

文章编号: 1003-0077(2020)09-0089-08

基于生成对抗网络的控辩焦点识别

杨 亮¹, 周逢清¹, 张 琨², 毛国庆³, 易 斌¹, 林鸿飞¹

(1. 大连理工大学 计算机科学与技术学院, 辽宁 大连 116024;

2. 北京计算机技术及应用研究所, 北京 100854; 3. 北京国双科技有限公司, 北京 100083)

摘 要: 近年来,随着深度学习技术的不断发展,自然语言处理作为人工智能的一个重要分支,在许多垂直领域有了广泛的应用,如司法、教育、医疗等。在司法领域的庭审过程中,控辩双方往往围绕案件的争议焦点持有不同观点,而该焦点也是影响案件最终判决和量刑的关键所在,该文旨在识别并生成电子卷宗中的控辩焦点。由于控辩焦点的构成大多依赖对案情文本的分析概括,受此启发该文尝试将文本摘要的思想迁移到该任务中,结合生成对抗网络构建控辩焦点的生成模型,进而获得案件的控辩焦点。在裁判文书网的真实司法数据的基础上,开展了相关的实验。实验结果显示,所提出的模型对控辩焦点的识别精度有了一定幅度的提升。因此,该文对检察机关办案人员的庭前预案及案件审理有着一定的辅助作用与应用价值。

关键词: 控辩焦点; 对抗生成网络; 智慧司法

中图分类号: TP391

文献标识码: A

Argument Recognition Based on Generative Adversarial Networks

YANG Liang¹, ZHOU Fengqing¹, ZHANG Li², MAO Guoqing³, YI Bin¹, LIN Hongfei¹

(1. College of Computer Science and Technology, Dalian University of Technology, Dalian, Liaoning 116024, China;

2. Beijing Institute of Computer Technology and Application, Beijing 100854, China;

3. Beijing GridSum Technology Co., Ltd., Beijing 100083, China)

Abstract: In the process of trial in the field of justice, the prosecution and the defense often hold different views around the argument of the case, which is also the key factors to the final judgment of the case. To identify the arguments in the cases, this paper introduce the text summarization model since the composition of the argument mostly depends on the analysis and summary of the case text. We construct the generation model of the argument by combining the generative adversarial network, and then obtain the argument of the case. Experimented on the real judicial data obtained from the website of China Judgements Online, the results show that the proposed model improves the accuracy in the task of argument recognition. This method can be applied as an auxiliary role in the pre-court preplan and trial of the case for procuratorial personnel in real application.

Keywords: argument of prosecution and defense; generative adversarial networks; smart justice

0 引言

随着科技的进步和发展,人工智能开始在学术界和工业界引领一股新的技术革命,在诸多垂直领域落地和应用,如司法、教育和医疗等,给人们的工作和生活带来极大便利。我国司法机关目前存在案

多人少、资源分配不均的问题,解决方案之一便是转变思维方式,通过人工智能和大数据等手段,进行审判方式的信息化改革,减少所需的人力成本,提高审判流程的科技含量。2013 年裁判文书平台——“裁判文书网”上线,除特殊情形外,我国最高人民法院生效裁判文书全部在“裁判文书网”上公布。据统计,截至 2017 年 8 月 23 日,“裁判文书网”总访问量

收稿日期: 2019-09-17 定稿日期: 2020-01-28

基金项目: 国家重点研发计划(2018YFC0830604);国家自然科学基金(61702080,61632011);中央高校基本科研业务费专项资金(DUT19RC(4)016);中国博士后基金(2018M631788)

突破 100 亿次,公开裁判文书超过 3 247 万篇^[1],这一平台为我国“智慧司法”建设提供了有力的数据支持。自 2015 年 7 月最高人民法院正式提出“智慧法院”概念以来,司法领域就开始依托大数据、深度学习等技术让司法系统更加客观、便捷与智能。智慧司法体系的建立,不但能够为法官判案提供智能辅助,提高判案质量与效率,促进公平正义,也能够让人民群众享受便捷化的诉讼服务。

控辩焦点是在法院庭审阶段基于控告和辩护方围绕证据与案情所持有不同观点,由法官最终总结出的影响案件的争议点。控辩焦点是案件的关键所在,对案件的判决和量刑起着重要作用。但是从一个案件中总结出控辩焦点并不容易。这需要对案情相关的大量证据及案例文书进行分析,进而总结归纳出控辩焦点。这个过程需要大量的专业知识以及行业经验。

近年来,关于对话领域及辩论文本中的观点或论点的研究逐渐增多,如 Hua 等人^[2]将维基百科中的外部知识结合至 Seq2Seq 模型中,在论坛讨论的数据集上生成高质量、论点明确的陈述。Durmus Esin 等人^[3]构建了一个辩论数据集,并对角色进行建模,提高了预测辩论胜方任务的性能。Haddadem 等人^[4]在政治演说数据集中提取丰富的特征,在论点句判断以及论证成分识别两个任务上都取得了不错的结果。Jo 等人^[5]提出了一个神经系统架构来模拟辩论对话,对意见持有者和挑战者的论点的相互作用进行建模,预测论点是否被成功地改变。

上述研究中多数工作针对对话式类型的数据集,如政治演说或者辩论文本等,而本文工作中的案例文书为非对话式文本,主要包含案情描述以及庭审过程中的记录等,属于司法领域的数据。此外,上述研究的实验部分,多数是进行分类任务,如判断辩论的胜方,或者判断一个句子中是否包含论点或者论证过程等,而本文工作属于生成任务,需要根据案例文书生成与之相匹配的控辩焦点。因此,本文将通过深度学习模型构建控辩焦点识别模型,辅助司法人员总结归纳控辩焦点,这对缓解司法资源缺乏、分配不均等有重要意义。本文的研究主要针对给定案例文书,如何得到文书对应的控辩焦点。从技术的角度来看,需要搭建一个“控辩焦点识别模型”,其输入是案例文书,输出则是对应的控辩焦点。根据问题的输入与输出特点,可以将该任务建模为一个文本摘要的问题。

1 相关工作

文本自动摘要主要为了帮助读者更快地获得有效信息,以避免无谓的资源浪费。20 世纪 50 年代末,Luhn 等人^[6]提出了文本自动摘要技术,也就是利用计算机自动地将文本或者是文本集合转换成简短摘要的一种信息压缩技术。其利用词频进行文本自动摘要,选择权重最高的句子作为摘要输出。到了 20 世纪 90 年代,机器学习方法在自然语言处理领域得到了广泛应用,基于统计的文本自动摘要技术也逐渐兴起。这类方法是将文本自动摘要问题简化为一个二分类问题或者是回归问题,前者判断一个句子是否是摘要,后者给句子打分,分数最高的句子即是摘要。Kupiec 等人^[7]提出了用贝叶斯方法进行文本自动摘要,结合回归方法,生成长度为原文 25% 的摘要,抽取准确率有一定的提升。Kanapala 等人^[8]提出了 GSA 模型,结合句子的长度、位置、TF-ISF(Term Frequency-Inverse Sentence Frequency)等特征,在 FIRE-2014 法律专题数据集上实现了良好的摘要结果。以上文本摘要方法均属于抽取式方法,其特点是将文章切分为句子集合,通过一定的规则,计算句子集合中每个句子的权重,抽取权重最高的句子形成摘要。这种自动摘要方式比较成熟,且易于实现,其中 TextRank 排序算法^[9]以简洁、高效的特点被业界广泛运用。

文本摘要的另外一种方法是生成式方法,这种方法主要依靠深度神经网络结构实现。2014 年由谷歌大脑团队提出的 Seq2Seq 方法^[10]在机器翻译、文本生成等领域成果颇丰,开启了自然语言处理领域中端到端网络的火热研究。这种方法包含两个深度学习网络,一个负责对原文进行编码生成中间向量,另一个负责对中间向量进行解码。文本自动摘要可以看作是文本到文本的映射,可以采用这种端到端的模型。生成式的文本自动摘要主要使用这种模型,通过深度学习模型训练大量的数据,对文本进行编码、解码,产生摘要内容,摘要的内容不限于原文内容。Facebook 的 Rush 等人^[11]使用卷积神经网络(CNN)作为编码器,使用上下文相关的注意力前馈神经网络作为解码器,取得了良好的结果。而 Chopra 等人^[12]使用了循环神经网络作为解码器,提高了摘要的准确度。Hochreiter 等人^[13]提出的长短期记忆模型(LSTM)在一定程度上解决了基本循环神经网络无法进行长期记忆的问题。Song 等

人^[14]提出了一种基于 LSTM 和 CNN 的文本自动摘要框架,提高了生成摘要的质量。

上文中提到的两种类型的文本摘要方法中,前者是传统的抽取式摘要方法,易于实现且不受限制,常用于新闻摘要,但是该方法仅使用文本表层的数据信息,如词频、句子位置、句子长度等,没有使用句子语义之间的联系,导致其结果具有一定的局限性。后者为生成式方法,使用最新的深度学习方法,是目前公认较好的文本自动摘要方法,它可以自动学习文本中的高阶特征,从深层的语义角度进行摘要生成,不过生成的摘要多是一句话,在生成长文本任务中会出现反向传播困难、模型难以收敛的问题,所以适合生成较短的摘要。

近年来,一种新的对抗训练理论——生成对抗网络(generative adversarial network, GAN)^[15]引发了许多学者的研究兴趣。生成对抗网络分为两个网络——判别网络和生成网络,生成网络致力于生成更好的数据来混淆判别网络,而判别网络致力于最大化地区分真实数据和生成网络生成的数据,两个网络相互对抗训练,最后达到平衡。理想化的结果是判别网络无法将真实数据和生成数据区分开来。生成对抗网络在图像处理领域取得了不错的成绩,包括图像生成、超分辨率、风格转换等。生成对抗网络也能很好地迁移到文本生成相关的任务上,但是应用于文本生成这类任务时主要存在梯度无法传递到生成器的缺点。Kusner 等人^[16]使用 Gumbel-softmax 采样来近似离散采样操作,可以使网络输出转化为一个类似于独热编码的形式,避免了参数难以更新的问题。Yu 等人^[17]提出的 SeqGAN 将生成对抗网络和强化学习结合起来解决梯度难以传递的问题,使用蒙特卡洛搜索的方法解决部分生成序列的难以评价的问题。

我国当前司法资源相对缺乏,为响应国家对智慧司法建设的号召,庭审控辩焦点识别任务具有很强的现实意义。本文结合生成对抗网络对控辩焦点识别任务进行探究。

2 方法

本文基于对抗生成网络构建控辩焦点识别模型,模型结构图如图 1 所示。下面用形式化语言介绍模型的各个部分。

2.1 输入层

给定案例文书 $\mathbf{x} = [x_1, x_2, \dots, x_n]$ 和对应的控

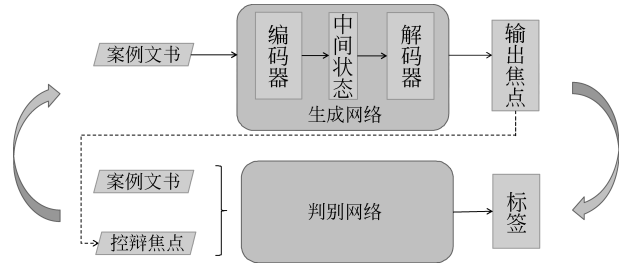


图 1 模型结构图

辩焦点 $\mathbf{a} = [a_1, a_2, \dots, a_m]$, 其中 n 和 m 分别为案例文书和控辩焦点的文本长度。输入层首先通过词向量矩阵 $\mathbf{W} \in \mathbf{R}^{V \times d}$ 将词语序列映射成词向量序列的形式, 其中 V 为词表长度, d 为词嵌入维度, 如输入案例文书映射为 $\mathbf{x}^e = [x_1^e, x_2^e, \dots, x_n^e]$, 其中 $\mathbf{x}^e \in \mathbf{R}^{n \times d}$ 。

2.2 判别网络

判别网络由两部分卷积神经网络 D_φ 构成, 其中 φ 是网络参数。第一个卷积神经网络作用于控辩焦点 \mathbf{a}^e , 主要抽取控辩焦点的深层特征, 首先对输入进行卷积操作, 如式(1)所示。

$$c_i^a = f(\mathbf{W} \cdot \mathbf{a}_{i:i+h-1}^e + b) \quad (1)$$

$$\mathbf{c}^a = [c_1^a, c_2^a, \dots, c_{m-h+1}^a] \quad (2)$$

其中, $\mathbf{a}_{i:i+h-1}^e$ 表示输入第 i 到第 $i+h-1$ 个词向量的连接, h 为卷积核的宽度, \mathbf{W} 和 b 为参数, f 为非线性激活函数。进而将每个步骤的卷积结果 c_i^a 拼接得到 \mathbf{c}^a , 如式(2)所示, 接着进行最大池化操作, 生成焦点对应的特征映射, 如式(3)所示。

$$\tilde{\mathbf{c}}^a = \max(\mathbf{c}^a) \quad (3)$$

对应地, 第二个卷积网络作用于案例文书 \mathbf{x}^e , 生成特征映射 $\tilde{\mathbf{c}}^x$ 。最后将 $\tilde{\mathbf{c}}^a$ 与 $\tilde{\mathbf{c}}^x$ 拼接后通过全连接层, 如式(4)所示, 判断案例文书和控辩焦点是否对应。

$$y = \text{softmax}(\mathbf{W} \cdot ([\tilde{\mathbf{c}}^x, \tilde{\mathbf{c}}^a]) + b) \quad (4)$$

最后通过定义交叉熵损失函数并结合梯度下降的方式对模型参数进行更新。

2.3 生成网络

生成网络 G_θ 是由 RNN 网络构成的 encoder-decoder 架构。其中 θ 为网络参数, 案例文书 \mathbf{x}^e 输入 encoder 网络后被编码成中间向量。首先计算每个时刻的隐层状态向量, 如式(5)所示。

$$h_t = f(\mathbf{U}\mathbf{x}_t^e + \mathbf{W}h_{t-1} + b) \quad (5)$$

其中, h_t 和 h_{t-1} 分别表示第 t 和第 $t-1$ 时刻的

隐层状态; \mathbf{x}_t^e 表示案例文书的第 t 个词对应的向量; \mathbf{U} 和 \mathbf{W} 为权重向量, \mathbf{b} 为偏置向量。接着将隐层状态向量转换成中间语义向量 \mathbf{c} , 如式(6)所示。

$$\mathbf{c} = f(h_1, h_2, \dots, h_n) \quad (6)$$

其中, f 为非线性激活函数。最后通过 decoder 将语义向量输出对应结果, 如式(7)所示。

$$y_t = f(\mathbf{U}\mathbf{c} + \mathbf{W}y_{t-1} + \mathbf{b}) \quad (7)$$

其中, y_{t-1} 和 y_t 分别表示第 $t-1$ 和第 t 时刻的输出。

2.4 对抗训练

判别器和生成器的对抗训练过程采用 Yu 等人^[17]提出的结合策略梯度的训练方法, 以解决梯度无法传递的问题。结合强化学习, 通过固定判别器变量传递至生成器的梯度, 计算过程如式(8)所示。

$$\nabla J(\theta) \approx \frac{1}{T} \sum_{t=1}^T E[\nabla_{\theta} \log G_{\theta}(y_t | Y_{1:t-1}) Q_{D_{\varphi}}^{G_{\theta}}(Y_{1:t-1}, y_t)] \quad (8)$$

其中, T 为生成序列的长度, t 表示第 t 个时刻, $Y_{1:t-1}$ 表示从第 1 时刻到第 $t-1$ 时刻已生成的序列, $Q_{D_{\varphi}}^{G_{\theta}}(Y_{1:t-1}, y_t)$ 表示在已生成 $Y_{1:t-1}$ 序列的情况下, 生成器第 t 个时刻生成 y_t 这个词语所带来的动作值函数。生成器参数 θ 更新如式(9)所示, 其中 α 为学习率。

$$\theta \leftarrow \theta + \alpha \nabla J(\theta) \quad (9)$$

3 实验

3.1 数据集

本任务的数据集来自裁判文书网, 共有 1 275 条标注数据(由北京国双科技有限公司负责标注)。其中 503 条数据的案由为“走私、贩卖、运输、制造毒品罪”, 500 条诈骗罪以及 272 条寻衅滋事罪, 数据集的具体构成情况如表 1 所示。

表 1 数据集构成情况

罪名	数量
走私、贩卖、运输、制造毒品罪	503
诈骗罪	500
寻衅滋事罪	272

每一条数据主要由案例文书的描述和标准控辩焦点构成, 部分案例有关键词信息, 关键词信息是对控辩焦点的进一步凝炼和抽象。

表 2 是一条关于未成年人贩卖毒品案的数据样例, 案件的焦点在于法院如何判断是否要对未成年人适用缓刑。

表 2 数据样例

案例文书	控辩焦点
<p>谢某某贩卖毒品案</p> <p>(一) 基本案情</p> <p>2012 年 10 月至 11 月间, 被告人谢某某(1995 年 6 月 1 日生), 伙同赵某(另案处理)以零包贩卖的方式, 多次向他人贩卖毒品氯胺酮。案发后, 公安机关抓获被告人谢某某及赵某, 并从被告人谢某某随身携带的物品中依法搜查出毒品氯胺酮 3.356 克, 随后在被告人谢某某及赵某住处依法搜查出毒品氯胺酮 65.604 克。</p> <p>(二) 裁判结果</p> <p>一审法院于 2013 年 8 月 22 日作出刑事判决, 认为被告人伙同他人多次贩卖毒品氯胺酮, 情节严重, 其行为已触犯刑律, 构成贩卖毒品罪。被告人谢某某犯罪时不满十八周岁, 依法应当从轻或减轻处罚。决定判处被告人谢某某有期徒刑二年八个月, 缓刑四年, 并处罚金人民币五千元。</p>	法院如何判断是否对未成年人适用缓刑

3.2 对比实验

本文选取以下几种方法进行对比。

(1) **CaseName**: 选取案件名称作为案情描述的摘要信息, 由于在辩论领域中论题通常会正是反方的争论焦点, 类似地, 案件名称中可能会涉及控辩焦点信息。

(2) **TextRank**^[9]: 在案情描述文本中以句子为单位计算 TF-IDF(Term Frequency-Inverse Document Frequency)指标, 并以此计算句子间的相似度矩阵, 通过 PageRank 算法得到权重最高的句子作为抽取焦点。

(3) **GSA**: 参照 Kanapda 等人^[8]的工作, 选取案例文书中的句子长度、位置、及 TF-ISF(Term Frequency-Inverse Sentence Frequency)等特征, 计算得到权重最高的句子作为焦点。

此外, 本文依据控辩焦点识别任务的特点, 设计了两种启发式抽取式方法作为对比实验。

3.2.1 基于词频统计的抽取式模型

首先对数据集中的标准控辩焦点分词并统计词频, 根据词频对词进行降序排序, 选取前 100 个词

语-词频对构成字典。然后对输入的某个案例文本按句子切分得到句子集合。具体流程如算法 1 所示。

算法 1 基于词频统计的抽取算法

输入：词语-词频字典 words_dict；案例文书的句子列表 sentences；中文分词函数 get_words()。

1. 初始化候选句 candidate_sentence 为空
2. 初始化最大权重 max_weight 为 0
3. 遍历 sentences 中的每个句子
4. 初始化当前句权重 weight 为 0
5. 利用 get_words() 对当前句分词得到 words
6. 遍历当前句的词列表 words 中的每个词
7. 如果当前词在字典 words_dict 中
8. 则 weight 的值加上当前词的词频
9. 完成 words 遍历后得到当前句子的总权重
10. 通过 $\text{weight} = \text{weight}^2 / \text{当前句词语个数}$
11. 对当前句子权重进行长度归一化
12. 若 weight 不小于 max_weight
13. 则将当前句子作为 candidate_sentence
14. 并将当前句子权重作为 max_weight

输出：案例文书中权重最大的句子 candidate_sentence 以及权重值 max_weight

3.2.2 基于关键词的抽取式模型

将数据集中存在的案件关键词抽取出来构成关键词集合。然后对输入的某个案例文本按句子切分得到句子集合。具体流程如算法 2 所示。

算法 2 基于关键词的抽取算法

输入：关键词集合 key_words_set；案例文书的句子列表 sentences；中文分词函数 get_words()。

1. 初始化候选句 candidate_sentence 为空
2. 初始化最大权重 max_weight 为 0
3. 遍历 sentences 中的每个句子
4. 初始化当前句权重 weight 为 0
5. 利用 get_words() 对当前句分词得到 words
6. 遍历当前句的词列表 words 中的每个词
7. 如果当前词在集合 key_words_set 中
8. 则 weight 的值加 1
9. 完成 words 遍历后得到当前句子的总权重
10. 通过 $\text{weight} = \text{weight}^2 / \text{当前句词语个数}$
11. 对当前句子权重进行长度归一化
12. 若 weight 不小于 max_weight
13. 则将当前句子作为 candidate_sentence
14. 并将当前句子权重作为 max_weight

输出：案例文书中权重最大的句子 candidate_sentence 以及权重值 max_weight

3.3 实验设置

基于对抗生成网络构建的控辩焦点识别模型中

的超参数设置为：判别器的输入序列长度统一为 2 038(2 000 为案例文书；38 为控辩焦点)，卷积核宽度为 1~10、15 和 20。生成器的 encoder 的隐层节点数为 128，输入序列长度为 2 000，decoder 的隐层节点数为 256，输出长度为 38。词向量维度统一为 64。

对抗训练过程采用 Yu 等人^[17]提出的 SeqGAN 模型中使用的策略梯度方法。首先生成器和判别器先进行预训练，然后再结合对抗训练过程进一步提升模型的控辩焦点识别效果。具体的算法流程如算法 3 所示。

算法 3 基于对抗生成网络的控辩焦点识别

1. 初始化对抗网络中判别器 D_φ 和生成器 G_θ 的参数 φ, θ
2. 初始化超参数 m, n, k, p, q
3. 遍历 n 轮数据集
4. 预训练判别器，通过交叉熵损失函数更新参数 φ
5. 遍历 m 轮数据集
6. 预训练生成器，用极大似然估计更新参数 θ
7. 循环 k 轮对抗训练
8. 遍历 p 轮数据集
9. 由当前的 G_θ 产生样本
10. 使用 roll out 策略进行采样
11. 将采样得到的样本送入 D_φ 计算得分
12. 结合策略梯度更新生成器参数 θ
13. 遍历 q 轮数据集
14. 通过交叉熵损失函数更新参数 φ

在本文的实验预训练过程中，负样本通过随机采样与案例文书不匹配的控辩焦点进行构造。预训练阶段判别器的预训练轮数 $n = 30$ ，损失函数为交叉熵，生成器预训练轮数 $m = 50$ ，优化目标为 MLE。在对抗训练阶段，对抗训练轮数 $k = 30$ ，每一轮中，生成器先训练 $p = 1$ 轮，将当前生成器生成控辩焦点匹配上案例文书作为负样本，再以交叉熵损失函数训练判别器 $q = 10$ 轮。在对抗训练过程中判别器的训练轮数大于生成器是防止由于判别器训练不充分，导致梯度指导生成器在错误的方向上生成控辩焦点的情况。

3.4 评价指标

评价指标主要选取 ROUGE-N 指标，该指标是评价文本摘要任务中与参考摘要的 n-gram 相似度，计算方法如式(10)所示。

$$\text{ROUGH-N} = \frac{\sum_{S \in \{\text{REF}\}} \sum_{n\text{-gram} \in S} \text{Count}_{\text{match}}(n\text{-gram})}{\sum_{S \in \{\text{REF}\}} \sum_{n\text{-gram} \in S} \text{Count}(n\text{-gram})} \quad (10)$$

其中, n -gram 表示句子 S 中的 n 个连续词语, $\{\text{REF}\}$ 为标准摘要集合, $\text{Count}_{\text{match}}(n\text{-gram})$ 为生成摘要和标准摘要中相同的 n -gram 个数, $\text{Count}(n\text{-gram})$ 为标准摘要中的 n -gram 个数。由于本实验中生成的摘要最大长度为 38, 因此主要采用 ROUGE-1 和 ROUGE-2 两个指标。

另外一个评价指标为 TF-IDF 相似度, 将由算法抽取或者生成的摘要和参考摘要通过 TF-IDF 转化为向量后, 计算二者的 cosine 相似度, 计算方法如式(11)所示。

$$\text{TF-IDF} = \frac{V_1 \odot V_2}{\|V_1\| \cdot \|V_2\|} \quad (11)$$

其中, V_1 和 V_2 表示生成摘要和参考摘要的向量表示, \odot 表示向量的内积操作, $\|\cdot\|$ 表示向量的 L_2 范数。以上两种类型评价指标的值越接近 1 就说明生成的摘要与参考摘要的相似性越高, 在本文中任务中也就对应着生成控辩焦点识别越接近案例文书标准焦点。

3.5 实验结果与分析

表 3 为各个算法的实验结果, WF 为基于词频统计的抽取式模型, KW 为基于关键词的抽取式模型, GAN 为本文提出的基于对抗生成网络的控辩焦点识别模型。从表 3 中的结果可以看出: GSA 方法的结果数值比较低, 主要原因是所使用的 TF-ISF 特征没有对句子的长度进行正则化, 所以模型倾向于给较长的句子赋予较大权重; CaseName 方法仅选择案例的标题作为焦点, 这种方式显然过于粗糙, 并且同一种案由的案例标题格式相似, 导致泛化性不足, 因此各指标的结果数据也都比较低; TextRank 方法通过句子间的相似度矩阵, 迭代算法找出与各句子关联性最大的句子, 比较符合控辩焦点的特征, 因此其 TF-IDF 指标明显高于 CaseName 和 GSA 方法。

表 3 各算法的实验结果

方法	ROUGE-1	ROUGE-2	TF-IDF
CaseName	0.149	0.090	0.060
TextRank ^[9]	0.180	0.056	0.131
GSA ^[8]	0.118	0.022	0.088
WF	0.359	0.103	0.153
KW	0.336	0.090	0.144
GAN	0.505	0.458	0.430

此外, 根据控辩焦点识别任务的特点, 本文构建的两种启发式抽取式模型 WF、KW, 以及生成式模型 GAN, 在各项指标上都优于上述三种方法。同时, WF 和 KW 方法在 ROUGE-N 和 TF-IDF 指标上低于生成式方法 GAN, 原因可能有如下两点: 首先, 启发式抽取规则不够全面, 导致没有选中在案例文书中更加关键的句子; 其次, 是抽取出的句子取自原文, 而案件的控辩焦点通常是专业人员在通读案例文书后, 用自己的话归纳总结出的焦点内容, 这是抽取式这类启发式算法所无法学习到的, 因此 TF-IDF 指标较低。同时, 抽取的句子叙述方式与案件参考控辩焦点存在一定的差异性, 导致 ROUGE 指标也相对比较低。

表 4 为各算法在数据集上输出的控辩焦点示例。前两种抽取式为原案例文书的句子, 虽然通过各自算法找到了权重最高的句子, 上述的两个例子也抽取到与参考摘要较相关的句子, 但是句子表达方式以及词语用法存在一定差异, 并且抽取出来的句子存在过长、信息冗余的情况。在抽取式方法的更一般的结果中, 从案例文书中抽取出的句子偏离焦点的情况占多数, 导致整体的 ROUGE-N 指标和 TF-IDF 指标都比较低。

表 4 各算法控辩率焦点输出样例

方法	参考焦点	输出焦点
WF	在犯罪人实施运输毒品行为之前的其他行为, 如何定性	运输毒品犯罪的实行行为系被告人运输行为, 在被告人实施运输毒品事实之前的行为, 应认定为运输毒品制造条件, 系犯罪预备
KW	法院如何判断是否对未成年人适用缓刑	被告人谢某某犯罪时不满十八周岁, 依法应当从轻或减轻处罚
GAN	如何认定通过电信诈骗活动, 骗取他人钱财, 数额巨大的行为	如何认定通过他人诈骗的

通过本文构建的基于对抗生成网络的控辩焦点模型生成的焦点中, 可以看到模型能够从案例文书中学习到的信息汇集和归纳生成能力, 根据案例文书的内容生成一些关键信息, 并且句式表达也和参考焦点较为一致。但是通过生成对抗网络模型生成的焦点存在语言流畅性不足, 且信息量不够丰富等缺点。

4 结论

本文将文本自动摘要任务的基本思想迁移到控辩焦点识别任务中。在传统的 Seq2Seq 模型的基础上结合对抗训练的方式, 构建了控辩焦点生成模型。同时构造了两种启发式的控辩焦点抽取算法作为对比。在由裁判文书网构建的数据集上的实验结果表明, 本文提出的基于生成对抗网络的控辩焦点识别模型在 ROUGE 和 TF-IDF 指标上都优于启发式抽取的方法, 在一定程度上提高了从案例文书中提取控辩焦点的能力, 在司法领域中具有现实应用价值。但是该模型仍存在生成焦点信息量不够充分、流畅性欠缺等不足, 在未来的工作中将尝试改善生成器的架构, 或者结合不同的对抗训练方式和架构如 WGAN^[18], LeakGAN^[19] 等, 来进一步提升控辩焦点识别能力。此外, 适合控辩焦点识别研究任务的数据格式应该是庭审过程中控告方和辩护方的辩论对话序列, 但是在目前司法领域的研究中, 多数工作所使用的数据收集自中国裁判文书网, 仅包含案情描述以及判决量刑等信息。因此, 受限于数据内容和形式, 本文采用文本摘要任务的思想来研究控辩焦点识别的问题。随着我国司法领域信息化建设的不断深入, 庭审公开网也有部分庭审视频放出, 但仍缺少在庭审过程中高质量的控辩双方对话文本数据, 且在收集方面仍然存在一定的困难, 如音频转文本、方言标准化、背景音去噪等, 未来的工作会尝试以此为出发点, 进一步展开相关研究。

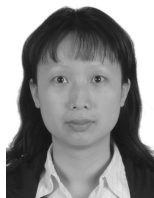
参考文献

- [1] 新华网. 中国裁判文书网总访问量突破 100 亿次[N/OL]. [2017-08-24]. http://www.xinhuanet.com/2017-08/24/c_1121538482.htm.
- [2] Hua X, Wang L. Neural argument generation augmented with externally retrieved evidence[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, 2018: 219-230.
- [3] Durmus Esin, Cardie Claire. A corpus for modeling user and language effects in argumentation on online debating[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019: 602-607.
- [4] Shohreh Haddadan, Elena Cabrio, Serena Villata. Yes, we can! Mining arguments in 50 years of US presidential campaign debates[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019: 4684-4690.
- [5] Jo Y, Poddar S, Jeon B, et al. Attentive interaction model: Modeling changes in view in argumentation[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2018(1): 103-116.
- [6] Luhn H P. The automatic creation of literature abstracts[J]. IBM Journal of Research and Development, 1958, 2(2): 159-165.
- [7] Kupiec J, Pedersen J, Chen F. A trainable document summarizer[C]//Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Seattle, Washington, United States, 1995: 68-73.
- [8] Kanapala A, Jannu S, Pamula R. Summarization of legal judgments using gravitational search algorithm[J]. Neural Computing and Applications, 2019, 31(12): 8631-8639.
- [9] Mihalcea R, Tarau P. TextRank: Bringing order into text[C]//Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. Barcelona, Spain: Association for Computational Linguistic, 2004: 404-411.
- [10] Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks[C]//Proceedings of Advances in Neural Information Processing Systems. Montreal, Canada, 2014: 3104-3112.
- [11] Rush A M, Chopra S, Weston J. A neural attention model for abstractive sentence summarization[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal: Association for Computational Linguistic, 2015: 379-389.
- [12] Chopra S, Auli M, Rush A M. Abstractive sentence summarization with attentive recurrent neural networks[C]//Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016: 93-98.
- [13] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [14] Song S, Huang H, Ruan T. Abstractive text summarization using LSTM-CNN based deep learning[J].

- Multimedia Tools and Applications, 2019, 78(1): 857-875.
- [15] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Proceedings of the International Conference on Neural Information Processing Systems. Montreal, Canada : MIT Press, 2014: 2672-2680.
- [16] Kusner M J, Hernández-Lobato J M. Gans for sequences of discrete elements with the gumbel-softmax distribution [J]. arXiv preprint arXiv: 1611.04051, 2016.
- [17] Yu L, Zhang W, Wang J, et al. SeqGAN: Sequence generative adversarial nets with policy gradient[C]//Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco : AAAI Press, 2017: 2852-2858.
- [18] Arjovsky, Martin, Soumith Chintal, et al. Wasserstein gan. [J]. arXiv preprint arXiv: 1701.07875, 2017.
- [19] Guo, Jiaxian, et al. Long text generation via adversarial Training with leaked information.[J] arXiv preprint arXiv: 1709.08624, 2017.



杨亮(1986—),博士,讲师,主要研究领域为自然语言处理、文本情感分析。
E-mail: liang@dlut.edu.cn



张琍(1973—),硕士,高级工程师,主要研究领域为司法数据分析。
E-mail: zhangliw1023@sina.com



周逢清(1995—),硕士研究生,主要研究领域为自然语言处理、文本情感分析。
E-mail: zhoulengqing@mail.dlut.edu.cn