



语言智能是人工智能皇冠上的明珠。如果语言智能实现突破,跟它同属认知智能的知识和推理就会得到长足的发展,进而将会推动整个人工智能体系的进步,也会实现更多应用场景的落地。

让机器说话怎么这么难

■本报记者 贡晓雨



语言智能是人工智能皇冠上的明珠。

图片来源:百度图片

“相对于计算机视觉等其它人工智能学科,自然语言处理是尚未受到大众关注的学科,但是自然语言理解却是当今人机自然对话的最大瓶颈。”第二届语言与智能高峰论坛上,华为技术有限公司诺亚方舟实验室主任李航号召大家多关注自然语言理解(NLP),称这是巨大的蓝海。

语言智能是人工智能皇冠上的明珠,用自然语言与计算机进行交流,获取合适的信息,得到满意的服务,是人们长期以来所追求的。如果语言智能实现突破,跟它同属认知智能的知识和推理就会得到长足的发展,进而将会推动整个人工智能体系的进步,也会实现更多应用场景的落地。

只有通过自然语言理解,才能实现智能与人类的无缝对接,实现真正意义上的人工智能。然而要实现这一技术,让机器人完全如人类一般“开口说话”,并不容易。

自然语言理解为何难

“语言是一种复杂的现象。”李航解释说,自然语言有自身的重要特点,使得计算机实现自然语言处理很困难。

语言是不完全有规律的,即使有,也是错综复杂的。“语言是经过上万年的时间演化而来,知识库的建立过程一定会出现功能冗余、逻辑不一致等现象。但是语言依旧有一定的规律,若不遵循规范,交流会比较困难。”李航说。

语言是可以组合的,其重要特点是能够将词语组合起来形成句子,能够组成复杂的语言表达;语言是一个开放的集合,我们可以任意地发明创造一些新的表达。“约定俗成的话语一旦形成之后,大家都会使用,形成固定说法。语言本质的发明创造就是通过比喻扩展出来的。”李航说。

不仅如此,语言还需要联系到实践知识,使用也要基于环境。如果在外语的语言环境里去学外语,人们就会学得非常快,理解得非常深。李航认为,在计算机里去实现与人一样的语言使用能力是一件非常具有挑战性的事情。

“首先,语言的不完全规律性和组合性,就意味着如果在目前的计算机上去实现,会产生组合爆炸;还有,如果需要语言做比喻,去联系到实践环境,就意味着要做全局的、穷举的计算。如果通过现代计算机来做,非常复杂,几乎不太可能。”李航说。

所以,如果想让计算机像人一样使用语言,原理上需要完全不同的、与人脑更接近的计算机体系架构。

其本质原因是,目前在计算机上去实现东西一定需要数学模型。也就是说,计算机能够做的事情要通过数学形式化。但是,到目前为止,语言的使用还不清楚是否能够用数学模型去刻画。

“语言是不精确的,字面意思背后还有太多太多。”美国伊利诺伊大学芝加哥分校的计算机科学教授刘兵说。这也是为什么相较于一对一表征的图像和语音,自然语言处理是一个如此艰巨的问题。

让机器终身学习

“从目前的理论和技术现状看,通用的、高质量的自然语言处理系统,仍然是较长期的努力目标。”李航说,但是针对一定应用,具有相当

自然语言处理能力的实用系统已经出现。

目前,所有的自然语言处理的问题都可以分类成为五大统计自然语言处理的方法或者模型,即分类、匹配、翻译、结构预测、马尔可夫决策过程,这些技术主要采用统计机器学习的方法来解决。

“各种各样的自然语言处理的应用,都可以模型化为这五大基本问题,基本能够涵盖自然语言处理相当一部分或者大部分的技术。”李航说。“自然语言处理的不纯粹是语言的问题。语言可能还要与视觉、听觉连在一起,让机器能够了解世界。”刘兵表示,深度学习的未来是终身学习。

刘兵解释说,不论是神经网络还是其他算法,都是在孤立地学习,现在的模型还无法积累信息。“我们想做的是像人类那样,让机器可以一直不停地学习。这个学习机制与现在的不同:要学,要记,还得能适应。”

深度学习是否代表了一种算法,将智能包

含在了里面,刘兵并不敢肯定,他倾向于认为不是。深度学习还是一种函数逼近和映射,通过一些列举的例子发现关系,真正的智能不应该是如此算法。

刘兵更关心终身机器学习还须“适应”相关的内容,深度学习里面有太多的参数,但怎么将知识真正充实在算法里,将来做微小的变化就能解决新的问题,“我不知道深度学习能不能做到这一点,我现在感觉目前的深度学习似乎还没有做到”。

技术不成熟无碍应用

人工智能的更高挑战是自然语言理解。现实中,虽然自然语言理解非常困难,但我们仍希望计算机能够越来越智能化,能够部分使用语言。

实际上,技术虽然不完善,应用却从未停止。就如问答这类应用,自然语言处理做的第一件事就是把问题简化。“先分析一下问句,接着去检索相关的知识或者信息,然后产生答案。”李航解释说。

现在的自然语言处理,本质是用数据驱动的方法去模拟人,通过人工智能闭环去逼近人的语言使用能力。但是,这种技术并没有真正实现人的语言理解机制。

“每项技术,尽量达到上界固然好,但应用中对于下界的要求是不一样的。”李航说。比如问路,其实听懂对方几个单词就能明白意思,“让机器翻译给我几个单词就行了,这时我们对翻译器的性能要求是相对较低的”。

不同的应用,用户对使用性能的要求不同,如果下界达到这个水平,就能满足用户使用。“但是每一个应用是否都能够达到我们要求的性能下界,就不好说了,要看未来的发展。”李航说。

“我们做的是以结果为导向,从市场论证技术,推动它的成熟。”京东商城智能通讯部总监刘丹认为,技术始终处在不稳定,“当技术发展另一个阶段又会不成熟,但是因为大数据的存在,就能帮助我们的产品提升,我们只要保证做的产品能够给消费者带来价值就好”。

“产品和技术是可以互补的,收缩和控制人们的使用预期,把产品的可能性发挥到最大,这需要技术研究者 and 产品生产者共同解决。”百度自然语言处理部门副主管赵士奇说。

酷技术



一种轻薄的低过敏性电子传感器。

图片来源:材料网

会呼吸的健康监测器

近日,日本东京大学的一个科研团队研发出了一种低过敏性电子传感器,这种很轻薄的传感器可以在皮肤上连续佩戴一周,且不会造成任何不适,甚至用户都忘记了它们的存在。这种弹性电极由透气纳米网片构成,有望用作长期持续监测人体健康的非侵入性电子皮肤装置。

近年来,可穿戴电子产品取得了重大进展,下一代产品采用轻质、高弹性材料,直接附着在皮肤上,以实现更灵敏、精确的测量。然而,尽管这些装置中使用的超薄薄膜和橡胶贴片与皮肤黏附性很好,但是对于长期使用而言,其缺乏透气性这一点被认为是安全隐患——皮肤病学试验显示,细小的可拉伸材料会妨碍出汗并阻塞皮肤周围的气流,造成刺激和炎症,最终可能对生理和心理都造成长期的影响。

“目前,在医疗和体育行业,急需一些可穿戴并持续监测达一周以上的设备。”东京大学工程系教授染矢高雄说,该教授的研究团队在此之前已经研发出了一种测量血液含氧量的皮肤贴片。

现在该团队开发的这种电极,由含有水溶性聚合物、聚乙烯醇(PVA)和与人体生物相容性良好的金属层状材料的纳米网片所构成。在该设备上喷洒少量的水,将PVA纳米纤维溶解,可以使其很容易就黏附到皮肤上——该设备无缝贴合人体皮肤的曲线表面,包括毛孔和指纹。

研究人员对20名受试者进行了皮肤贴片试验,在佩戴该设备一周后,没有受试者产生炎症。该团队还评估了该纳米网片的水蒸气渗透率,和超薄塑料薄膜以及橡胶薄片对比,其多孔网状结构具有更优异的气体渗透性。

此外,研究人员将该电极贴附在手指上,通过1万次以上的反复弯曲和伸直试验,证明了该装置的机械耐久性;同时,该装置对肌肉电活动的读数与常规凝胶电极获得的读数相差无几,也证实了其作为肌电图记录电极的可靠性。

“该设备可以做到监测患者的生命体征,而不引起任何压力或不适。”染矢高雄说,除了护理和医疗应用外,该设备还能够对运动员的生理信号和身体运动进行持续而精准的监控,而不会妨碍他们的训练或表现。(李惠钰)

Lp(a)检测方法标准化 帮助准确评估 ASCVD 残留风险

■薛坤

日前,在西安举办的“2017中国医师协会检验医师年会”上,首都医科大学附属北京安贞医院/北京市心肺血管疾病研究所王绿娅教授就血脂管理在ASCVD中的应用及治疗新进展进行了深入分享与探讨,并阐释了新兴危险因素脂蛋白(a)(Lp(a))的临床应用价值。王绿娅指出:“血脂检测是ASCVD风险评估的关键。在血脂七项检测中,低密度脂蛋白胆固醇(LDL-C)增高可提升ASCVD发生风险;但其它心血管病残留风险也不容忽视,即使LDL-C达标,血压和血糖控制良好,但心血管疾病的发病和风险事件依然存在。即使多种危险因素在同时治疗,残留风险也不能消除。其中Lp(a)是ASCVD显著的、独立的残留风险因素,对于提高ASCVD风险评估准确率、改善患者诊疗具有重要意义。”

Lp(a)是预测ASCVD残留风险的显著因素

Lp(a)最早于1963年由挪威遗传学家Kare Berg在血浆中发现,其蛋白颗粒的结构类似于低密度脂蛋白(LDL),但含有独特的载脂蛋白(a)。目前对于Lp(a)生理作用的研究尚缺乏充足证据,但其致病机制已被广泛认可,即促炎、促动脉硬化作用和促血栓抗纤维溶解。

JUPITER研究是一项探讨强效他汀治疗降低心血管事件的随机对照临床试验,共纳入17802人。该研究的后续分析发现,接受他汀治疗前基线Lp(a)水平高的个体,与具有正常Lp(a)水平的个体相比,心血管疾病风险升高。Lp(a)是残留风险的重要决定因素,在识别无症状个体CVD风险上有潜在用途。

LIPID研究是一项探讨他汀治疗冠心病的随访随机对照试验,其延长研究评估了7863例患者的基线血浆Lp(a)水平。研究同样发现,基线Lp(a)与未来心血管事件有关,当Lp(a)水平处于Lp(a)最高四分位数的上限(即>50mg/dL)时,发生后续事件的风险高,因此提出将Lp(a)>50mg/dL作为强大临界值。

值得注意的是,不同种族的Lp(a)临界值可能不同,有研究提示我国人群Lp(a)切分点低于高加索人种,可能是20mg/dL。

流行病学也支持将Lp(a)作为重新评估风险的指标,一项对826名患者进行的15年前瞻性随访研究发现:被Framingham危险评分模型和Reynolds危险评分归为中等风险的患者中,根据Lp(a)水平,约39%中危组应进行重新分层。

此外,Lp(a)升高是急性脑梗复发事件的危险因素。柏林C&S研究对250名急性缺血性脑卒中患者进行了12个月随访并测定Lp(a)水平。在157名Lp(a)水平正常的患者中,有11名(7%)在中位时间161天发生心脑血管事件,而在93名Lp(a)水平升高的患者中,15名(16%)在48天的中位时间发生事件。研究表明,Lp(a)水平升高的患者复发事件风险显著升高。研究还提示血清Lp(a)升高与脑卒中发病呈正相关,是卒中事件复发的独立危险因素。因此,Lp(a)升高被认为是一个明确的危险因素,可为临床诊断和鉴别诊断提供有力的依据。

2016年《柳叶刀》(Lancet)上发表的一项前瞻性队列研究发现,Lp(a)是家族性高胆固醇血症(FH)的独立危险因素。FH患者中有1/4的患者Lp(a)水平升高,如果Lp(a)水平大于50mg/dL要高度怀疑是否是FH。研究建议所有临床诊断明确的FH患者均应检测Lp(a),以识别心梗风险。现有数据提示我国FH患者中Lp(a)升高的比例高于国外,且Lp(a)升高与FH患者冠心病严重程度相关。

王绿娅指出:“众多研究均表明了Lp(a)在预测ASCVD残留风险中的重要作用。为了给临床提供治疗的策略,已有很多心血管相关指南推荐将Lp(a)作为筛查指标。”欧洲动脉粥样硬化协会(EAS)推荐以下人群应该至少接受一次Lp(a)检测:有心脏病和中风历史者;FH患者;有心血管疾病及Lp(a)升高家族史者;有反复心脏病史,长期服用他汀者等。此外,随着Lp(a)检测临床应用的正规化,EAS在2012年发布的《Lp(a)筛查管理临床手册》帮助医生正确应用和管理Lp(a)。

Lp(a)检测方法标准化 改善心血管疾病风险评估

受个体遗传基因的影响,Lp(a)呈现相对分子量大不一的多态性。Lp(a)早期采用的方法是检测每单位体积的质量(即mg/dL),用这种质量结果评估患者心血管疾病风险可能导致错误分层。美国一项调查也证实了这点,在易患冠心病的高危人群中仅有37%的人Lp(a)水平增高,而在低危人群中也有14%的人Lp(a)水平增高。因此,质量检测会造成检测值偏差,影响风险评估,该方法不能为临床医生提供评估个体心血管疾病风险的正确值。

EAS和美国心肺血液研究所(NHLBI)一致推荐用摩尔浓度(即nmol/L)来表达Lp(a)颗粒数,替代mg/dL质量浓度,新的方法应可溯源至IFCC/WHO国际临床化学学会/世界卫生组织的参考物质SRM 2B,可正确评估心血管疾病风险。同时指出,Lp(a)水平在75nmol/L以上的个体,提示心血管患病风险上升。

目前,Lp(a)标准化的难点是apo(a)大小的多态性。因不同标准品中apo(a)大小随机,故其不能代表所有待测标本中apo(a)颗粒的大小;同时也会影响与抗原决定簇结合的抗体的反应性,导致检测标本中的apo(a)结果过高或过低。

王绿娅表示:“过去受到检测技术及用药的限制,很多临床医生不看好Lp(a)这个新兴的检测指标,导致Lp(a)临床应用研究受到很大的限制。但近十年来,标准化检测方法已经出现,使得Lp(a)可以得到更精准和更科学的检测。”

罗氏诊断Tina-quant® Lp(a)二代试剂盒是首个遵循EAS指南、采用nmol/L标准化单位的产品,溯源至IFCC/WHO SRM2B参考物质及ELISA参考方法,采用颗粒增强免疫比浊法,该方法可准确检测Lp(a)微粒数量而非质量单位,可以提供不受多态性影响的可靠检测结果,帮助临床医生更准确地评估CVD疾病风险。

科学家发现功能活性更强的间充质干细胞

前沿点击

干细胞的应用性研究日新月异,发展迅速。在干细胞诸多类型中,间充质干细胞是目前研究较为充分的干细胞类型。间充质干细胞对人类健康的益处已得到充分证明,尤其在促进血管新生方面的作用。能否将间充质干细胞进一步分群,找到优化的干细胞,进一步提升其促进血管新生的能力?基于这一思路,对间充质干细胞亚群的功能进行细分研究,成为新的研究热点。

中国医学科学院北京协和医学院教授韩忠朝研究团队首次将胎盘间充质干细胞中功能活性最强的细胞亚群鉴别出来,并申请了专利,这就是CD106分子阳性的胎盘间充质干细胞。

CD106,也称为血管细胞黏附分子1,在间充质干细胞发挥免疫调节和促血管新生的功能中起到了关键作用。韩忠朝院士团队首次发现CD106阳性细胞在胎盘间充质干细胞占70%左右的比例。进一步的研究证实,相比其它间充质干细胞,胎盘CD106阳性间充质干细胞具备更强的促血管新生的活性。

早在2003年和2005年,韩忠朝带领的研究团队就发现间充质干细胞在糖尿病血管病变中的作用,其机制就是促进血管新生。韩忠朝研究团队证明了其优异的促血管新生的活性,相关研究成果已发表在《国际刊物》上。

研究工作主要是通过体外实验和体内实验来验证。在体外实验中,将CD106阳性的胎盘间充质干细胞接种到一种特殊的凝胶上,称为基质胶。接下来,干细胞一边自我扩增,一边分化形成血管。与CD106阴性间充质干细胞相比,CD106阳性间充质干细胞在基质胶上形成了更丰富的新生血管。

实验小鼠的血管被结扎后,在小鼠下肢注射CD106阳性的胎盘间充质干细胞。实验表明,过一段时间后,新的血管慢慢生长了出来,与注射生理盐水和CD106阴性间充质干细胞的小鼠相比,注射CD106阳性间充质干细胞的实验小鼠下肢生长出了更丰富的新生血管。

此外,通过采用凝胶做载体,韩忠朝研究团队还将CD106阳性胎盘间充质干细胞用于严重烧伤和糖尿病伤口的愈合。目前,这一新型专利产品已进入正式临床试验,取得了令人振奋的结果。(王伟强)