第2周汇报

——张溢炉

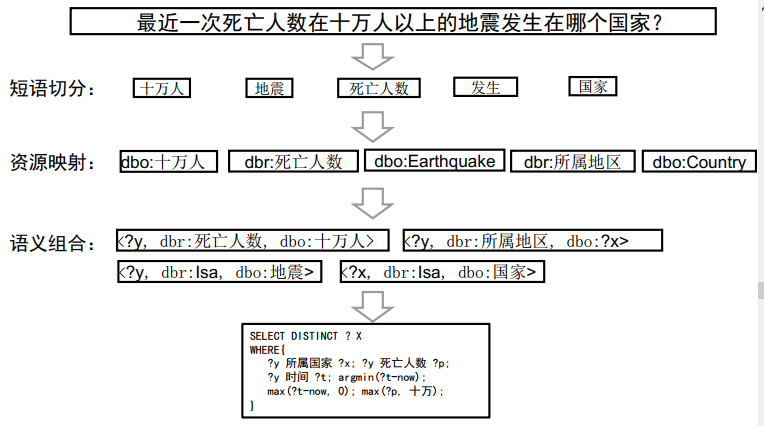
1. 《深度问答技术》总结
2. 内容概述

本文是第三届中文知识图谱研讨会的会议记录，会议的主题是深度问答技术。会议的主要内容有深度问答中的知识表示、关键技术、资源整合和对未来的展望。

1. 总结及理解
2. 涉及人工智能难点问题：自然语言理解、知识表示、知识推理。
3. 发展历程：图灵测试、专家系统、问答式检索、基于知识图谱的深度问答、社区问答和阅读理解式问答。知识图谱本质上是语义网络的知识库。从实际应用的角度出发其实可以简单地把知识图谱理解成多关系图。由节点（Vertex）和边（Edge）构成，一般包含多种类型的节点和多种类型的边。
4. 答案的来源：结构化知识（知识库、抽取事实）、非结构化知识（文本）。结构化知识一般只需要简单预处理即可以作为后续AI系统的输入，但非结构化知识一般需要借助于自然语言处理等技术来提取出结构化信息。
5. 知识表示：符号表示（one-hot）、分布表示。分布表示能将自然语言表示为计算机擅长处理的数值计算。
6. 基于符号语义的深度问答：基于复述的语义匹配（事实库问答）、基于句法树的语义匹配（文本库问答）、基于篇章语义图的语义演算（篇章阅读理解）。
7. 事实问答步骤：问句复述（将问句以计算机易于理解的方式复述）->问句解析（将问句映射到事实库）->查询转换（事实的不同表示）->执行查询，回答问题。
8. 符号表示：准确度好，覆盖度低，泛化能力弱，不方便扩展，推理效率低，用户可理解，流水线方式；

分布表示：准确度弱，覆盖度好，泛化能力强，容易扩展和融合到其他系统，效率相对高，不可解释，端到端方式。

1. 关键技术：

* 知识抽取：包括实体识别、关系识别、实体链接、事件识别。通过基于卷积神经网络的实体关系分类方法可实现抽取实体间语义关系的文本表示。
* 问句语义解析：通过短语切分、资源映射和语义组合可实现自然语言问句到结构化查询语句的映射，如下图所示。规则和模板驱动的传统方法无法处理大规模链接数据上的真实问句。歧义严重是面临的难点，传统pipeline策略可逐步解决每一步的歧义问题，但会造成错误的积累与传递。可尝试采用联合提取的方法解决歧义问题，如基于马尔科夫逻辑网络的联合消歧。联合消歧巧妙运用了语句的潜在联系。
* 推理规则学习：通过推理规则学习可对未知知识进行补全，如基于高斯分布表示学习的补全方法、基于动态映射矩阵的补全方法（可表示知识的多样性）。

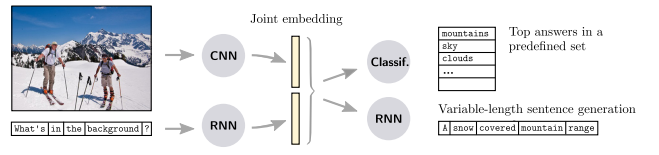
1. 通过基于表示学习对知识图谱、抽取事实和文本统一学习，不仅能够进行知识补充， 还可以学习语言符号与知识符号间的关联、匹配和推理。
2. 启发

通过本文让我对深度问答有了一定的了解。要想让计算机像人一样进行问答，利用传统学习方法手工选择特征很难做到，只能利用深度学习来自动选择特征。

1. 《Visual question answering A survey of methods and datasets》总结
2. 内容概要

本文是对图像问答（Visual question answering，VQA）的概述。首先介绍了一下图像问答，再从动机、方法、性能和局限介绍了VQA的4个主要方法：联合嵌入方法、注意力机制、组合模型、使用外部知识库的模型。然后按时间线介绍了典型的数据库及性能评估方法。之后介绍了注释图像的方法，最后对VQA作了一些讨论、指出了其发展方向并作了简单总结。

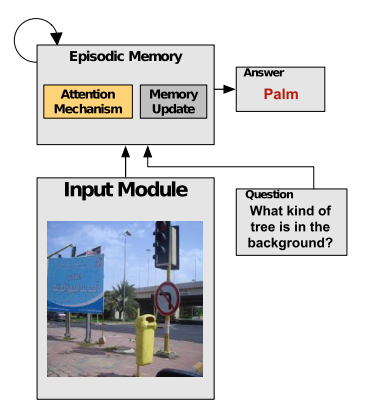
1. 总结及理解
2. VQA有机结合了计算机视觉和自然语言处理（NLP）。它与其它计算机视觉任务最大的区别是问题没有预先设定，是在运行时设定的。图像问答与文本问答有很大相似之处，文本问答的很多方法也能用在图像问答上。图像问答相较于文本问答，将输入的文本换成了图像，这就带来了很多挑战，比如维度变高、噪声变多、图像中包含现实生活的很多信息。相较于图像描述，图像问答可能涉及不包含在图像中的常识。
3. 图像问答用于解决不同形式、开放性的问题。
4. 联合嵌入方法：从文中 图中可看出，大多数VQA模型都要用到联合嵌入方法。

* 动机：使用CNN和RNN模型将图像和问题嵌入到同一个低维空间，方便后续处理。
* 方法：Neural-Image-QA（如下图）、Multimodal QA（mQA）、dynamic parameter layer (DPPnet)、MCB pooling、multimodal residual learning (MRN)等。
* 性能和局限：随着方法的改进，性能不断提高，也说明方法还有待继续改进。

1. 注意力机制

* 动机：找出与问题相关的重要局部特征，用重要的局部特征替换全部特征，减少模型处理的维度，同时减小预测层输入的无关信息或噪声。
* 方法：word-guided attention、question-guided attentionmap(QAM)、stacked attention networks(SAN)、multi-hopimage attention scheme(SMem)、hierarchical co-attention model(HieCoAtt)、Multimodal Compact Bilinear pooling(MCB)等。HieCoAtt会同时关注图像和问题，而在它之前方法仅仅关注图像。
* 性能和局限：注意力机制在所有VQA数据集使用时的准确度都能得到提升，但面对2个选项（是/否）的问题表现很差，可能是这种问题需要更多推理。

1. 组合模型（Compositional models）:到目前为止，所讨论的方法表现出与用于提取图像和句子的表示的CNN和RNN的整体性质相关的限制，可以通过模块化结构的人工神经网络解决这一问题，它为特定所需功能（如记忆功能、特定类型的推理）设计不同的模块。它的两个特别模型如下：

* 神经模型网络（Neural Module Networks，NMN)
* 动机：充分利用问题的语言构成结构，针对每个问题实例即时组合合适模型解决问题。
* 方法：使用NLP中使用的问题的语义解析工具，面对不同的问题组合不同的模块，其中各模块之间相互独立。
* 性能和局限：它通常在具有一定组成结构的问题上表现较好，但在简单问题上表现不好。
* 动态记忆网络（Dynamic Memory Networks，DMN）
* 动机：通过多次传递对数据的多个部分之间的交互进行建模来解决需要复杂逻辑推理的任务。
* 方法：如下图所示，DMN一般由输入模块、问题模块、注意力机制和记忆更新模块组成。
* 性能和局限：与NMN方法相比，在2个选项（是/否）问题上表现差不多，在计数问题上表现差一些，但在其他问题上明显优于NMN方法。

1. 使用外部知识库的模型

* 动机：联合嵌入的方法存在2个局限：只能学到训练集中的知识；神经网络的容量有限。将外部知识库作为由图像读取内容的有效补充，可以有效提高回答超出图像信息的问题的准确率，这里的外部知识包括常识、百科知识等。
* 典型知识库（Knowledge Bases，KB)：DBpedia、Free base、YAGO、OpenIE、NELL、WebChild、ConceptNet。
* 方法：Ahab、FVQA。
* 性能和局限：这种方法在需要超出图像信息的问题表现比较好，但现存的很多数据集中的很多问题都能从图像中找到答案，不需要额外知识。要训练这种网络需要开发新的数据集。

1. 数据集：VQA主要运用的是深度学习算法，同样也需要大数据来驱动，而VQA数据集的标注相对于其他任务要复杂一些，需要考虑问题的类型、问题的长度、答案的形式等，

* 数据集的主要因素：图像的类型，如自然图像、剪贴画、合成图像；问答的形式，如开放性问题、多选项问题；数据集的大小，如图片、问题和表示的不同概念的数量；所需推理的复杂程度；所需超出图像信息的知识。
* 常用数据集
* 自然图像数据集：DAQUAR、COCO-QA、VQA-real、FM-IQA、VQA-real、Visual Genome and Visual7W、Visual Madlibs等。
* 剪贴画数据集：VQA abstract scenes、Balanced dataset等。
* 知识库增强数据集：KB-VQA、FVQA等。
* 其他数据集：Diagrams、Shapes等。
* 写这篇论文时最大的数据集是Visual Genome dataset。
* 问题和答案的生成方式：人工标注、使用模型自动生成。
* 人工标注的数据集会带有一些偏见，如人会倾向于关注一些突出的物体，如桌椅。
* 7W：who, what, where, when, why,how, and which。
* 如何将外部知识库与VQA有机结合还是一个有待解决的问题

1. 结果评估方法：对于答案为选项的问题比较好判断，但对于开放性问题，评估会比较困难，以下是文中提到的一些评估方法：

* 使用字符串匹配简单地测量关于地面实况的准确性。只有完全匹配才被认为是正确的。
* Wu-Palmer similarity (WUPS)：评估分类树中它们共同子序列之间的相似性。 当两个单词之间的相似度超过阈值时，候选答案被认为是正确的。
* 图灵测试。

1. 展望：

* VQA将更好地与知识库相结合。
* 更多计算机视觉和自然语言处理的方法会更好地运用到VQA上。
* 更复杂多样的数据集将会出现。

1. 启发

通读全文后，自己对VQA有了初步的了解。VQA将图像描述、文本问答等计算机视觉和自然语言处理结合起来，CV和NLP的很多方法也能用在VQA上，能在一定程度上代表AI进步，同时也面临巨大挑战。不同类型问题和问题输入及不同类型答案输出需要模型处理更多不同的情况，也意味着需要更大的数据集。而VQA的数据集相比其他数据集标注更复杂，也存在更多偏见和歧义，比如不同的人对相同的问题有不同的答案。个人觉得很难找到单一的模型解决所有问题，需要通过模块化的思想建立串行和并行网络和接入知识库来处理各种不同情况，如通过激活函数的阈值控制不同并行通道的开关来实现各通道的切换，处理不同问题。另外，体会到大数据的量级，如150000个问题，感觉训练集最好尽可能多的包含现实中的各种情况，如各种姿态下的人或猫。VQA方向发展是个新方向，论文更新很快，要关注新的论文。

1. 《中文信息处理前沿技术进展》总结
2. 内容概述

本文首先对自然语言处理（NLP）进行了简单概述。从感知到认知，由运算智能到感知智能，再到认知智能，而NLP正是认知只能的核心，可见NLP的发展前景优越。然后依次介绍了NLP的阶段性发展、所面临的技术挑战及应对之道、行业应用。

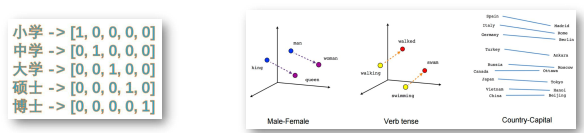
1. 总结及理解
2. 下表有条理、清晰的展现了NLP的一些应用，其中形式=名；语义=名+实；推理=名+实+知；语用=名+实+知+人。另外名为语言符号；实为客观事实、主观事实；知为知识；人为语言使用者。
3. NLP中常用的深度学习模型：循环神经网络、卷积神经网络、递归神经网络。
4. 分布表示较符号表示有了较大改进，由高维变为低维，可以减小训练数据集，可以根据两个词的距离分辨词义相似度。

图 1 符号表示 图2 分布表示

1. 知识图谱的发展将为NLP的发展奠定坚实的基础，将为算法提供大量可靠的先验知识。
2. 模型的预训练相当于把整理好的数据提供给训练模型，可以减小噪声干扰、特征维度等，可大幅提高多任务的准确率。
3. 文本生成的应用意味着计算机在一定程度上拥有了像人一样的创作力，有了思维。
4. 自然语言处理平台的开放很可能像开源软件（如安卓、opencv等）一样极大促进NLP领域的发展。
5. 汉语相较于英语处理更困难。
6. 对话系统从应用到平台化意味着技术的落地，逐步走向成熟。
7. 带标数据不足和常识知识不足是NLP面临的2个主要技术挑战，相信随着NLP的发展，带标数据会越来越多，随着事理图谱和知识图谱的发展，常识不足的问题也会得到解决。
8. 事理图谱和知识图谱的区别：事理图谱研究对象是谓词性事件及关系，谓词一般是动词，而知识图谱研究对象是名词性实体及关系。
9. 智慧医疗将极大提高医院的工作效率，有利于解决看病难等问题，医学大数据和使用经验也将推动NLP的发展；同样智慧教育、智慧司法都可提高工作效率，进一步促进NLP的发展，另外智慧司法也可有效避免徇私枉法等，使法律做到真正意义上的公平、公正、公开。
10. 启发

通过本文对自然语言处理有了初步了解。作为人工智能皇冠上的明珠，NLP的前途无限光明。除了学习NLP相关算法模型外，还应注意分析和理解词、句的含义，将两者有机结合起来

1. 《机器学习在能源与电力系统领域的应用和展望》总结
2. 内容概述

本文介绍了机器学习在能源与电力系统领域的应用和展望，同时，以此为支点，在文中和附录中很好的总结了强化学习（RL）、深度网络（DL）、迁移学习（TL）、平行学习（PL）、混合学习（HL）、对抗学习（AL）、集成学习（EL）。

1. 总结及理解
2. 深度强化学 习 （DRL）作为AI2.0代表在无需先验知识、降低资源消耗、提高训练速度等方面取得了可喜进步,可提高机器对无标签的未知环境的感知能力。
3. AI概念：一般认为AI是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门集控制论、信息论、计算机科学、数理逻辑、神经生理学等学科于一体的新兴交叉学科。其终极目标是让计算机能像人一样思考和行动，并进一步提升人的智能。
4. 推动AI发展的三大关键因素：海量数据资源、超常计算能力、核心算法模型。
5. 人工智能（AI）、机器学习（ML）、深度学习（DL）三者的层次关系是：AI>ML>DL。常见ML算法主要分四大类：有监督式学习、无监督学习、半监督式学习和强化学习。
6. 强化学习（RL）：RL属于主动学习的一种，即本质上是一种在环境中“审时度势”地学习策略的过程，强调如何基于状态而行动，以取得最大化的预期收益。近20年来，以 马尔可夫决策过程（MDP）为严格数学基础的RL算法成为ML领域一个新的突破。算法中的几个关键因素为：状态集、动作集、状态转移分布、贴现因子、奖励函数。常用的优化方法有：值迭代和策略迭代。其中，值迭代更常用一些。
7. 深度学习（DL）：本质上是具有多层隐藏层的神经网络，事实上，4层以上的神经网络就可称之为DL。其基本思想是通过多层的网络结构和非线性变换，组合低层特征，形成抽象的、易于区分的高层表示，以发现数据的分布式特征表示。其避免了手动提取特征的过程。近年来，DL逐渐与其他ML方法结合形成了功能更强的高级混合ML算法。周志华教授提出的“深度森林”未来有可能取代DNN。
8. 迁移学习（TL）：强调的是在不同但相似的领域、任务和分布之间进行知识的迁移。
9. 平行学习（PL）：把“小数据”炒成“大数据”，再把“大数据”提炼成解决具体问题的精准知识的“小智能”的过程，即“小数据、大定律”到“大数据、小定律”的过程（即从牛顿定律到默顿定律的升华），使得ML从已知训练样本集（有限小数据）走向了自我探索获取海量假想训练样本（无限大数据）的时代，这也是AI超越人类智力的分水岭。这样能在一定程度上增大数据集，促进AI的发展。使用预测学习解决如何随时间发展对数据进行探索，使用集成学习解决如何在空间分布上对数据进行探索，使用指示学习解决如何探索数据生成的方向。
10. 混合学习（HL）: 将至少两种优势互补的ML方法集成或综合为一种性能更强的ML。将DL作为“感知”与RL作为“决 策”结合形成HL的一对“黄金组合”。
11. 对抗学习（AL）或生成式对抗网络（generative adversarial network，GAN）：通过构造相互竞争的生成器和判别器来提高学习的效率。
12. 集成学习（EL）: 严格意义上说并不算一种ML算 法，而更像是一种优化手段或策略，其通常结合多个简单的弱ML算法有策略地生成一些基础模型，然后有策略地将它们集成模型组，去更可靠地完成学习任务并做出最终决策。注意其与混合学习的差别。
13. 发展瓶颈：数据，需要大量数据；泛化能力，有时实际泛化能力比训练时泛化能力明显下降；能耗，程序需长时间运行，比较耗电；语义鸿沟，缺乏语言理解能力；可解释性，现有AI过于依赖训练数据，缺乏深层次数据语义挖掘，导致可解释性差；可靠性，现有 AI系统可靠性较差，某些错误识别结果可能会带来致命后果,其为黑箱模型，内部结构难解释。
14. 启发

通读全文让我对机器学习有了初步了解。同时，也让我意识到“没有免费的午餐”定理，每种算法都有自己的优点与不足，找不到通用的算法，要根据实际问题选择合适的算法，也可以考虑把几种算法融合起来，取长补短。算法在不断优化，朝着更好、更快、操作更简单的方向发展。然后阐述了NLP面临的最大的几个开放性问题。

1. 《Frontiers of Natural Language Processing》总结
2. 内容概述

本文是一篇会议记录，会议主要围绕以下三个问题展开：1）什么是NLP？最近几年NLP的重要发展有哪些？2）NLP中最大的开放性问题是什么？3）鼓励从业者了解各地社区并加强合作。首先回顾了最近几年（2001年-2018年）NLP的发展历史。最后回答了2个问题：1）是什么导致该领域走向了错误的方向？2）对于开始学习NLP的在读研究生，有什么建议？

1. 总结及理解
2. 神经语言模型（2001年）：

* 语言模型：通过前一个词预测下一个词。
* 经典语言模型：光滑的n-grams。举例来说：将“informationretrieval”视为一段文本，它的5-grams的items依次为：

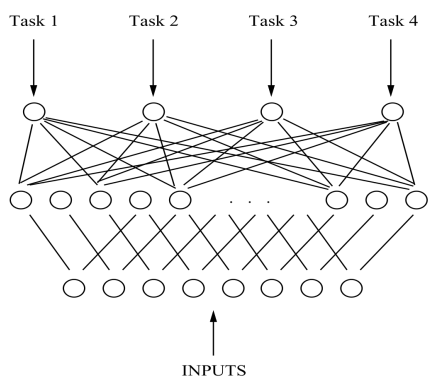
infor,nform,forma,ormat,rmati,matio,ation,tion,ionr,onre,nret,retr,retri,etrie,triev,rieva,ieval

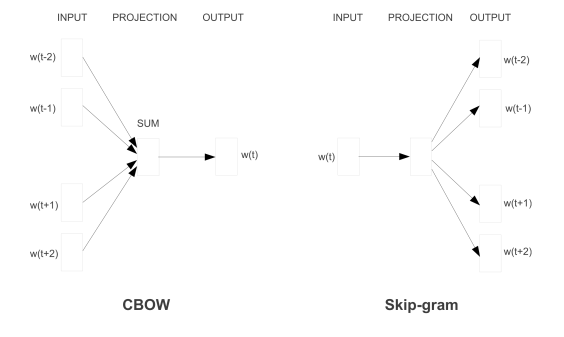
有时为了便于分析，还会在前面加空格，这就多出4个items：

\_\_\_\_i,\_\_\_in,\_\_inf,\_info。

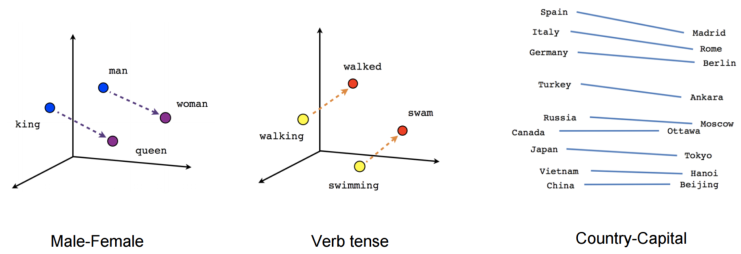
参考网站：<https://www.cnblogs.com/cdsj/p/5720391.html>

* 第一个神经语言模型：基于前n个词的前馈神经网络。
* 因为每个词对应于初始查找层的权重矩阵的一个向量，所以我们经常把这个矩阵作为词嵌入矩阵。
* 之后的语言模型有：RNN、LSTM。

1. 多任务学习（2008年）：模型在不同任务的训练中共享权重，如下图所示，各任务神经元相互连接。
2. 词嵌入（2013年）：自然语言处理（NLP）中语言模型与表征学习技术的统称。概念上而言，它是指把一个维数为所有词的数量的高维空间嵌入到一个维数低得多的连续向量空间中，每个单词或词组被映射为实数域上的向量。

* 主要创新点是：可在很大的未标记数据集通过预训练得到词嵌入矩阵。
* CBOW和skip-gram：两者模型如下图所示。CBOW模型的训练输入是某一个特征词的上下文相关的词对应的词向量，而输出就是这特定的一个词的词向量。Skip-Gram模型和CBOW的思路是反着来的，即输入是特定的一个词的词向量，而输出是特定词对应的上下文词向量。CBOW对小型数据库比较合适，而Skip-Gram在大型语料中表现更好。

参考网站：<https://www.jianshu.com/p/471d9bfbd72f>

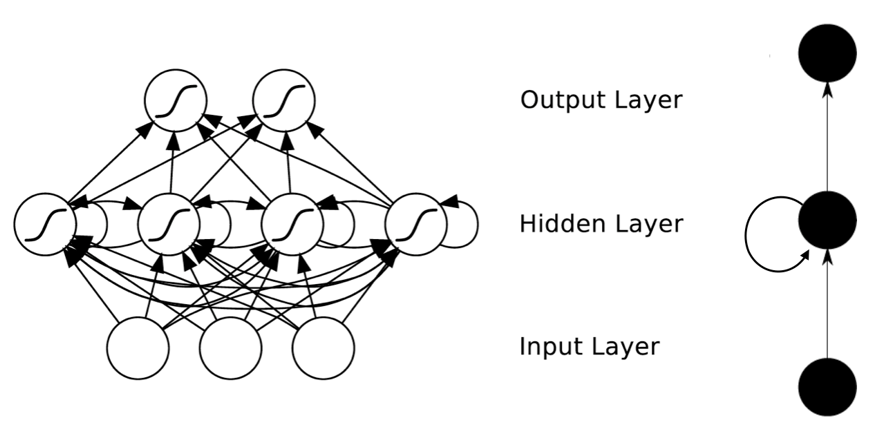
* 词嵌入可以通过对无标记数据集的预训练发掘词之间的某些关系，如下图所示：

1. NLP相关神经网络（2013年）：

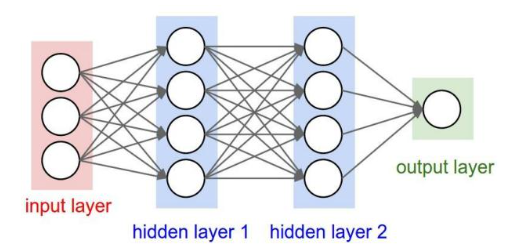
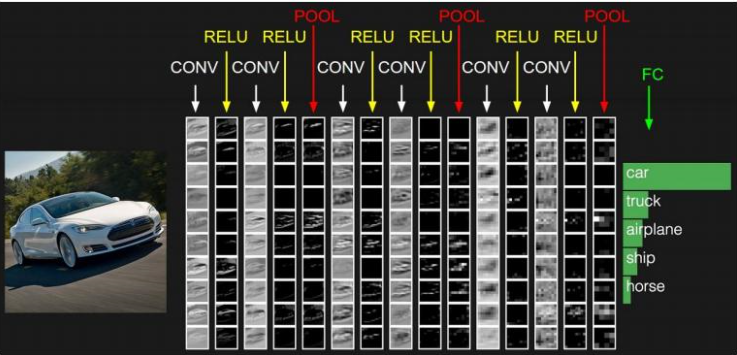
* 主要挑战：处理动态输入序列。
* 三个主要模型：
* 循环神经网络（Recurrent neural networks，RNN）：结构如下图，与CNN不同的是：不仅考虑前一时刻输入,而且赋予了网络对前面的内容的一种'记忆'功能。

之所以称为循环神经网路，即一个序列当前的输出与前面的输出也有关。具体的表现形式为网络会对前面的信息进行记忆并应用于当前输出的计算中，即隐藏层之间的节点不再无连接而是有连接的，并且隐藏层的输入不仅包括输入层的输出还包括上一时刻隐藏层的输出。

参考网站：<https://blog.csdn.net/qq_32241189/article/details/80461635>



* 卷积神经网络（Convolutional neural networks，CNN）：结构如下图，CNN一般包含卷积层、激励层、池化层。卷积层中每个神经元可看成一个滤波器，只关注输入的某个特征；激励层采用函数一般为ReLU，其收敛快，求梯度简单，但较脆弱；池化层用于压缩数据和参数的量，用于图像处理的主要作用是压缩图像。

参考网站：<https://www.cnblogs.com/skyfsm/p/6790245.html>

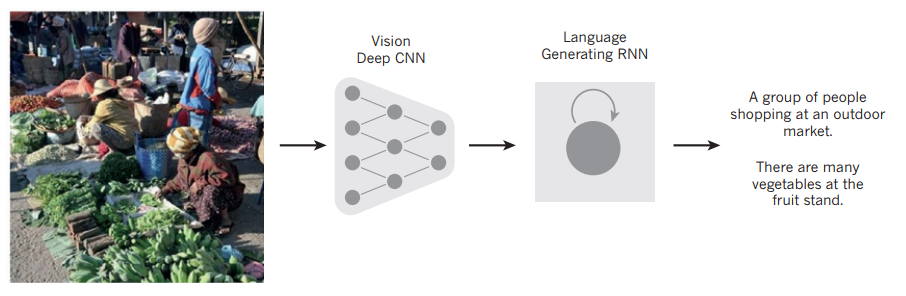
* 递归神经网络（Recursive neural networks，RNN）: 为了能够让模型区分出两个不同的意思，我们的模型必须能够按照树结构去处理信息，而不是序列，这就是递归神经网络的作用。当面对按照树/图结构处理信息更有效的任务时，递归神经网络通常都会获得不错的结果。

尽管递归神经网络具有更为强大的表示能力，但是在实际应用中并不太流行。其中一个主要原因是，递归神经网络的输入是树/图结构，而这种结构需要花费很多人工去标注。想象一下，如果我们用循环神经网络处理句子，那么我们可以直接把句子作为输入。然而，如果我们用递归神经网络处理句子，我们就必须把每个句子标注为语法解析树的形式，这无疑要花费非常大的精力。很多时候，相对于递归神经网络能够带来的性能提升，这个投入是不太划算的。

参考网站：<https://blog.csdn.net/u010164190/article/details/78172289>

* 长短时记忆网络（LSTM）：它在原始RNN的基础上增加了单元状态c来保存长期的状态。解决了梯度消失和梯度爆炸的问题。它用两个门来控制单元状态的内容，一个是遗忘门（forget gate），它决定了上一时刻的单元状态有多少保留到当前时刻；另一个是输入门（input gate），它决定了当前时刻网络的输入有多少保存到单元状态。LSTM用输出门（output gate）来控制单元状态有多少输出到LSTM的当前输出值。

参考网站：<https://zybuluo.com/hanbingtao/note/581764>

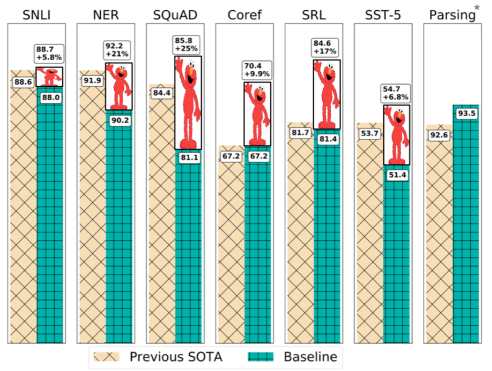
1. 序列-序列模型（2014年）：其主要应用是机器翻译。除了句子-句子的模型，还有图像-句子的模型，如下图所示，其中CNN用作编码器，RNN用作解码器。
2. Attention（2015年）：Attention-based Model其实就是一个相似性的度量，当前的输入与目标状态越相似，那么在当前的输入的权重就会越大，说明当前的输出越依赖于当前的输入。严格来说，Attention并算不上是一种新的model，而仅仅是在以往的模型中加入attention的思想，所以Attention-based Model或者Attention Mechanism是比较合理的叫法，而非Attention Model。

参考网站：<https://blog.csdn.net/u014665013/article/details/82619808>

<https://blog.csdn.net/hahajinbu/article/details/81940355>

1. 基于记忆的神经网络（2015年）：attention具有模糊的记忆，为了弥补这个缺陷，提出了具有更加清晰记忆的模型：神经图灵机、记忆网络、end-to-end记忆网络、动态记忆网络、神经可微计算机、循环实体网络等。一个记忆网络包括了记忆m，还包括以下4个组件：I (input feature map)、G (generalization)、O (output feature map)、R (response)。

参考网站：<https://blog.csdn.net/u014248127/article/details/84894739>

1. 预训练语言模型（2018年）：它仅用于初始化的第一层，在大数据集的预训练中可以获取很多附带的信息。如下图所示，应用预训练模型，在处理很多不同的任务中获得很大的提升。
2. NLP面临的最大问题：自然语言理解；低资源场景的应用；大文本或多文本的推理；数据集、定义问题和评估方法。
3. 启发

跟着会议记录的时间线阅读，让自己对NLP近几年的发展有了初步了解。会议最后的两个问题也给了自己较大的启发：要敢于尝试不同的模型和方法，成功和失败的案例将一起推动NLP螺旋式发展；可以通过阅读大量机器学习、深度学习、强化学习文章学习相关知识，并将它们与NLP有机结合起来；要勤于动手实践，学会调节模型和设计评估标准，用代码来验证理论的可行性；不要过分关注某篇文章，若不能验证其可行性的情况下，可等待更多相关论文的发表。