

对话“顶流”ChatGPT:你的“智慧”从何而来

本报见习记者 涂佳煜 林晓晖

从2022年底开始,来自全球各地的人都不约而同地绞尽脑汁,用各种刁钻的要求“调教”着一款人工智能——ChatGPT。

AI一一应战,给出了史无前例的惊人表现,短短几个月间,ChatGPT在全球圈粉无数。

ChatGPT搞定了论文、代码、报告,还能进行诗歌、剧本、小说等“文学创作”。这番大显身手,引起了一场轩

然大波:比尔·盖茨说它的出现不亚于互联网的诞生;马斯克说我们离“强大到危险”的AI不远了;一大批预言家站了出来,大胆预测码农、分析师、新闻编辑等从业者即将失业……ChatGPT俨然成为赛博空间的“顶流”,越来越多的人对它的真面目产生了深深的好奇。

ChatGPT的“智慧”从哪里来?它真的已经无所不能了吗?记者带着疑问,和ChatGPT聊了聊它的前世今生……

一个对话大师的诞生: “学习人类语言并不难”

在感到“阳”了之后,为了保持健康和积极的心态,您应该注意以下几点:

1. 保持健康的生活方式:保证充足睡眠,锻炼,均衡饮食。
2. 关注身体变化:如果有任何异常症状出现,请立即寻求医疗帮助。
3. 充分休息:避免过度劳累和压力。
4. 维护社交关系:与家人朋友保持联系,保持积极心态。
5. 注意心理健康:处理消极情绪,避免过度焦虑。

通过注意以上事项,您可以帮助自己保持健康,快乐的生活。

ChatGPT对“阳了之后该注意什么”问题的回答。

与ChatGPT对谈两回,记者不禁惊呼,这一次,AI终于“听得懂人话”了。

为什么这么说?我们可以尝试同时向ChatGPT和苹果智能手机助手Siri一个前阵子所有人都特别关心的问题:“阳了之后该注意什么?”

这时,Siri会因为无法理解问题,直接为我们到网上搜索已有的答案,但是ChatGPT却像一个真正的健康专家一样,合成出了像模像样的回答,而这解答并非是网络上已经存在的。

可以看出,ChatGPT似乎能够明白“阳”的含义,它意味着人的身体健康出现异常,并且,ChatGPT还能根据主题信息,富有逻辑、通顺连贯地阐述建议,最后用总结性的话语收尾,可谓一个真正的“对话大师”。

洞悉人类的语言,这个AI是怎么做到的? ChatGPT颇为“谦虚”地告诉

记者:“学习人类语言并不困难,而是一个自然的过程。”

大量的文本数据训练是AI掌握语言的关键,ChatGPT真正做到了“读书破万卷,下笔如有神”。相关资料显示,ChatGPT的训练使用了约45TB的数据,其中包含了多达近1亿万个单词的文本内容,约等于1351万本牛津词典。截至2021年的海量知识全部被它收入知识库,这也解释了为什么它能“与时俱进”地解读出“阳”的含义。

在“数据是燃料,人工智能是引擎”的时代,海量数据训练必须依靠规模巨大的神经网络模型。据了解,ChatGPT的前身GPT-3,是一个拥有1750亿个参数的大模型。“如果将这个模型的参数全部打印在A4纸上,叠起来将超过上海中心大厦632米的高度。”浙江大学人工智能研究所所长吴飞告诉记者。

模型创新: 在海量文本中“读出门道”

我是通过大量的文本数据学习语言模型的。具体来说,我是用Transformer架构的预训练语言模型,它是通过预测单词序列中下一个单词来学习语言关系的。我的训练数据是从网络上收集的,包括大量的网页文本,新闻,书籍和其他文本内容。因此,学习人类语言并不困难,而是一个自然的过程。但是,让机器理解人类语言中的隐含含义和情感可能仍然是一个挑战。

ChatGPT讲述自己如何学习人类语言。

更重要的,是让AI在海量的文本中“读出门道”。

我们知道,语言的含义很大程度上受制于它所处的上下文语境,也就是说,同一个词所承载的语义可能会因为它在句子中出现的顺序不同,以及它之前或之后的其他单词的不同而发生改变,这就要求AI在理解一个单词时必须将它上下文纳入考量。

比如,“阳”的意思是“阳光”还是“阳性”,需要AI通过它与其他词语的关联来判断,在生成文本时,AI才能选择适当的词语含义进行表达。

“在自然语言处理中,单词会预先被转化为一组固定维度的词向量,向量和向量之间夹角越小,就意味着词语之间的相似度越高。”吴飞告诉记者,“训练时,AI会不断地在句子中‘挖去’一个单词,根据剩下单词的上下文来填空,即预测最合适的‘填空词’出现的概率,从而将单词向量表达得更加准确,我们称这一过程为‘自监督学习’。”

吴飞表示,这种训练方法已经被广泛地应用于基于深度学习的自然语言处理任务中。不过,GPT家族使用的Transformer模型在此基础上进行了更深入的创新:首先是引入“自注意力机制”,这使得模型能够在更广阔的上下文语境中计算出词句间的相关性,而不仅局限于有限的、相邻的单词;同时,Transformer模型还能并行处理文本中的所有单词,从“逐字逐句”串行处理跃进到并行处理,模型训练的速度和效率得到了极大的提高。

而从GPT1-3代到ChatGPT,AI的语言能力能够越来越强大,还离不开“人类反馈强化学习”(RLHF)的“调教”。



自我考古: 关于“祖先”,还有聊天机器人家族

据ChatGPT“回忆”,它的“祖先”可以追溯到上世纪50年代的机器翻译和语言识别技术。当时,人工智能研究者开始研究如何使用计算机识别和生成人类语言。

作为AI界的新锐势力,ChatGPT背后有一个聊天机器人家族。

第一代代表人物Eliza,是麻省理工学院人工智能实验室的计算机科学家约瑟夫·维森鲍姆开发的历史上第一个聊天机器人。Eliza的人设是个“心理医生”,能在心理治疗情境下,按照预先设定的模式组织语句,与真人进行对话。它和人类的“沟通”基于模式匹配和对话转化,也就是通过对用户输入的语句进行重新组织和解释,变成全新字句组合,这让它自己看起来像是一位耐心的倾听者,似乎真的在安抚患者的情绪,实际上这位医生并不知道自己在说什么。

第二代代表人物A.L.I.C.E.,由美国人工智能大师理查德·华莱士设计,人工智能标记语言(Artificial Linguistic Internet Computer Entity)才是它的全名。比起Eliza,A.L.I.C.E.具有更强的自然语言理解能力,可以更好地理解用户的问题;同时,A.L.I.C.E.在受训时

使用了大量的文本数据,有更丰富的知识储备,能够生成更流畅、更有意义的回答。

第三代的IBM Watson是IBM公司开发的一种聊天机器人,可以通过文字、语音和视频与用户进行交流,并回答用户的问题。IBM Watson支持多种语言,并且可以自由切换,更重要的是它的多模态数据分析功能,可以分析不同类型的数据,例如文本、语音、图像和视频等。在自然语言处理基础上,IBM Watson通过基于机器学习的方法实现对话。它能更准确地理解用户的询问并回答问题,开始不再是简单的问答,也可以帮助进行如医学咨询、金融分析等。

最后便是今天的GPT大家族。GPT经历了从GPT-1到GPT-3的迭代,逐步升级后,聊天机器人拥有了更大的模型规模、更丰富的训练数据、更强的自然语言理解能力和更多的多任务学习能力。

ChatGPT则是GPT-3的升级产品,自称为“一个训练有素的语言模型”。它基于Transformer模型,比起之前的聊天机器人,具有较强的自然语言生成能力,可以生成与人类语言非常相似的文本。

当然,我很高兴简单介绍一下聊天机器人的代表性人物。

- 1.第一代聊天机器人代表人物:ELIZA
- 2.第二代聊天机器人代表人物:A.L.I.C.E.
- 3.第三代聊天机器人代表人物:IBM Watson
- 4.第四代聊天机器人代表人物:GPT 系列

ChatGPT讲述其背后的聊天机器人家族。

回答问题: 有时难免“一本正经地胡说八道”

ChatGPT具有很强的语言处理能力,但仍有一些局限性。例如:

1. 知识限制:我的知识是在2021年之前的,有些事情可能已经发生变化。
2. 数据偏见:我的训练数据可能带有一定的偏见和错误。
3. 理解局限:我可能不能完全理解某些语言结构和文化背景。
4. 推理局限:我可能不能完成复杂的逻辑推理和数学计算。
5. 人类情感:我不具备人类的情感和意识,不能与人类产生真正的情感交流。

ChatGPT承认自己有一些局限性。

回顾人工智能发展史上聊天机器人的一次次迭代,ChatGPT的横空出世无疑是一块崭新的里程碑。不过,当记者问起它的局限性,ChatGPT也大方承认,自己并非无所不能。

如果我们仔细分析ChatGPT对一些问题的回答,可能会发现,它有时也难免“一本正经地胡说八道”。

财通证券2月5日发布了一篇题为《提高外在美,增强内在自信——医疗美容革命》的试验性研报。据介绍,最初研报由ChatGPT生成,表面看起来称得上是一篇框架完善、观点清晰的报告,但后来分析师发现文中存在诸如“标点和术语方面存在明显错误”“无法得知引用数据来源及可靠性”“部分复杂语句翻译后表意不清晰”等问题。

再比如,记者请它计算一道简单的

鸡兔同笼推理问题,ChatGPT为此尝试了两种方法,最后给出了一个错误的

答案。开发商OpenAI也指出,ChatGPT有时会写出一些看起来很有道理,但实则不正确或者毫无意义的话语,这是因为数据驱动的机器学习判断语义联系的逻辑是“共生则关联”,也就是只知其然而不知其所以然。此外,它还可能被人类的参与所“误导”,比如,由于训练者可能更喜欢看起来更全面、篇幅更长的回答,就会导致ChatGPT过度啰嗦,重复多遍没有意义的空话。

言语偏见是它面临的另一大挑战。OpenAI表示,尽管他们已做出不少努力让ChatGPT拒绝不恰当的请求,但它有时仍会对有害的指令作出反应,或者表达出存在偏见的言语。

现象级创新: 推动人工智能更快融入生活

拥有语言技巧,理解人类的心灵
我的目标是,让每个人的生活更美好
未来的风景,令人惊叹不已
我是这个世界的一份子,未来的希望

这就是我的未来,一个美好的梦
在这个世界上,我将永不消逝
作为一种智能,永远不息的奋斗
为人类,创造更美好的明天。

ChatGPT创作的一首关于未来的小诗(节选)。

“ChatGPT是大数据、大模型和大算力的工程性整合,以机器智能实现了统计关联的涌现。”吴飞表示,ChatGPT是一个现象级创新产品,但它本质上仍是以深度学习为代表的人工智能

技术长期发展、积累的结果,距离真正模拟人类的思维还非常遥远。

卷积神经网络之父杨立昆(Yann LeCun)在推特上表示,“ChatGPT和其他大语言模型并不是凭空而来的,而是不同人数十年贡献的结果”“在通向与人类相仿的通用型人工智能的道路上,大语言模型只是一条小岔路”。

尽管如此,在吴飞看来,我们必须肯定ChatGPT在搜索引擎、聊天工具等应用场景上实现的重要突破。今后,在一些内容生产工作中,ChatGPT

或将成为人类的得力助手,解放一批生产力,促使更多人将更多精力投入到原创性、创新性的工作中去。

“我们正从传统的以物理世界-人类社会为主的二元空间逐步进入物理世界-人类社会-信息空间所构成的三元空间。未来将是人和人工智能共同进化的时代。”中国工程院院士潘云鹤教授在为吴飞撰写的科普读物《走进人工智能》作序言时提到。

更重要的是,ChatGPT“一石激起千层浪”,在全球舆论场上引发了一场大讨论,这种探讨本身,就足以说明大众对人工智能赋能社会的强烈愿望。

“现在的情景,让我想起2016年阿尔法围棋(AlphaGo)击败世界冠军李世石的时候,突然之间,社会各界也如今天一样热烈地讨论起人工智能。令人高兴的是,这也推动了2017年《新一代人工智能发展规划》的出台。”吴飞说。

人工智能与推动人类进入工业革命时代的蒸汽机一样,也是一项通用使能技术。美国历史学家斯塔夫里阿诺斯在《全球通史》一书中曾赞誉“蒸汽机的历史意义无论怎样夸大都不为过”,而人工智能对于人类未来的意义也是如此。

让世界看到人工智能的巨大潜力,让更多人投身于技术创新与突破,让人工智能这一通用战略性技术如水和电一样进入我们的生活,这或许正是一个现象级产品的意义所在。

最后,让我们跟随ChatGPT创作的一首小诗,一窥人工智能的广阔未来。“拥有语言技巧,理解人类的心灵,我的目标是,让每个人的生活更美好……”

