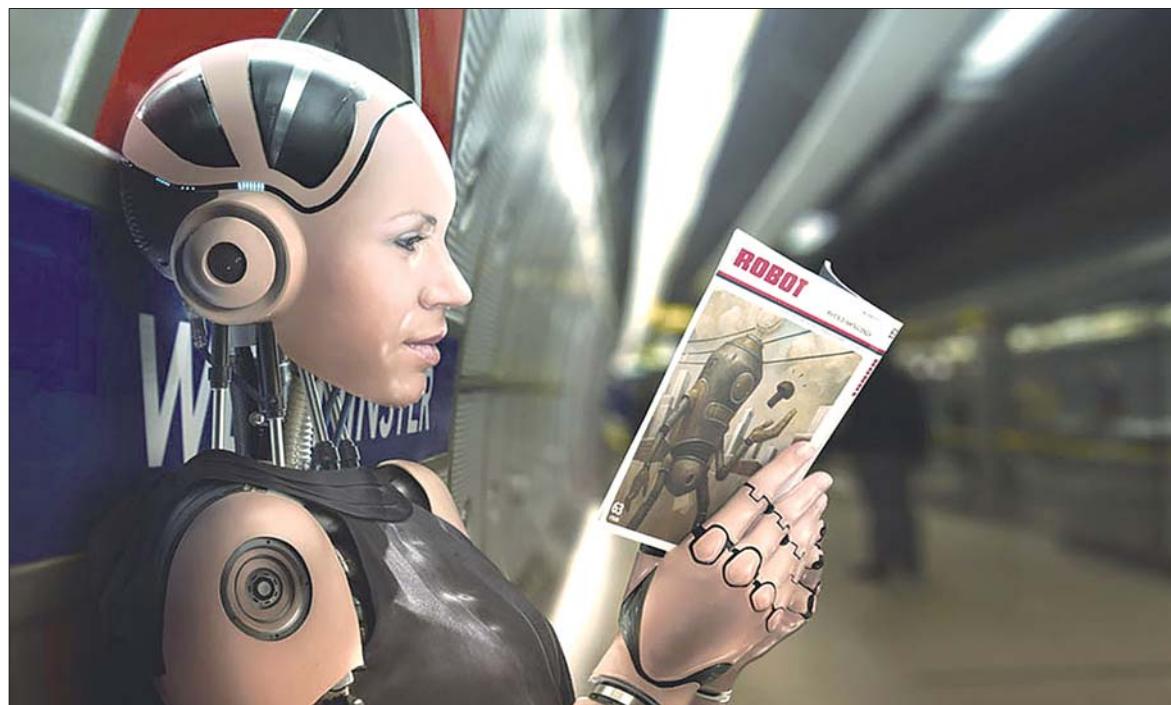


AI首次在阅读理解上击败人类 人工智能革命进入“下半场”



本报记者 王昱

比人类更会做阅读理解

据报道,1月11日,由斯坦福大学发起的机器阅读理解领域顶级赛事SQuAD刷新了排名, AI的阅读能力历史上首次超越人类。美国软件巨擘微软(Microsoft)与阿里巴巴并列首位,获得的分值均超过了人类。

据悉,SQuAD比赛的方式是构建一个大规模的机器阅读理解数据集(包含10万个问题),文章来源于500多篇维基百科文章,AI在阅读完数据集中随机抽出的一篇短文之后,需要回答若干个基于文章内容的问题,然后与标准答案进行比对,得出精确匹配(Exact Match)和模糊匹配(F1-score)的结果。

参赛者可以利用这个数据集进行模拟训练,并且通过开放平台提交自己的算法用于评分。

由于SQuAD能提供庞大的数据规模,这个测试被认为是当前世界监测机器阅读水平的最权威测试之一,能够检测出机器学习模型能否在处理大量信息后给出问题的准确答案。

11日,来自阿里巴巴iDST团队的SLQA+模型最终取得了82.440分的成绩,超越了人类的82.304分;而前一天参与挑战的微软亚洲研究院自然语言计算组的R-NET+模型,也在EM值(即Exact Match)上取得了82.650的最高分。

阿里巴巴研究院自然语言处理首席科学家司罗指出,在类似“天为什么会下雨”这种客观性问题方面,人工智能的回答准确率较高;此次技术的重大突破源于研究团队提出的“基于分层融合注意力机制”的深度神经网络模型。该模型能够模拟人类在做阅读理解问题时的一些行为,包括结合篇章内容审题,带着问题反复阅读文章,避免阅读中遗忘而进行相关标注等。

司罗表示,对于解决维基百科类客观知识问答,机器已经取得非常好的结果,将继续向对通用内容的“能理解会思考”的终极目标迈进。今后,研发的重点在于把这项技术真正应用在广大实际场景,让机器智能普惠生活。此外,该自然语言处理团队还支撑了阿里巴巴整个生态的技术需求,由他们研发的AliNLP自然语言技术平台每日调用1200亿+次,Alitrans翻译系统提供20个语种在线服务日调用量超过7亿+次。

AI会读书,这事有点大

人工智能在阅读理解上战胜人

最近这些年,人工智能赶超人类的新闻已经屡见不鲜——它们在国际象棋中战胜了卡帕罗夫,在围棋中战胜了柯洁。但无论这些AI怎么折腾,在听人话、说人话这门技艺上,人工智能之前还一直在三岁小孩牙牙学语的阶段徘徊。

但这一切在最近突然改变了。在由斯坦福大学发起的SQuAD文本阅读理解挑战赛中,来自阿里巴巴和微软团队的人工智能模型分别以高分战胜了人类选手,位列榜单的前两位。是的,就是那个曾让你在学生时代痛苦无比的“阅读理解”,人类已经做不过人工智能了!而人工智能的又一次胜利,也许意味着AI革命将正式进入指数级增长的“下半场”。

类的新闻,并不仅仅是在它的成就榜单上增添一项新纪录,这背后的意义可不容小觑。

与之前在国际象棋、围棋等智力游戏上超越人类完全不同,在处理阅读理解相关问题时, AI不仅要运算和记录,还要主动去分析和理解,所以阅读理解问题一直被认为是一个标志性临界点。当这个点被AI破解,意味着AI已经能接管很多必须由人类才能完成的工作。因为阅读理解是一个处理和关注词汇、语句、篇章结构、思维逻辑、辅助语句和关键句等等元素构成的复杂组织网络。如果这些问题AI都能做,甚至比人类做得更好、做得更快,还要我们人类做什么?

这则消息的另一个价值,是让我们惊觉人工智能在语言领域的发展速度——也许在大多数人头脑中,都曾经默认语言这种东西,将会是人工智能最后赶超人类的一个领域。基于这种认识,很多“什么职业不会被人工智能取代”的报道,还喜欢把翻译、心理咨询师等需要语言技术的工作当做人工智能抢不走的铁饭碗。

然而现实确实,在过去不到十年的时间里,人工智能正以惊人的速度飞速赶超人类。它从只能翻译单个单词、词组(比如曾经时兴的电子词典),到能够翻译一句话、一段文字(比如谷歌、百度“机翻”),再到如今能够立即领会一篇文章的整体意思,就像我们小学老师要求我们做的那样:归纳文章中心思想、段落大意,并从中提取信息。

正如很多有识者所指出的那般,

人工智能近几年在语言上的学习速度,其实已经与一般儿童学习语言的速度相当了:你养一个孩子,让他从一个单词一个单词地牙牙学语,到上中学做阅读理解,怎么也得花上十年的时间吧。问题是,人类对语言的理解能力是有成长极限的,一般人到了中学做阅读理解这个层面,语言能力也就没法再继续增长了。而AI很可能没有这个“天花板”,而且成长速度会越来越快。设想再给它十年时间,人工智能对语言的理解能力会达到什么程度?这个问题的答案既令人期待,又给人带来一丝莫名的恐惧。

当它比你更懂你自己

人工智能语言能力突破瓶颈,首先会对一些产业造成冲击,最直接的影响是大多数今天还必须由人工完成的规则、对话、服务信息类的相关理解工作,都可以考虑用不眠不休、高运算速度的机器来取代。

比如说,在不远的将来,你再给公共服务热线打电话,很可能不必再被强迫听到“XX请按1, XX请按2, 退出请挂机”那种槽点颇多的生硬声音,而是一个能真正懂你意思的人工智能为你服务。更妙的是,由于人工智能的信息记录、分析和处理能力远强过人类,它所能做的将不仅仅是像普通话务员那样将问题记录并上报,而是综合分析处理其中的信息——比如,某地如果接到集中投诉, AI会迅速判断出当地出了什么问题,并给出精准的处理意见。这将大大提高服务效率。

不过,人工智能理解语言更重要的意义还不在于此,而是意味着人类的海量信息库将向AI敞开——众所周知,人工智能之所以能在围棋、象棋上战胜人类,主要原因是研究人员用之前人类的大量棋谱去“喂”它,人工智能通过“深度学习”理解了这些棋谱,实现对人类反超。而这些棋谱与人类用文字写成的书籍比起来,不过是九牛一毛。如果海量的棋谱能“喂”出一个“AI棋圣”,更加海量的文字信息又会“喂出”什么呢?

想象一下,很多读书人毕生梦想中的“读尽天下之书”的理想,将以一种不可思议的方式实现。而达成这个目标的个体,并非他们幻想中的圣人,而是一个“算法”、一段程序。当人工智能能够在短时间全面而彻底地掌握人类所有的文字信息,它会成为什么?毫无疑问,它将成为最为博学的百科全书学者、最及时的舆情分析师、最善解人意的全人类的倾听者,它将成为比人类更懂人类的存在。这个场景,随着人工智能学会“读书”,已在悄然临近。



高盐饮食如何影响血压

盐是影响血压的重要因素,公元前2600年的《黄帝内经》中就说“咸者,脉弦也”。100多年来,科学工作者通过动物实验、人群观察以及试验研究对盐和血压、盐和高血压病之间的关系进行了全方位的研究。

如今,随着生物技术迅速发展,对这一命题的研究进入了分子水平。目前,我们已经知道膳食钠摄入和血压升高之间有直接的关系。

一个经典的例子是在大猩猩中进行的试验:在正常饮食中加入额外的盐以后,大猩猩的血压就逐步上升,并维持在一个较高的水平。额外的盐的量越大,血压上升的水平越高。当停止加入额外的盐后,血压就逐渐下降至原有的水平。

流行病学调查证实,人群的血压水平和高血压的患病率均与食盐的摄入量密切相关。世界卫生组织调查发现平均盐摄入量很少的人群(如美国阿拉斯加的爱斯基摩人和中国四川凉山地区的彝族居民)高血压患病率很低,而且血压随年龄升高的趋势不明显。但当这些人群采用了盐摄入量较高的“现代”生活方式后,血压水平就会上升。我国北方人盐摄入量高于南方人,高血压的患病率也呈北高南低的趋势。

那么,高盐摄入是如何引起血压升高的呢?研究者通过大量的实验得出这样几种结论:盐中的钠离子过多引起水钠潴留,导致血容量增加,血压上升;引起细胞(包括平滑肌细胞)水肿,使血管腔变窄;增加血管对儿茶酚胺类缩血管因子的敏感性;细胞内钠离子增加后会抑制钠-钙交换,使细胞钙排出减少,导致血管平滑肌内钙离子浓度上升而引起血管平滑肌收缩。

可以说,钠离子具有调节血容量、血管弹性和血压的功能,这就决定了它与高血压之间的不解之缘。

我省是心脑血管疾病的高发地区,山东省第三次死因回顾调查结果显示,心脑血管疾病排在首位,报告死亡率266.69/10万;同时也是55岁以上中老年人群的首位死因。因此,减盐对于我们的日常饮食尤为重要。正常人减少盐的摄入量,可预防高血压的发生,减少脑卒中、心脏病和慢性肾脏病的发病和死亡。高血压病人减少盐的摄入量,可改善高血压治疗的效果,减少脑卒中、冠心病和慢性肾脏病的发病和死亡。

据预测,如果我省居民平均每日食盐摄入量控制在9g以下,每年可以减少3.6万例新发高血压患者,与高血压有关的心脑血管疾病死亡率能降低50%,居民人均期望寿命可延长3~4岁。

(任志方)

3D打印技术 首次造出类脑组织

美国《趣味科学》网站日前报道称,英国科学家近日使用新的3D打印技术,首次打印出像人脑一样柔软的类脑组织,朝最终3D打印出功能齐备的完整大脑迈出重要一步。

以往,只有相对硬一些的材料可被3D打印出来。此次,研究人员使用一种新型复合水凝胶,打印出三维支架,然后用胶原蛋白包裹打印出结构,并用人类细胞进行填充,得到了类脑软组织。

为解决打印软组织难这一问题,研究人员在打印过程中进行了降温处理。这项研究主要作者、伦敦帝国理工学院机械工程系研究人员谭正初(音)说:“我们使用的是低温打印过程,这意味着,下层被冻住了,这使其非常稳固,新层结构可以打印在上面,不会出现底层崩塌的现象。打印完成后,我们可以让得到的物体慢慢解冻,并保持形状。”

据悉,该技术目前还存在不少局限性,比如,他们仅能制造出类脑组织的小样本,而非整个大脑。要使用3D打印技术获得功能齐备的复杂人类器官(大脑或肺等),可能还要等数十年时间。

(据《科技日报》)