

以下文章来源于公众号：新智元



新智元报道

编辑：袁榭 好困

【新智元导读】对英语文本做完形填空，是折磨数代中国学生的噩梦。现在DeepMind宣布，他们的AI连两千年前的古希腊文完形填空都能做出70分以上的成绩，简直吊打了几亿人类。

最近，Nature的常客DeepMind似乎有些「不务正业」？

2月才刚学会如何控制核聚变，这次又制霸了古希腊文，还顺便上了封面。



DeepMind表示，学都学了，不如论文也用古希腊文写算了？

Ανακαλύπτοντας το παρελθόν με την «Ιθάκη»

Αποκατάσταση, εντοπισμός τόπου γραφής και χρονολόγηση αρχαίων κειμένων με τη συνεργασία Ιστορικών και Τεχνητής Νοημοσύνης

Κύριοι ερευνητές: Γιάννης Ασσαέλ^{1,*}, Thea Sommerschild^{2,3*}

Επιστημονική ομάδα: Brendan Shillingford¹, Mahyar Bordbar¹, Ιωάννης Παυλόπουλος⁴, Μαρίτα Χατζηπαναγιώτου⁴, Ίων Ανδρουτσόπουλος⁴, Jonathan Prag⁵, Nando de Freitas¹

¹ DeepMind, United Kingdom; ² Ca' Foscari University of Venice, Italy; ³ Center for Hellenic Studies, Harvard University, USA; ⁴ Athens University of Economics and Business, Greece; ⁵ University of Oxford, United Kingdom

古希腊文版本: <https://t.ly/ZKYW>

看到这里，当年试图拿洛布丛书（Loeb Classical）自学古希腊语、从入门到放弃的中年编辑一声长叹：真是人不如AI啊.....

不如改天试试至今都无人能够解读的「斐斯托斯圆盘」？



「伊萨卡」AI复原佚文

2500多年前，希腊人开始在石头、陶器和金属上书写，以记录从租契、法律到日历、神谕的所有内容，这是历史学家详细了解泛地中海区域历史的根基。

由于年代久远，这些记录多数都遭到了损坏，内容并不完整。

更不幸的是，当代史学界的文本修复技术，与18世纪的技术核心几乎没差别：用目标文本材料比对其他更详尽、可解读的传世记录材料，用近于侦探推理的思路来推测残断文本。

修复的铭文记录了一个雅典卫城在公元前485或484年颁布的法令

DeepMind与威尼斯大学人文系、牛津大学古典学院和雅典经济大学信息学系合作，探索机器学习如何帮助历史学家更好地解读这些残佚文本，从而让人们更深入地了解古代历史，并释放AI与历史学家之间的合作潜力。

在发表于《自然》杂志上的论文中，这些机构共同推出了名为「伊萨卡」(Ithaca)的AI，这是第一个可以恢复受损铭文缺失文本、识别其原始位置并确定其书写时间的深度神经网络AI。

论文链接：<https://www.nature.com/articles/s41586-022-04448-z>

「伊萨卡」以史诗《奥德赛》中的希腊岛屿命名，基于DeepMind在2019年提出的「皮媞亚」(Pythia)文字修复系统，并在性能上有了史诗级的提升。

目前，DeepMind已将「伊萨卡」的代码、训练前的模型等全部开源。

项目链接：<https://github.com/deepmind/ithaca#restoring-and-attributing-ancient-texts-using-deep-neural-networks>

修复准确率加成47%

「伊萨卡」是一种基于Transformer的人工神经网络。

其数据训练集是帕卡德人文学院的希腊铭文文本数码库，这是世界上最大的古希腊铭文数据库，其中包含大约78,608条已解读好的古希腊铭文。

根据评估，「伊萨卡」的头20条预测表现比前任「皮媞亚」优化了1.5倍。

在单独恢复受损文本方面，「伊萨卡」的字符错误率为26.3%、比单独人工预测的结果好2.2倍，最确定预测结果的准确率达到了61.8%。

「伊萨卡」在识别铭文原始位置方面的准确率达到了70.8%，并且可以将文本的书写时间确定在其真实时间区间的平均值差异29.3年、中位值差异3年的范围内。而人工确定时间的精度是在真实时间区间的平均值差异144.4年、中位值差异94.5年的范围内。

伊萨卡的架构

测试中，DeepMind将短语「δῆμο το αθηναίων」（雅典的人民）中的「δημ」被隐去。伊萨卡不仅修复了文本，而且还预测了写作的时间和地点。

此外，「伊萨卡」也显示出了人机合作在推进历史解释、确定历史事件的相对年代、甚至为当前的历史方法论探讨做出贡献的潜力。

历史专家在单独修复古代文本时达到了25%的准确率。但当使用「伊萨卡」时，字符错误率降低到18.3%、准确率提高到71.7%，在字符错误率上改善了3.2倍、在预测准确率上改善了2.8倍，超过了模型或个人各自的单独表现。

在测试集上对文本修复、地理（地区）和时间归属（日期）进行了评估

AI当助手，历史学得好

自然语言处理模型通常使用单词进行训练，因为它们在句子中出现的顺序、以及它们之间的关系提供了额外的上下文背景和含义。

例如，「很久很久之前」（once upon a time）的意思远不是这个短语中的每个单字的意涵加总。

然而，历史学家想分析的许多铭文都已损坏，其中大块的文字佚失。

研究团队为了确保AI模型在解读这种缺损铭文时仍有效，就要用单词/词组和单个字母同时输入，对AI进行训练。

伊萨卡的输出

(a) 对雅典铭文中6个缺失字符的修复预测，绿色表示正确的预测结果（συμμαχία，「联盟」）。

注意这些字眼的预测输出（ἐκκλησία，「公民大会」；προξενία，「城邦与异邦人之间的条约」），被AI自动用红色标出。

这些字眼常会出现在雅典的政治谕令中，AI特别注意到这些字符，体现了伊萨卡对文本背景的理解能力。

(b) 来自阿摩尔戈斯的铭文的地域归属。伊萨卡的头一个预测是正确的，其他最接近的预测也在邻近地区。

(c) 德洛斯岛铭文的日期分布。公元前300-250年的真实日期区间为灰色；伊萨卡的预测分布为黄色，其平均值为公元前273年（绿色）。

(d) 「伊萨卡」AI将铭文句子中的个人姓名(Νικίας, 「尼基亚斯」)和个人头衔(στρατηγοίς, 「将军」)自动变色标出。尼基亚斯对西西里岛的远征是古典雅典史大事，这佐证了「伊萨卡」判读句子的准确性。

为了最大限度地发挥「伊萨卡」作为研究工具的价值，DeepMind还创建了视觉辅助工具，以确保「伊萨卡」的研究结果易于被历史学家解读：

- 恢复假说：「伊萨卡」为文本恢复任务生成了几个预测假说，供历史学家利用他们的专业知识进一步精选。
- 地域归属：「伊萨卡」通过为历史学家提供所有可能预测的概率分布，来显示其判读结果的不确定性，而非输出单一的笃定结论。因此，它给出铭文可能出现的84个不同古代地区的概率分布，并在数字地图上将这些结果可视化，以阐明古代世界可能存在的潜在地理联系。
- 年代归属：在对文本进行时代估测时，「伊萨卡」会生成从公元前800年-公元800年间的所有可能时点的分布。这可以使历史学家调整特定日期范围的置信度。
- 显著性映射：为了将结果更好地传达给历史学家，「伊萨卡」使用计算机视觉的常用技术，来识别哪些输入序列对预测的贡献最大。然后在输出结果时，以不同颜色显示出让伊萨卡预测缺失文本、位置和日期的关键线索单词。

图中处理的古希腊文书记录，是公元前361年雅典城邦与色萨利城邦间结盟的记录。使用显著性映射方法，DeepMind开发者让「伊萨卡」AI在输出佚失的「盟约」字样时自动对「雅典人」和「色萨利人」字眼改变

颜色。

铭文为古希腊的思想、语言和社会发展提供了第一手的证据

为了更加深入地测试伊萨卡的性能，DeepMind用一个不包含指向精确时间的字符的铭文数据集上对AI进行了重新训练，然后提交了包含早期字符的文本让「伊萨卡」再进行分析。

结果，「伊萨卡」对法令文本的平均预测日期是公元前421年，这与最新的史学界研究突破相一致，展示了机器学习如何为围绕古希腊史上最关键时刻之一的研究发挥助益。

「伊萨卡」AI对处理的古希腊铭文的年代估测值、帕卡德人文学院的旧有铭文年代估计值、与历史学界最新的铭文年代精确估计值对比

学界旧有估测与最新成果之间差了平均27年，而AI估测与学界最新成果之间只差平均5年。

为何选择古希腊？

DeepMind选用古希腊铭文开发「伊萨卡」AI，是因为残断的古希腊铭文材料在种类与文字量上都足够变化多样，有测试AI的空间。并且古希腊字母与文本的今存语料数据也足够详实，方便用作AI的基础训练集。

与此同时，古希腊史在人类理解地中海世界方面有重大意。

目前的历史学界对一系列重要的雅典法令的确切颁布日期存在分歧，这些法令是在苏格拉底、伯里克利等名人生活的时代制定的。

长期以来，这些法令一直被认为是在公元前446/445年之前制定的，而最新的证据表明真实的日期其实为公元前420年代。

虽然这些时间看起来差异很小，但精确日期对于学界理解古典雅典的政治演变关键至为重要。

如果简体中文读者要更容易理解，不如想象微信公众号走红的日期是1991年、2001年、2011年还是2015年，对于百代后的传媒史研究的意义如何。

除此之外，古希腊史只是广阔的全球文明图景的一部分。

DeepMind表示，目前正在研究使用其他古代语言训练不同版本的「伊萨卡」，历史学家已经能在当前架构中使用多种数据集来研究其他古代语言，从阿卡德语到古埃及语民书体、从希伯来语到玛雅语，都包括在其中。

而「伊萨卡」，只是一个开始。

参考资料：

<https://deepmind.com/blog/article/Predicting-the-past-with-Ithaca>

<https://www.nature.com/articles/s41586-022-04448-z>