

教育新观察

人工智能评判学生作文，靠谱吗？

在中考、高考这样的高利害考试中，套题写作是一个突出问题。实际评分时，有时难以认定并客观给分。

今后，如果有统计型的人工智能参与，那就可以找出相似度很高的作文进行对比、评判，并且给出精确的重复率百分比。在记叙文评价时，甚至可以更精细地通过自动摘要技术、关键词技术找出同样事例来一起评分。

这些技术处理，无疑将为防范抄袭、套题等不良行为提供有效措施。



徐默凡

近来，人工智能成了炙手可热的话题：在“阿尔法围棋”横扫人类顶尖棋手之后，又被自己的升级版“阿尔法元”无情碾压；一个名叫索菲亚(Sophia)的“女性”机器人被授予沙特阿拉伯国籍；而著名科学家霍金则表示，人工智能的崛起可能是人类文明的终结……

在这样一种背景下，我们不禁想要探讨一下：对于语文教育，人工智能能做什么？

深度学习
目前最接近人类智能

要回答上述问题，需要先了解一下人工智能在自然语言处理中的工作模式。

所谓自然语言处理，简单地说，就是利用计算机对人类语言进行分析，以完成自动分词、词性标注、语音识别、自动文摘、机器翻译、人机对话等一系列由简到繁的语言任务。在自然语言处理技术的发展过程中，经历了三种研发模式：

第一种是基于规则的自然语言处理模式，主要通过对话语进行语法分析和语义分析，然后转换成计算机程序以实现自然语言的理解和表达。这种工作模式是最容易想到也是最早进行广泛研究的，它依赖于语言学家和计算机专家的通力合作。但是，这种模式很快就遇到了无法突破的瓶颈，因为人类的语言理解过程实在太复杂，而语言学家对自然语言的分析很不充分，无法提供充足的语法规则和语义规则，计算机专家就陷入了“巧妇难为无米之炊”的窘境。

第二种是基于统计的自然语言处理模式，主要是对语言表达进行概率统计。这种模式下的人工智能，不需要了解话语的句法结构和语义关系，只需考察它被人类说出的可能概率就行，被说出的概率越大，相关话语就越合理。而概率的计算，可以通过大语料库基础上的词频统计来实现。这种工作模式不需要语言学家提供复杂的规则，让计算机搞统计正是它最拿手的工作。统计模式的广泛运用，在语音识别、机器翻译等领域产生了革命性变化，使很多技术从实验室走向了实际应用。

第三种是深度学习的自然语言处理模式。深度学习依赖的是大规模人

工神经网络，也就是利用大量电脑处理单元对人类大脑的神经系统进行模拟，然后让这个人工神经网络通过不断自我学习和自我调整来完成相应的工作。这可能是目前最接近人类智能的一种人工智能模式，目前的发展态势惊人，全面超越“阿尔法围棋”的“阿尔法元”利用的主要就是深度学习技术。

把作文评分交给电脑？
高利害考试中无法实施

三种工作模式下的不同人工智能能不能应用于语文教育呢？我们不妨以作文评分为例来分别加以说明。

如果让人工智能给学生作文评分，按照基于规则的工作模式，就必须把评价一篇作文好坏的要素都找出来，如语言、结构、内容、思想等等。最关键的工作还要把这些评分因素量化，比如给一篇作文的“语言”项目打10分，你就得告诉电脑，这10分的依据是什么？是词汇量多少，还是句子的复杂度，还是句式的不同类型？在第二语言教学中，类似的评分系统已经得到较为广泛的应用，因为仅是“语言”项目的还比较容易量化，但在母语作文评价中其可行性显然不大。因为对于母语作文评分来说，结构、思想等项目更为重要，之前人类阅卷者的评价主要依赖整体感知，但这种感觉很难分解，更无法量化。因此，基于规则的人工智能模式很难在作文评分上有用武之地。

如果是基于统计的工作模式，那我们就必须掌握足够数量的作文语料，然后构建大型语料库，分析其中各类作文的各种数据。比如，优秀作文和一般作文在词汇量和句子结构上有什么统计差异；比如，议论文平均用几个例子，平均引用多少句名人名言；又如，记叙文写了几个细节，每个细节平均多少字……在统计的基础上，把每篇作文在各方面的表现与平均值进行对比，然后评分。

显然，基于统计的人工智能模式可以详细描述作文的各方面数据，也可以根据这些数据对作文进行等级排序，但是统计哪些数据、这些数据的解释意义，这些数据与作文分数之间的关系，仍然需要语文专家提供意见，而这方面的研究仍然非常薄弱的。

如果是深度学习的工作模式，那就需要有大量事先标注好的作文对机器进行训练，这些已经精准给分的作文，被称为“训练语料”。将训

练语料输入到人工神经网络，由其分解为一组向量，再通过分层计算得出评分，然后将机器评分与已经标注好的得分相比较，得到误差值。再根据误差值，调整人工神经网络的计算方法和各个向量的权重，这样反复训练后最后可以达到理想效果：人工神经网络的评分结果和事先标注的作文分数高度一致。这样，就算在训练语料的封闭环境里获得了成功，然后就可以推向开放环境，也就是提供没有正确评分的陌生作文，直接由机器打分。如果经验证后和人类专家的评分结果一致，那么我们就得到了一个能自动评分的人工智能。

深度学习的自动评分在理论上是有可能会成功的，但是问题也很多。首先，即使评分结果可用，过程的可解释性也很差。人工神经网络把输入的语料当作一个数据序列，不考虑这篇作文的语言表现、思想内容，只是通过各种复杂计算的调试获得理想的结果。机器的分析过程是没有理解的，准确地说是和人类评价作文优劣的理据截然不同——它只是把语料干了，但是它是以机器的方式干的，人类无法理解。其次，我们刚才对研发过程的说明其实是过于轻描淡写了，真正要获得实用的理想结果，训练复杂度以及因此而要消耗的精力和财力在目前技术条件下恐怕是没法想象的。更重要的是，这里还存在一个“智能伦理”问题，如果代表人类智能结晶的作文交给机器去评分，而这机器又是以人类无法理解的方式评阅的，这恐怕会大规模引发公众的质疑甚至恐惧，在高利害考试中是根本无法实施的。

根据以上分析，要在作文评价领域完全使用人工智能，不说完全不可能，至少要走的路还很长很长。

作为语文教育的辅助工具
人工智能将大有作为

那么，在语文教育领域，人工智能是否毫无作为呢？当然不是。我们认为，人工智能可以成为一个很好的辅助工具，在诸多领域大有作为。

在阅读教学中，人工智能可以全面统计阅读材料的各种表现，为阅读材料的难度分级提供可量化的标准，从而为教材选文、编制课外阅读书目等教学实践提供有效的帮助。

在写作教学中，人工智能可以通过自动摘要、自动校对等技术对学生作文进行辅导，从而提升一些基本的语言技能

和写作技术。在练习系统中，人工智能可以构建封闭的专家系统，对一些有标准答案的知识自动出题、自动测试、自动讲解，极大提高参与性和趣味性……

在评价领域，我们还是以之前分析过的作文评分来详细谈谈人工智能的可能应用吧。在中考、高考这样的高利害考试中，套题写作是一个非常严峻的问题，在实际评分时，有时难以认定并客观给分。今后，如果有统计型的人工智能参与，那就可以找出相似度很高的作文来对比、评判，并且给出精确的重复率百分比。在记叙文评价时，甚至可以更精细地通过自动摘要技术、关键词技术找出同样事例来一起评分。这些技术处理无疑为杜绝抄袭、套题等不良行为提供了有效的防治措施。

此外，可以通过计算作文语句在大数据库中的概率值来对作文语言的规范性进行评价。作文语言的规范性越高，说明其用法越常规，越符合大众的语言规范。而概率值特别低则提示了两种可能性，一个是语言不规范，应给予负面评价，一个是语言新颖，应给予正面评价，机器初步筛选后的具体判断工作可以交给人类专家来执行，从而使作文语言项目的评分更加准确规范。

还有，统计型人工智能还可以对学生作文的整体表现进行系统评估。通过对历届考试作文的统计分析，可以得出一系列大数据，如作文平均词量和字量、平均句长和句法复杂度、虚词使用情况等进行统计分析，为基层的作文教学提供充足的科学数据。

最后想强调的是，人工智能除了在实际应用中对于语文教育有多方面的辅助作用，其开发和研究中遇到的一些问题也在倒逼语文教育工作者思考一些最基本的语文问题：要解决阅读材料难度分级的问题，就必须搞清楚构成阅读材料难度的因素有哪些；要解决人工智能自动命题自动训练的问题，就必须搞清楚语文学科必备的知识点有哪些；要解决作文自动评分的问题，就必须搞清楚作文的可操作的评价量表是怎么样的……这些基本的语文难题以前都是含混处理掉了，这些问题不解决，人工智能再先进也不可能对语文教育产生决定性的影响。

(作者为华东师范大学中文系副教授)



语言学专业，是时候告别“前牛顿时代”了

叶子

在人工智能兴起的今天，自然语言处理领域面临着大量的人才缺口。这对语言学来说似乎是前所未有的黄金时代，多少年来，看似“无用”的文科专业终于能够学有所用。

但“不幸”的是，目前国内语言学专业普遍的培养模式，为学生提供的知识储备，恐怕不足以帮助他们应对这一新挑战。能够填补眼下这个人才缺口的，只有计算语言学方向的毕业生。

计算语言学与语言学，仅两字之差，前者似乎只是后者的一个分支，实则天壤之别。举个例子：诺姆·乔姆斯基的《句法结构》是一本语言学经典名著，可多数文科生在阅读时，都会被“有限状态自动机”和“马尔可夫过程”等名词唬住。哪怕对大部分语言学研究生来说，这些

都显得艰深晦涩。可这些内容，本科二、三年级的理工科学生都有能力读懂，对计算语言学来说都算入门级的知识。

这一现象所折射的事实是：非计算方向的语言学专业学生，往往既没有在大学中接受过数学课程的训练，也没有在专业课中受过数学思想与方法的熏陶。目前经常出现的一则状况是，由所谓的“语言学家们”提出一些类似哲学思辨的想法，至于具体的、数学上的形式化或计算实现，都得依靠来自其他学科有数学背景的学者——双方之间存在着一种诡异的分工。

说语言学还处于“前牛顿时代”，严格来讲是有失偏颇的。事实上，在语言学中使用数学方法早在半个世纪以前就有过一阵风潮。

1954年，现代语言学之父约瑟夫·格林伯格就已经使用定量方法

定义了一些连续的指标，完善了语言分类方法。上世纪五六十年代，有大量探讨形式化语法理论的学者、文章和专著，将各种处于思想胚胎中的方案进行形式化，或者比较不同语法之间的强、弱生成能力。

可惜的是，在格林伯格之后的近半个世纪中，类型学界延续这个定量传统的学者非常少。遵循这条路的学者汇聚成了一个专门的学科——计量语言学，而这门学科之前只在德国盛行，一直未入“主流语言学界”的法眼。研究形式语法的传统同样在乔姆斯基早期著作之后的五十年中陷入了沉寂。

近三十年认知语言学和一些边缘学科的盛行，反映出大量文科出身的语言学者在逃离形式化方面的倾向。以认知语言学为例，直到最近才有一些学者开始使用机器学习的手段，通过计算模拟的方法来研究认知，走上了正轨。

因此，当下高校语言学专业的培

养模式，亟需进行两方面的改变：

首先，语言学基础学科，如音系学、句法学、语义学等应当加大形式化描述的比例，让学生在语言学实例中感受数学模型的运用。以理论语言学的核心课程“句法学”为例，目前国内往往将其狭隘地等同于“生成语法”。“句法学”或“形式句法学”课程教授的内容通常都是生成语法中“原则与参数”版本的理论，但偏偏这一时期的理论鲜见形式化方案。换句话说，它没有对应的数学模型。

其次，要加大基础课程中数学学科的分量。对语言学来说，由于支撑核心的句法语义现象的数学主要是离散数学，因此有必要将这门学科作为重点。

夯实数学基础，告别“前牛顿时代”，这也是语言学专业学生在人工智能大潮中不被淘汰的生存之道。(作者为上海外国语大学语言研究院硕士生)

校园牛人

开公众号、办电台、做直播……

“另类”辅导员，圈粉有妙招

■本报见习记者 李晨琰

在同济大学汽车学院，有一位毕业工作两年，就已经在学生中圈粉无数的明星辅导员：他自诩为一名“热衷于自媒体的潮男”，一名“热爱学生

工作、步入而立之年的‘中年人民’教师”。他如同一个不知疲倦的陀螺，24小时不停旋转。他就是曹卢。

从文字、图片到音乐、电台、视频、直播，曹卢均有所涉猎，他用“接地气”的方式做辅导员工作，为众多的“95后”大学生指点迷津。

“同三代”留校当辅导员，只因被一件小事感动

“我的父母、爷爷奶奶、外公外婆都是同济大学的教职工，作为一个成长在同济周边的‘同三代’，可以说我读大学之前，就没有走出过‘环同济经济圈’。”对于进入同济大学，曹卢将其称为“命中注定”。家住同济新村，从小上同济幼儿园、同济小学，初中就读于铁岭中学“同济班”，高中时，他就读的杨浦高级中学与同济大学南校区仅有一墙之隔。高考时，本着对汽车的热爱与父辈心中“工科为王”的执念，曹卢顺利考入同济大学汽车学院。

大学毕业后，曹卢只身前往日本东北大学留学，回国后研修汽车营销。得到同济保研机会后，他就读发动机结构设计专业。“什么都要试一下，才知道自己最适合什么。”拥有数技傍身、又在多家企业实践学习，最后让

曹卢决定留校担任辅导员工作的，恰恰是一件微不足道的小事。

在同济大学汽车学院一楼大厅，有一处别有洞天：那里的大学生创新创业基地，正是两支大学生车队的工作车间。学生们白天上课，晚上就在这里组装、拆卸汽车。所以，大家在这里通宵忙碌是常有的事。

有一天，忙到深夜12点的曹卢本以为自己是学院最后一名“留守人”，没想到走到下楼梯时，却发现还有数十个学生正忙着研究图纸、组装车辆。有这么多学生，为了追求自己的梦想坚持不懈地努力着，那一幕深深触动了曹卢。“我身边有一些比我早毕业的人，他们经常发牢骚，抱怨工作和生活。而我是一个比较推崇理想的人，或许只有在校园中才能一直保持这种充满激情的状态。”

运用自媒体平台，走到学生心坎去

当上辅导员没多久，曹卢发现，原先在校期间与他颇为亲近，叫着“学长”的学弟学妹们变得收敛起来，不愿与他走得近，更别提分享心事。“那时我才注意到，原来自己从‘学长’到‘老师’不只是身份的改变。”曹卢寻思着，“与其这样，不如搞个‘树洞’，让学生能够以匿名方式和我交流。”就这样，2015年初，曹卢的个人微信公众号“偶滴大森啊”应运而生。

“偶滴大森啊”如同一个“媒体矩阵”，文字、图片、音乐应有尽有。有一次，一名学生向曹卢哭诉，因为一门课“挂”了，痛失保研机会。“我问她为什么要选择读研，她说因为家里人希望她这样。”曹卢表示，这其实不是个例，而是许多大学生都会面临的困惑。该不该读研？为什么要读研？

当新生代心灵导师，是一件幸福的事

“不务正业”是曹卢常常用来形容自己的词汇，确实，开公众号、做直播，这怎么都不像个“正经”的辅导员，但学生对曹卢很认可。

在同济大学，曹卢开讲一门选修课《汽车文化》，学生们都很爱上的。曹卢说，他上这门课只有一个目的，就是让那些不懂汽车、不了解汽车行业的学生能够知道现在中国汽车工业有多厉害。

“每学期开课前，我都会问学生一个问题：你们觉得国产汽车怎么样？往往大家的第一反应都是不怎么样。通过一学期的讲解、介绍，上完这门课，大家认识了中国品牌，甚至会向亲朋好友推荐购买国产汽车。”曹卢很自豪，“国产汽车质量性能越来越好，我有这个义务让大家知道。”此外，他还带队“汽车文化公益宣讲团”，与学生走进幼儿园、中学、社区，和大家分享汽车知识。

为了开拓同济大学汽车学院学生的视野，他特意组织“汽车青年说”线下活动，邀请知名评车人及行业大佬与学生分享与互动。

前不久曹卢过生日，那天刚过零点，生日祝福纷至沓来，他手机的电量也瞬间从100%降到70%。曹卢说，

针对这种普遍情况，曹卢后来撰文《你本可以不读研的》，与学生们分享自己的观点。

“我只是讲出了我的看法，你可以选择不认同、不接受。”这样的亲和和同时，赢得了越来越多学生的青睐。通过微信公众号，曹卢阐释的大量关于人生、学习、生活等话题的见解，也激发了学生们对生活的观察与思考。

文笔活泼轻快、字里行间溢满真诚，又分寸拿捏适度，渐渐地，学生们愿意向曹卢敞开心扉。

鉴于学生们有不同的收视、收听习惯，曹卢还创办了网络电台，每周定期通过音频与学子进行另一种形式的交流，分享更多人生经验与感悟。目前他的微信公众号已拥有八万多“粉丝”；网络电台“大森电台”每期收听量均在十万以上。

这是来自“全球后援会”的祝福，“炫耀”的语气中透露出满满的幸福感。据说，曹卢的学生，可以上百人同时在微信群里换他的头像“捣乱”；学生们时常与他开玩笑，因为他不仅工作的时候高效且专注，而且也非常懂得生活……对于学生们的积极评价，曹卢自然是乐在其中。

作为新生代们的心灵导师，每年进进出出校园的年轻人也不断给曹卢带来灵感。做自己喜欢的工作，同时也被学生关心着，应该没有比这更幸福的了。



曹卢微信摘录

■大学能教会你些什么？每个人心中的答案都不一样。大学不是用来比智商的，未来由我做主。不要被自己的恐惧和现实禁锢住，生活不仅有诗和远方，还有你自己。

■对于更多人来说，改变是潜移默化的，“苟日新，日日新，又日新”，唯有宽恕自己的过错，反思自己的选择，才能让自己和过往和解，以时时刻刻崭新的自我继续走下去，而此时此刻就能成为一个很好的开端。

■要勇敢地相信，美好与希望总会存在。也许你心里没有天生足以温暖所有人的太阳，但请永远别放弃往黑夜中的火堆里增添木柴。

■我希望你今晚可以睡个好觉，我希望你明早起来可以吃到一顿称心如意的早餐，我希望你每一天都可以把自己要做的事情做完，有所成长，也希望你总是能够遇到一些人一些事让你看到这个世界善意的一面。对你们，“大森”要衷心地说一句：“我很高兴听到你很好。”