如何写一篇合格的NLP论文

刘知远 ✦
自然语言处理、机器学习、深度学习（Deep Learning）话题的优秀回答者

已关注

盛源车、jiangfeng、忆臻、Zewei Chu、zack 等 851 人赞同了该文章

**说明：本文欢迎转载，请注明作者、出处与原文链接。**

前几天刚过完ACL 2019投稿季，给不少同学的论文提供了修改建议。其中很多论文，特别是初学者的论文的问题都很相似。一想到未来还要给更多新同学重复这些话，决定索性把这些建议总结出来，不仅以后能少费一番唇舌，说不定还能帮助更多同学。于是就有了这篇短文。

本文题目取"合格"的论文，而不是优美的论文，或精彩的论文。一个原因是，我自知英文水平特别是词汇有限，从未写过自认精彩或优美的论文，并无资格提供这方面的建议。另一个原因是，下面会讲到，学术论文的关键目标并非辞藻优美而是清晰准确，我在这方面还积累了不少经验。凭借这些经验，相信"辞达已矣"不难，"言之有文"则各凭本事吧。

实际上，同组的刘洋老师对NLP学术论文写作做过非常全面而精彩的报告 [1]，强烈推荐所有NLP同学都仔细阅读这份报告，相信会让你少走不少科研的弯路。而本文可以看做对这个报告的脚注或补充。

论文在NLP学术研究中的意义

NLP是一门重视实践和应用的领域，创新成果可以是新的算法、任务、应用、数据、发现等，务求一个"新"字，其影响力则取决于它对该领域发展的推动作用。如下图所示，学术研究是一项系统工程，包括多个环节，共同完成对"创新"的追求：问题务求挑战，模型务求创新，实现务求准确，实验务求深入。
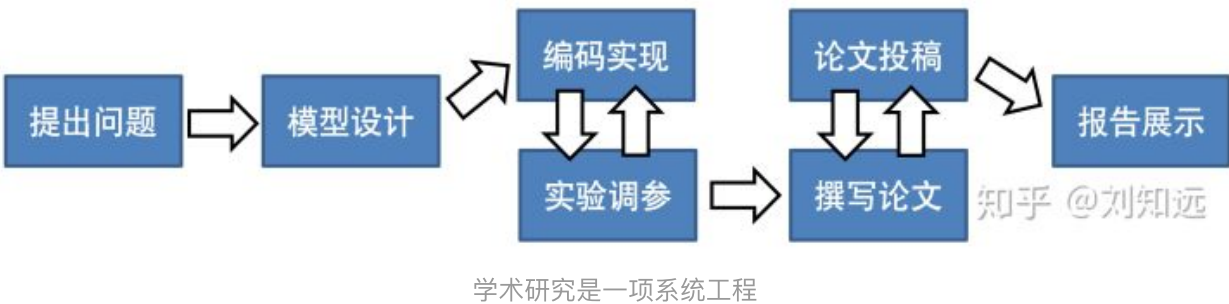
学术研究是一项系统工程

在这个系统工程中，论文的作用则是，向学术界同行清晰准确地描述成果的创新点、技术思路、算法细节和验证结果。明白这一点，才能正确的对待论文写作：一项乏善可陈的工作，很难通过写作变得众星捧月；一项充满创新的成果，却有可能因为糟糕的写作而无法向审稿人准确传递重要价值所在，延误成果发表。

一篇NLP论文的典型结构

NLP学术会议（甚至包括期刊）论文已经形成比较固定的结构。绝大部分论文由以下六大部分构成：摘要（Abstract）、介绍（Introduction）、相关工作（Related Work）、方法（Method）、实验（Experiment）、结论（Conclusion）。少数论文会根据创新成果形式不同而略有不同，例如提出新数据集的论文，可能会把Method部分调整为Dataset的标注与分析，但不影响论文整体构成。每个部分作用不同：

- 摘要：用100-200词简介研究任务与挑战、解决思路与方法、实验效果与结论。

- 相关工作：用0.5-1页左右篇幅介绍研究任务的相关工作，说明本文工作与已有工作的异同。
- 方法：用2-3页篇幅介绍本文提出的方法模型细节。
- 实验：用2-3页篇幅介绍验证本文方法有效性的实验设置、数据集合、实验结果、分析讨论等。
- 结论：简单总结本文主要工作，展望未来研究方向。

乍看这样每篇论文显得死板，实际上这正凸显了学术论文的真正意义，不追求在形式上给读者带来意外，而将读者注意力集中在论文介绍的研究成果上。

如前所说，论文的作用是向学术界同行清晰准确地描述成果的创新点、技术思路、算法细节和验证结果。由于学术界的**同行评审**制度，贯穿全文的线索和目标就是要论证这份工作的**创新价值**，每个部分都要各司其职为这个目标而服务。为了实现这个目标，需要作者特别注意以下几点：

（1）**学会换位思考**。要始终站在审稿人或读者的角度审视论文，思考如何更清晰地表达。这是初学者最容易忽视的问题：作为研究成果的亲历者，论文作者掌握所有细节，如果不多加留意，写作中就会出现新概念没有被明确定义就被使用等情况，很多描述和分析缺少逻辑衔接。对作者而言，这些省去的东西并不影响他对这些文字的理解；但对并不了解这份工作的读者而言，这无疑是一场噩梦，因为他们并没有作者脑中的那套背景信息。因此，写作时要时时留神，读者读这句时能否理解，所需要的背景知识前文是否已经介绍。

（2）**注意逻辑严谨**。严谨是学术论文的底色，从引用格式、公式符号到谋章造句，虽不至于美国法学期刊的Bluebook那么变态，都力求风格统一，行文严谨。引用、公式、拼写等方面都容易学，初学者更需要注意行文严谨，力求全文从章节、段落、句子等不同级别都逻辑严密，争取做到没有一句话没来由，没有一句话没呼应：

- 章节层面，Introduciton提到已有方法面临的几个挑战，就要对应本文提出的几个创新思路，对应Method中的几个具体算法，对应Experiment中的几个实验验证。
- 段落和句子层面，段间要注意照应，是并列、递进、转折还是总分关系，需要谋划妥当，要有相应句子或副词衔接。段内各句，有总有分，中心思想句和围绕论述句分工协作。

除了整体结构上的建议外，每个部分也各有定式，下面按各部分提供一些写作建议，同时用我们最近发表的一篇ACL 2018论文 [2] 作为例子。

Abstract和Introduction怎么写

Abstract可以看做对Introduction的提要，所以我们先介绍Introduction的写法，然后再说如何写Abstract。Introduction是对整个工作的全面介绍，是决定一篇论文能否被录用的关键。一般Introduction这么写：起手介绍**研究任务**和意义；随后简介面向这个任务的**已有方法**；接着说明已有方法面临的**关键挑战**；针对这些挑战，本文提出什么**创新思路**和具体方法；最后介绍**实验结果**证明本文提出方法的有效性。这几个部分各挡一面，同时又有严密的内在逻辑。每个部分也各有章法，下面分别介绍对各部分的建议：

（1）**研究任务**。介绍本文的研究任务及其在该研究领域的重要价值和意义。如果是领域公认的重要任务的话，则可以不用详细论述其研究价值/意义；如果是新提出的研究任务，则需要花费比较多篇幅论证该任务的价值。如下所示论文[2]的第1段集中说明阅读理解研究任务。

> Reading comprehension, which aims to answer questions about a document, has recently become a major focus of NLP research. Many reading comprehension systems (Chen et al., 2016; Dhingra et al., 2017a; Cui et al., 2017; Shen et al., 2017; Wang et al., 2017) have been proposed and achieved promising results since their multi-layer architectures and attention mechanisms allow them to reason for the question. To some extent, reading comprehension has shown the ability of recent neural models for reading, processing, and comprehending natural language text.

（2）**已有方法**。从研究任务递进一步，介绍这个任务的已有代表方法。如下所示论文[2]的第2段，开始介绍DS-QA。需要注意，这个已有方法需要是目前最好、最具代表性的，也是本文工作准备改进的。所谓站在巨人的肩膀上，一篇值得发表的论文需要找到那个最高的巨人。

> Despite their success, existing reading comprehension systems rely on pre-identified relevant texts, which do not always exist in real-world question answering (QA) scenarios. Hence, reading comprehension technique cannot be directly applied to the task of open domain QA. In recent years, researchers attempt to answer open-domain questions with a large-scale unlabeled corpus. Chen et al. (2017) propose a distantly supervised open-domain question answering (DS-QA) system which uses information retrieval technique

（3）**面临挑战**。已有方法一定仍然存在某些不足或挑战，才需要进一步研究改进。因此，需要总结已有方法面临的挑战。这是Introduction的关键部分，起着承上启下的作用。初学者特别注意，这部分涉及对已有工作的评价，务必保证精准客观。要知道，当论文投稿至NLP国际会议后，是通过同行评审决定是否录用发表，评审人一般是小同行，有很大概率是已有工作的作者。所以这部分论述一定要做到客观公正，让这些工作作者本人也能信服。

如下所示论文[2]的第3、4段，先介绍DS-QA的noisy labeling挑战，并且通过举例直观呈现。面对这个挑战，已有一些相关工作，还需说明他们各自有什么不足和挑战，为引出本文创新思路做好铺垫。
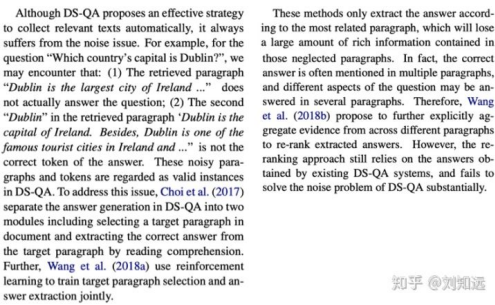
（4）**创新思路**。水来土掩，兵来将挡，既然已有方法有这些不足和挑战，就需要有新的创新思路和方法。这部分需要注意与上面的"挑战"部分严丝合缝，密切呼应，让读者清楚领会到这些创新思路与方法的确能够解决或缓解这些挑战问题。

如下所示论文[2]的第5段，就是介绍创新思路和方法。可以看到，一般"面临挑战"和"创新思路"部分还配图示，更直观地展示本文要解决的挑战问题和创新思路。例如论文[2]这张丑丑的图，比较直观地展示了创新方法包括Selector和Reader两个模块和作用。也可以随便看我们的其他论文[3]，大部分论文都会在Introduction中提供图示。
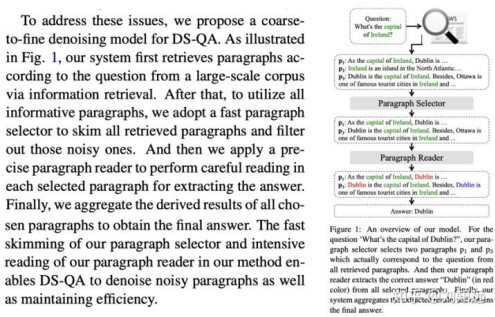
（5）**实验结论**。除了在"创新思路"部分图文两开花地说明本文创新工作外，还要通过合理的实验验证方法的有效性。一般要得到"our method achieves significant and consistent improvement as compared to other baselines"的结论，从而验证本文工作的创新性。
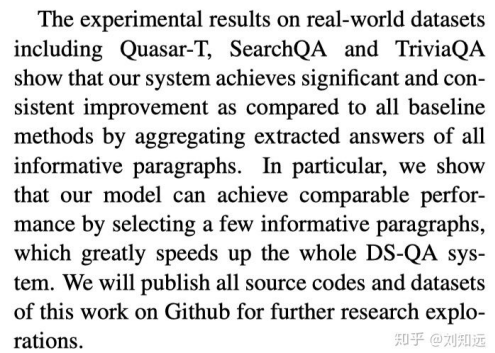
有些论文最后还会体贴的总结本文的主要贡献，一般说"In summary, the key contributions are x-fold: (1)...(2)...(3)..."。这样做的好处是，可以帮助审稿人总结本文的创新点放在审稿意见中，节省不少工作量。但需要注意，这些创新点要简洁明了，不能是前文的简单重复，也不能overclaim。如果要说"首次"提出或发现，一般也要前置"to the best of our knowledge"。此外还有论文最后一段会介绍接下来几个Section结构，个人感觉对一篇8页论文可能并不需要。

对于Abstract，可以看做对Introduction的简介，最简单的做法是，以上每部分都精简为1-2句话组成Abstract皆可。如下是论文[2]的Abstract内容，可以看出与Introduction的对应关系。

Abstract

Distantly supervised open-domain question answering (DS-QA) aims to find answers in collections of unlabeled text. Existing DS-QA models usually retrieve related paragraphs from a large-scale corpus and apply reading comprehension technique to extract answers from the most relevant paragraph. They ignore the rich information contained in other paragraphs. Moreover, distant supervision data inevitably accompanies with the wrong labeling problem, and these noisy data will substantially degrade the performance of DS-QA. To address these issues, we propose a novel DS-QA model which employs a paragraph selector to filter out those noisy paragraphs and a paragraph reader to extract the correct answer from those denoised paragraphs. Experimental results on real-world datasets show that our model can capture useful information from noisy data and achieve significant improvements on DS-QA as compared to all baselines. The source code and data of this paper can be obtained from `https://github.com/thunlp/OpenQA`.

Method怎么写

这部分要详细介绍本文创新方法的具体细节，由于涉及非常艰涩的细节，要采用"总-分"结构来介绍。

这部分起手"总"的部分要介绍本文任务的符号定义，以及本文方法的框架组成，或者按步骤来介绍或者按模块来写，让读者对本文方法有全景式的理解。如下所示论文[2]的Methodology"总"的部分，就先介绍一些符号，然后分别介绍了Selector和Reader两个模块的主要功能。

3 Methodology

In this section, we will introduce our model in details. Our model aims to extract the answer to a given question in the large-scale unlabeled corpus. We first retrieve paragraphs corresponding to the question from the open-domain corpus using information retrieval technique, and then extract the answer from these retrieved paragraphs.

Formally, given a question $q = (q^1, q^2, \cdots, q^{|q|})$, we retrieve m paragraphs which are defined as $P = \{p_1, p_2, \cdots, p_m\}$ where $p_i = (p_i^1, p_i^2, \cdots, p_i^{|p_i|})$ is the i-th retrieved paragraph. Our model measures the probability of extracting answer a given question q and corresponding paragraph set P. As illustrated in Fig. 1, our model contains two parts:

1. Paragraph Selector. Given the question q and the retrieved paragraph P, the paragraph selector measures the probability distribution $\Pr(p_i|q, P)$ over all retrieved paragraphs, which is used to select the paragraph that really contains the answer of question q.

2. Paragraph Reader. Given the question q and a paragraph p_i, the paragraph reader calculates the probability $\Pr(a|q, p_i)$ of extracting answer a through a multi-layer long short-term memory network.

Overall, the probability $\Pr(a|q, P)$ of extracting answer a given question q can be calculated as:

$$\Pr(a|q, P) = \sum_{p_i \in P} \Pr(a|q, p_i) \Pr(p_i|q, P)$$

然后进入"分"的部分，则需对应"总"中的框架，分别介绍各关键模块/步骤。例如，论文[2]的Methodology"分"的部分，就包括3.1 Paragraph Selector、3.2 Paragraph Reader、3.3 Learning and Prediction。读者在"总"的部分已经对方法有全景式的了解，有的放矢，就比较容易理解每个模块的具体细节。而每个"分"的部分中，又可以进一步采用"总-分"结构进行介绍，例如3.1小节做完总体介绍后，又会按照Paragraph Encoding和Question Encoding分别介绍。为了更清晰地体现"总-分"结构，可以将各"分"的部分命名并加粗。

3.1 Paragraph Selector

Since the wrong labeling problem inevitably occurs in DS-QA data, we need to filter out those noisy paragraphs when exploiting the information of all retrieved paragraphs. It is straightforward that we need to estimate the confidence of each paragraph. Hence, we employ a paragraph selector to measure the probability of each paragraph containing the answer among all paragraphs.

Paragraph Encoding. We first represent each word p_i^j in the paragraph p_i as a word vector $\mathbf{p}_i^j$, and then feed each word vector into a neural network to obtain the hidden representation $\hat{\mathbf{p}}_i^j$. Here, we adopt two types of neural networks including:

1. Multi-Layer Perceptron (MLP)

$$\hat{\mathbf{p}}_i^j = \text{MLP}(\mathbf{p}_i^j), \tag{2}$$

2. Recurrent Neural Network (RNN)

$$\{\hat{\mathbf{p}}_i^1, \hat{\mathbf{p}}_i^2, \cdots, \hat{\mathbf{p}}_i^{|p_i|}\} = \text{RNN}(\{\mathbf{p}_i^1, \mathbf{p}_i^2, \cdots, \mathbf{p}_i^{|p_i|}\}), \tag{3}$$

where $\hat{\mathbf{p}}_i^j$ is expected to encode semantic information of word p_i^j and its surrounding words. For RNN, we select a single-layer bidirectional long short-term memory network (LSTM) as our RNN unit, and concatenate the hidden states of all layers to obtain $\hat{\mathbf{p}}_i^j$.

Question Encoding. Similar to paragraph encoding, we also represent each word q^j in the question as its word vector $\mathbf{q}^j$, and then feed them into a MLP:

$$\hat{\mathbf{q}}^j = \text{MLP}(\mathbf{q}^j), \tag{4}$$

or a RNN:

$$\{\hat{\mathbf{q}}^1, \hat{\mathbf{q}}^2, \cdots, \hat{\mathbf{q}}^{|q|}\} = \text{RNN}(\{\mathbf{q}^1, \mathbf{q}^2, \cdots, \mathbf{q}^{|q|}\}), \tag{5}$$

where $\hat{\mathbf{q}}^j$ is the hidden representation of the word q^j and is expected to encode the context information of it. After that, we apply a self attention operation on the hidden representations to obtain the final representation $\mathbf{q}$ of the question q:

$$\hat{\mathbf{q}} = \sum_j \alpha^j \hat{\mathbf{q}}^j, \tag{6}$$

where α_j encodes the importance of each question word and is calculated as:

$$\alpha_i = \frac{\exp(\mathbf{w}_b \mathbf{q}_i)}{\sum_j \exp(\mathbf{w}_b \mathbf{q}_j)}, \tag{7}$$

where $\mathbf{w}$ is a learned weight vector.

Next, we calculate the probability of each paragraph via a max-pooling layer and a softmax layer:

$$\Pr(p_i|q, P) = \text{softmax}\left(\max_j(\hat{\mathbf{p}}_i^j \mathbf{W} \mathbf{q})\right), \tag{8}$$

where $\mathbf{W}$ is a weight matrix to be learned.

初学者特别注意，（1）Introduction中对创新思路与方法的介绍，不要在Method中简单重复，否则会让认真通读全文的审稿人颇感厌烦。要做到前后照应，有所递进，前略后详，不妨使用"as mentioned in Section 1"来做关联。（2）Method部分往往包含大量公式，需要保证公式风格和符号使用前后统一，新符号使用均需显式解释。

Experiment怎么写

这部分要详细介绍与实验相关的具体细节。一般先介绍实验数据、评测标准和比较方法等基本信息。以论文[2]为例，实验部分首先介绍实验数据与评测标准（4.1 Datasets and Evaluation Metrics）、实验比较的已有代表方法（4.2 Baselines）、实验方法的参数设置（4.3 Experimental Settings）等基本信息。

在介绍完实验基本信息后，主要开展两种实验：

（1）**主实验**。目的是证明本文方法与已有方法相比的有效性。一般需要选取业界公认的数据集合或已有工作采用的实验验证方式，提升实验的可信性。对于学术论文而言，并不需要比该任务上最好的方法相比，只要证明采用本文创新方法与不采用本文方法相比更有效即可，也就是说，实验中尽量控制其他变量，只聚焦于本文关注的挑战问题即可。当然，如果能够因为本文创新思路，得到该任务上的最好效果，会更有吸引力，但不必总是强求。

一般实验结果用图表展示，然后在正文进行观察分析。例如，论文[2]的主实验部分先介绍不同Selector和Reader对实验效果的影响（4.4 Effect of Different Paragraph Selectors、4.5 Effect of Different Paragraph Readers），接着介绍主实验结果和观察分析（4.6 Overall Results）。其中表格中会把最好效果加粗显示，一般应大部分位于本文提出的方法；为了更加清晰明了，观察分析结论可用（1）（2）（3）列出，其中第1条一般要得出主要结论，即本文方法要显著优于已有方法。

Datasets	Quasar-T		SearchQA		TriviaQA		CuratedTREC	WebQuestions	
Models	EM	F1	EM	F1	EM	F1	REM	EM	F1
GA (Dhingra et al., 2017a)	26.4	26.4	-	-	-	-	-	-	-
BiDAF (Seo et al., 2017)	25.9	28.5	28.6	34.6	-	-	-	-	-
AQA (Buck et al., 2017)	-	-	40.5	47.4	-	-	-	-	-
R³ (Wang et al., 2018a)	35.3	41.7	49.0	55.3	47.3	53.7	28.4	17.1	24.6
Our + AVG	38.5	45.7	55.6	61.0	42.6	48.2	28.6	17.8	24.5
+ FULL	**42.2**	**49.3**	**58.8**	**64.5**	**48.7**	**56.3**	**29.1**	**18.5**	**25.6**

Table 2: Experimental results on four open-domain QA test datasets: Quasar-T, SearchQA, TriviaQA, CuratedTREC and WebQuestions. TriviaQA, CuratedTREC and WebQuestions do not provide the leader board under the open-domain setting. Therefore, there is no public baselines in this setting and we only report the result of the DrQA and R³ baseline. CuratedTREC dataset is evaluated with regular expression matching (REM).

主实验结果

主实验分析

（2）辅助实验。目的是展示本文创新方法的优势和特点。例如，不同超参数对本文方法的影响（Hyper-Parameter Effect），不同模块对本文方法效果的贡献（Ablation Test），不同数据划分对本文方法的影响（如Few-shot Learning相关工作比较常见），本文方法的主要错误类型（Error Analysis），本文方法能够改进效果的典型样例（Case Study）等。这些实验需要根据论文创新工作特点而有针对性的设计，一切要为体现本文的创新价值而服务。

例如，论文[2]的辅助实验包括4.7 Paragraph Selector Performance Analysis、4.8 Performance with different numbers of paragraphs、4.9 Potential improvement、4.10 Case study等，从各方面呈现本文提出方法的特点。

Experiment部分的特点是要图文并茂，注重通过多个表格和图示来呈现本文方法的优势和特点，需要注意图表风格统一。初学者特别注意，要做到仅凭图表下方的说明文字就可以理解每张图表内容，不要让读者还要到跑到正文寻找相关说明。因为，很多有经验的审稿人在看完Introduction后，会直接跳到Experiment图表中寻找对比效果。

Related Work怎么写

这部分主要是介绍本文任务和方法的相关工作，目标是通过对已有工作的梳理，凸显本文工作的创新价值。对已有工作的梳理，不应是对每个工作的简单介绍，而应当注意汇总、分类、分析，或者按照时间发展顺序，或者按照技术路线划分，例如论文[2]就是按照时间脉络介绍。

在对相关工作的介绍中，要注意暗合本文创新思路要解决的挑战，不应是单纯的介绍，而是夹叙夹议，时刻注意与本文工作的照应。在Related Work的最后，应该落脚到本文工作与已有工作相比，有什么新的思路，解决了什么挑战问题。

初学者特别注意，Introduction和Related Work部分是特别需要导师或其他有经验学者帮助把关的。一是，不能遗漏重要相关工作，这点需要论文作者对相关领域工作保持跟踪；二是，与Introduction要求类似，对已有工作的评述务必精准客观。

Related Work一般放在Introduction之后，或者Conclusion之前，这一般取决于论文工作的特点。对于那些与已有工作联系紧密、创新精微的工作，一般建议放在Introduction之后，方便读者全面了解本文工作与已有工作的关系，然后开始在Method介绍本文方法。而对于有些框架性创新工作，如果主要是对已有方法的组合，一般建议Related Work放在Method、Experiment之后即可。这点并无成法，完全根据行文方便来定。

2 Related Work

Question answering is one of the most important tasks in NLP. Many efforts have been invested in QA, especially in open-domain QA. Open-domain QA has been first proposed by (Green Jr et al., 1961). The task aims to answer open-domain questions using external resources such as collections of documents (Voorhees et al., 1999), webpages (Kwok et al., 2001; Chen and Van Durme, 2017), structured knowledge graphs (Berant et al., 2013a; Bordes et al., 2015) or automatically extracted relational triples (Fader et al., 2014).

Recently, with the development of machine reading comprehension technique (Chen et al., 2016; Dhingra et al., 2017a; Cui et al., 2017; Shen et al., 2017; Wang et al., 2017), researchers attempt to answer open-domain questions via performing reading comprehension on plain texts. Chen et al. (2017) propose a DS-QA system, which retrieves relevant texts of the question from a large-scale corpus and then extracts answers from these texts using reading comprehension models. However, the retrieved texts in DS-QA are always noisy which may hurt the performance of DS-QA. Hence, Choi et al. (2017) and Wang et al. (2018a) attempt to solve the noise problem in DS-QA via separating the question answering into paragraph selection and answer extraction and they both only select the most relevant paragraph among all retrieved paragraphs to extract answers. They lose a large amount of rich information contained in those neglected paragraphs. Hence, Wang et al. (2018b) propose strength-base and coverage-based re-ranking approaches, which can aggregate the results extracted from each paragraph by existing DS-QA system to better determine the answer. However, the method relies on the pre-extracted answers of existing DS-QA models and still suffers from the noise issue in distant supervision data because it considers all retrieved paragraphs indiscriminately. Different from these methods, our model employs a paragraph selector to filter out those noisy paragraphs and keep those informative paragraphs, which can make full use of the noisy DS-QA data.

Our work is also inspired by the idea of coarse-to-fine models in NLP. Cheng and Lapata (2016) and Choi et al. (2017) propose a coarse-to-fine model, which first selects essential sentences and then performs text summarization or reading comprehension on the chosen sentences respectively. Lin et al. (2016) utilize selective attention to aggregate the information of all sentences to extract relational facts. Yang et al. (2016) propose a hierarchical attention network which has two levels of attentions applied at the word and sentence level for document classification. Our model also employs a coarse-to-fine model to handle the noise issue in DS-QA, which first selects informative retrieved paragraphs and then extracts answers from those selected paragraphs.

Conclusion怎么写

在论文最后会有总结展望，一般用一段来再次总结和强调本文的创新思路和实验结果，然后说明未来建议的研究方向和开放问题。这部分相对来讲比较固定。稍微留意的是，在准备论文最后阶段，如果发现论文有哪些应当做还没来得及做的，可以写作本文的未来工作。至少可以向审稿人表明你也想到这个问题了，赢得一点同情分。

5 Conclusion and Future Work

In this paper, we propose a denoising distantly supervised open-domain question answering system which contains a paragraph selector to skim over paragraphs and a paragraph reader to perform an intensive reading on the selected paragraphs. Our model can make full use of all informative paragraphs and alleviate the wrong labeling problem in DS-QA. In the experiments, we show that our models significantly and consistently outperforms state-of-the-art DS-QA models. In particular, we demonstrate that the performance of our model is hardly compromised when only using a few top-selected paragraphs.

In the future, we will explore the following directions:

(1) An additional answer re-ranking step can further improve our model. We will explore how to effectively re-rank our extracted answers to further enhance the performance.

(2) Background knowledge such as factual knowledge, common sense knowledge can effectively help us in paragraph selection and answer extraction. We will incorporate external knowledge bases into our DS-QA model to improve its performance.

其他建议

要想写出一篇合格的NLP论文，首先是**态度问题**，只有态度重视，才有可能不厌其烦地反复修改，才会"不择手段"地寻找各种办法来尽力改进论文（找学长找外教借助Grammarly工具等）。其次是**动手问题**，只有写下来，才可能不断改，只要改就能不断进步。最后是**经验问题**，要写得精彩可能需要天赋，而要写得合格，只要坚持写，不断根据评阅人和其他人的意见进行思考和修改，就可以进步。总之，坚持就是胜利。

实际上，我觉得论文写作，是对思维模式的训练。也许未来你并不会从事学术研究，但通过论文写作锻炼的凝练工作创新价值的能力、清晰传递复杂信息的表达能力，对未来工作中无论是工作沟通、成果展示等，都有重要帮助。所以还希望大家都能重视这个科研道路上难得的锻炼机会。加油！

小结

论文写作有很多需要注意的细节和技巧，很多领域甚至都有专门的厚厚一本指南来介绍写作技巧。这么一篇短文很难面面俱到，只是介绍一下我指导同学准备论文提的比较多的建议，希望对大家有用。以后想到任何新的建议，随时更新。也欢迎各位反馈建议和问题，共同进步。

相关链接

1. 刘洋. 机器翻译学术论文 写作方法和技巧. nlp.csai.tsinghua.edu.cn...
2. Yankai Lin, Haozhe Ji, Zhiyuan Liu, Maosong Sun. Denoising Distantly Supervised Open-Domain Question Answering. ACL 2018. nlp.csai.tsinghua.edu.cn...
3. 个人主页：Publications

编辑于 2019-03-12

「真诚赞赏，手留余香」

赞赏

5 人已赞赏
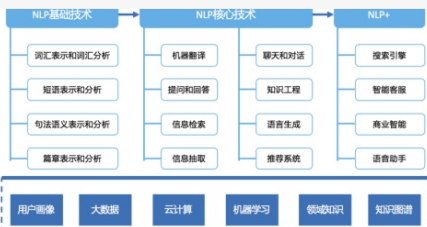
52nlp

自然语言处理

文章被以下专栏收录

NLP日知录   已关注

推荐阅读

NLP深度学习最新趋势
文章链接：
https://arxiv.org/pdf/1708.02709.p
深度学习方法利用多层处理结构来
学习数据中的层次化表示，在很多
领域都得到了SOTA的结果。最近几
年，NLP领域涌现了大量的深度学...

Simpl...   发表于自然语言理...

练手|常见30种NLP任务的练手项目

cstghitpku

近期值得读的20篇AI论文（NLP/CV/ML方向）

张俊   发表于Paper...

关于NER中bi-LSTM+CRF
题
作为NLP邻域的初学者，最近
命名体识别，用的是比较流行
lstm+crf模型，模型来源于
Github。在测试训练效果时我
了一个怎么都无法解决的问题
识别的东西能够识别。不该识

Daniel.D

28 条评论

切换为时间排序

写下你的评论...

夏飞   3 天前
创新点好难想啊
👍 赞

herbzq   3 天前
学到了[赞同]
👍 赞

▲已赞同 851 ▼   💬 28 条评论   ✈ 分享   ★ 收藏   ···