义共 540 个。

在框架语义知识库中,对句子的框架语义标注是一种类似于“谓词-论元”结构的“目标词-框架语义”(Target-Frame)结构。每个句子可能包含一个或多个“目标词-框架语义”结构。从图 4 中可以看出,在两个事件实例 Event1 和 Event2 中, Amy 和 lady 对应的框架语

义为 *Person*, *code* 与 *password* 对应 *Message* 框架语义, 而 *anesthetic* 和 *wine* 对应 *Drag* 框架语义。虽然两个事件具有不同的语义描述, 但却存在相同的框架语义。这就说明, 通过框架语义, 能够有效的挖掘语义相近的事件, 辅助事件关系推理。

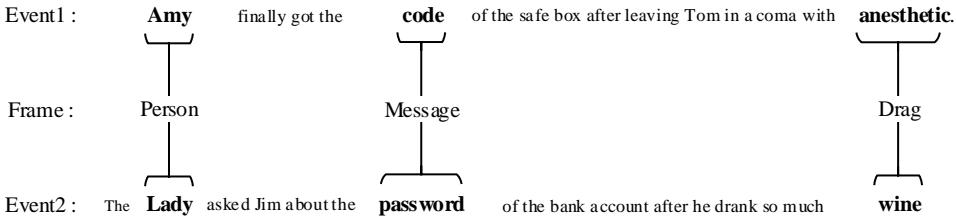


图 4 Event1 与 Event2 中相同的框架语义

3.2 未标注样本集合数据获取

本文使用的外部数据资源为 GIGAWORD 纽约时报新闻语料 (LDC2003T05), 共包含 4,111,240 篇文本。本文从该语料中抽取符合下述要求的未标注事件关系样例:

- 条件 1. 切分后的文本由事件片段 1 (EventArg1) 和事件片段 2 (EventArg2) 组成;
- 条件 2. 事件片段 2 中包含 Golden 连接词, 借此将 Golden 连接词对应的事件关系作为未标注事件片段对的先验知识, 通过后续方法进一步提高标注准确率;
- 条件 3. 事件片段对中须包含能触发事件的框架语义, 从而保证挖掘到的未标注样本是一个事件样本。

例 1 EventArg1: *This is a great award here in Sweden,*

EventArg2: [*so*] *we will remember it all our lives.*

例 1 为符合挖掘规则的未标注事件关系样例, 该样例为一个事件关系对, 且包含两个具有独立语义信息的文字片段, 两者之间通过连接词 “so” 进行关联。图 5 给出挖掘到的以 Golden 连接词为关系类别先验知识的各事件关系类型样本分布比例。从图中可以看出, 四种事件关系类型中, Expansion 与 Comparison 关系类型分布比例较大, 而 Temporal 关系的分布比例较小。

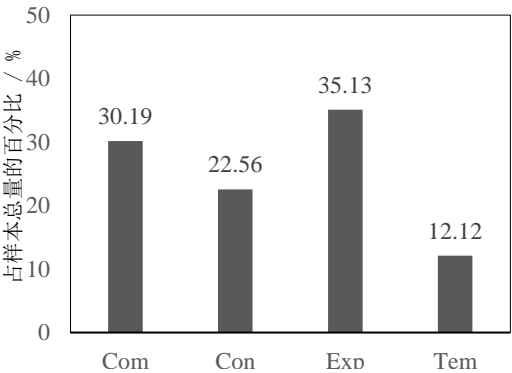


图 5 以 Golden 连接词对应的关系作为先验知识的各关系类型样本分布

4 基于Tri-Training的事件关系样本扩展方法

图 6 给出本文基于 Tri-Training 的事件关系分类方法流程图。该方法以 GIGAWORD 语言学资源 (LDC2003T05) 作为外部数据资源, 从中挖掘包含 Golden 连接词的事件关系对, 将 Golden 连接词对应的关系类别作为先验知识。在此基础上, 抽取框架语义、事件触发词、触发词词性和事件类别四种特征生成未标注事件关系数据集。在对少量人工标注的事件关系

数据集抽取相同的特征后生成已标注集。本文将其中一半的已标注集作为训练集，另一半作为测试集。然后利用 Tri-Training 的方法从未标注集中选取较高置信度的事件对加入到训练集中，不断迭代训练模型，直到未标注集为空或未标注集大小不再变化为止。最后根据学习好的模型对测试集进行分类，生成最终的分类结果。

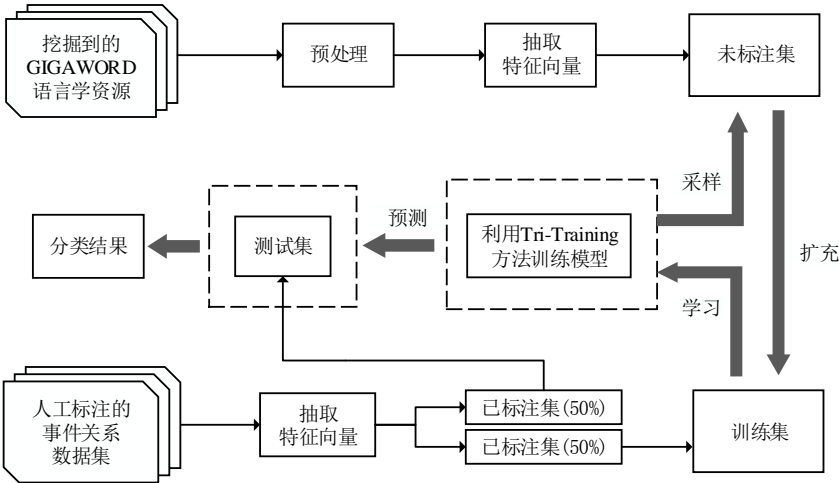


图 6 基于 Tri-Training 的事件关系分类方法流程图

Tri-Training 算法旨在对少量已标注数据集进行重复采样，训练三个不同的分类器 X , Y , Z ，通过三个分类器之间的一致性判断，对大量未标注样本数据进行分类标注。详细而言，在训练过程中，任意一个分类器（例如分类器 X ）所获得的新标记样本由另外两个分类器（分类器 Y 和分类器 Z ）协作判定，如果两个分类器对同一个未标记样本 x 给出相同的分类标签 L , $Y(x) = Z(x)$ 那么就将该未标记样本加入当前分类器 X 中。该方法的主要优点是无需使用不同的分类方法，就可以实现大量未标注样本的标注任务^[8]。由于本文的已标注事件关系集很小，可重复采样效果并不是很好。因此，本文将 S. Goldman 和 Y. Zhou^[9] 的 Co-Training 思想应用到 Tri-Training 方法上，即在同一份已标注集上，使用三种不同算法训练三个不同类型的分类器。表 1 为基于 Tri-Training 的事件关系样本扩展流程。

表 1 基于 Tri-Training 方法的事件关系样本扩展流程

初始化	
准备实验数据，将已标注的数据集按照一定比例分为训练集 L 和测试集 T ，对于未标注数据，根据连接词标上先验关系，记作 U	
训练	
1	根据训练集 L 分别训练三个分类器 $C1$ 、 $C2$ 和 $C3$
2	分别用分类器 $C1$ 、 $C2$ 、 $C3$ 对未标注数据集 U 进行预测，得到预测结果 $R1$, $R2$,
$R3$	
3	对 $R1$ 、 $R2$ 和 $R3$ 采取多数投票的方式得到带有标记的样本，若该样本类别与先验关系类别一致，则将其加入到第三个分类器的训练集中
4	重复 1-3 操作，直到 U 为空，或者 U 的大小不在变化训练停止

值得注意的是，在第 3 步中，对于一个未标注事件关系对，在两个分类器给出的分类结果相同的情况下，还需满足与先验关系类别相同的条件，该事件对才会被加入到第三个分类器的训练集中。例如，某未标注事件对先验关系类别为 Expansion，而分类器 $C1$ 与 $C2$ 给出的类别均为 Expansion，那么该事件对就会被加入到已标注集合 L 中，并且将类别标记为 Expansion，否则就不加入 L 中。

5 实验与分析

5.1 实验数据集

本文选取杨等^[1]在 2014 年标注的事件关系语料作为已标注数据集, 该语料是以 FrameNet-1.5 中的新闻语料作为标注集, 由两名具有一定领域知识的标注人员, 对每篇新闻中描述的事件以及事件关系类型进行标注, 标注结果的 Kappa 值为 0.78。该语料共包含 968 个事件对, 各关系分布情况如表 2 所示。本文主要研究第一层四种关系类型, 包括 Comparison (Com)、Contingency (Con)、Expansion (Exp) 和 Temporal (Tem)。其中, 本文将 468 个关系样例作为训练集, 剩余的 500 个样例作为测试集, 具体实验数据分布情况如表 3 所示。

表 2 已标注的关系类型的分布

关系类别	Com	Con	Exp	Tem
数量/个	126	209	525	108

表 3 训练集与测试集

关系类别		Com	Con	Exp	Tem
训练集	正例/个	76	109	275	58
	负例/个	70	100	250	50
测试集	正例/个	50	100	250	50
	负例/个	50	100	250	50
未标注集	正例/个	10000	8000	10000	5000
	负例/个	10000	8000	10000	5000

本文并未将挖掘到的全部外部资源作为实验的未标注集, 而是根据四种关系的分布比例, 从各关系中随机等分布的选取一定数目的事件对, 作为最终的未标注集。未标注集中各关系的分布如表 3 所示。所采用的评价方法是标准的准确率 (Precision)、召回率 (Recall) 和 F_1 值。

5.2 分类特征

本文选用框架语义 (Frame)、事件触发词 (Trigger)、触发词词性 (Trigger POS) 和事件类别 (Event Type) 作为分类特征。下面分别对这四种特征予以介绍。

框架语义 (Frame): 即本文在 3.1 节中介绍的框架语义, 其中具有事件属性的框架语义共 540 个。本文选择具有事件属性的框架语义作为特征之一。

事件触发词 (Trigger): 事件触发词直接引发事件的产生, 是决定事件类别的一个重要特征。根据 ACE (Automatic Content Extraction) 2005^[10]中对触发词的定义, 本文选择触发词作为分类特征之一。

触发词词性 (Trigger POS): 根据事件触发词抽取对应的词性 (POS), 并且将其作为分类特征。

事件类别 (Event Type): ACE2005 定义了 8 种事件类别以及 33 种子类别。本文将事件类别与事件子类别统称为事件类别, 并将此作为一种分类特征。

本文通过 SEMAFOR² 工具抽取框架语义特征, 根据 Li 等^[11]提出的基于结构化感知机的联合事件抽取方法分析一个事件的触发词、触发词词性及事件类别。例如, 对于事件片段 *EventArg: Fifteen people died in the traffic accident.* 所抽取的各特征如表 4 所示。

² <http://www.ark.cs.cmu.edu/SEMAFOR>

表 4 事件片段 EventArg 特征抽取结果

特征名称	特征内容
框架语义	Death, Coincidence
触发词	died
触发词词性	VBD
事件类型	Die

5.3 实验系统

为了验证本文所提方法的有效性, 本文设计如下 6 个实验系统。详细介绍如下,

Baseline1: 不使用 Golden 连接词所对应的事件关系作为先验知识, 按比例随机从未标注集中抽取事件对作为最终的扩充样本, 训练 LIBSVM³ 分类器对测试集进行预测; 实验共进行 10 次, 最终性能取平均值。

Baseline2: 以标有先验事件关系类别的未标注集作为训练语料, 不使用本文提出的基于 Tri-Training 的事件关系分类方法, 训练四个二元 LIBSVM 分类器, 对测试集进行预测。

YangSystem: 重现杨^[1]的方法, 即利用“框架语义对”构建事件场景, 然后计算待测事件对与已挖掘到的事件对之间的事件场景相似度, 将拥有最大相似度的事件对中连接词所对应的事件关系作为待测事件对的最终事件关系。

SelfTraining: 基于 Self-Training^[12]方法的对比实验。在实验过程中, 本文通过融合 Golden 连接词对应的先验关系类别与分类器预测结果, 对待标记样本的事件关系类型进行标记, 进而扩充训练集。例如, 在使用 Self-Training 方法从未标注集中选择扩充样本时, 若分类器对某事件对预测的类别为 Expansion, 那么只有该事件对的先验关系类别为 Expansion 时, 该事件对才会被加入到已标注集中。实验选择的分类器为 LIBSVM。

CoTraining: 基于 Co-Training^[9]的半监督学习方法的对比实验。同样的, 在实验过程中, 本文通过融合 Golden 连接词对应的先验关系类别与分类器预测结果, 对用以扩充训练集的待标记样本的事件关系类型进行标记。例如, 在使用 Co-Training 方法的过程中, 对于某一待标记样本, 当且仅当两个分类器给出的预测结果与先验事件关系类别均相同时, 该样本才会被加入对方的已标注集中。实验选择的分类器为 LIBSVM 和 MaxEnt⁴。

TriTraining: 本文所提出的基于 Tri-Training 的事件关系分类方法实验系统。分类器选择为 LIBSVM、MaxEnt 与 NaiveBayes⁵。

5.4 实验结果与分析

表 5 给出各系统分别在四种事件关系上的分类性能。从表 5 中可以看出, 采用随机采样的 Baseline1 系统在四类关系上的分类性能均在 50% 左右, 这与理论的情况也基本吻合。而 Baseline2 相较于 Baseline1, 在四大事件关系类别分类性能上都有着明显的提升, 这一现象说明, 利用连接词所指向的关系类别作为未标注语料的先验事件关系, 有利于提高各关系的分类性能的。因此, 本文在使用 Self-Training、Co-Training 和 Tri-Training 三种半监督学习方法的实验中, 均以 Golden 连接词对应的事件关系作为该事件对的先验关系类别。

实验结果表明, 杨等^[1]所提出的方法在 Expansion 关系上达到较好的性能 (60.74%), 然而在其他三种事件关系上的结果都不是很理想。原因在于, 杨等^[1]所采用的方法是通过计算事件场景的相似度, 找出包含连接词的未标注样本中与待测事件对最相似的事件对, 以该事件对中连接词所属的关系类别作为该待测事件对的事件关系类别。但是这种通过计算相似度的方法, 不仅要求语料规模足够大。而且, 直接将 Golden 连接词所对应的事件关系作为

³ <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

⁴ <http://github.com/lzhang10/maxent>

⁵ <http://mallet.cs.umass.edu/>

最终的关系类别很难保证准确性。

表 5 各系统在四种事件关系上的分类性能

实验系统	实验结果 (P R F ₁) %											
	Comparison			Contingency			Expansion			Temporal		
Baseline1	48.44	45.60	46.76	50.42	57.60	53.81	52.80	63.95	57.80	44.45	40.27	43.39
Baseline2	68.57	48.00	56.47	62.07	54.00	57.75	55.06	74.00	63.14	63.16	48.00	54.55
YangSystem	83.33	10.00	17.86	65.85	27.00	38.30	58.80	62.80	60.74	46.15	24.00	31.58
SelfTraining	65.12	56.00	60.22	64.29	54.00	58.70	62.14	69.60	65.66	62.79	54.00	58.06
CoTraining	63.46	66.00	64.71	63.95	55.00	59.14	62.87	68.40	65.52	54.90	56.00	55.45
TriTraining	71.15	74.00	72.55	67.03	61.00	63.87	60.87	67.20	63.88	57.41	62.00	59.62

本文采用 Self-training 和 Co-Training 方法的两个对比实验在事件关系的分类任务上相比于 Baseline2 都有着很好的分类性能提升。说明将半监督学习方法运用到事件关系分类任务中，对提升分类性能有着明显的作用。本文提出的基于 Tri-Training 的事件关系分类方法，在四大类事件关系上均获得了不同程度的性能提升。图 7 显示了 Tri-Training 方法在各事件关系上学习曲线，可以发现随着迭代次数的增加，F₁ 值呈总体上升趋势，并最终达到稳定。由于迭代过程中会引入噪音，因此 F₁ 值有一定的波动。

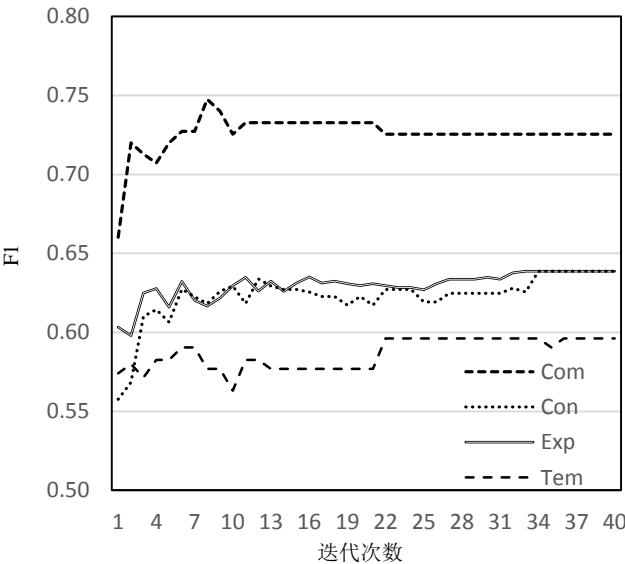


图 7 Tri-Training 方法在四种事件关系上的学习曲线

对比 Baseline2, TriTraining 系统在这四种事件关系分类性能上分别提高了 16.08%、6.12%、0.74% 和 5.07%。而且相比于 SelfTraining 与 CoTraining，该系统在 Comparison、Contingency 和 Temporal 三个关系的分类性能上分别提高了 12.33%、5.17% 和 1.56%，与 7.84%、4.73% 和 4.17%，均达到最高的 F₁ 值。从表 6 也可以看出，Tri-Training 方法在四种关系上的综合分类性能也是最高的，F₁ 值为 64.36%。这充分说明了本文将 Tri-Training 方法与 Golden 连接词所对应的先验关系类别相结合来辅助事件关系分类的合理性与有效性。

表 6 各系统的综合分类性能

实验系统	P (%)	R (%)	F ₁ (%)
Baseline1	51.12	57.87	54.29
Baseline2	57.86	63.78	60.68
YangSystem	59.12	44.67	50.89
SelfTraining	62.89	62.89	62.89
CoTraining	62.26	63.78	63.01
TriTraining	62.79	66.00	64.36

6 总结

本文主要研究基于 Tri-Training 的半监督学习方法在事件关系分类任务中的应用。该方法针对事件关系分类任务中训练语料不足的现象, 利用 Tri-Training 方法在少量已有人工标注的事件关系数据集上训练三个不同的分类器, 然后从挖掘到的包含 Golden 连接词的大量未标注数据集中, 通过简单投票的方式对训练语料进行扩充, 继而提高事件关系分类性能。由实验结果可以看出, 相比于其他几种方法, 基于 Tri-Training 的事件关系分类方法, 在各事件关系分类性能上均有不同程度的提高。且该方法在四种关系上的综合分类性能上也达到了最高的 F₁ 值, 为 64.36%。这就说明, 在事件关系标注语料很少的情况下, 使用基于 Tri-Training 的半监督学习的方法扩充事件关系训练语料能够有效地提高事件关系分类的性能。在下一步工作中, 我们将尝试使用更多的分类特征, 如事件上下文和词典信息等特征, 以此进一步的提高事件关系分类性能。

参考文献

- [1] 杨雪蓉, 洪宇, 陈亚东等. 基于跨场景推理的事件关系检测方法[J]//中文信息学报, 2014, 28(5): 206-214.
- [2] Fillmore C J, Johnson C R, Petruck M R L. Background to framenet[J]//International journal of lexicography, 2003, 16(3): 235-250.
- [3] Hashimoto C, Torisawa K, Kloetzer J, et al. Toward future scenario generation: Extracting event causality exploiting semantic relation, context, and association features[C]//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, USA, June. Association for Computational Linguistics. 2014.
- [4] Chklovski T, Pantel P. Global path-based refinement of noisy graphs applied to verb semantics[M]//Natural Language Processing-IJCNLP 2005. Springer Berlin Heidelberg, 2005: 792-803.
- [5] Pantel P, Pennacchiotti M. Espresso: Leveraging generic patterns for automatically harvesting semantic relations[C]//Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2006: 113-120.
- [6] Harris Z S. Mathematical structures of language[J]. 1968.
- [7] Lin D, Pantel P. DIRT@ SBT@ discovery of inference rules from text[C]//Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2001: 323-328.
- [8] Zhou Z H, Li M. Tri-training: Exploiting unlabeled data using three classifiers[J]//Knowledge and Data Engineering, IEEE Transactions on, 2005, 17(11): 1529-1541.
- [9] Goldman S, Zhou Y. Enhancing supervised learning with unlabeled data[C]//ICML. 2000: 327-334.
- [10] Linguistic Data Consortium. ACE (Automatic Content Extraction) Chinese Annotation Guidelines for Events[J]. 2009-09-08. Version 5.5.1. <http://www ldc.upenn.edu/Projects/ACE>.

- [11] Li Q, Ji H, Huang L. Joint Event Extraction via Structured Prediction with Global Features[C]// Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, 2013: 73-82.
- [12] Zhu X. Semi-supervised learning literature survey[J]. 2005.

作者简介：丁思远（1992—），男，硕士研究生，主要研究领域为自然语言处理，事件关系检测。Email: dsy.ever@gmail.com; 朱珊珊（1992—），女，硕士研究生，主要研究领域为自然语言处理，篇章分析。Email: zhushanshan063@gmail.com; 洪宇（1978—），男，博士，副教授，主要研究领域为信息检索，话题检测与跟踪，篇章分析。通讯作者。Email: tianxianer@gmail.com; 姚建民（1971—），男，博士，副教授，主要研究领域为机器翻译。Email: jmyao@szkj.gov.cn; 朱巧明（1963—），男，博士，教授，博士生导师，主要研究领域为中文信息处理。Email: qmzhu@suda.edu.cn。

