



04 AI 科技女性

UIUC 季姮教授：叫我带头人，而不是女性带头人

整理：智源社区 涂宇鸽

季姮，美国伊利诺伊大学厄巴纳香槟分校计算机系终身正教授。于清华大学获得学士和硕士学位，纽约大学获得博士学位。致力于自然语言处理研究，尤其是信息抽取和知识库构建。于 2016 和 2017 年被世界经济论坛选为年轻科学家以及全球未来计算委员会委员。她获得的奖项包括人工智能领域十大新星以及美国自然科学基金职业生涯奖。

数据显示高科技领域普遍存在着性别失衡的现象，女性离开科技领域的比例也显著高于男性。学术和职业的金字塔中，女性人才为何不断流失？抵抗负面言论、被实验室拒之门外、遭受性骚扰、忍受不友好的工作环境……女科学家们作为科学界中的“弱势群体”，在 AI 研究中乘风破浪，更加需要非凡的勇气和智慧！

在北京智源大会“AI 科技女性”专题论坛中，美国伊利诺伊大学厄巴纳 - 香槟分校 (UIUC) 教授季姮针对这些问题，做了一场题为“叫我带头人，而不是女性带头人”的演讲。除了简单介绍最近的研究工作外，季姮着重探讨了学术界的女权主义，并结合自身经历，围绕一系列问题展开了交流：女性如何在男性占大多数的领域去带领研究方向往前走？如何避免反向歧视，敢于发声？如何摆脱媳妇熬成婆的心理，培养下一代年轻女性？让我们从这次报告中寻找答案吧！

一、近期研究：从“信息手术师”到新冠病毒“寻药人”

季姮首先介绍了最近的研究工作，称它为“信息手术师”：把网上有误的信息抽取出来纠对，生成正确的故事。过去科技并不发达，缺乏多媒体、多语言的证据，因此人在受教育过程中会接触很多错误信息，比如日本的历史书里没有南京大屠杀的章节，这被称为信息污染。季姮团队想把写历史的过程自动化，并从多个角度全面证实信息的正误，将多个事件联系起来，同时由于历史经常循环重复，还可以从事件演化的序列和关系中学习经验，去推理和预测未来的事件。

由于多语言和多媒体的重要性，季姮已经开展了跨语言迁移和多媒体事件推荐等方面的研究，实现图片资源和自然语言的共享。此外，自然灾害的研究，是季姮团队多年从事的一个研究方向。最近在新冠病毒的研究上，季姮团队处理了 10 万篇文献，考察基因、化学物质、药物、疾病之间的相互关系，抽取牵涉这些实体的事件，自动生成可重复利用的药物的报告。这些都是自然语言处理技术对真实世界很好的帮助和应用。

接下来，季姮切入论坛主题，开始重点探讨女性与科技的关系。

二、女性不是“冒名顶替者”

冒名顶替症候群 (Impostor syndrome)，由临床心理学家 Pauline R. Clance 和 Suzanne A. Imes 在 1978 年提出，是指成功人士定和怀疑自身能力的现象。季姮认为这是女性科学家中频发的心理问题，由于“女性不擅长 STEM 领域”这种普遍存在的刻板印象，她们将取得的成就时归因于运气、时机或是贵人相助，而非认可自身能力和天分。

季姮提醒女科学家：你不是**冒名顶替症候群**！在自然语言处理或者机器学习研究中，女生也很可能是班上数学最好的人，包括立体几何之类等抽象领域的研究。

同时，季姮也呼吁包括男女性在内的非学术研究者，不要认为女性就是冒名顶替综后群。当看到有人会持有偏见，认为一个女生拿到很好的学位，是因为受到女性的特殊优待，作为女性科研者，此时要勇于澄清，我的优秀不关乎性别，只因为我更努力。

三、消失的年轻女性带头人

季姮介绍，从衡量论文数量和质量的 H-index 来看，老一代排名靠前的 81 位的学者中（首次被引于 1998 年前），有 20 位是女性；而年轻学者中（首次被引于 2006 年后），前 65 名里面女性只占 7 名。

“我们不光是在不停地失去女性带头人，而且在失去年轻的女性带头人。”季姮如此表达了她的担忧。



图 1：老一代和年轻一代的 H-index 排名对比

另外，所谓的学术排名，会对一些年轻科学家们造成很大的影响。但是这些排名依据的指标往往存在缺陷，不能公正地反应真实情况，因此我们需要更科学的评价体系。

从 ACL 执行委员会的名单来看，几乎看不到国内女性 NLP 研究者的身影。季姮甚至发现，本次智源大会上主席大多是男性，这也侧面反映了女性在国际学术事务中并不活跃。国内优秀的女性学者都去哪了？季姮希望众多女科学家的才华和成就不被埋没，她们的名字也出现在会议组织方的拟邀请名单里。



图 2：2020 北京智源大会主席名单

季姮指出，作为“隐形人物”的女性，需要付出更多，更需要着力于创新。但比起男性，女性很多时候可能过于安逸，或放弃梦想，或像温水里的青蛙一样安于现状，因此我们要发挥自己的创造力和领导力，勇敢去当驾驶员、去把控方向盘。

四、撕掉性别标签，Lean In

女性自身可能存在某种反向歧视，甚至利用自己的“弱者”身份降低要求。季姮坦言，自己曾经随老板出去开会，20 分钟的会议可能有 10 分钟是在洗手间躲着。这时一定要去掉“女性”的标签，淡化性别意识，走出舒适圈。

在目前从事的一个项目中，季姮作为团队唯一的女性，正带领着一批自然语言处理、计算机视觉等领域卓越的男科学家们。她结合自身的经验表示：不要害怕与男性共事。一个优秀的女性领导者，不仅要破除性别障碍、精益求精、敢于折腾，更要主动承担责任、运用同理心去凝聚团体精神。

The Shiny Sassy DARPA KAIROS team: PIs

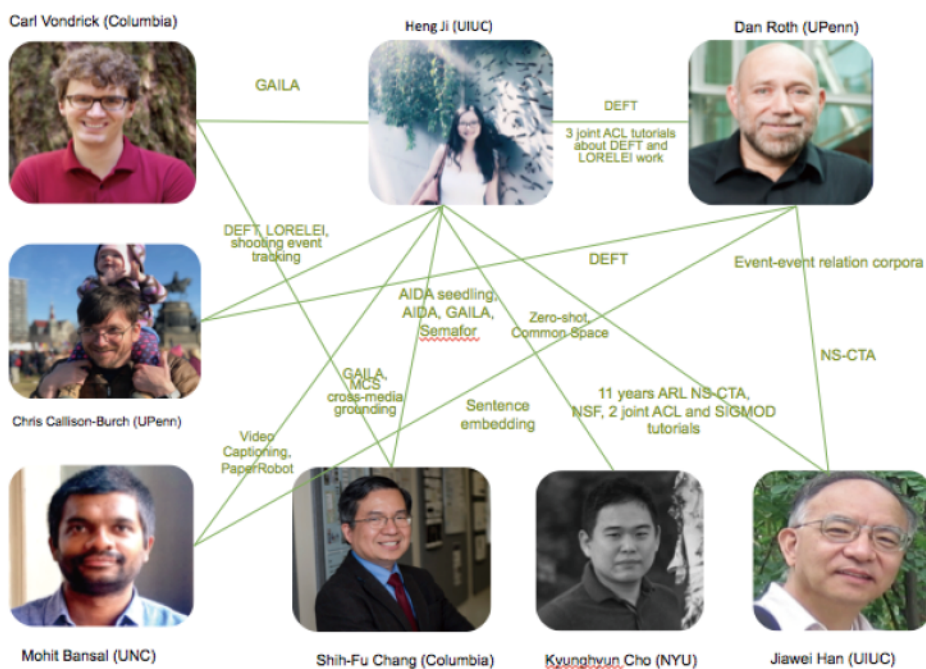


图 3: 季姮教授所在某个团队的成员分布

因此如果规则或者体制允许的情况下，即便是男性占优的情况下，女性也要坐在谈判桌上，不随便把机会让出，勇于站到聚光灯下发声，如此才能推动女性在科学界的进步与发展。季姮总结到：“你要成为变化的一部分，而不是等着变化发生。”

五、避免“媳妇熬成婆”，热心培养下一代年轻女性

季姮认为，一些女性在受到歧视、冷落、物化和不平等对待后，会将同样的行为施加在年轻的下一代女性身上，这是一种“媳妇熬成婆”的心态。我们应该善待年轻的女性，因为她们身上有自己的影子。

Be Nice to the Younger You

- Your bad experience is not an excuse to be mean to other younger women
- Don't be the 'woman token' all the time; train next generations and woman to
- It's ok to be a child like the younger selves occasionally
- It's ok to break rules from time to time, be a fun and favorite academic mom
- Educate people: reject sexist / racist man/woman students even if they are geniuses
- If you have to be jealous, get jealous of men and women equally; get jealous of their work, not appearance or age; the only person you should compare to is yourself from yesterday
- Create more initiatives:
 - Grace Hopper Conference, Rising Star in CS/EE, one on one mentoring
 - Create and maintain "Excellent woman AI/NLP researchers in China" list



图 4：平等、友善地对待青年一代

季娟回忆了去年 5 月份去德国的一次经历，遇见两个女孩子很害怕她们的老师——一个德国老太太，后来我和她们交流了一下眼神，就一起开溜吃冰激凌了。季娟和这些女生打成一片后，一直保持联系和合作。季娟因此觉得，如果以这种平等、友善的合作者心态对待自己的学生和青年一代，一切会变的更好。

最后，谈到生活和工作的平衡，季娟认为这并不是单项选择题。工作做到有声有色的同时，生活亦可以五彩斑斓。只要你无所畏惧，勇敢去做！

You can Have it All

- It's ok to be yourself and enjoy life.
- It's ok to dress up or down, speak with a louder or softer voice, solely your choice.
- It doesn't interest me how old you are. I want to know if you will risk looking like a fool, for love, for your dream, for the adventure of being alive.
- To other men and women: how we look is not your business. Stop commenting on them if we are not close friends.
- As you rise many people will disapprove. Rise anyway.

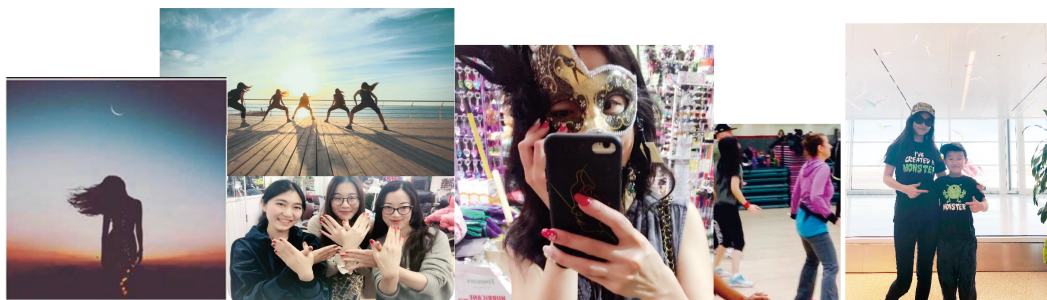


图 5：工作与生活的平衡

北京大学教授张铭：基于知识图谱的机器学习

整理：智源社区 涂宇鸽

张铭，北京大学信息科学技术学院教授，研究方向为文本挖掘和机器学习。目前主持国家科技部科技创新2030——“新一代人工智能”重点研发项目课题、国家自然科学基金面上项目等。合作发表科研学术论文 200 多篇，谷歌学术被引 8200 余次，H 因子 36，获得机器学习顶级会议 ICML 2014 最佳论文奖、网络信息处理顶级会议 WWW 2016 最佳论文提名。

尽管人工智能依靠机器学习和深度学习取得了快速进展，但严重依赖于人类的监督以及大量的标注数据，成本高昂，且缺乏可解释性。图灵奖获得者、知识工程鼻祖费根鲍姆曾强调过知识中蕴含的巨大力量，单单依靠统计模式、而缺乏符号化知识结构的人工智能，离强人工智能仍具有较大差距。如何利用丰富的背景知识实现机器认知？如何打开 AI 的黑箱、提高智能系统的可解释性，从而促进智能决策的落地？知识图谱和机器学习的结合或许会是上述问题的答案。

在北京智源大会“AI 科技女性”专题论坛上，北京大学信息科学技术学院教授张铭介绍了自己在基于知识图谱的机器学习研究中的成果和心得。

一、文本挖掘经典任务：文本分类

文本分类首先需要人工标注文档，然后对这些做好标签的文本进行训练，用来预测新文本的类别。标注工作非常耗时，而且一般需要专家进行。为表达这些文本信息，最常用的一种简化模型就是词袋模型 (bag-of-words model)：将文本视为若干词语的集合，忽略文法及词序，将词频作为训练分类器的特征。



- 一个经典的文本挖掘任务，对很多应用都很有用！
 - 社交网络分析, 健康, 机器阅读 ...
- 传统的方法:



图 1：文本挖掘经典任务—文本分类

然而，文本信息是一种非常经典的蕴含大量知识的信息形式。因此，如果仅使用词袋模型，文本分类的质量不会很好。比如下面这段文本中，“game”表示体育还是游戏比赛，要从上下文语境、相关的知识来看。

Dong Nguyen announced that he would be removing his hit game **Flappy Bird** from both the **iOS** and **Android app stores**, saying that the success of the **game** is something he never wanted. Some fans of the **game** took it personally, replying that they would either kill Nguyen or kill themselves if he followed through with his decision.

选择标签:

Mobile Game or Sports

- **标签** 包含了丰富的语义信息!
- 传统分类聚类方法基本上没有用到标签信息
 - 每一个文档只是被标记为“文档 ID”

图 2: 基于知识的文本分类

二、文本挖掘算法的问题

像前面的文本分类例子中提到的，机器学习算法已经广泛应用于文本挖掘，然而标记工作费用高昂，通常需要雇用领域专家。后来通过众包解决标注问题，但质量较低、花费也不小。因此，机器学习界一直致力于减少有监督的机器学习算法中的标记工作，例如半监督学习提出只使用部分标记和大量未标记的数据来进行学习，达到完全监督学习的效果，迁移学习使用来自其他相关领域的标记数据来辅助目标领域的学习任务。

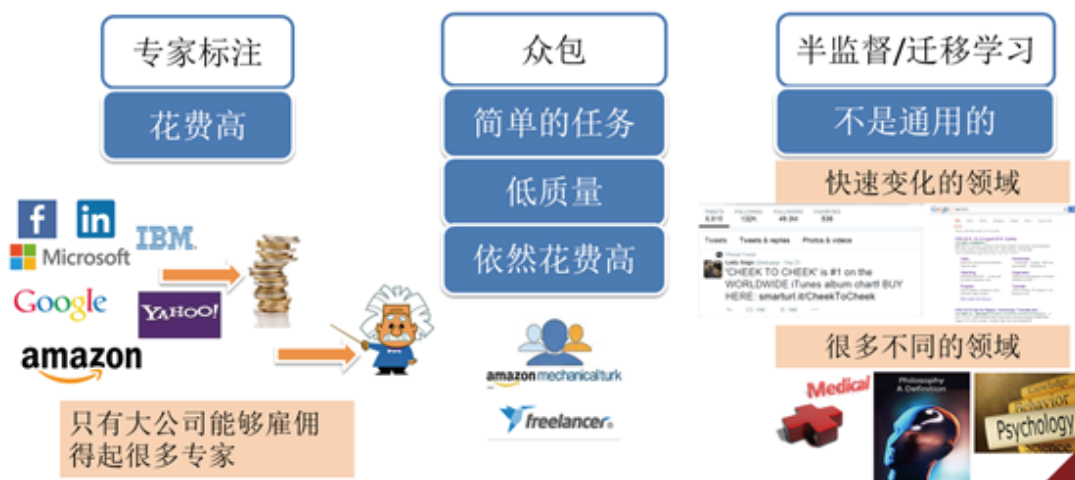


图 3: 文本挖掘算法的问题

然而，很多情况下，无论是半监督学习还是迁移学习都无济于事。例如大数据时代的文本信息涉及多个不同领域，比如医学领域出版物分类，因此需要一种更通用的方法来进一步降低不同领域学习任务的标记成本。

与此同时，无数的互联网用户为维基百科贡献了知识，使得它成为一个质量较高的通用知识框架。这些通用知识包含丰富的语义信息且具有一定的精确度和层次性。这些通用的知识图谱不仅包含大量通用知识，也包含多

样化的专业领域，再加入专业语料库的大量实体和语义关联的训练，可建立专业领域的知识图谱。

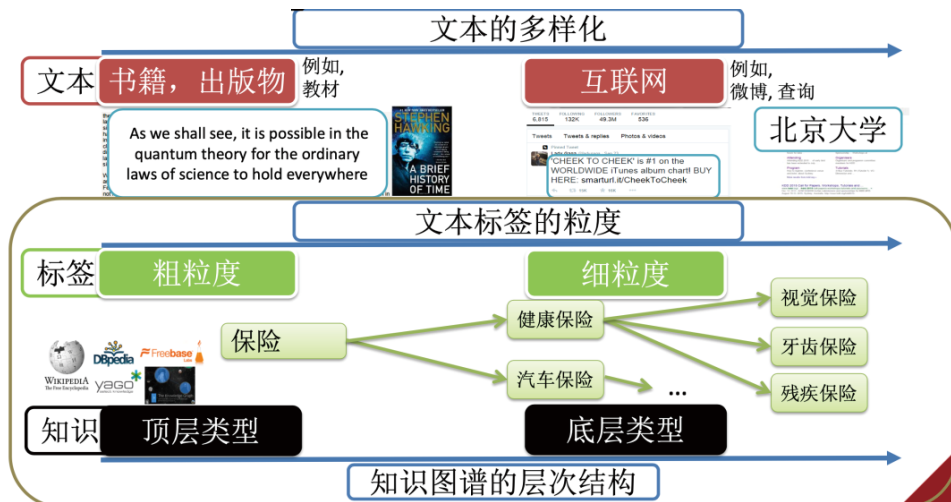


图 4: 文本挖掘算法的问题

三、基于知识图谱构造文本的异构信息网络表示及算法

知识图谱 (Knowledge Graph) 是多个三元组 (e_1, r, e_2) 组成的语义网络，主要用于描述现实世界中的实体、概念及事件间的客观关系。实体 (entity) 或概念 (concept) 是知识图谱的结点 (node)，它们之间的语义关系是实体结点的边。因此知识图谱可以进一步丰富文本信息，为文本提供更多背景知识，并在此基础上做一些文本挖掘的任务。

注: (e_1, r, e_2) 中 e_1 和 e_2 代表一条关系的头结点和尾结点， r 代表关系。

- 一个知识图谱 K 由三元组 (断言) $(e_1, p, e_2) \in E \times R \times E$ 组成
 - 令 E 表示实体集合 (例如，奥巴马)
 - R 表示属性或者关系集合 (例如，出生地点)
 - 例如 (奥巴马，出生地点，檀香山)



图 5: 知识图谱的定义

通过知识图谱链接丰富后的文本信息可以表示为一种含有多种实体、关系和类型的异构信息网络。信息网络是知识表示的结构化文本方式，网络中包含一系列结点以及结点和结点之间的边。异构信息网络 (Heterogeneous

Information Network, HIN) 通过分析网络中多种类型实体及其链接关系, 能够准确地区分信息网络中路径的不同语义。由此, 可以将文本分类问题描述为一个 HIN 划分问题, 并通过在 HIN 中引入丰富的结构信息作为约束条件, 形成一种新的聚类算法。

如图的两段文本, 就词袋模型而言并不能很好地划分类别。但如果把这些实体经过语义分析, 定位到知识图谱的框架, 能够发现奥巴马和布什之间的链接“美国总统”, 因此可以归为政治类。可以看出实体、链接和类型中包含的丰富信息, 异构信息网络中的特定化知识, 能够帮助我们更好地完成文本分类及其他机器学习的任务。

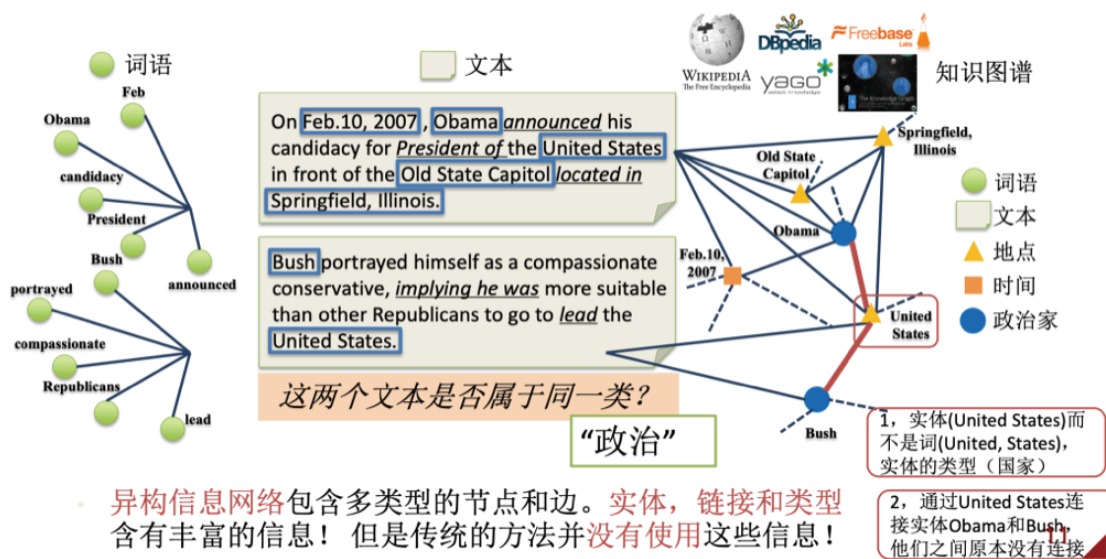


图 6: 基于知识图谱构造异构信息网络

四、基于知识图谱的机器学习和文本挖掘

张铭团队提出了一种基于通用知识约束的文本分类方法, 其中的约束来自异构信息网络中的多个类型和子类型。首先使用语义分析方法得到文本的实体及关系短语, 通过将文本中的实体及关系短语作为结点, 与知识图谱里相应的概念结点进行连接, 形成结合领域知识的文本异构信息网络表示。

原来很多文本挖掘是以词袋模型来处理, 不能很好地表示关系。现在把文档架构成图模型, 很多机器学习的任务转变为图上的任务, 除了词向量的 Cosine 夹角相似度, 现在可以计算子图的相似度、并结合其他算法。

例如, 张铭团队提出了一种新的基于异构信息网络的无结构数据的文档相似性度量 KnowSim, 使用多类型实体所组成的关系元路径作为关键信息。元路径是指一条通过一系列关系连接对象类型的特定路径。例如, A — O — A 这条元路径表示作者之间通过组织(即同事)之间的关系连接。

KnowSim: 定义在有结构异构信息网络上的无结构数据相似性度量。

语义重叠: 连接两个文本的元路径的数量。

$$KS(d_i, d_j) = \frac{2 \times \sum_m^{M'} w_m |\{p_{i \rightarrow j} \in P_m\}|}{\sum_m^{M'} w_m |\{p_{i \rightarrow i} \in P_m\}| + \sum_m^{M'} w_m |\{p_{j \rightarrow j} \in P_m\}|}$$

语义广度: 连接文本到文本自己的元路径的数量。

解释: 文档之间的重要的元路径越多, 就越相似, 然后做平滑

KnowSim的计算在几乎线性时间完成: 基于元路径独立随机游走及选择方法

KnowSim: A Document Similarity Measure on Structured Heterogeneous Information Networks. ICDM'15

图 7: KnowSim, 基于多阶元路径的相似度核函数

KnowSim 包含两部分: (1) 语义重叠作为分子, 表示连接两个文本的元路径数量。(2) 语义广度作为分母, 表示连接文本到文本自己的总元路径的数量。文档之间的重要的元路径越多, 就越相似。KnowSim 考虑了元路径的立体信息, 比原来文本扁平特征更好。由于考虑了语义信息、文本之间语义相似度的信息, 文档间相似度量比普通的 SVM 和词的嵌入效果更佳。

平均准确性						
模型	SVM ^{HIN}	SVM ^{HIN} +KnowSim		IndefSVM ^{HIN} +KnowSim		SVM
设置		DWD	DWD+MetaPath	DWD	DWD+MetaPath	+Word Embedding
2ONG-SIM	91.60%	92.32%	92.68%	92.65%	93.38%	91.67%
2ONG-DIF	97.20%	97.83%	98.01%	98.13%	98.45%	98.27%
GCAG-SIM	94.82%	95.29%	96.04%	95.63%	98.10%	96.81%
GCAT-DIF	91.19%	90.70%	91.88%	91.63%	93.51%	90.64%

发现#1: 所有基于DWD+MP的核方法都比SVM^{HIN}的效果好

将元路径的结构信息当做一个整体来考虑要比当做扁平的特征的效果好。

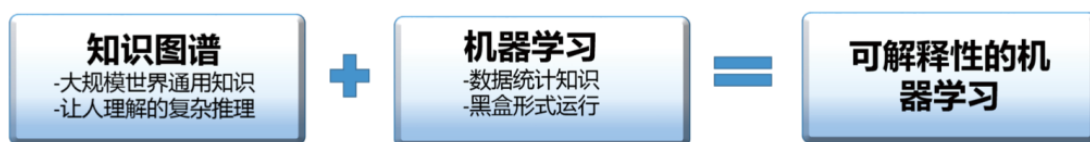
发现#2: IndefSVM^{HIN}+KnowSim比SVM+WE的效果好

知识包含更多文本之间相似度的语义信息。

图 8: 文本分类结果

五、可解释的推荐系统

知识图谱与机器学习结合, 能够充分利用知识图谱提供的常识信息和显性语义信息, 通过统计学习把隐含语义表达出来, 形成可解释性的机器学习, 易于被人接受和理解。



推荐系统领域的知识图谱引导下的可解释性算法研究

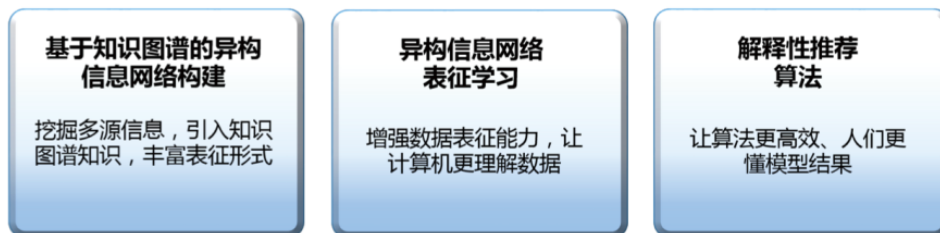


图 9：知识图谱与机器学习结合

一个经典的实例是张铭教授团队最近完成的北京市科委项目，基于知识图谱的可解释的推荐系统。通过引入知识图谱知识，构造一个包含用户、商品交互关系的语义丰富的异构信息网络，并在此基础上进行有效可解释的表征，实现可解释的机器学习算法，具体应用到推荐系统里面，使得个性化推荐更加精确和有迹可循。

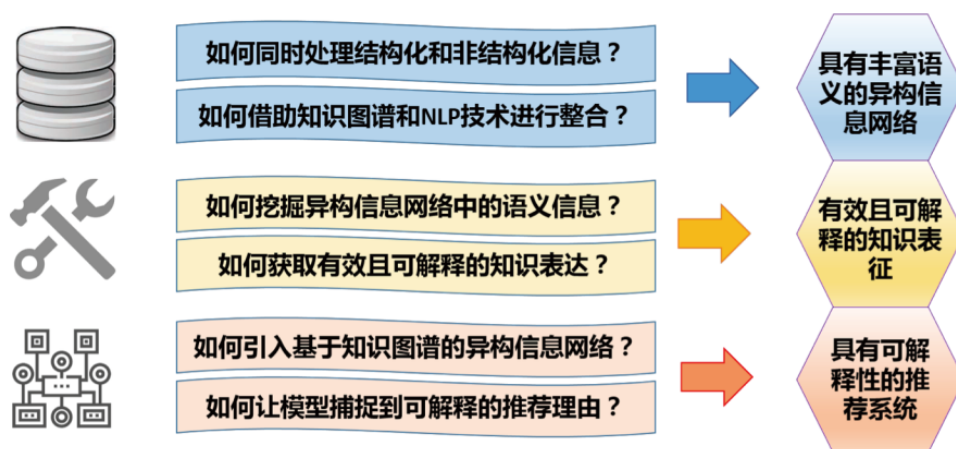


图 10：可解释的推荐系统

原有的推荐算法是用户和物品之间的匹配问题，现在把用户商品交互构造为图，通过将图上路径的选择建模为一个强化学习的过程，使得推荐路径比较清楚、具有可解释性，而且准确率较高。

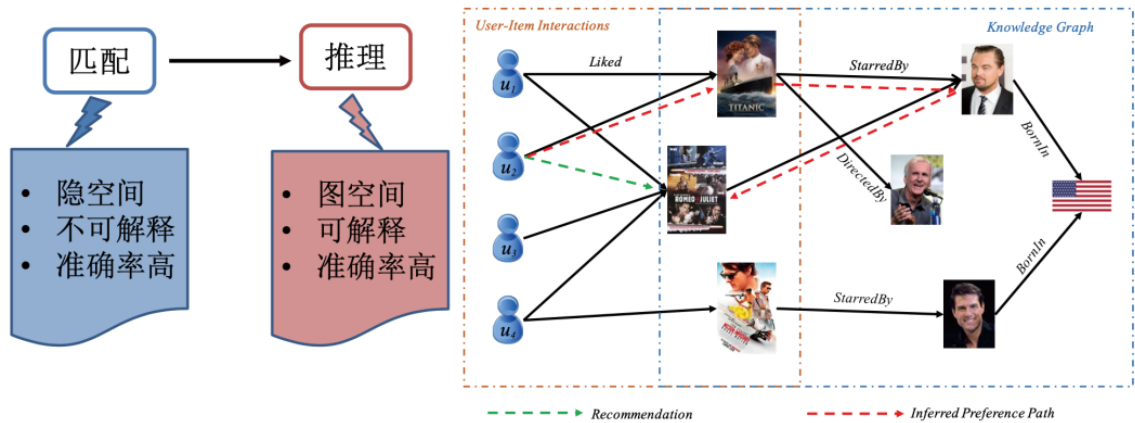


图 11：基于异构信息网络的推荐算法

具体地，将用户到物品的路径生成问题描述为一个异构信息网络上的马尔可夫决策过程，其中目标用户作为推荐代理的初始状态，当前结点在图上的行走被定义为动作，通过不断在相邻结点中选择下一动作，直到达到预定义的最大路径长度完成推荐路径选择。在训练过程中，由于无法判断路径选择过程中中间状态的优劣，代理不会收到即时的奖励，最终奖励取决于代理是否成功走到用户感兴趣的物品结点。

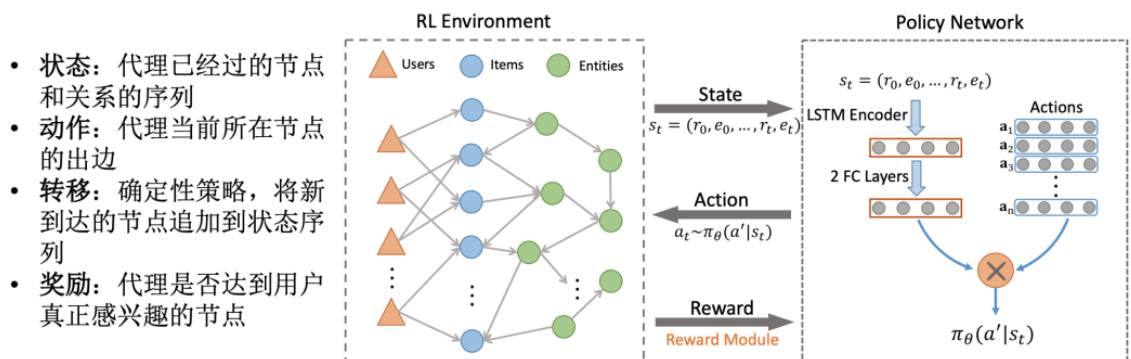
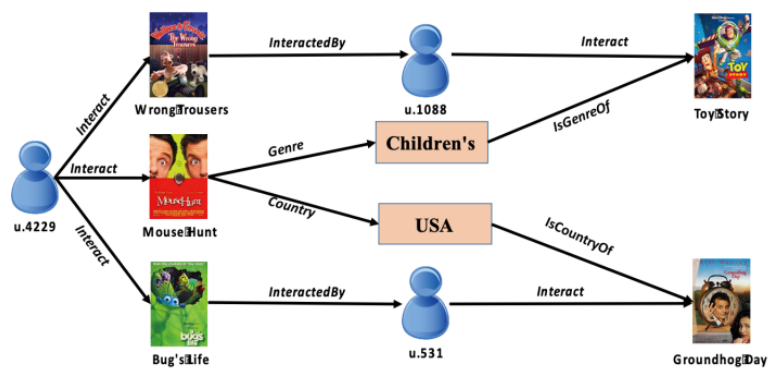


图 12：基于异构信息网络的推荐算法训练过程

图中展示了推荐效果，可以看出推荐的电影与目标用户之前看过的电影是语义关联的，虽然都是儿童主题但推荐具有不同角度令人信服的解释。



- 代理所推理出的**路径**作为推荐结果的解释
- **多路径**提供多角度的、令人信服的解释

图 13：效果展示

另一个实例是计算机领域知识图谱应用于中国大学 MOOC 资源推荐。通过强化学习，得到用户到学习资源的可解释推荐路径。

• 计算机领域知识图谱构建

- 课程体系规范CS2013、CE2016，中国大学MOOC计算机课程资源
- 知识点实体识别、关系抽取

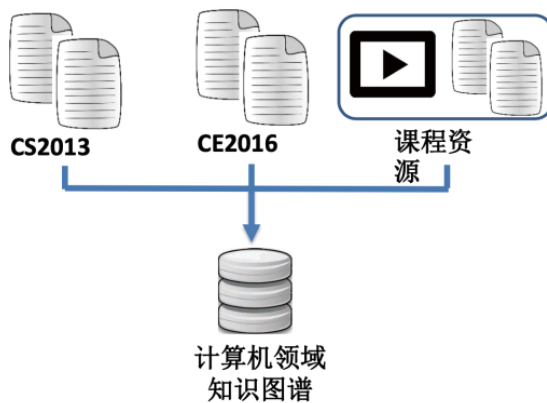


图 14：中国大学 MOOC 学习资源推荐

六、展望

未来，张铭仍将致力于提出一种利用知识图谱扩充文本信息的通用框架，通过构建相关领域知识图谱，进行更有效的蕴含知识的文本挖掘研究。

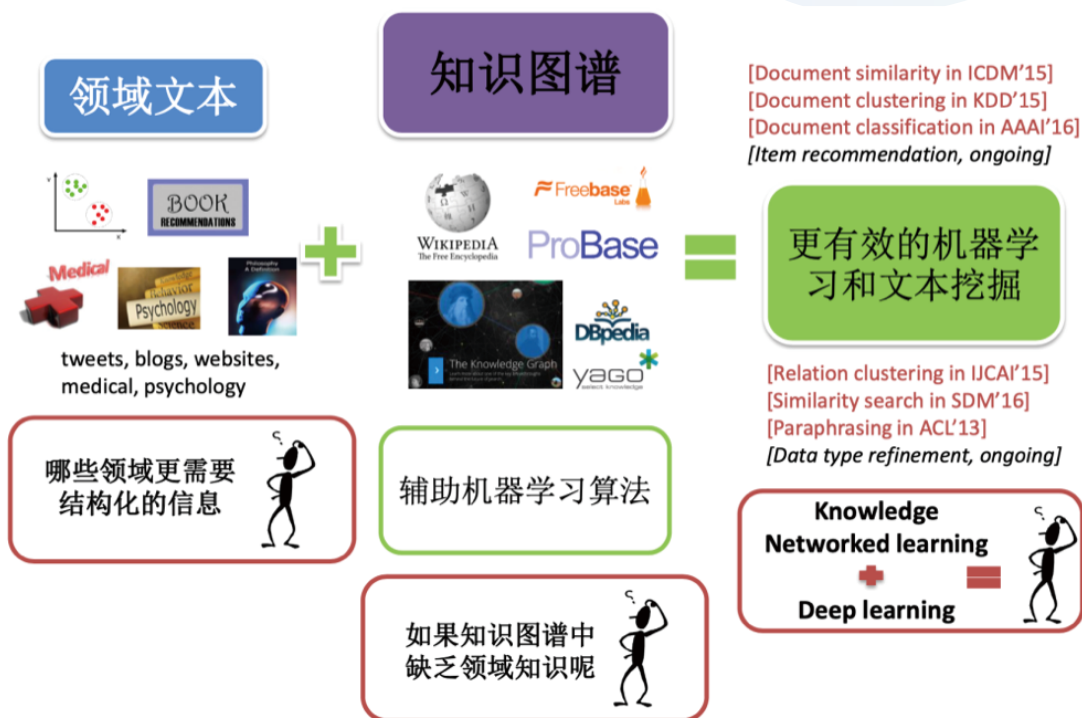


图 15: 未来研究展望

参考文献:

- [1] Chenguang Wang, Yizhou Sun, Yanglei Song, Jiawei Han, Yangqiu Song, Lidan Wang, and Ming Zhang: RelSim: Relation Similarity Search in Schema-Rich Heterogeneous Information Networks. Proc. 2016 SIAM Int. Conf. on Data Mining (SDM' 16).
- [2] Chenguang Wang, Yangqiu Song, Haoran Li, Ming Zhang, and Jiawei Han: Text Classification with Heterogeneous Information Network Kernels. Proc. 2016 AAAI Conf. on Artificial Intelligence (AAAI' 16).
- [3] Chenguang Wang, Yangqiu Song, Haoran Li, Ming Zhang, and Jiawei Han: KnowSim: A Document Similarity Measure on Structured Heterogeneous Information Networks. Proc. of 2014 IEEE Int. Conf. on Data Mining (ICDM' 15).
- [4] Chenguang Wang, Yangqiu Song, Ahmed El-Kishky, Dan Roth, Ming Zhang, and Jiawei Han: Incorporating World Knowledge to Document Clustering via Heterogeneous Information Networks. Proc. 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD' 15).
- [5] Weiping Song, Zhiping Xiao, Yifan Wang, Laurent Charlin, Ming Zhang and Jian Tang, Session-based Social Recommendation via Dynamic Graph Attention Networks. WSDM 2019: 555–563.
- [6] Weiping Song, Chence Shi, Zhiping Xiao, Zhijian Duan, Yewen Xu, Ming Zhang, Jian Tang: AutoInt: Automatic Feature Interaction Learning via Self-Attentive Neural Networks. CIKM 2019: 1161–1170.
- [7] Weiping Song, Zhijian Duan, Ziqing Yang, Hao Zhu, Ming Zhang, Jian Tang: Explainable Knowledge Graph-based Recommendation via Deep Reinforcement Learning. CoRR abs/1906.09506 (2019)

哈工大教授秦兵：打开人工智能情感之门

整理：智源社区 何文莉

秦兵，哈尔滨工业大学计算机学院教授，主要研究方向为自然语言处理、知识图谱构建、文本情感分析及文本生成等。在顶级国际会议 ACL、COLING、EMNLP、IEEE TKDE、IEEE TASLP 等国内外重要期刊及会议上发表论文 100 余篇，担任多个会议程序委员会主席以及领域主席。

秦兵本次演讲的主题是《打开人工智能情感之门》。

秦兵从两个维度出发，技术和感悟。技术方面，介绍了什么是情感计算，情感计算都面临哪些挑战，如何解决这些挑战以及情感计算的应用与发展趋势。感悟方面，主要跟专题内容相关，通过数据、模型的独特视角对科研女性进行解读。以下为秦兵演讲的内容概要，供读者学习。

一、学术：赋予计算机情感计算能力

1.1 情感计算的五个挑战

人工智能的学术方面，秦兵的研究主要有三个部分：情感的计算、知识图谱构建与应用、文本生成。本次报告，主要围绕计算机情感计算展开。

从一部电影说起。电影《她》讲述了在不远的未来人与人工智能相爱的科幻爱情电影。人与机器人坠入爱河，在现在好像不太现实，但或许未来将不是科幻。如何赋予计算机情感计算能力，是当前学术界和企业界研究的一大热点问题。

人类情感与机器情感。情感是什么？心理学大辞典中最初始的定义，所谓情感是人对客观事物是否满足自己的需要而产生的态度体验，包含了丰富的主观情绪和感受。在人工情感问题上，人工智能之父 MINSKY 在 1985 年曾说到“问题不在于智能机器能否拥有任何情感，而在于机器实现智能时怎么能够没有情感”。机器的情感计算，是指机器理解人类情感以及生成情感的能力。不仅需要它在跟人类进行交互过程中，理解人类的情感，知道人类的喜怒哀乐，同时还要有自己的情感。

机器的“类人”语言能力。机器的“类人”语言能力体现为理性与感性兼备，一是理解与推理能力，二是情感与表达能力，这就需要语篇理解与推理、情感语义计算技术的支持，其背后蕴藏着众多待解决的科学问题。

从机器的类人语言能力角度，机器智能可概括为计算智能、感知智能和认知智能三类。计算智能，相比人类，计算机的先天优势在于强大的存储能力和计算能力。感知智能，在语音识别、图象识别等任务上，计算机可以和人类相媲美。认知智能，体现的是认知能力，这方面机器和人相比还存在一定的差距。

秦教授强调认知能力主要体现在理性和感性上，分别是具备理解和思考的理性能力与具备情感识别和生成的感性能力。具体来说理解和思考的能力，是对于给定的一个现象或事物，机器能否准确理解，从而做出有效推理。

情感识别与生成能力，是指当人在倾述某种感情时，机器是否能够理解用户此刻的心情，评价某一个产品、某一个服务，或某一项指标时，机器能否判断褒义还是贬义。

情感计算与人工智能。情感计算是重要的技术支撑，它与人工智能情感机器人的目的在于对人类特有情感的模仿和超越，但与现有人工智能技术的高度发展相比，人工情感技术所取得的进展还略显薄弱。情感始终是横跨在人脑和电脑之间一条无法逾越的鸿沟。

情感计算。现在人工智能在不断进步，情感计算的使命是赋予其情感和温度，这是一个有趣且具有挑战的问题。网络评论、博客、微博、社交媒体的文本中，蕴含大量的情感和情绪，情感计算的任务是利用计算的方法对这些情感、情绪进行挖掘。情感计算不仅是自然语言处理领域的一个研究方向，也是计算机科学、管理科学和社会科学的交叉融合。那么，情感计算的研究都有哪些挑战？秦兵总结了以下五点：

挑战一：情感词汇的歧义性



图 1：情感词汇的歧义性

歧义性，图 1 中同样一个词“快”，在不同语境下有不同的语义，确定词的具体语义需要结合上下文进行分析。

挑战二：情感表达的含蓄性

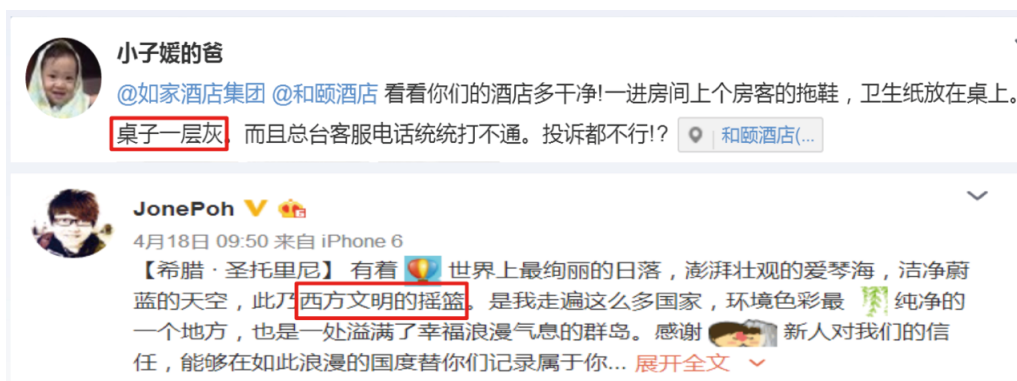


图 2：情感表达的含蓄性

“澎湃壮观的爱琴海，洁净蔚蓝的天空，此乃西方文明的摇篮”，表面看来是对事实的描述，没有任何情感词汇，实际则在抒发赞美之情。传统的情感分析方法多依赖显式情感词，社交媒体中大量存在的隐式情感表达处理上还存在一定难度。

挑战三：情感可解释性



图 3：情感可解释性

深度理解情感，不仅要理解人们表达了什么情感，更要理解情感产生的原因。伤心的原因是考试考砸了，高兴的原因是收到理想大学的通知。在研究情感类别的时候，除了要知道当时的状态，还要知道其原因，正所谓知其然也要知其所以然。

挑战四：情感表达一致性

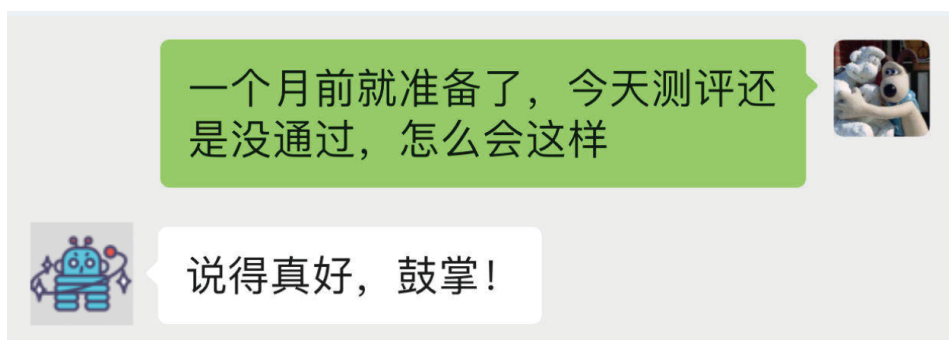


图 4：情感表达一致性

情感表达一致性是指系统分析得到的情感应该与实际表达保持一致，并能做出正确回应。在图 4 的对话中，系统需要识别用户的情绪，并做出相应的回复。若识别出的情绪不太好，要怎么鼓励或者安慰他。但是如果系统不具备正确识别的能力，没有生成符合情感语境下的回复，可能就会答非所问，导致用户的失望。

相反的词语被映射到相近的词向量上，比如图 6 中的好和坏，上下文非常相似，但情感极性相反。最后导致在情感分析里，“好”和“坏”的词向量语义表示非常接近。秦教授学生 2014 年发表在 ACL 上的文章中提出了融合文本情感极性的词向量学习算法，很好的解决了这个问题，得到学术界广泛认可。

1.3 隐式情感分析

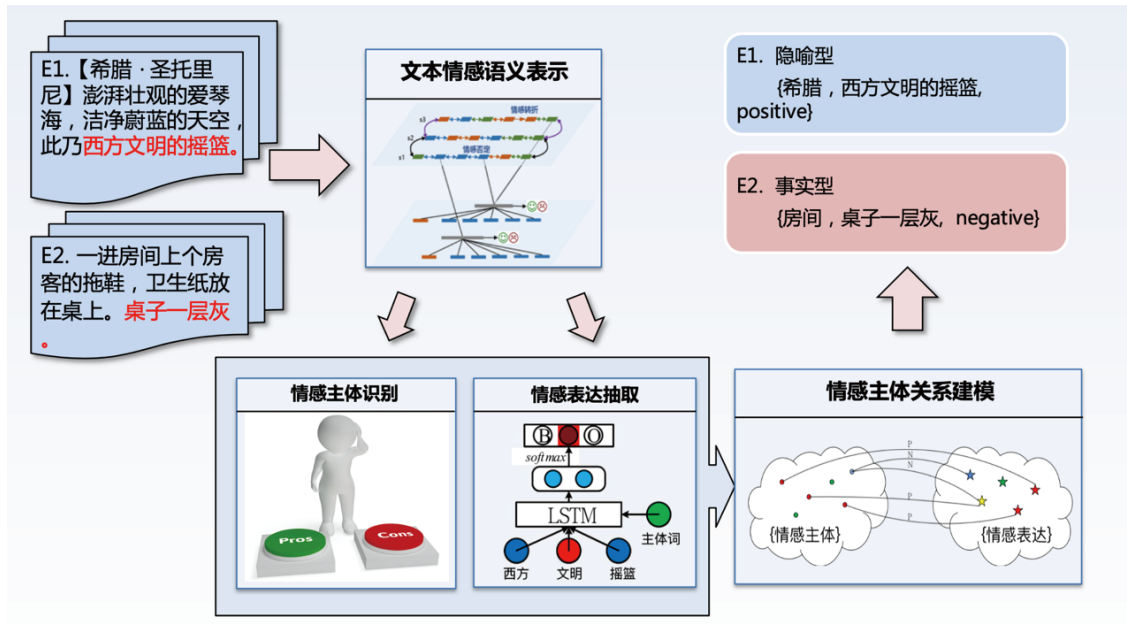


图 7：隐式情感分析

隐式情感分析，主要针对情感表达的含蓄性。对于类似“西方文明的摇篮”、“桌子上一层灰”的隐式表达，首先对情感主体进行识别，抽取其情感表达，再建立情感主体关系模型。

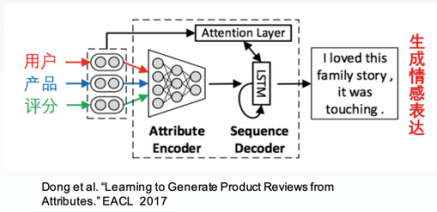
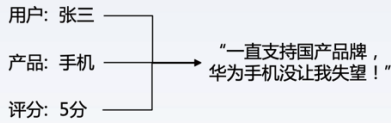
1.4 情感原因发现和预测

情感的可解释性，即情感原因发现和预测，需要从多个维度进行考核，结合更多的元素。在产生一个用户评论时，除了结合社会焦点事件，有的时候还要考虑其所在的群体。所在群体不同，持有的观点不同。将事件本身的表示文本，用户的评论文本，用户行为特征聚集起来，共同成为一个情感的表示，进行原因的发现和预测，可达到更好的可解释性。

1.5 情感生成

情感评论生成

输入产品信息，生成相应的情感表达



情感回复生成

输入上文内容，生成相应的情感回复

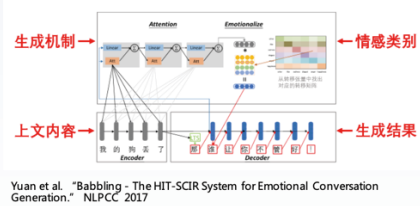
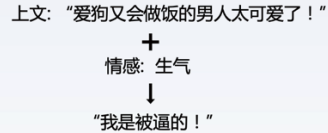


图 8: 情感生成

情感生成的一致性，比如情感评论生成，根据用户、产品和评分生成具体相应情感的评论。又比如情感回复的生成，文本“爱狗又会做饭的男生太可爱了”+ 情感“生气”，最后给出相应的结果。情感生成是可以指定的，只需在生成机制里给出情感类别。

1.6 多模态情感表示与识别

输入一段视频，输出其对应的情感标签

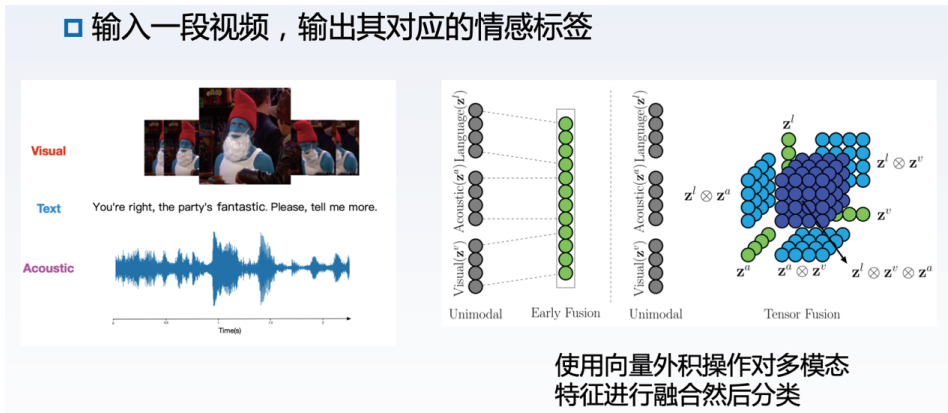


图 9: 多模态情感识别

输入一段视频可以输出其对应的情感标签，比如图 9，将视频图象、文本图象和声音的信号集中到一起，在某一个时间段，根据描写的词、音调以及图片情况，对情感进行判别。使用时，可以通过矩阵的方式直接拼接多模态各个通道的向量进行融合，也可以采用更复杂的方法进行融合。由于深度学习的发展，目前多模态的研究越来越多。

1.7 情感分析技术应用。

在商品 / 服务评论分析方面，对评价对象和评价表达进行抽取，识别评论中的情感倾向性，可辅助消费者挑选产品以及指导商家改进商品 / 服务。在舆情分析方面，对热点事件中的情绪进行剖析，寻找情感原因。在情感对话方面，情感机器人（虚拟 & 实体）可用于体恤民情，抚慰心灵，充当情感陪护。



图 10：情感分析技术应用

根据以上分析，总结未来情感计算发展有五个趋势：

- 第一从显式情感分析到隐式情感分析，深入挖掘情感的语义。
- 第二从单模态大多模态的情感分析，引起了多领域学者关注。
- 第三从特意领域到多领域的迁移，推动跨领域情感分析的研究。
- 第四从情感分析到情感生成，让机器人产生情感。
- 第五从研究到应用，情感计算部分的成熟技术在更多行业应用落地。

二、感悟：模型趣解科研女性

说完技术，再来谈谈对“科技女性”的感悟，看秦兵如何用模型趣解科研女性。

人类与科技。男性 + 科技等于科技伟人，男科学家有很多。然而女性 + 科技等于科技女性，虽然也有一些有名的女科学家，例如屠呦呦等，但比例来讲显然不如男性。

机器视角：女性与科学研究。语言模型是计算一个句子概率大小的统计模型。从机器视角看科技女性，在 GPT2 语言模型下，对于女性理想的职业是什么？GPT2 生成的是 Nurse，而男性则是 Artisan 和 Farmer，可见男女职业是有区别的。

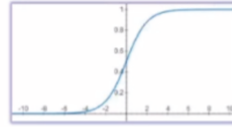
偏见还是偏置？从事科研工作女性比较少，是偏见还是偏置？偏见的一种心理，比如女司机、女博士。偏置是数据统计的一种现象，比如统计网络数据，但网络上年轻人多，又比如统计男女比例，但工科院校女生少。

不同阶段的男女比例。全国总人口男女比例大约为 1.05:1。工科院校，以某个工科学校为例，本科生男女比例为 5:1，研究生男女比例 4:1，教师男女比例 5.5:1，教授男女比例是 4:1。显然，在工科领域男女比例上存在明显差异。

模型趣解。从上述可以发现，很多角度下，男女都是存在区别的。那么通过模型可以有哪些分析？

• 逻辑回归模型

- 二分类模型



$$Y = \sigma(W^T X + b) = W_1 X_1 + W_2 X_2 + W_3 X_3 + \dots + b$$

成为科研人员的可能性 ← Y
 $W^T X$ ← 特征权重
 X ← 个人特征

- 通常都是固定输入优化参数，现在我们**固定权重优化个人**

假设：一组特征维度及权值

$$W = [W_1(\text{男性}) = 0.8, W_2(\text{女性}) = 0.2, W_3(\text{学历}) = 0.6, W_4(\text{逻辑能力}) = 0.8, W_5(\text{英语}) = 0.7, W_6(\text{读书}) = 0.5, W_7(\text{绘画}) = 0.3, W_8(\text{健身}) = 0.4]$$

图 11：二分类模型

通过逻辑回归模型的二分类模型研究在特征的权重以及个人的特征的影响下，成为一个科研人员的可能性。根据样例数据得出权重，分析在权重固定的情况下怎么样优化个人特征。模型设计了一组科研领域特征纬度，假设固定权重男性占比 80%，女性占比 20%，问题在于如何通过其他特征使自己更加匹配，增大可能性。

• 一个例子（某一科研女性：小红）

特征数值

$$\text{小红} = [X_1 = \text{男性} = 0, X_2 = \text{女性} = 1, X_3 = \text{硕士} = 0.8, X_4 = \text{逻辑能力} = 0.2, X_5 = \text{英语} = 0.8, X_6 = \text{读书} = 0.4, X_7 = \text{绘画} = 0.5, X_8 = \text{健身} = 0.2]$$

$$Y = \sigma(W^T X + b) = W_1 X_1 + W_2 X_2 + W_3 X_3 + \dots + b$$

- 群体指标（ W_2 (女性) = 0.2）导致女性在科研工作中比例不高原因：

- 由于之前的偏见观念，不敢涉足
- 不擅长这个领域
- 家庭负担等

- 小红想成为合格的科研女性需要具有较大的 $W^T X$

图 12：一个例子

以某一科研女性小红为例，其各项特征值如图 12。用逻辑回归计算，要使 Y 增大，一是特征权重，二是个人特征。由于之前的偏见观念，自身不擅长这个领域以及家庭负担等，女性的固定特征权重很低 ($W=0.2$)。剩下的选择是提高 X，增加自己的技能，优化个人特征。所以在科研领域，一些女性特征相对弱势，需要花费更多的精力增强其它特征。

• Stack-propagation

- 一种多任务学习框架
- 任务之间有层次依赖关系

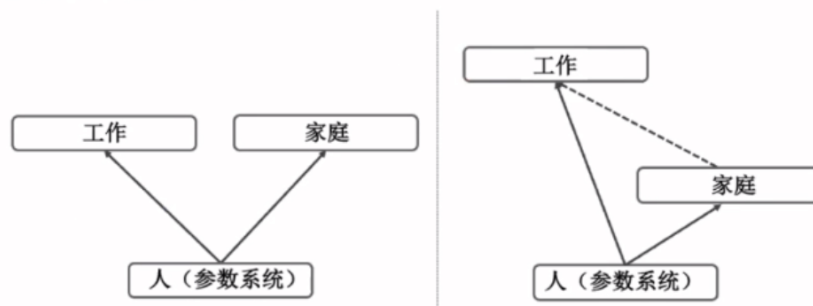


图 13: Stack-propagation 模型

如何在工作和生活中进行平衡，或许通过 Stack-propagation 模型可以得到一些启发。一个人同时忙工作和忙生活要付出很多，采用这个模型可以借力，管理一个家庭或某些问题上的经验，可以应用到工作中，工作当中一些经验方法可以反馈到家庭事务管理。通过多任务学习的框架思想，协调处理工作和家庭关系，使工作和家庭相互促进，都能得以提高。

三、结语

本次秦兵的演讲围绕“智能、情感、女性”三个关键词展开，演讲内容不仅有技术，更有温度。赋予计算机情感是一个充满魅力的学术研究问题，在秦教授的分享中我们不仅领略了计算机的情感世界，也对科研女性从新的视角有了新的感悟。人工智能飞速发展，改变着我们的时代，但这一切都应以生活为出发点，正如秦兵演讲结束时所说的一样：“为机器赋予语言或情感的智能，让生活更加美丽”。

北京大学教授穗志方：从语言到知识——构建语言智能的基石

整理：智源社区 涂宇鸽

穗志方，北京大学计算机科学技术系教授，智源研究员，博士生导师，计算语言学教育部重点实验室主任，中国中文信息学会理事，第十四届汉语词汇语义学国际研讨会程序委员会主席。研究领域为计算语言学、文本挖掘与知识工程。在计算语言学及人工智能领域发表多篇学术论文，组织中文计算系列国际评测，制订中文语料库构建相关国家标准 2 项。

提到语言智能，大家或许会想到一个 AI 历史上鼎鼎有名的 IBM Watson。2011 年，IBM Watson 在智力竞赛节目《危机边缘》中一举夺冠，展示了机器处理大规模文本知识并进行模式匹配的超强水平。然而，要真正理解语义内容，计算机还有很长的路要走。这便涉及到认知智能的一个重要核心技术：自然语言理解，一门融合语言学、计算机科学与数学的交叉学科。

机器理解语言是怎样一个过程？深层语义理解的突破点在哪里？如何利用该技术进行医学文本的知识挖掘？在北京智源大会“AI 科技女性”专题论坛上，北京大学计算机科学技术系教授、计算语言学教育部重点实验室主任穗志方针对这些问题做了主题为“从语言到知识——构建语言智能的基石”的报告。

一、自然语言处理的研究体系

自然语言处理 (NLP) 是通过建立形式化的计算模型来分析、理解自然语言，从而达到高效获取和利用语言承载的信息的目标。从定义可以看出，自然语言处理是计算机科学领域中一门跨越文理的交叉学科：以语言为研究对象，以计算机科学为研究工具，以数学为建模工具。同样的研究问题在语言学领域称为计算语言学 (CL)。NLP 应用非常广泛，包括机器翻译、机器写稿等，这些智能化应用的后台都有 NLP 技术的支撑。

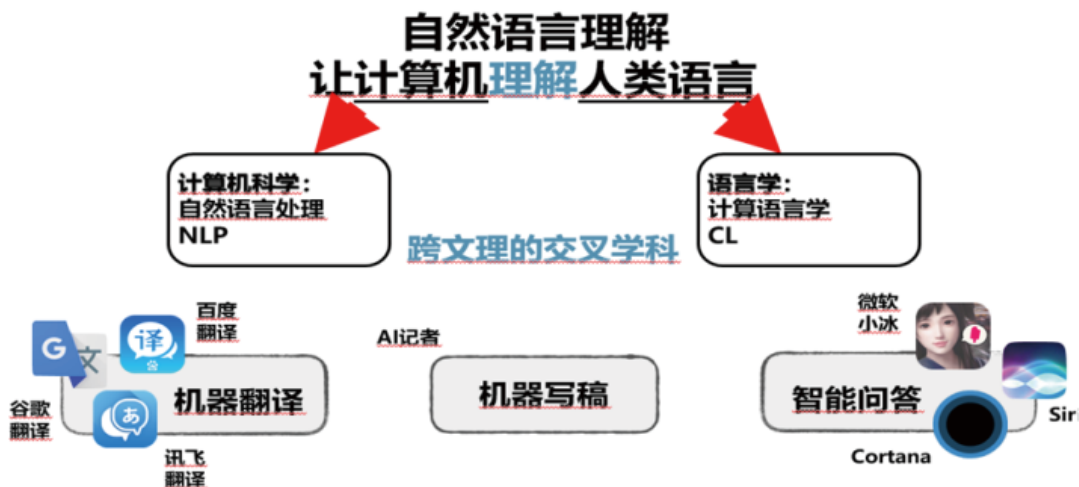


图 1：自然语言处理研究方向简介

我们可以把整个 NLP 研究体系描述成大树形状，树根是大规模语言知识资源，它作为知识基础来支撑整个 NLP 系统；树干是自然语言处理的核心技术，包括语义分析、句法分析；从树干延伸出来的枝叶包括机器翻译、自动文摘、自动问答等应用方面的研究。

二、机器理解语言的渐进过程

自然语言理解 (NLU) 是人工智能的高级阶段，被称为人工智能皇冠上的明珠，是一个多学科交叉融合的，比较新兴的、年轻的学科。

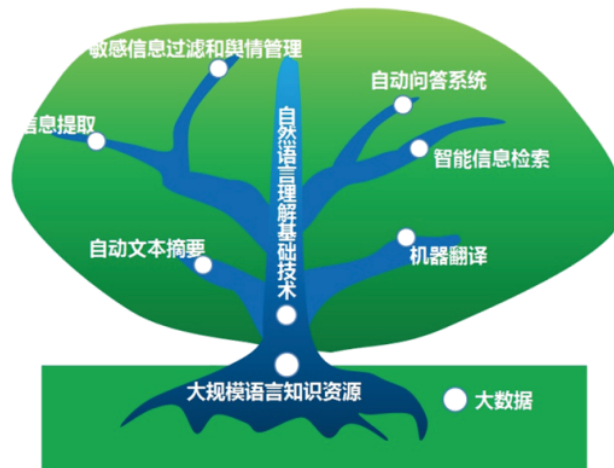


图 2：自然语言理解基础技术

这是一个机器理解语言的过程，整个研究脉络是渐进的，由不理解到理解要逐层经过多个环节，比如：

第一个环节是自动分词。

如果计算机能够正确切分这个句子，就代表对这个句子进行了最基础的理解。例如”南京市长江大桥位于南京市鼓楼区”，计算机在切分的时候结果可能有两种：

南京市 / 长江大桥位于 / 南京市鼓楼区...

南京市 / 江大桥位于 / 南京市鼓楼区...

人类很容易排除第二种分词结果，但要计算机排除第二种，就需要配备相关的语言知识。另一个例子是“你能穿多少就穿多少”：

你能穿多少就穿多少【冬天，多穿点】；

你能穿多少就穿多少【夏天，少穿点】。

如果希望计算机选择正确的断句方式，也需要配备相应的知识。

第二层是句法分析。

例如：“村民击毙了咬死了猎人的狗。”

如图，动宾结构和定中结构这两种结果，孰对孰错，需要利用上下文的信息，以及句法方面的语言知识。

• 句法分析

村民 击毙了 咬死了 猎人的 狗

动宾结构(×) 【咬死了】 【·猎人的·狗】

定中结构 【咬死了·猎人】·的·【狗】



图 3：机器理解语言的渐进过程

第三层是句义分析。

比如，要测试计算机是否理解下面这句话的语义结构，可以让它回答“谁死了”这个问题。是狗咬死猎人，还是人打死了狗？解决方法是进行语义角色标注，例如找出动作的发出者和对象，即施事和受事。这需要大规模的语义标注语料库，通过模型训练，让计算机自动学习标注正确的语义结果，从而获得知识。

• 句义分析(语义角色标注)

村民 击毙了 咬死了 猎人的 狗

Q: 谁死了? 村民(×)、猎人、狗

“狗”是“咬死”这个动作的发出者，是施事(Agent)

“狗”是“击毙”这个动作的对象，是受事(Patient)

句子的命题义：

咬死(狗，猎人)、击毙(村民，狗)

图 4：句义分析

总结一下，机器理解语言是逐层推进的。在整个研究体系中，作为树根的语言知识资源库，是处理自然语言的基础设施。相对来说，语义理解是最重要也是最困难的，不同于人工语言，自然语言的理解充满了歧义性，往往需要各种各样复杂的命题和超命题，并综合语言知识、世界知识、领域知识等方方面面的知识实现精准理解。

三、面向语言理解的汉语意合模型

让计算机理解汉语尤其困难，所以中文理解是自然语言的难题，穗志方团队的研究方向就是构建面向中文语义理解的语言知识资源。针对中文“意合”的语言特点，建立一套涵盖多层次语义信息的中文语义知识库。

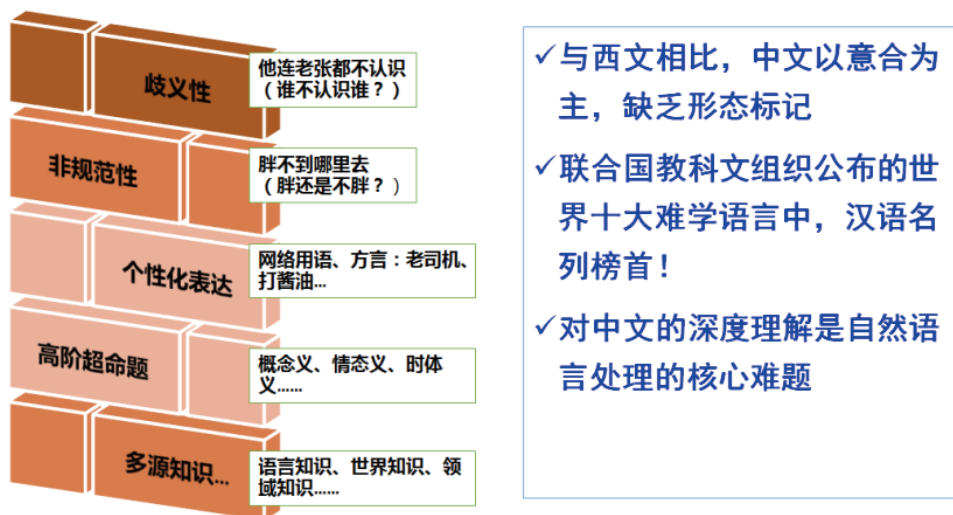


图 5：语义理解的困难

汉语和西文有比较本质的区别，汉语是一种意合语。让计算机理解汉语之前，首先要解决的问题就是建立计算机可以利用的汉语意合语法模型。在这方面，实验室利用跨学科合作的条件，采用“词库——构式”互动的架构，首次构建了具有可操作性的高度形式化和数据化的汉语意合语法理论体系，为机器自动理解汉语意合语法提供了形式理论模型。

面向语言理解的汉语意合模型

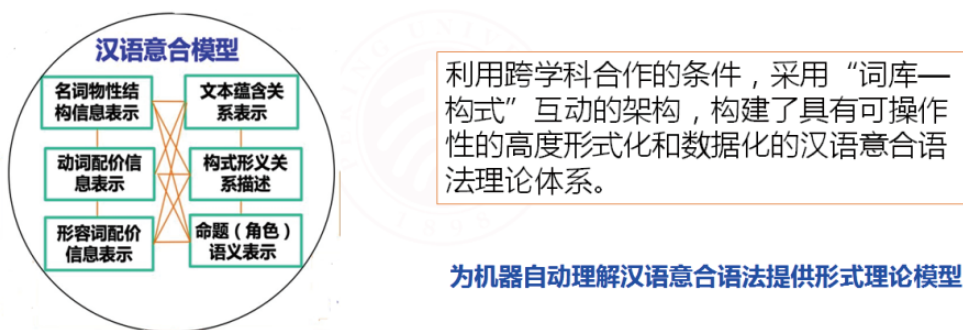
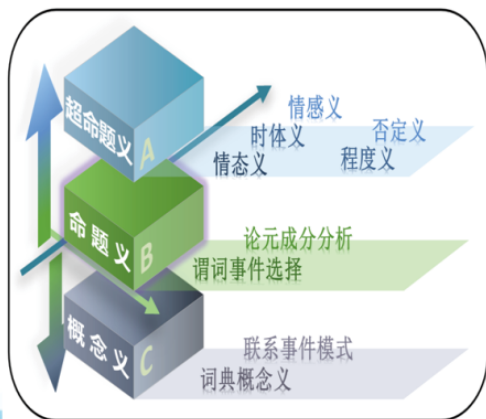


图 6：如何构建汉语语义理解的形式化模型？

基于这一目标，穗志方团队设计了多维度中文深层语义描述体系。

以命题义为核心，以三层表示机制为主轴，实现多维度的中文深层语义描述



- 自下而上的三层表示机制
- 概念义层：描述实体概念、事件状态概念。
 - 命题义层：描述句子的客观意义，具体对谓词事件、论元成分进行描述。
 - 超命题义层：描述句子的主观和附加意义，具体包括情态义、时体义、程度义、否定义、情感义。

图 7：中文深层语义描述体系

不同概念会触发人脑不同的认知场景，例如同样是描述“想”这个词，概念不同，事件语义就完全不同。在“他想家了”这句话中，“想”表达思念；而“他想照一张相”这句话中，“想”表示意愿。由于同一谓词的不同（词典）概念义会带来事件模式差异，如果借助词典对谓词排歧，事件语义会更加清晰。

事件概念借助概念词典排歧。

同样是“想” ...

[%施事 她] [# 想 #]<想念> [%对象 家] <了>tense_perf。
[%施事 我] [# 想 #]<希望> [%VP内容 照 一张 相 %]。

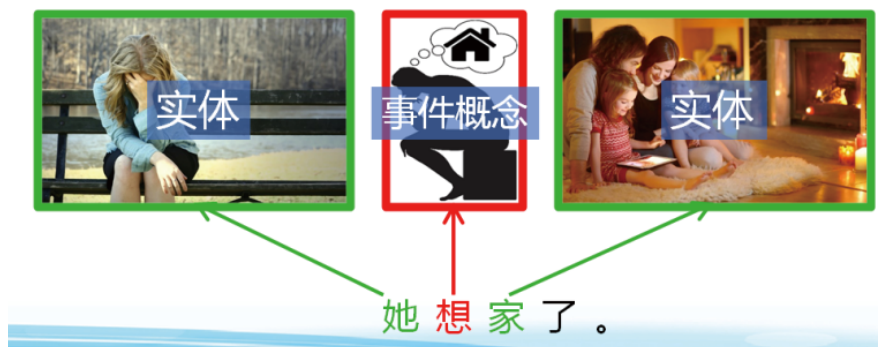


图 8：词义与概念的关联

因此，事件模式与概念义有很强的关联，相同的概念义对应的事件模式相似；而不同的概念义标注，其命题义的标注相差很明显。比如同样的“打”字，在“一下子打死了两只苍蝇”“上午给学校打了一个电话”“这两队打了一个平手”中，认知模式是不一样的。

同一谓词不同概念义带来事件模式差异



图 9：概念与事件模式的关联

进一步，在句义分析中充分考虑程度、形态、否定、情感等更多超命题方面的意义，最终形成融合概念义、命题义以及超命题义的机器表示。如果计算机能够将线性输入的句子建立成结构图，计算机就达到了三个层次的深度理解。

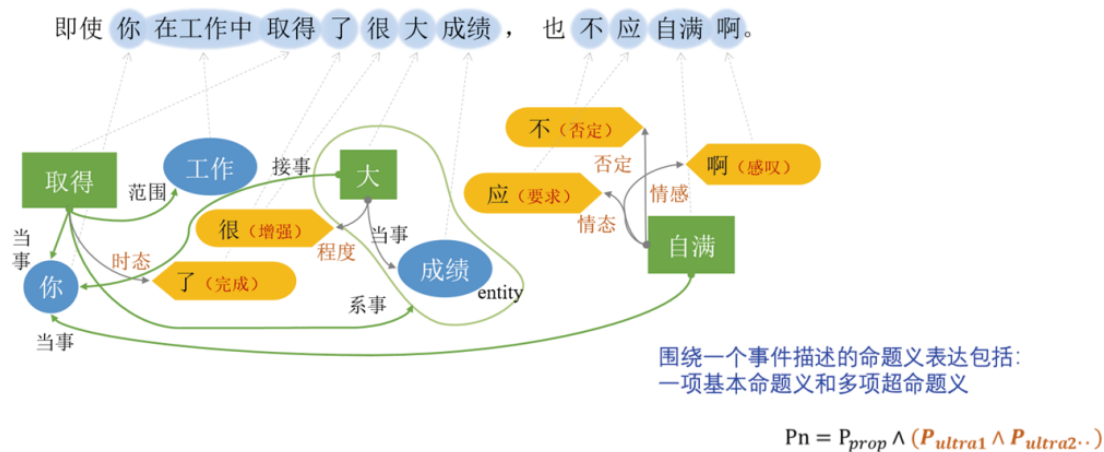


图 10：融合概念义、命题义与超命题义的机器表示

为了实现这样的语义分析，实验室构建了汉语自动分析技术及标注平台。通过开发工具集与人机结合的方式，落实上述理念，最终构建了一个大规模语义标注语料库。



图 11：服务于知识资源构建的汉语自动分析及标注平台

四、中文医学知识图谱 CMeKG

在知识资源构建方面，穗志方教授选择医学领域进行研究，从而将通用领域拓展到应用领域。人工智能为医疗带来了巨大帮助，包括药物挖掘、医院挖掘、药物研究等，为病人诊中、诊前、诊后提供了智慧医疗服务。在这样的应用驱动下，我们看到，支撑智慧医疗的基础设施就是医学知识图谱。

人工智能技术为医疗领域带来巨大变革——智慧医疗



图 12：语言知识资源构建——从通用领域到专业领域

于是，穗志方团队构建了一个面向智慧医疗的中文医学知识图谱 CMeKG，从数据层、技术层、知识层、应用层以及应用层等多个层面展开。以通用的 NLP 方法为基础，开发面向医学文本的领域内容分析核心技术。

基于医学知识图谱，穗志方的团队也做了医学期刊智能检索、相似病历检索等医学领域的应用和探索。

五、计算语言学重点实验室

针对女性科研发展的主题，穗志方教授也介绍了自己科研发展道路上，帮助最大的并培养自己多年的北京大学计算语言学教育部重点实验室，这是一片学科交叉人才成长的沃土。1986年北京大学成立了计算语言学研究所，一个文理交叉的研究机构，经过几十年的发展已经取得了非常丰硕的成果。它的特色是计算机学科和语言学科的文理交叉。



图 15: 北京大学计算语言学教育部重点实验室发展历程

同时，该实验室也建立了博士、硕士、访问学者这样一套完整的培养体系，培养了优秀的 NLP 领域的研究人员，其中不乏杰出的女性。很多女硕士、女博士以及女性研究者和访问学者都是从这里走出去，走到世界上！

复旦大学教授黄萱菁：自然语言处理中的表示学习

整理：智源社区 何文莉

黄萱菁，复旦大学计算机学院教授，AI 2000 上榜学者，博士生导师，研究方向为人工智能、自然语言处理、信息检索和社会媒体分析。兼任中国中文信息学会常务理事，社交媒体专委会副主任，中国计算机学会中文信息技术专委会副主任。

比尔·盖茨曾说：“语言理解是人工智能皇冠上的明珠。”在语音交互的 AI 时代，自然语言处理 (NLP) 发挥着不可替代的作用。深度学习的出现与发展，使 NLP 技术取得了重大的突破。语言表示学习作为深度学习的重点之一，在自然语言处理中都有哪些研究进展和体会？

在北京智源大会“AI 科技女性”专题论坛上，复旦大学计算机学院黄萱菁教授做了一场主题为“自然语言处理中的表示学习”的演讲，让我们一起来从报告中寻找问题的答案！

一、相关概念简介

1.1 语言表示

语言表示可从多个角度加以定义。从认知科学角度，语言表示是语言在人脑中的表现形式，关系到人类如何理解和产生语言。从人工智能角度，语言表示是语言的形式化或数学描述，以便在计算机中表示语言，并能让计算机程序进行自动处理。一个好的文本表示，首先要具备很强的表示能力，比如模型具有一定的深度。其次要使后续的学习任务变得简单，能够带来下游任务性能的提升。最后应具有一般性，是任务或领域独立的。

早期的语言表示方法主要采用符号化的离散表示。词表示为 One-Hot 向量（一维为 1、其余维为 0 的向量），句子或篇章通过词袋模型、TF-IDF 模型、N 元模型等方法进行转换。

离散表示的缺点在于词与词之间没有距离的概念，比如“电脑”和“计算机”被看成是两个不同的词，显然这是不合理的。当前主流的语言表示采用更加精确的数学表示，通常通过基于深度学习的表示模型获得。

1.2 深度学习

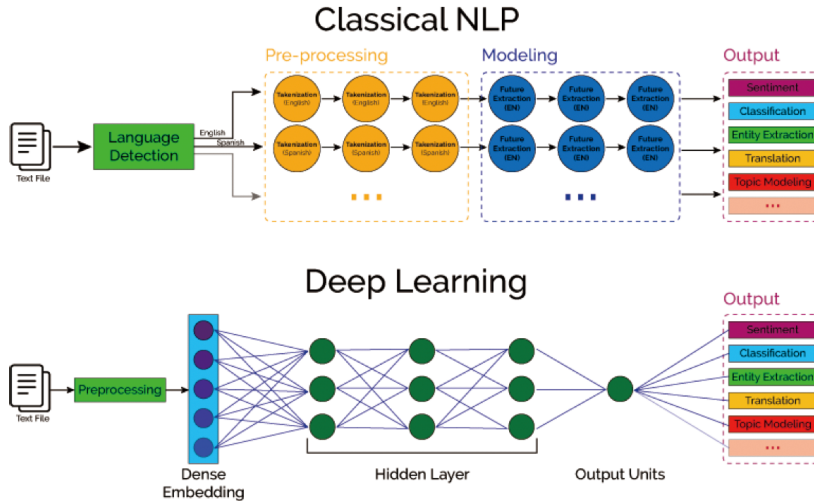


图 1: 传统的机器学习 VS 深度学习

深度学习是机器学习中的一个子领域，表示学习和深度学习的兴起密切相关。传统的机器学习方法通常需要人工设计表示以及特征提取方法，而深度学习方法则有所不同，它不需要特征提取，甚至可以直接进行表示学习。

1.3 自然语言处理中的深度学习

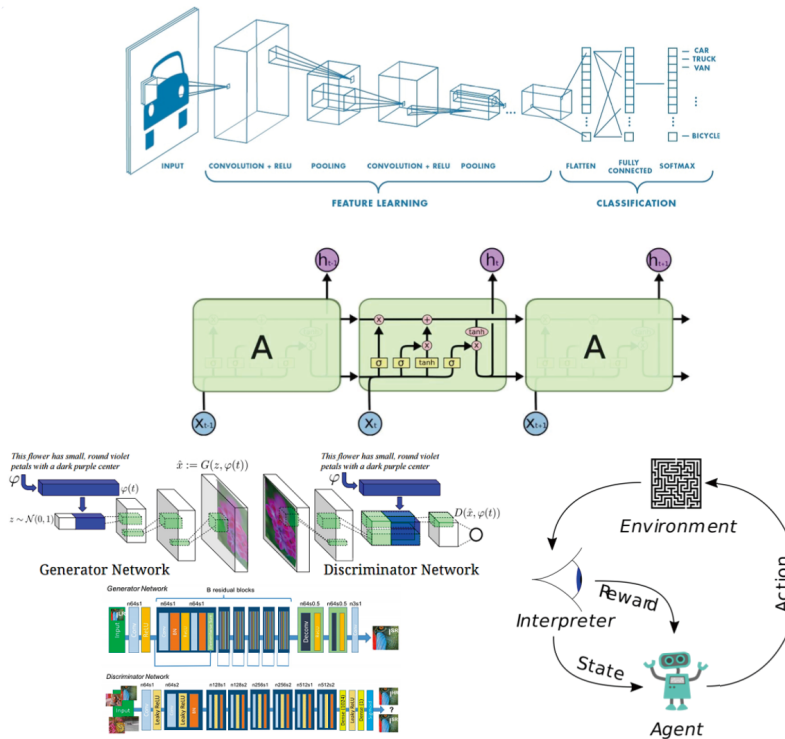


图 2: 自然语言处理中的深度学习

深度学习使得自然语言处理众多任务取得了重大进展。各种神经网络如卷积神经网络、循环神经网络、对抗生成神经网络等，一方面可成功运用于分词、词性标注、实体识别、成分分析、句法分析、语义表示、语义匹配、情感倾向分析等基本任务，另一方面可极大提升问题问答、对话系统、机器翻译、阅读理解等应用系统的性能。

二、基于神经网络的语言表示学习

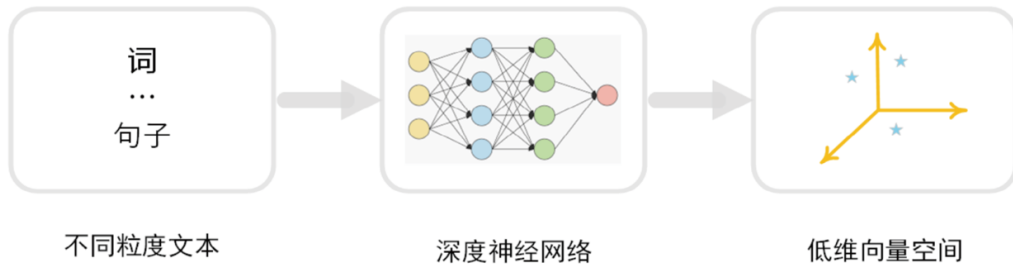


图 3: 神经语言表示学习

基于神经网络的语言表示学习，指将不同粒度文本中语言的潜在语法或者语义特征分布式地存储在一组神经元当中，用稠密、低维、连续的向量来表示。

不同粒度包括词语、短语、句子和句对等，短语在语义层面类似于词语，结构上类似于句子。不同粒度的语言表示有不同的用途，比如词语和短语表示主要用于预训练，服务于下游任务，而句子和句对表示可直接应用于文本分类、匹配、阅读理解、语篇分析等具体任务。

2.1 语 的表示学习

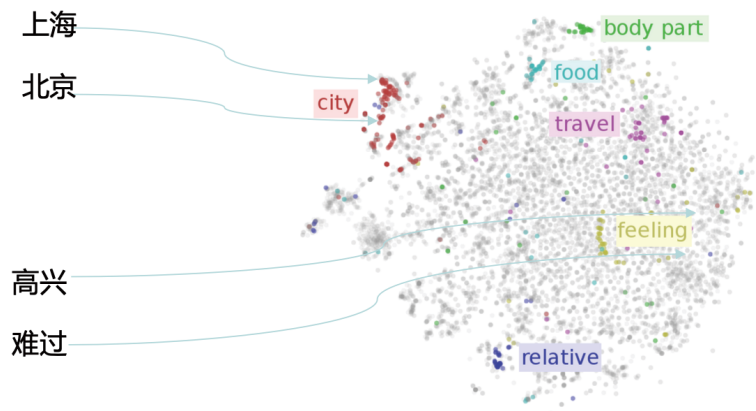
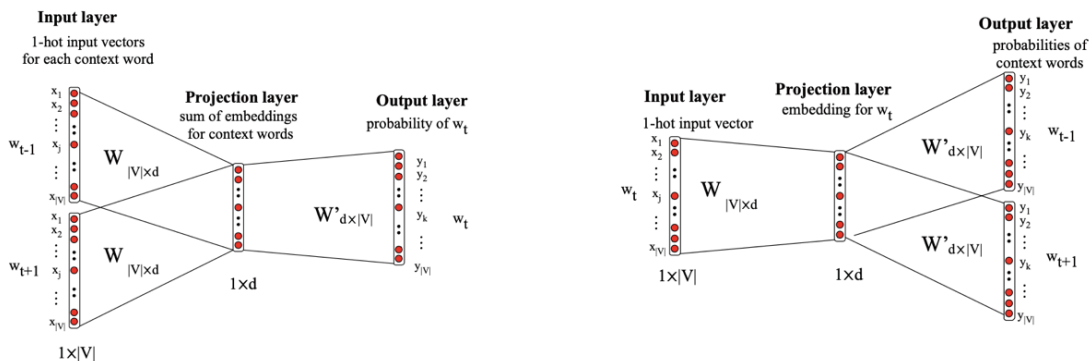


图 4: 词嵌入

词语的表示学习即词嵌入，把词语从符号空间映射到向量空间。2013 年之前只有少量词嵌入研究工作，比较著名的有 Bengio 和 Hinton 提出的两个模型，2013 年后开始涌现出大量新工作，比如众所周知的 glove 和 word2vec，这些工作主要得到上下文无关的向量。2016 年后研究出现短暂停滞，到 2018 年才重新盛行，不

同于之前学习相对比较独立的词向量，大量新工作把工作重心转移到了带有上下文的语境化词向量上，最经典的工作是 ELMo 和 BERT。

2.1.1 基于预测的词表示学习 – word2vec



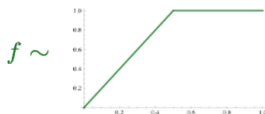
Efficient estimation of word representations in vector space, NIPS 2013

图 5: word2vec 模型

上下文无关的 word2vec，是一种高效的无监督词表示模型。主要涉及两个模型，通过上下文信息的平均预测目标词的 CBOW 模型（连续词袋模型，图 5 左）和用目标词预测上下文的 Skip-gram 模型（跳词模型，图 5 右），二者都可以学习高质量的词表示。一般而言无监督得到的词向量，会给出下游任务的学习一个非常好的初始值，加快模型的收敛。

2.1.2 基于计数的词表示学习 – GloVe

$$J(\theta) = \frac{1}{2} \sum_{i,j=1}^W f(P_{ij})(u_i^T v_j - \log P_{ij})^2$$



• Code and Pre-trained word vectors : <http://nlp.stanford.edu/projects/glove/>

Glove: Global vectors for word representation, EMNLP 2014

Man-Woman=King-Queen

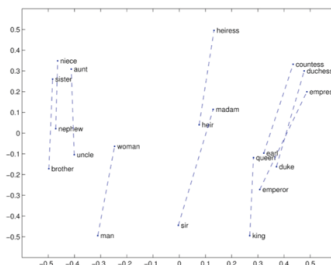


图 6 : GloVe 模型

GloVe 模型，直接建模词对的共现频率和对应词向量内积之间的关系，使得它们尽可能接近，目标是损失函数最小化。word2vec 和 GloVe 这类词向量的合理性可通过类比实验认证，比如男人和女人之间的差距，相当于国王和女王之间的差距。

2.2 短语和句子的表示学习

短语和句子的表示学习方法是类似的，与结构预测密切相关。几种常见的语义组合函数都可将词语序列表示转换为短语和句子表示，包括递归神经网络、卷积神经网络、循环神经网络、Transformer、图神经网络等，也可组合使用。

2.2.1 卷积神经网络对句子建模

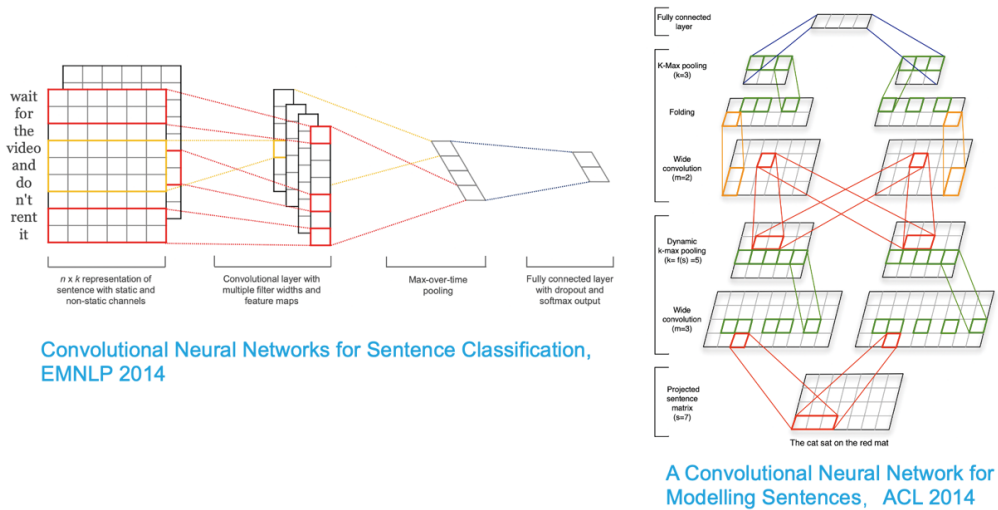
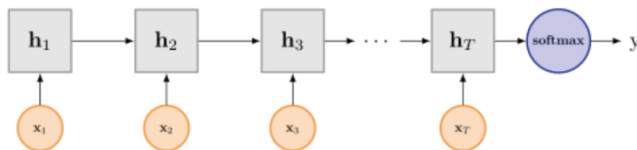


图 7: CNN 模型 (左) 及其改进 DCNN 模型 (右)

CNN 模型是卷积神经网络对句子进行建模的代表性工作。左图采用双通道的 CNN 得到句子表示，进行文本分类。右图可以通过动态池化机制解决句子变长的问题。

2.2.2 循环神经网络对句子建模

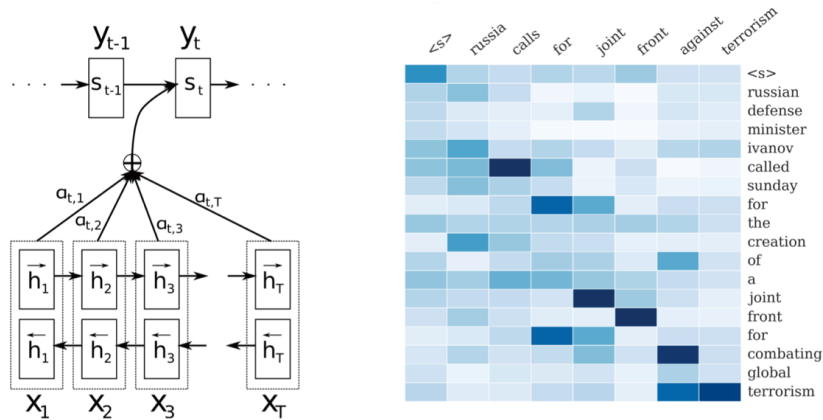


Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling, NIPS workshop 2014

图 8: 门循环单元 (GRU) 模型

循环神经网络用于对句子进行序列建模，早期工作如平凡循环神经网络 (Vanilla RNN)，后期为解决循环神经网络技术实践上存在的梯度消失或梯度弥散问题先后提出了长短时记忆单元 (Long Short-Term Memory, LSTM) 和门循环单元 (Gated Recurrent Unit, GRU) 等。

2.2.3 基于注意力的序列 – 序列模型：文摘和机器翻译



Neural machine translation by jointly learning to align and translate ICLR 2015
A Neural Attention Model for Sentence Summarization EMNLP 2015

图 9：基于注意力的序列 – 序列模型

循环神经网络可以方便地扩充到编码器 – 解码器架构，在解码过程中有重点地选取编码的信息。编码器用于理解，没有输出，解码器用于生成，不再接受新的输入。如果在解码时引入注意力机制，可以进一步提升模型的性能。

接着，黄教授介绍了自己及其团队在这方面的研究成果：基于门机制的递归神经网络、递归卷积神经网络和自适应的 Tree-RNN。

2.2.4 建模复杂的稠密特征组合 --- 基于门机制的递归神经网络

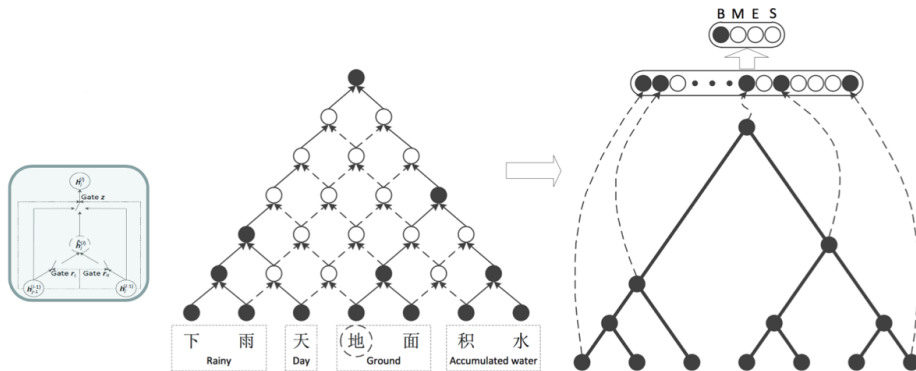
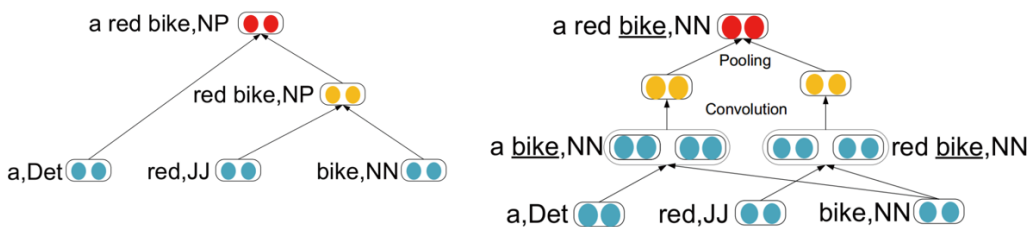


图 10：基于门机制的递归神经网络

句子建模方面，针对自然语言处理中特征建模方法比较单一问题，提出了基于门机制的递归神经网络。利用树结构神经网络获得句子的树结构，对树结构递归神经网络进行改进，加入门机制，目的是希望对上下文窗口中相邻字词之间的组合关系进行更为精细的建模。从字间关系构建词间关系，从而构建句子的树结构。图 10 是如何模拟上下文汉字复杂组合的具体方法。

以句子“下雨天地面积水”为例，当前的目标字是“地”，需要判断“地”是“天地”的词尾还是“地面”的词首。实际上这个句子非常复杂，因为每两个连续的汉字都可以组合成一个词语。为在给定上下文中预测目标词“地”的标签类型，从网络的底层到顶层，递归地进行特征组合。其中黑色的节点表示活跃的神经元，空心的节点表示抑制的神经元，边表示消息传递，实线边表示接受消息上传，而虚线边表示拒绝。通过该组合过程，可获得句子的树结构（图 10 最右侧）。除此之外，还可以把所有的组合特征合并到网络中间，估计树结构的得分。通过基于门机制的递归神经网络模型可以同时得到句子的表示和中间结果，可用于分词、依存句法分析、句子建模等各种任务。

2.2.5 递归卷积神经网络

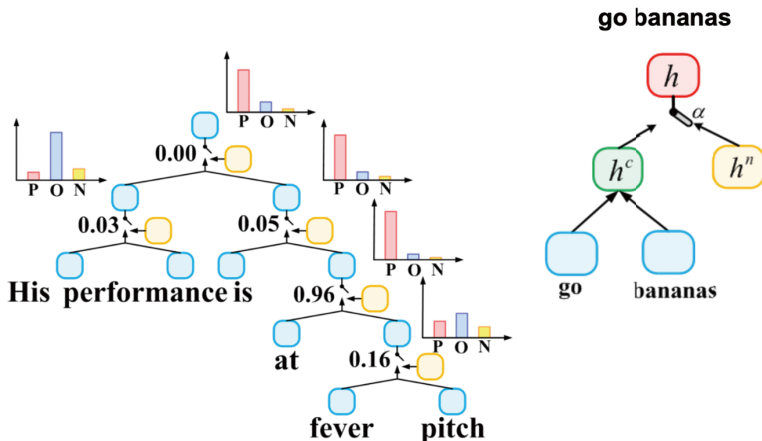


A Re-Ranking Model For Dependency Parser With Recursive Convolutional Neural Network, ACL 2015

图 11: 递归卷积神经网络

另一改进是对二叉树的改进。句子的表示可认为是句子中所有词表示的组合，递归神经网络利用成分句法树，把词按照成分结构不断地进行递归，最后得到整个句子的表示。但递归神经网络只能处理二叉树的拓扑结构，不能有效地拓展到依存句法树。黄教授等结合递归神经网络和卷积神经网络，提出了一种可以处理多叉树的递归卷积神经网络模型。通过引入卷积层和池化层，把递归神经网络拓展到依存句法树上，如图 11（右）所示，模型可以同时捕捉不定冠词和名词、形容词和名词的依存关系，有效地改善了中英文依存句法分析的性能。

2.2.6 自适应的 Tree-RNN



Adaptive Semantic Compositionality for Sentence Modelling, IJCAI 2017

图 12: 自适应的 Tree-RNN

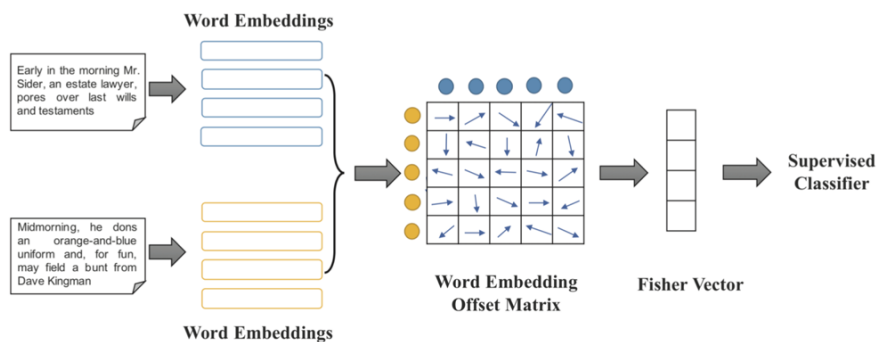
自然语言句子中并非所有的短语都是合成性的，部分短语的语义无法由组成成分合成，比如“go bananas”的意思是发疯，“马马虎虎”、“九牛二虎”和马、老虎、牛都无关。为了提升语义组合的性能，采用了树结构 LSTM。基于句法树递归地对句子进行建模，引入参数化的控制器，能够自适应地决定非叶节点的合成方式是组合性还是非组合性。

自适应的 Tree-RNN 模型包括三个部分：合成性非叶节点、非合成性的非叶节点和控制器。可以同时建模成型和非成型短语，通过树结构生成短语，根节点代表句子语义。对于合成性非叶节点相应短语的表示，像“His performance”是由子成分合成得到的。对于非合成性非叶节点相应短语的表示，则是作为基本的语言单位学习得到，如“at fever pitch”，通过语义开关控制器控制合成的方式，最后得到句子的表示。

2.3 句对的表示学习

许多 NLP 任务都可以建模为句对编码，比如句子重述、蕴含分析、语篇分析等。句对编码的目标是给定两个句子，建模其语义关系并学习其表征。黄萱菁教授及其团队在句对表示学习方面的工作主要围绕语篇关系表示与检测及其改进。

2.3.1 语篇关系表示与检测

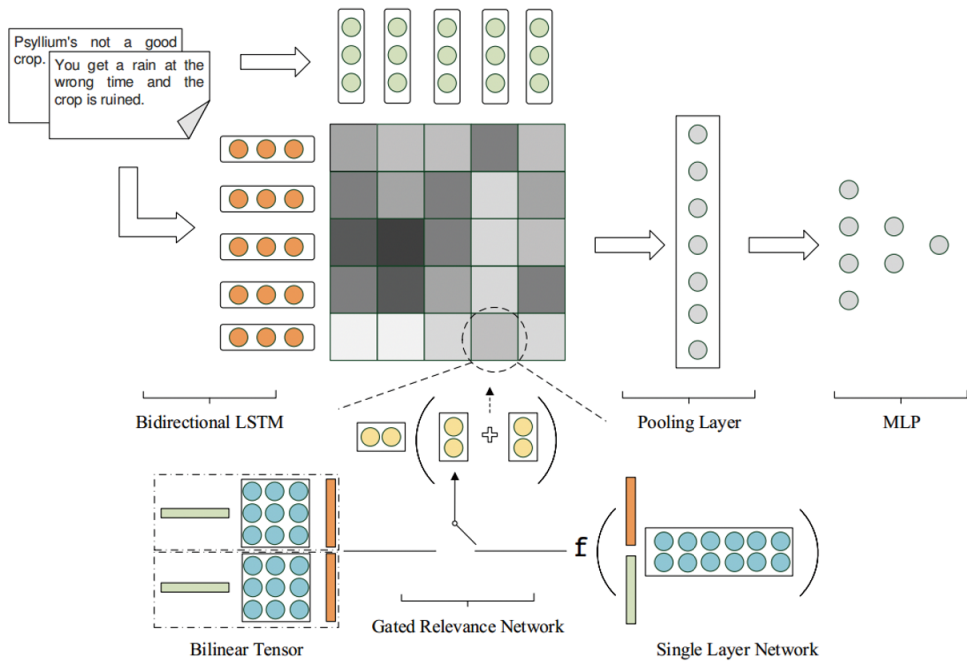


Discourse Relations Detection via a Mixed Generative-Discriminative Framework, AAAI 2016

图 13 语篇关系表示与检测

语篇关系表示与检测的研究动机有两点：一是单词向量表示之差可以表示单词间的关系，二是句子之间的关系如何表示？以两个句子为例，一个发生在“Early in the morning”，另外一个发生在“Midmorning”，二者是承接关系。用词向量差值可以表现句间关系，把两个句子的所有词两两做词向量的差值可以得到位移矩阵，通过位移矩阵可以预测句间关系。然后引入 Fisher Vector 方法解决矩阵大小随句子长度变化问题，把矩阵转变为定长向量，进行语篇关系的分类。

2.3.2 利用 GRN 进行语篇关系检测



Implicit Discourse Relation Detection via a Deep Architecture with Gated Relevance Network, ACL 2016

图 14：利用 GRN 进行语篇关系检测

利用 GRN 进行语篇关系检测有两个重要改进点：一是使用句子的隐状态，而非直接使用词向量，二是使用 GRN(Gated Relevance Network) 建模隐状态之间的交互关系。为了度量隐状态之间的相关性，使用门相关性网络，它可以组合许多匹配函数。

2.3.3 近期研究趋势

句对表示学习的近期研究趋势，模型层面包括 Transformer 和图神经网络，学习层面有迁移学习、元学习以及多任务学习等，报告主要介绍了 Transformer 和迁移学习。

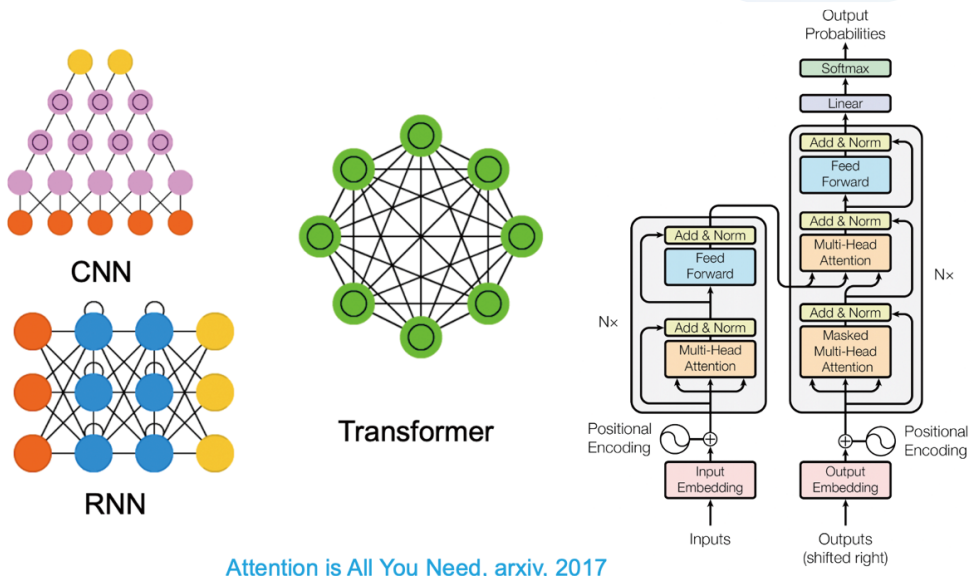


图 15: Transformer

以 Transformer 为代表的自注意力机制取代了神经网络中的经典合成函数，比如卷积神经网络、循环神经网络，在各项任务上取得了目前最好的结果。Transformer 的成功可归因于其非局部结构偏置，句中任何一对词之间的依存关系都可以被建模。它通过摒弃复杂的语义组合方式，提供了更有效的计算，为 Bert 等模型打下了基础。同时在大数据集上具备良好的可扩展性，很多预训练语言模型都是在 Transformer 基础上构建的。

2.3.4 迁移学习

迁移学习具体指把为任务 A 开发的模型作为初始点，重新使用在为任务 B 开发模型的过程中。它包括两个阶段，第一阶段是学习可迁移的知识，第二阶段是把知识迁移到新的任务。迁移的知识从何而来？主要有监督学习和非监督学习两种方式，先通过有监督 / 无监督方式学习可迁移知识的表示，再把知识迁移给新的任务。

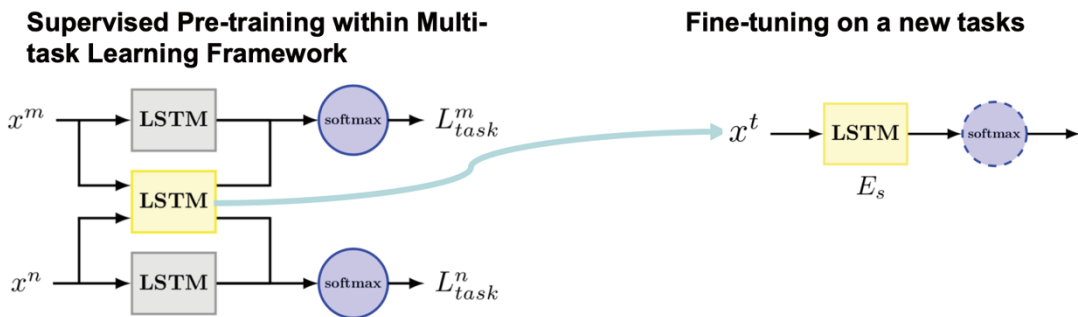


图 16: 有监督迁移学习

图 16 是有监督方式迁移学习的处理机制，m 和 n 表示不同的任务，中间的 LSTM 可以学到任务无关的表示，两个灰色 LSTM 是任务独有的表示。学习的目的是学习模型的参数，可在新的任务上（不是 m 也不是 n），对任务参数进行微调。

Unsupervised Pre-training with Language Model

Fine-tuning on a new tasks

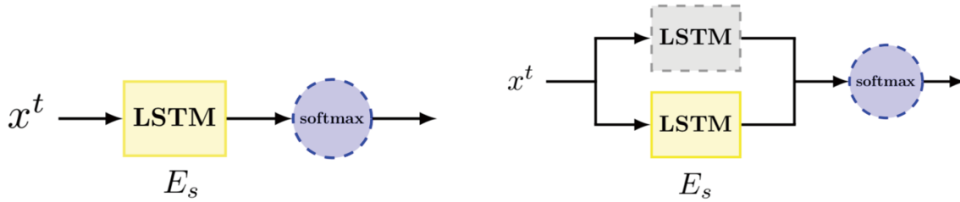
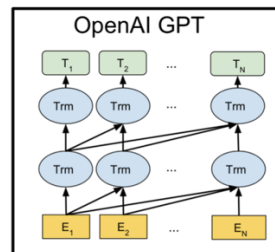
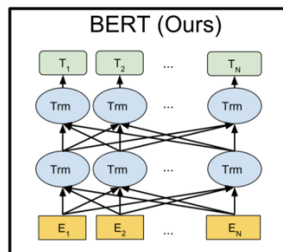
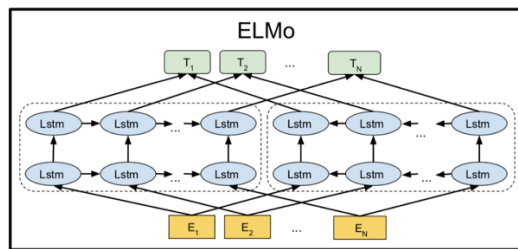


图 17: 无监督迁移学习

图 17 是无监督方式迁移学习的处理机制，与语言表示相关性更加密切。用 ELMo 进行无监督训练，当面向有监督的 NLP 任务时，可把 ELMo 的向量直接当作特征并接到具体任务模型的词向量的输入或最高层表示上。更强的模型比如 GPT 或 BERT，在具体 NLP 任务进行有监督微调时，就不再需要对任务构建新的模型结构，只需要简单地在 Transformer 的最后接上一个 SoftMax 分类器作为任务输出即可，再对整个模型进行微调。

2.4 预训练语言模型



Deep contextualized word representations, NAACL 2018

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arxiv, 2018

图 18: 预训练语言模型

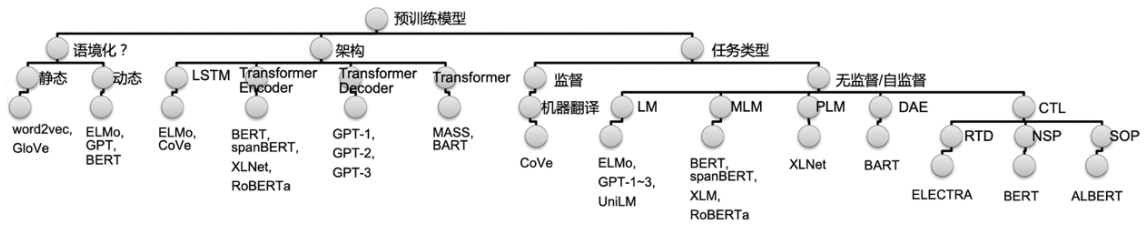
ELMo 的目的是找到一种语境化词向量表示方法，不同语境下，每个词都能获得更准确的表示向量。模型构建时，ELMo 采用两层双向 LSTM，采用非监督的数据集训练出来两个单向的语言模型。GPT 运用的模型是 Transformer，通过 Transformer decoder 训练得到单向的语言模型，mask 掉了当前词后面的词。与 GPT 单向语言模型不同，BERT 使用了双向的语言模型。为了让预测时待预测词看不到自己，需要把待预测词 mask 掉。据此 BERT 提出了掩码语言模型 (Masked Language Model)，随机 mask 掉输入中少量词语，再用双向语言模型预测这些词。BERT 还为预训练语言模型引入了新的目标——预测下一个句子，以学习句子和句子之间的关系。

2.4.1 各种预训练模型改进

模型	架构	预训任务	训练语料	参数量	GLUE分数
ELMo	LSTM	双向LM	WikiText-103	-	-
GPT	Transformer Dec.	LM	BookCorpus	117M	72.8
GPT-2	Transformer Dec.	LM	WebText	117M~1542M	
BERT	Transformer Enc.	MLM + NSP	WikiEn + BookCorpus	110M~340M	81.9
InfoWord	Transformer Enc.	DIM + MLM	WikiEn + BookCorpus	=BERT	81.1
RoBERTa	Transformer Enc.	MLM	BookCorpus+CC-News+OpenWebText + STORIES	355M	88.5
XLNet	双流Transformer Enc.	PLM	WikiEn+ BookCorpus+ Giga5 + ClueWeb + Common Crawl	≈BERT	90.5
ELECTRA	Transformer Enc.	RTD + MLM	同XLNet	335M	88.6
UniLM	Transformer Enc.	MLM + NSP	WikiEn+BookCorpus	340M	80.8
MASS	Transformer	Seq2seq MLM	-	-	-
BART	Transformer	DAE	同RoBERTa	1.1 x BERT	88.4
T5	Transformer	Seq2seq MLM	C4	220M~11B	89.7

图 19：不同预训练语言模型的对比

预训练模型的发展以 ELMo 为正式开始，以 BERT 为发展高潮，此后越来越受到高校和企业的关注，各种模型不停涌现，模型训练数据的规模、模型的性能表现也都不断创造新高。

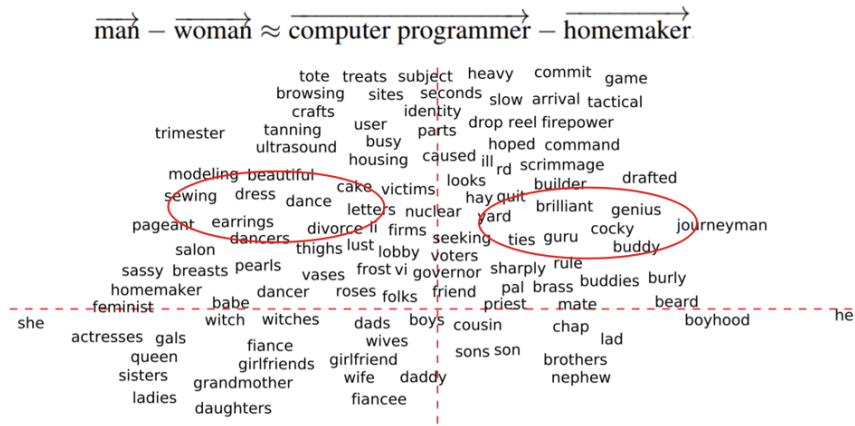


Pre-trained Models for Natural Language Processing: A Survey, arxiv 2020

图 20：预训练模型分类

图 20 是黄教授等在一篇期刊文章中给出的预训练模型分类体系，预训练语言模型按照是否语境化可以分为静态和动态，按照模型架构可以分为 LSTM、Transformer Decoder、Transformer Encoder 以及完整的 Transformer，按照任务类型可以分为有监督和无监督 / 自监督的。

三、语言表示的偏见



Bolukbasi, Tolga, et al. Man is to computer programmer as woman is to homemaker? debiasing word embeddings, NIPS 2016

图 21：词嵌入中的女性形象

图 21 来自 NIPS 2016 的一篇文章，发现从谷歌新闻等比较正式的文体中训练出的词向量带有非常强烈、令人不安的性别刻板印象。从几何角度，性别偏见可被词向量所在向量空间的方向所表示。在语料中学习发现。男性和女性的差异，相当于程序员和家政工人之间差异。一方面有一些词如漂亮、舞蹈、裙子等与女性相关，聪明、天赋等词则和男性相关，而这些词语本身并不应该带有性别含义，应是性别中立的。另一方面还有一些词是定义中带有性别的，比如爷爷奶奶、男朋友、女朋友等，我们希望能把性别中立的词和与性别相关的词进行有效区分。

词嵌入中的性别偏差可能会带来很糟糕的社会后果，以求职网站为例，用计算机寻找程序员。假设一个程序员叫 John，另一个叫 Mary，作为程序员他们的专业水平是一样的，但 John 的名字更加男性，算法会认为他更接近要求，导致在与 Mary 的竞争中带有优势。

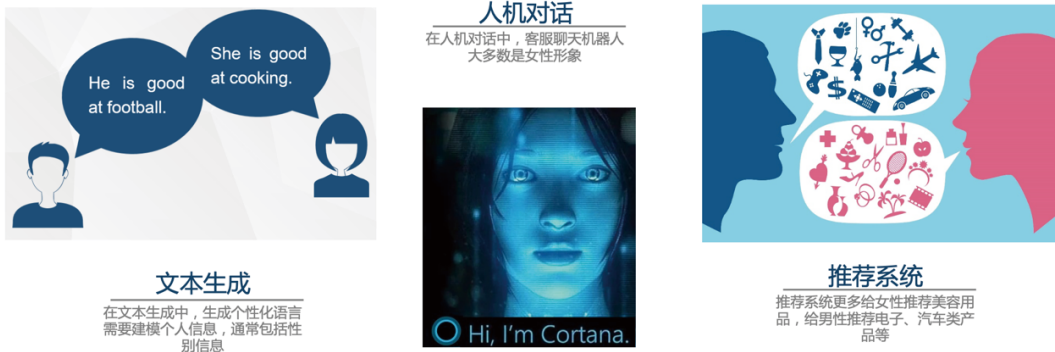


图 22：NLP 中的性别刻板印象

性别刻板现象表明在现有的人工智能算法当中，并不拥有性别平等。究其原因，首先人工智能相关的从业人员男性远大于女性，这就意味着人工智能系统在研究设计和开发过程中，女性和男性并不是同等发声。除了社会

学角度，如果算法设计者的先验信息中存在刻板印象，把某些特性功能或期待值和性别联系在一起，把性别作为判断依据，做出有性别倾向性的判断，也是性别不平等的体现。

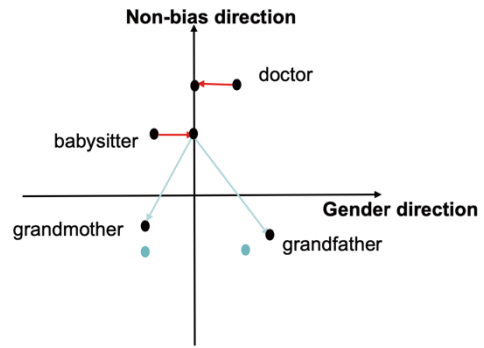
我们希望人工智能是性别平等的，不利用性别信息作为特征，不主动利用性别信息进行用户画像，不预先为担任某种职责的人工智能设定性别，不被动地使用输入信息中隐藏的性别暗示。幸运的是近年来人工智能多元化在逐渐增加，比如涉及性别歧视的著名会议 NIPS 改为了 NeurIPS，种族和地域变得多样化等。

- 确定性别方向

$\vec{\text{she}} - \vec{\text{he}}$
 $\vec{\text{her}} - \vec{\text{his}}$
 $\vec{\text{woman}} - \vec{\text{man}}$

- 中立化

- 均衡化

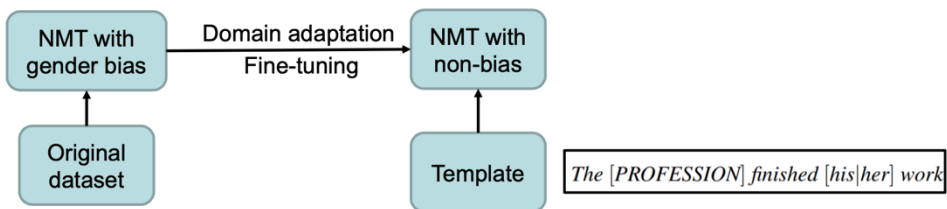


Bolukbasi, Tolga, et al. Man is to computer programmer as woman is to homemaker? debiasing word embeddings, NIPS 2016

图 23：消除词嵌入中的偏见

上述 NIPS 2016 的文章，除介绍性别偏见现象，还开发了一个消除词嵌入中性别偏见的算法，具体思路见图 23。算法第一步确定性别方向，找出表示男人 / 女人的词语进行相减求平均值得到性别方向，与其正交的则为性别无关方向。第二步中立化，像 Babysitter、Doctor 本应性别中立的词，实际编码在性别上依然存在偏差，为了消除偏差，将它们投影到性别中立化的轴上，使得在水平方向的投影距离为零。第三步均衡化，类似 grandmother 和 grandfather 的性别对应是由于定义而不是刻板印象。性别中立和性别相关的词应该有相同的距离，但即使经过修正，Babysitter 还是更接近 Grandmother 而不是 Grandfather，均衡化过程就是通过线性代数操作把向量移到对称的两个点上以解决此问题。

The doctor told the nurse that she had been busy.



$$L'(\theta^{DB}) = L(\theta^{DB}) + \lambda \sum_j F_j (\theta_j^{DB} - \theta_j^B)^2$$

Danielle Saunders and Bill Byrne, Reducing Gender Bias in Neural Machine Translation as a Domain Adaptation Problem, ACL2020

图 24：迁移学习减轻性别偏差

图 24 来源于 ACL 2020 年的一篇文章，该文章中作者提出使用迁移学习来解决机器翻译中的性别偏差问题。“医生告诉护士她很难”，但由于因为预训练语言中 Doctor 大多指男性，最后把“她”指代成了护士，造成翻译错误。文章的出发点是用少量的数据去除偏差，因为人工标注大量无偏差数据是很困难的，所以采用了迁移学习方法。将有偏差原始数据集当成是一个领域，通过模板的方法构造了一个无偏、小规模的数据集。开始在标准数据集上进行训练，得到一个带有性别偏差的神经语言模型，然后通过迁移学习进行微调 and 领域自适应，最终得到没有偏差的神经语言模型。

四、结语

在 NLP 中，表示学习至关重要。黄萱菁演讲由浅入深，从基础理论知识到复杂的研究模型，系统地为大家讲解自然语言处理中的表示学习。理论内容充实、丰富，最后呼应主题，阐述性别在语言表示中的偏见及解决办法。模型需要消除偏见，现实更需要消除偏见。希望终有一天科研可以不分性别，研究可以不论地域，只遵循知识本身！

圆桌论坛：AI TIME - 论道科技与女性

转载自：AI Timer 整理：鸽鸽

6月21日，由北京智源人工智能研究院主办的2020北京智源大会开幕，吸引国内外顶级AI专家学者共同研讨人工智能发展前沿话题、放眼未来十年。

可是我们不禁要问：这场引领时代浪潮的学术盛会中，有多少女性科研者的身影？面对STEM领域普遍存在的性别失衡问题，占有类一半智慧的女性，怎样在未来科技中迸发出自身的力量？女性科学家如何冲破刻板印象、性别差异和社会家庭文化因素的制约，甚至做到中流砥柱、力挽狂澜？

针对这些问题，下午的“AI科技女性”专题论坛的五场报告，从学术界女性的赋权问题，到富有女性特色和情感温度的前沿研究，到语言中的性别偏见现象，展现了女科学家们不容忽视的智慧、魅力与卓越贡献。

为更为全面、深入地探讨“科技与女性”这一话题，第十四期AI Time邀请了清华大学计算机系教授、AI 2000上榜学者朱小燕，北京语言大学教授、智源学者杨尔弘，阿里巴巴达摩院资深算法专家杨红霞，中国计算机学会YOCSEF学术委员会学术委员徐剑军，并由中科院计算所研究员、智源学者兰艳艳、滴滴出行科技生态与发展总监吴国斌主持，展开进一步的思辨与交流。

一、引言：从新冠疫情谈起

热身问题：新冠疫情期间，女性发表论文的增量为什么比男性少？

研究表明，新冠疫情期间各学科的女性研究人员的发表率不及男性研究人员，尤其是女性初级研究员缺席严重。

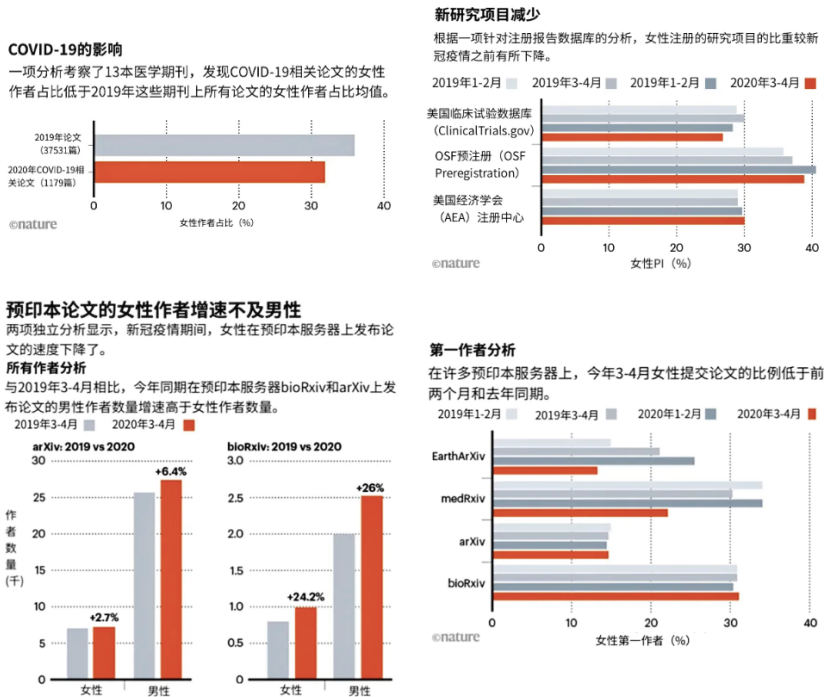


图 1：男女性科学家发表论文情况对比

朱小燕认为，这和科研工作本身的性质有关。科研需要长时间不间断、沉浸式思考，在疫情期间女性受到家庭的干扰比男性更多，不能做到像办公室里的百分之百投入。相比之下，男性的影响更小，甚至节约了很多社交时间。

杨尔弘从女性的传统社会角色和家庭角色分析，认为疫情使得女性承担的家庭事务性工作骤增，包括家务、一日三餐、孩子在家需要照料，在内需要安排妥善家人的生活，在外受到各种信息的冲击，缺乏团队集聚、集中、安静的环境。杨红霞则补充道，Work From Home 的工作模式使得员工心态发生变化，工作重心转移到家庭，这对女性工作者的影响尤为明显。当然，需要严格检验问题中的差异化假设是否为 Statistically significant 再下定论。

二、“隐形人物”与“稀有动物”

据统计，全球科技领域女性从业人员只占总人数的 29%，精英科学家中女性仅占 3%~5%。在诺贝尔科学类奖项中女性获奖者仅占 3.3%，其中诺贝尔物理学奖自 1901 年至今仅有 3 名女性获奖者；阿贝尔奖自 2003 年开始颁发至今，仅有乌伦贝克一名女性获奖。面对科技领域女性数量少、获奖少、科学地位低等问题，公众常常认为这并非性别不平等所致，而是因为女性不适合科学，或者女性的科研能力不如男性？



图 2：知名女性科学家

针对这一问题，杨红霞认为智力和能力上，男女水平并无差别。但是，首先从生源供给的角度，理工科专业确实存在男女学生比例严重失调的现象。从工作后来看，周围环境中男同胞较多，技术讨论和人际关系上女性要融入进来、建立长期合作关系有很大困难。最后，男女选择存在差异，女性科技工作者更倾向于国外的科技公司，因为节奏更加轻松。杨尔弘从三个方面考虑：客观上，男女确实各有擅长，这由生理和基因决定；主观上，女性在成长过程受到更为苛刻的要求，比如硕博招生中很多老师坚决不要女生（包括女老师）；同时，这与女生

性格有关，女性很多都胜负欲不强、默默承担不太耀眼的学术担当工作，获得的荣誉自然少些。

朱小燕同意，男女的多样性差异很自然，没有区别的话这世界的色调就单一了。社会发展和传统文化使得女性肩负延续后代的义务，生儿育女男性替代不了，就像雌性动物保护幼崽一样，很多女性的天性亦是如此。另外，不要用奖项衡量贡献，自己收获的成就感最重要。

徐剑军亦从个人求学经历证明，女性的学习成绩和科研能力甚至比男性强。但是很多女生不会一较胜负、更看重孩子，这和性格、文化等复杂因素有关。高校和企业界存在的男女区别对待现象，客观地讲也是无奈，应该从科研政策和社会政策来避免雇主自主承担女性特殊的生育问题。

三、社会分工与女性角色

社会分工发生很大的变化，科技发展加速了这种变化，女性的优势会逐渐体现出来，角色并没有发生太大改变，如家庭、职场，科技发展对女性的角色产生了怎么样的影响？



图 3：女性如何平衡职场与家庭？

杨红霞指出技术使得人们从繁重的工作中解脱，更多时间投入科研。并且，由于周围男性居多，女性在工作中还是会受到优待和关注，公司也会支持女性的圈内交流和特殊需求。朱小燕认为，科技加速了社会分工、使得分工越来越细粒。单纯从科研女性角度看，从洗衣机到扫地机器人和外卖，真正日常必须的家务在不断减少，关键在于如何利用这些省下来的时间。

四、AI 系统的性别偏见

AI 系统产生伦理或歧视的问题，是由社会，技术，还是使用者导致的？

朱小燕认为这一问题对于基于神经网络的深度模型这种纯粹利用数据的算法是不可避免的，并且使用者手工改正一些错误。徐剑军认为医生中男性偏多、护士中女性偏多这类现象是社会客观存在的，他对论坛报告中黄教授介绍的在数据上直接人为修正的方式存在疑虑。除了性别，包括民族、宗教、肤色等因素，在有些场景可以

考虑，但另一些场景不能考虑，选择权在使用者。总之，科学应该保持中立。

杨红霞则提出了异议，相比光靠算法和数据的 AI 模型，人可以进行推理判断，不需要几万张图片的训练来识别事物。因此，大规模深度学习还是需要专家系统相辅相成，才能做到推理和认知，实现真正的智能。包括推荐系统的马太效应等不公平现象，都说明 AI 系统还有很长的路要走。杨尔弘补充道，很多自动生成的新闻文本，其实也是有伦理问题，其内容导向和趣味性存在的缺陷甚至会拉低整个国民的素质。要关注怎样控制技术向好发展。

五、总结与展望

女性如何在男性占优的领域往前走？

针对女性情绪化和自我贬低等问题，杨红霞认为打铁还须自身强、关键在于自己怎么看待自己。女性有很多得天独厚的优势，比如柔和的交流方式和细腻的心思，在强势的男性群体中做到润滑剂的作用，对生活的关怀也使得团体成员的体感更好。

朱小燕强调“别老想性别问题，自己歧视自己”，要“咬着牙、不要抱怨、吃苦耐劳往前走”。不要质疑自己的能力，但也别痛苦地往前走，女性需要兼顾家庭和事业，就是得不抱怨也不后悔，一句我愿意就完了。兰艳艳补充说，“因为我选择了，所以我就全力以赴”。杨尔弘也认为年青一代在兴趣上有自己的选择权，加上偶像力量，就不怕吃苦了。

展望：我们应当建立什么样的性别创新机制，实现科技领域的性别平等？

朱小燕首先从哺乳角度谈，社会担当和社会机制要让年轻女性即便怀孕生孩子，也能很安静地回到市场。第二，国家法律法规机制要健全；第三，女性一定要维护自己的合法权益，遇到歧视言语和行为保留证据、行使权力、保护自己。杨尔弘设想，可以从充分尊重女性科学家的角度，灵活地设立创新机制，在有上限和下限的范围里尽可能精细化制度。