知识指导的预训练语言模型

Knowledge-Guided Pre-Trained Language Models

韩旭/HAN Xu，张正彦/ZHANG Zhengyan，刘知远/LIU Zhiyuan
（清华大学，中国 北京 100084）
(Tsinghua University, Beijing 100084, China)

DOI：10.12142/ZTETJ.202202003
网络出版地址：https://kns.cnki.net/kcms/detail/34.1228.tn.20220410.1321.002.html
网络出版日期：2022-04-12
收稿日期：2022-02-20

摘要：作为典型的数据驱动工具，预训练语言模型（PLM）仍然面临可解释性不强、鲁棒性差等难题。如何引入人类积累的丰富知识，是改进预训练模型性能的重要方向。系统介绍知识指导的预训练语言模型的最新进展与趋势，总结知识指导的预训练语言模型的典型范式，包括知识增强、知识支撑、知识约束和知识迁移，从输入、计算、训练、参数空间等多个角度阐释知识对于预训练语言模型的重要作用。

关键词：自然语言处理；PLM；知识图谱

Abstract: As a typical data-driven method, pre-trained language models (PLMs) still face challenges such as poor interpretablility and robustness. Hence, it is important to introduce human knowledge into these models for better performance. The latest progress and trend of knowledge-guided PLMs are introduced and the paradigm of knowledge-guided PLMs is summarized, including knowledge augmentation, knowledge support, knowledge regularization, and knowledge transfer.

Keywords: natural language processing; PLMs; knowledge graphs

1 知识的重要作用

20世纪90年代前，研究人员将大量的精力投入到语法理论[1-2]和专家系统[3-4]的研究中。无论是语法理论中的语言规则还是专家系统中的知识库，其背后的核心思想均为使用符号体系来表示语言理解所需的各类知识。这些离散稀疏的符号系统有利于抽象丰富的人类知识，并通过人为设计的精密规则实现语言理解中的知识推理。

近些年来，陆续构建的大型知识图谱（知识库），诸如Wikidata、YAGO和DBpedia，就采用了结构化的符号形式来存储海量的世界知识，并在语言理解中发挥重要作用。近些年的研究也证明，大规模知识图谱中的丰富知识可以有力驱动一系列人工智能和自然语言处理的应用，例如问答系统、对话系统、文本检索和推荐系统。

符号知识的一大痛点在于难以发挥机器所擅长的数值计算优势。此外，早期的语法规则与专家系统在泛化性上也存在问题。这就需要一套基于数值计算且具有一定泛化性的知识表示框架。统计学习[5-6]也由此被应用于自然语言处理任务中。20世纪90年代后，支持向量机[7]、决策树[8]、条件随机场[9]的诸多经典统计模型被广泛应用，在各类自然语言处理任务上取得了一系列突破。这些统计方法用模型参数来隐式地表示各类知识，并基于概率计算来进行推理。相对于符号知识的“人类友好”，这种连续数值化的模型知识更加“机器友好”。

统计模型拉开了从符号知识到模型知识的序幕，开启了用数值表示知识的新纪元，但统计模型本身的性能是十分有限的。近年来，神经网络蓬勃发展，它为数值化的知识表示及语义理解提供了更强大的工具。浅层神经网络首先被应用于知识表示中。分布式词向量表示旨在利用低维连续向量来表示词汇相关的语言知识，并通过海量无标签文本的自监督学习来学习词向量[10]。得益于分布式词向量中蕴含的丰富语言知识，词的向量化表示已经成为当前完成各类自然语言处理任务的标准范式，也有效地填补符号知识与数值计算间的鸿沟。

随着神经网络的深度与参数量的增加，大规模预训练语言模型（PLM）被提出，这推动了一系列自然语言处理任务的发展。预训练语言模型的主要特点在于其两阶段的构建方法：第1阶段，与分布式词向量表示类似，在海量无标签文本上进行自监督学习，以学习通用的语言特征和规则（即预训练）；第2阶段，将预训练模型在具体的自然语言处理任务上进行小规模、有标注数据的二次训练（即微调），以快速提升模型在这些任务中的性能，最终形成可部署应用的模型。研究表明，在自监督学习过程中，预训练语言模型可以捕捉到丰富的词法知识、句法知识、语义知识、世界知识，并通过庞大的参数将这些知识存储起来。这样一来，微调模型的参数可以有效地将模型知识迁移到具体的任务上。

图1显示了自然语言处理技术的发展脉络，清晰地表明

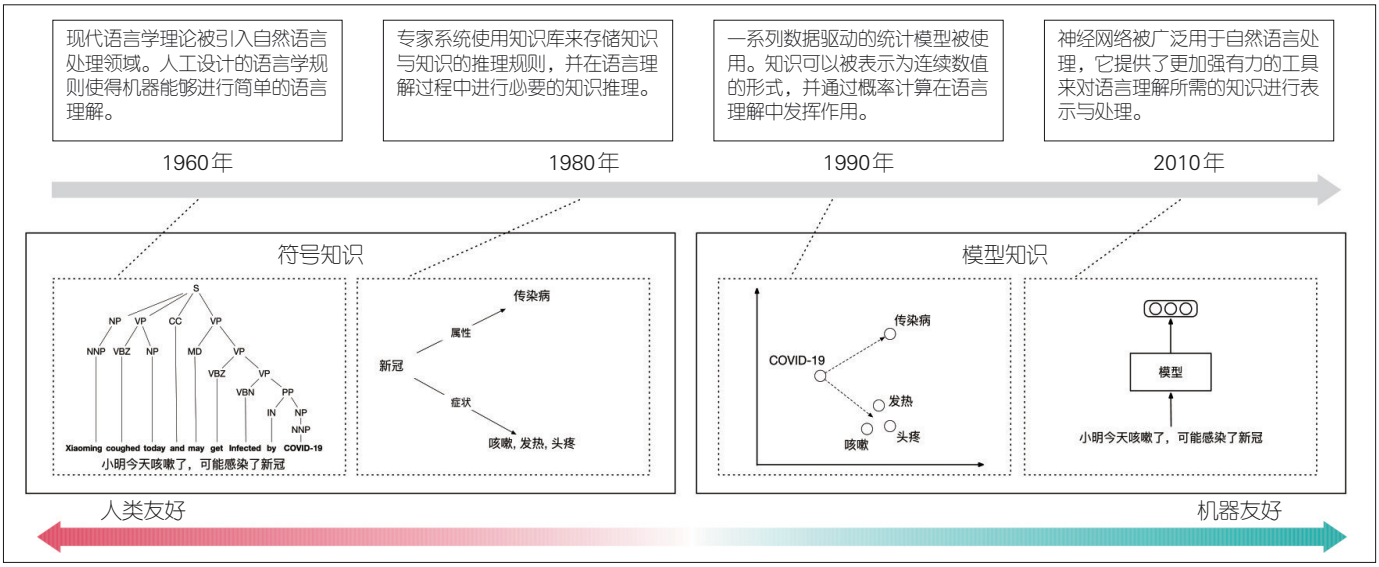
了各个时期知识是如何表示的，以及是如何被运用于语言理解的。在使用上，符号知识与模型知识也各有优势。尽管预训练语言模型已经在当前诸多自然语言处理任务上取得了很好的效果，但大量数据驱动下的预训练语言模型依然在可解释性、鲁棒性上存在不足。数据驱动的预训练语言模型具有善于学习的语义特征，同时符号表示的结构化知识有着善于认知推理的特征。综合发挥以上两个优势，形成知识指导的预训练语言模型，对于揭示自然语言处理机理，实现智能语言理解，具有重要的理论意义与实用价值。

2 知识指导的预训练语言模型范式

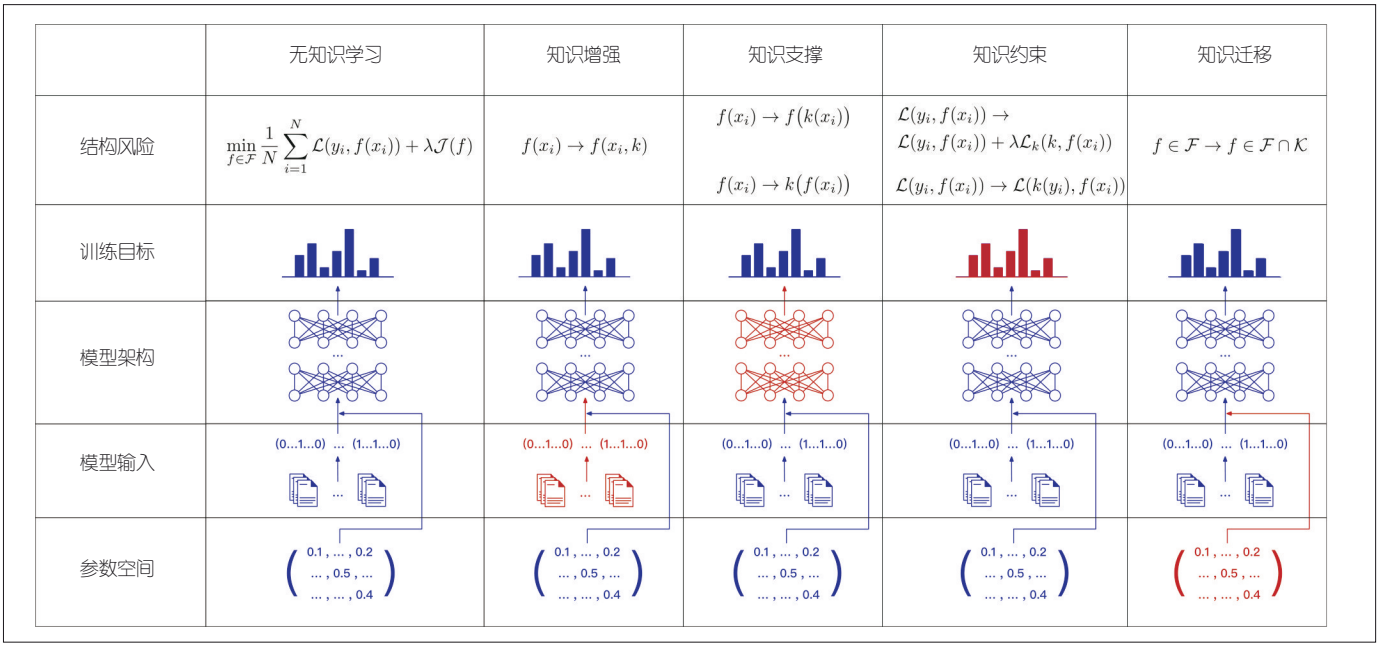
对于如何将知识有效地应用在预训练语言模型中，我们已在文献[11]中做了简要介绍。本文中我们进一步扩展并提出了知识指导的预训练语言模型。如图2所示，一般来讲，预训练语言模型有4个要素：模型输入、模型架构、训练目标和参数空间。

• 对模型输入而言，知识是输入的重要补充，为文本中的关键词句提供更加有效的语义解释和语义特征；

• 对模型架构而言，知识可以引入先验指导模型内部的特征处理流程，进而提升模型性能；



▲图1 自然语言处理技术发展脉络[11]



▲图2 知识指导的预训练语言模型范式[11]

• 在训练目标上，知识可用于构造新的训练任务，提供更加丰富的训练目标，促进预训练语言模型能力的多样化；

• 在参数空间里，相比于随机初始化，用引入知识的方式来约束参数空间可以提供一个更好的参数空间初始点，有利于加速收敛，优化出更好的模型参数。

正如图2所示，知识可被应用于其中任意一部分，以起到强化预训练模型性能的作用。接下来，我们将介绍这个框架的具体内容。在图中，我们给出了结构风险函数在知识指导前后的变化。其中，x、y是样本的输入输出，k是引入的知识信息或者知识驱动的模块，f是预训练语言模型本身，$\mathcal{F}$、$\mathcal{K}$分别是参数空间、知识约束的参数空间。

2.1 知识增强

在语言表达过程中，人们习惯省略一些众所周知的背景知识。这并不影响人类对语言的理解，却不利于机器对语言的理解。知识增强旨在将这部分背景知识显式地作为补充输入，丰富上下文信息，以帮助模型更好地进行文本理解。

知识增强的方式主要有两种。第一种是直接将知识转换成文本形式，并拼接到已有文本中作为输入。最简单的做法就是将相关的结构化图谱信息转换为文本内容[12]。在此过程中，如何找到和输入相关的知识就是一个主要挑战。基于信息检索的预训练语言模型是一个有效的解决方案，例如REALM[13]和RAG[14]。其预训练一个文本检索器，用于构建输入文本和背景知识文本的关联，使用时再将检索到的知识文本与输入文本拼接起来，给模型提供更加丰富的信息。

知识增强的另一种方式则是通过设计特定的知识融合模块，将文本的表示向量和相关知识向量融合在一起[15]。这与上述文本拼接有明显不同：知识不再以符号形式进行表达，而是被蕴含在模型参数中。ELMo[16]是该方向的代表性工作。由于ELMo是一个在超大规模语料上训练的语言模型，其表示向量可以提供丰富的语言知识，解决一词多义等问题。人们通常使用ELMo来代替传统词向量，以提升模型的基本文本理解能力。更进一步地，不少工作[17−20]将知识图谱中的实体与关系表示为向量，并将这些向量输入到预训练语言模型以进行知识融合，这也是非常有效的知识增强方法。

2.2 知识支撑

知识支撑可以利用大量已有的知识来构建更好的结构先验。具体而言，在模型底层，知识支撑可以作为一种数据预处理模块；而在模型顶层，知识支撑可以指导模型的预测。

知识记忆网络[21]是数据预处理模块的代表技术。根据输入特征，底层的网络结构会动态调整，以连接对应的记忆区域，从而将记忆模块中的知识注入到模型的推理计算中。在此过程中，知识的表示形式通常为低维稠密向量，也就是所谓的模型知识。采用了记忆机制的预训练语言模型[22−23]在多跳推理、长文本处理等需要长距离语义关系处理的任务上有显著效果。

当知识支撑作为顶层的预测指导模块时，其目标是借助知识的先验信息，构建答案之间的关联，更好地对备选答案进行筛选。在此过程中，知识的表示形式通常是符号化、层次化的。结构化知识库支撑的语言模型是该方向具有代表性的研究工作[24−26]。在生成句子的过程中，语言模型可以利用知识库信息生成更加适合当前语境的词。

2.3 知识约束

对于知识约束，我们既可以基于已有输入数据并结合相关知识来构建训练目标，也可以直接使用外部知识来构建新数据和新目标。

知识蒸馏是一种代表性的知识约束方法[27]，也是知识结合已有输入数据来构建训练目标的典型案例。知识蒸馏能够利用大模型对已有数据进行预测，从而提供新的监督信号，帮助小模型学习取得更好的效果。具体而言，知识蒸馏要求小模型的中间计算结果和大模型的中间计算结果尽可能保持一致，包括隐层表示以及预测的标签分布。相比于单一的人工标注标签，知识蒸馏能提供更加丰富的模型知识信息。知识蒸馏已被广泛用于预训练语言模型以提升其计算效率与模型表现[28−31]。

远程监督是另一种具有代表性的知识约束方法[32]，能够根据已有知识图谱和无监督文本自动生成大量新训练数据。远程监督在信息抽取领域获得广泛应用，大大降低了数据标注成本，显著提升了模型性能。我们给出了一个远程监督的简单示例：给定知识图谱中的三元组（包含头实体、尾实体及其关系），找出同时包含头尾实体的文本，并将其标注为该关系类型的样例。基于上述启发式规则，我们可以自动获取大量知识相关的文本分类数据来训练预训练语言模型。尽管这种自动标注方式存在噪音，如标注的样例可能并不反映头尾实体间的标注关系，但不少工作表明[17−18,33−35]，远程监督数据依然能够有效地帮助模型的训练。这些使用远程监督数据增强的预训练语言模型被验证具有强大的实体关系理解能力。

2.4 知识迁移

知识迁移的目的在于利用知识进行参数空间的约束，以降低参数空间的搜索代价，提升最终模型的性能。知识迁移

技术已被广泛应用于自然语言处理。迁移学习和自监督学习都是知识迁移的重要研究方向[36]。各种预训练语言模型的微调阶段本身就是一种知识迁移，旨在将预训练阶段获取的丰富模型知识迁移到具体任务上。

对于预训练过程而言，最近的一些工作尝试以已有的预训练语言模型为基底来训练新的预训练模型。部分工作[30,37]侧重于利用较小的预训练语言模型的模型知识，来降低大规模预训练模型的训练代价；而另一些工作[38-39]则基于已有预训练语言模型的通用知识，来指导更多垂直领域的知识。

无论是对于预训练语言模型的预训练还是下游任务适配，充分迁移已有的模型知识相较于毫无基础的重新学习，在计算效率和模型效果上均有显著优势。

总之，我们从预训练语言模型的模型输入、模型架构、训练目标和参数空间4个方面入手，构建了全面的知识指导的预训练语言模型框架。在该框架下，符号知识和模型知识均可以得到充分利用，有效提升预训练模型的学习能力和模型表现。

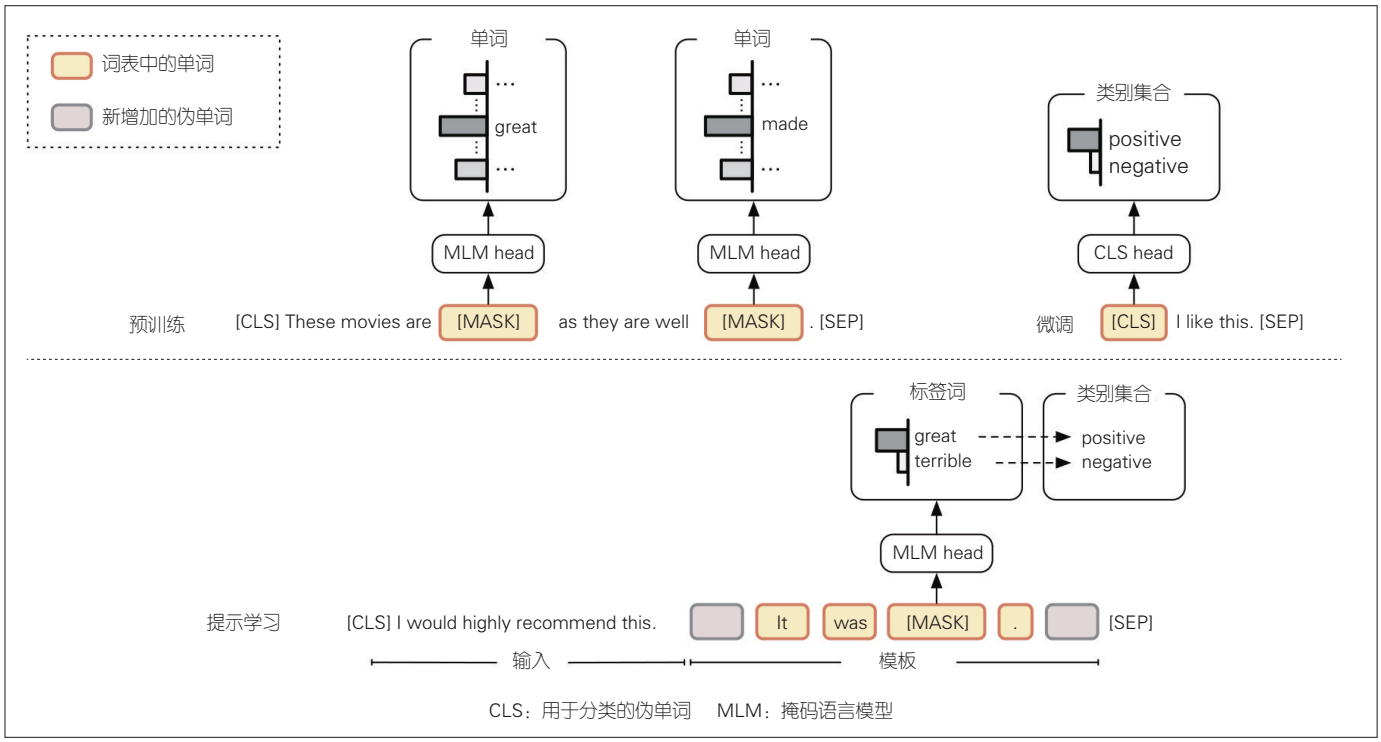
3 预训练语言模型的知识激发

在上一章节中，我们关注的是如何将知识注入预训练语言模型之中。在这一章节中，我们将简单介绍如何激发预训练模型中的知识。这对于应用知识指导的预训练语言模型具有重要意义。

预训练语言模型能够通过微调显著提升下游任务性能，却仍然面临着两个重要挑战：（1）预训练和微调之间的任务形式存在较大差别，预训练只考虑语言建模，但下游任务目标形式可能各有不同，这种差别会显著影响知识迁移的效能。（2）随着预训练模型参数规模迅速增加，即使进行模型微调，也需要大量技术资源。为了解决这些问题，最近学术界提出了一种新的微调技术，即提示学习（Prompt Tuning）。该技术能够有效利用大规模的模型知识，日益获得广泛关注。

提示学习的目的是将下游任务转化为类似于预训练目标的填空任务。采用相同的优化目标有利于在下游任务中更好地激发预训练模型中的知识。以情感分类的提示学习为例（图3），模型的输入由两部分组成：输入数据以及提示学习所需的提示模板。基于该输入，预训练语言模型在一组标签词中选择概率最高的词进行填空，再将预测的词映射到相应的分类标签上。图3中，提示模板为“It was [Mask]”，“[Mask]”代表需要进行填空的位置。标签词为“great”和“terrible”，“great”对应正向情感，“terrible”对应负向情感。提示微调也在一系列自然语言处理任务上取得了成效，包括文本分类[40-43]、序列标注[44-45]、文本生成[46-47]等任务。

为了在下游任务上取得成功，提示模板和标签词（提示



▲图3 预训练、微调、提示学习示意图

语）需要进行精细的设计和选择。为了避免费力而复杂的提示语设计，自动搜索高质量的提示语成为目前工作的一个重点：研究者探索使用梯度优化来搜索最佳提示语[48]，或使用生成模型来提供多个候选提示语[42]，然后逐一评估其有效性，以选择最佳提示语。目前，自动搜索提示语的成本仍然很高，这限制了这些自动方法的使用场景。为此，也有研究者提出用逻辑规则指导提示学习[49]。这种方法将先验知识编码到提示语中，降低搜索以及训练难度，使模型知识可以更好地为下游任务服务。为了避免复杂的提示设计，一些工作[50–52]采用了可学习的提示向量来驱动预训练语言模型进行提示微调，无须变动预训练模型的任何参数，只须调整提示向量即可。

不少知识探测工作[53–55]表明，通过设计提示模板，预训练语言模型甚至可以补全结构化知识信息。上述研究表明，除了知识模型的性质外，预训练语言模型也有一定的符号知识特性。输入提示能充分激发出预训练语言模型中各个层面丰富的知识信息，以解决具体问题。预训练语言模型在推动自然语言处理中模型知识的使用方面有着重要作用。从某种程度上而言，预训练模型也将影响自然语言处理中符号知识的使用范式。尽管预训练语言模型仍需符号知识进行强化，但其本身也是一种符号知识的优秀载体，有利于符号知识与模型知识的融合与统一。

4 结束语

在文章中，我们围绕知识对于自然语言处理的重要性、知识指导的预训练范式、预训练语言模型的知识激发3个方面，介绍了知识指导的预训练语言模型的相关技术。在各个方向上，尽管目前均已获得一些成果，但仍有许多尚未解决的重要问题。这需要研究者进一步努力，以取得突破。

致谢

清华大学姚远、李涓子和孙茂松在文章的撰写过程中，给出了宝贵的建议，在此表示感谢。

参考文献

[1] CHOMSKY N. Syntactic structures [M]. Germany: Walter de Gruyter,1957
[2] NOAM C. Aspects of the theory of syntax [M]. USA: The MIT Press, 1969
[3] WILSON S, BARR A, COHEN P R, et al. The handbook of artificial intelligence [J]. Leonardo, 1984, 17(4): 299. DOI: 10.2307/1575114
[4] ROTH H F, WATERMAN A D. Building expert system [M]. USA: Addison–Wesley, 1983
[5] LANDAUER T K, DUMAIS S T. A solution to Plato´s problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge [J]. Psychological review, 1997, 104(2): 211–240. DOI: 10.1037/0033–295x.104.2.211
[6] PENNINGTON J, SOCHER R, MANNING C. Glove: global vectors for word representation [C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, 2014: 1532–1543. DOI: 10.3115/v1/d14–1162
[7] BOSER B E, GUYON I M, VAPNIK V N. A training algorithm for optimal margin classifiers [C]//COLT´92: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. ACM, 1992: 144–152. DOI: 10.1145/130385.130401
[8] BREIMAN L, FRIEDMAN J, STONE C J, et al. Classification and regression trees [M]. USA: CRC press, 1984
[9] LAFFERTY J D, McCALLUMM A, FERNANDO C N. Conditional random fields: probabilistic models for segmentingand labeling sequence data [C]//Proceedings of the 18th International Conference on Machine Learning (ICML 2001). ICML, 2001: 282–289
[10] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality [C]//Proceedings of the 26th International Conference on Neural Information Processing Systems (NeurIPS 2013). NIPS, 2013: 3111–3119
[11] HAN X, ZHANG Z, LIU Z. Knowledgeable machine learning for natural language processing [J]. Communications of the ACM, 2021, 64(11): 50–51
[12] LIU W J, ZHOU P, ZHAO Z, et al. K–BERT: enabling language representation with knowledge graph [J]. Proceedings of the AAAI conference on artificial intelligence, 2020, 34(3): 2901–2908. DOI: 10.1609/aaai.v34i03.5681
[13] GUU K, LEE K, TUNG Z. REALM: integrating retrieval into language representation models [EB/OL]. [2022–01–10]. https://arxiv. org/abs/2005.11401
[14] LEWIS P, PEREZ E, PIKTUS A, et al. Retrieval–augmented generation for knowledge–intensive NLP tasks [EB/OL]. (2021–04–12) [2022–01–10]. https://arxiv.org/abs/2005.11401
[15] LIU Z H, XIONG C Y, SUN M S, et al. Entity–duet neural ranking: understanding the role of knowledge graph semantics in neural information retrieval [C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 2018: 2395–2405. DOI: 10.18653/v1/p18–1223
[16] PETERS M, NEUMANN M, IYYER M, et al. Deep contextualized word representations [C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). Association for Computational Linguistics, 2018: 2227–2237. DOI: 10.18653/v1/n18–1202
[17] ZHANG Z Y, HAN X, LIU Z Y, et al. ERNIE: enhanced language representation with informative entities [EB/OL]. (2019–06–04)[2022–01–10]. https://arxiv.org/abs/1905.07129
[18] WANG X Z, GAO T Y, ZHU Z C, et al. KEPLER: a unified model for knowledge embedding and pre–trained language representation [J]. Transactions of the association for computational linguistics, 2021, 9: 176–194. DOI: 10.1162/tacl_a_00360
[19] SU Y S, HAN X, ZHANG Z Y, et al. CokeBERT: Contextual knowledge selection and embedding towards enhanced pre–trained language models [J]. AI open, 2021, 2: 127–134. DOI: 10.1016/J.AIOPEN.2021.06.004
[20] PETERS M E, NEUMANN M, LOGAN R, et al. Knowledge enhanced contextual word representations [EB/OL]. (2019–10–31) [2022–01–10]. https://arxiv.org/abs/1909.04164
[21] WESTON J, CHOPRA S, BORDES A. Memory networks [EB/OL]. (2014–10–15)[2022–01–10]. https://arxiv.org/abs/1410.3916
[22] DING M, ZHOU C, CHEN Q B, et al. Cognitive graph for multi–hop reading comprehension at scale [C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2019. DOI: 10.18653/v1/p19–1259
[23] RAE W J, POTAPENKO A, JAYAKUMAR S M, et al. Compressive transformers for long–range sequence modelling [EB/OL]. (2019–11–13) [2022–01–12]. https://arxiv.org/abs/1911.05507
[24 LOGAN R, LIU N F, PETERS M E, et al. Barack´s wife Hillary: using knowledge graphs for fact–aware language modeling [EB/OL]. (2019–06–17)[2022–01–10]. https://arxiv.org/abs/1906.07241
[25] AHN S , CHOI H, PARNAMAA T, et al. A neural knowledge language model [EB/OL]. (2017–03–02) [2021–12–12]. https://arxiv. org/pdf/1608.00318.pdf
[26] HAYASHI H, HU Z C, XIONG C Y, et al. Latent relation language models EB/OL]. (2017–03–02)[2022–01–12]. https://arxiv.org/abs/1908.07690

[27] HINTON G E, VINYALS O, DEAN J. Distilling the knowledge in a neural network [EB/OL]. (2015-03-09) [2022-01-10]. https://arxiv. org/abs/1503.02531
[28] SUN S Q, CHENG Y, GAN Z, et al. Patient knowledge distillation for BERT model compression [EB/OL]. (2015-03-09)[2022-01-10]. https://arxiv.org/abs/1908.09355
[29] RASHID A, LIOUTAS V, REZAGHOLIZADEH M. MATE-KD: masked adversarial TExt, a companion to knowledge distillation [EB/OL]. (2021-05-12)[2022-01-10]. https://arxiv.org/abs/2105.05912v1
[30] QIN Y, LIN Y, YI J, et al. Knowledge inheritance for pre-trained language models [EB/OL]. (2021-05-28) [2022-01-12]. https://arxiv. org/abs/2105.13880v1
[31] JIAO X Q, YIN Y C, SHANG L F, et al. TinyBERT: distilling BERT for natural language understanding [EB/OL]. (2019-09-23) [2022-01-10]. https://arxiv.org/abs/1909.10351v4
[32] MINTZ M, BILLS S, SNOW R, et al. Distant supervision for relation extraction without labeled data [EB/OL]. [2022-01-10]. https://paperswithcode.com/paper/distantly-supervised-ner-with-partial
[33] BALDINI SOARES L, FITZGERALD N, LING J, et al. Matching the blanks: distributional similarity for relation learning [C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2019. DOI: 10.18653/v1/p19-1279
[34] SUN Y, WANG S H, LI Y K, et al. ERNIE: enhanced representation through knowledge integration [EB/OL]. (2019-04-19) [2022-01-10]. https://arxiv.org/abs/1904.09223v1
[35] SUN Y, WANG S H, FENG S K. Ernie 3.0: large-scale knowledge enhanced pre-training for language understanding and generation [EB/OL]. (2019-04-19)[2022-01-10]. https://arxiv.org/abs/2107.02137
[36] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [C]//Proceedings of the 2019 Conference of the North. Association for Computational Linguistics, 2019. DOI: 10.18653/v1/n19-1423
[37] GU X T, LIU L Y, YU H K, et al. On the transformer growth for progressive BERT training [C]//Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics, 2021. DOI: 10.18653/v1/2021.naacl-main.406
[38] GURURANGAN S, MARASOVIĆ A, SWAYAMDIPTA S, et al. Don't stop pretraining: adapt language models to domains and tasks [EB/OL]. (2020-04-23)[2022-01-10]. https://arxiv.org/abs/2004.10964v2
[39] PFEIFFER J, RÜCKLÉ A, POTH C, et al. AdapterHub: a framework for adapting transformers [EB/OL]. (2020-10-06) [2022-01-10]. https://arxiv.org/abs/2007.07779
[40] LIU X, ZHENG Y N, DU Z X, et al. GPT understands, too [EB/OL]. (2021-03-18)[2022-01-11]. https://arxiv.org/abs/2103.10385v1
[41] LIU X, JI K X, FU Y C, et al. P-tuning v2: prompt tuning can be comparable to fine-tuning universally across scales and tasks [EB/OL]. (2021-10-18) [2022-01-10]. https://arxiv.org/abs/2110.07602v2
[42] GAO T Y, FISCH A, CHEN D Q. Making pre-trained language models better few-shot learners [EB/OL]. (2021-06-02) [2022-01-12]. https://arxiv.org/abs/2012.15723v2
[43] HU S D, DING N, WANG H, et al. Knowledgeable prompt-tuning: incorporating knowledge into prompt verbalizer for text classification (2021-08-04) [2022-01-11]. https://paperswithcode. com/paper/knowledgeable-prompt-tuning-incorporating
[44] DING N, CHEN Y, HAN X, et al. Prompt-learning for fine-grained entity typing [EB/OL]. (2021-08-24) [2022-01-10]. http://121.199.17.194/paper/1430587541732179968?adv
[45] MA R, ZHOU X, GUI T, et al. Template-free Prompt Tuning for few-shot NER [EB/OL]. (2021-09-28) [2022-01-10]. https://paperswithcode. com/paper/template-free-prompt-tuning-for-few-shot-ner
[46] DATHATHRI S, MADOTTO A, LAN J, et al. Plug and play language models: a simple approach to controlled text generation [EB/OL]. [2022-01-10]. https://paperswithcode. com/paper/plug-and-play-language-models-a-simple
[47] ZOU X, YIN D, ZHONG Q Y, et al. Controllable generation from pre-trained language models via inverse prompting [C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. ACM, 2021. DOI: 10.1145/3447548.3467418
[48] SHIN T, RAZEGHI Y, LOGAN R L, et al. AutoPrompt: eliciting knowledge from language models with automatically generated prompts [C]// Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, 2020. DOI: 10.18653/v1/2020.emnlp-main.346
[49] HAN X, ZHAO W L, DING N, et al. PTR: prompt tuning with rules for text classification [EB/OL]. (2021-05-24) [2022-01-10]. https://paperswithcode.com/paper/ptr-prompt-tuning-with-rules-for-text
[50] LESTER B, AL-RFOU R, CONSTANT N. The power of scale for parameter-efficient prompt tuning [C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2021. DOI: 10.18653/v1/2021.emnlp-main.243
[51] GU Y X, HAN X, LIU Z Y, et al. PPT: pre-trained prompt tuning for few-shot learning [EB/OL]. (2021-09-09) [2022-01-12]. https://paperswithcode. com/paper/ppt-pre-trained-prompt-tuning-for-few-shot
[52] VU T, LESTER B, CONSTANT N, et al. SPoT: better frozen model adaptation through soft prompt transfer [EB/OL]. (2021-10-15)[2022-01-12]. https://paperswithcode. com/paper/spot-better-frozen-model-adaptation-through
[53] PETRONI F, ROCKTASCHEL T, RIEDEL S, et al. Language models as knowledge bases? [EB/OL]. [2022-01-12]. https://paperswithcode. com/paper/language-models-as-knowledge-bases
[54] PETRONI F, LEWIS P, PIKTUS A, et al. How context affects language models´ factual predictions [EB/OL]. (2021-10-15) [2022-01-10]. https://paperswithcode.com/paper/spot-better-frozen-model-adaptation-through
[55] JIANG Z B, XU F F, ARAKI J, et al. How can we know what language models know? [J]. Transactions of the association for computational linguistics, 2020, 8: 423-438. DOI: 10.1162/tacl_a_00324

作 者 简 介

韩旭，清华大学计算机系2017级博士研究生；研究方向为预训练语言模型及知识图谱；已在ACL、EMNLP等会议及期刊发表论文50余篇，出版专著1部。

张正彦，清华大学计算机系在读博士研究生；研究方向为预训练语言模型及其加速；发表论文20余篇，出版专著1部。

刘知远，清华大学计算机系副教授、博士生导师，北京智源人工智能研究院青年科学家；研究方向为知识图谱、预训练模型等；获得多项国家自然科学基金资助；曾获中文信息学会青年创新奖，入选国家青年拔尖人才支持计划、中国科协青年人才托举工程；发表论文80余篇，出版专著5部，Google Scholar统计引用超过1万次。