

## 论文"太长不想读"怎么办? 自动生成一句话摘要的人工智能新软件问世

2021/3/18 15:58:40 新闻来源: 中国仿真学会仿真算法专业委员会

论文太长不想读怎么办? 一个软件就可以搞定。

据《自然》报道,一位科学搜索引擎的创造者推出了一款软件,可自动为研究论文概括生成一句话摘要,他们说,这有助于科学家更快地浏览论文。

这个免费工具被创建它的团队称为TLDRs("太长,没读过"的缩写),已于日前在科学搜索引擎"语义学者"上投入使用。"语义学者"是美国艾伦人工智能研究所(AI2)创建的搜索引擎。

AI2语义学者小组管理者Dan Weld表示,目前,TLDR软件只用于语义学者所涵盖的1000万篇计算机科学论文,软件经过微调后——预计在12月左右,其他学科的论文应该也能通过该软件自动生成一句话摘要。

初步测试表明,与查看标题和普通论文摘要相比,TLDR软件能够帮助读者更快对搜索结果进行排序,尤其是在手机上。"人们似乎真的很喜欢这个工具。"Weld说。

Weld创造TLDR软件的灵感来源于同事在推特上分享的一些标记文章的简短句子。与其他语言生成软件一样,该工具采用基于大量文本的深层神经网络学习。研究团队创建了数万篇与标题匹配的研究论文,这样网络就可以学会生成简洁的句子。然后,研究人员对该软件进行微调,训练一个新的数据集,将论文和摘要匹配起来,这些数据集由几千篇计算机科学论文组成,其中一些论文的摘要由论文作者撰写,另一些论文摘要则由本科生撰写。该团队已经收集了一些培训实例,以提高该软件在其他16个研究领域的性能,其中生物医学可能是第一位的。

一份描述TLDR软件的论文4月1日在预印本平台arXiv上首次发布,并在11月接受同行评审后发表。研究人员已经免费提供了代码,还有一个可以使用该软件的演示网站,任何人都可以试用这个工具。

"我预测这种工具在不久的将来会成为学术搜索的标准功能。事实上,考虑到需求,我很惊讶过了这么久才将其投入实际应用。"华盛顿大学信息科学家Jevin West说,他应《自然》的要求对该工具进行了测试。"这个工具并不完美,但绝对是朝着正确方向迈出的一步。"

来源: 中国科学报

相关链接:

政府机构

行业网站

国际网站

友情链接

地址: 北京市海淀区学院路37号工程训练中心637室 电话: 010-82317098 传真: 010-82317098

中国仿真学会 版权所有 电子邮箱: cassimul@vip.sina.com

京ICP备17016611号-1; 技术支持: 北京中捷京工科技发展有限公司(010-88516981)

业界动态

业界动态

图片中心

点击排行

自主知识产权的仿真软件需求

航天系统仿真重点实验室2020年科

中国仿真学会: "智汇"仿真科技人

湖南省系统仿真学会成立

北京仿真中心航天系统仿真重点实

航天科工集团二院北京仿真中心科

纯粹数学的雪崩效应: 庞加莱猜想

航天系统仿真重点实验室 召开20

青年工作联合会"智能人机交互领

"大数据分析与应用技术国家工程