

众创机制带给气象青年人才的机遇与挑战

——对气象干部培训学院“2022江苏气象AI算法挑战赛”参赛团队的采访

采访人：于丹、唐伟（中国气象局气象发展与规划院）

受访人：中国气象局气象干部培训学院人工智能（AI）教学团队

赵阳（中国气象局气象干部培训学院气象网络教育中心）

王晴旭（中国气象局气象干部培训学院信息技术部）

张悦含（中国气象局气象干部培训学院新技术培训部）

采访方式：线下 + 线上

采访时间：2022年8月

DOI: 10.3969/j.issn.2095-1973.2022.05.026



采访现场合影（从左至右依次为于丹、王晴旭、赵阳、张悦含、唐伟）

【编者语】

2022年5—8月，由江苏省科学技术协会、江苏省气象局作为指导单位，江苏省气象学会、江苏省气象台和阿里云天池联合组织了2022年江苏气象AI算法挑战赛。此次挑战赛以“AI助力强对流预报”为主题，旨在聚焦短临强对流预报技术研发及应用，探索利用人工智能技术突破短临预报关键性技术瓶颈，推进人工智能技术在气象领域的学术合作、人才培养、技术交流以及多学科交叉融合应用。这种不限年龄、国籍、职业与专业的新型众创机制，既可以为短临强天气预报技术提供新思路新方案，同时也为气象青年人才创造了更多机遇与挑战。中国气象局气象干部培训学院组成的团队，在这次竞赛中，从1874支参赛代表队中脱颖而出，取得了第25名的好成绩。本刊委托来自中国气象局气象发展与规划院的于丹和唐伟，对团队进行采访，共享这支朝气蓬勃的青年教师队伍通过竞赛经历的锻炼和感悟。干部学院AI教学团队负责人钟琦博士的点评也启发我们如何面对AI引领的新创新时代。

采访人：组建团队参加本次气象AI算法挑战赛的
动力及目的是什么？

赵阳：“气象+AI”是当下发展非常快的热门领域之一，近年来中国气象局气象干部培训学院（以下简称“干部学院”）新创立了人工智能教学团队，吸引了许多对“气象+AI”感兴趣的青年教师。本次气象AI算法挑战赛，瞄准的基于观测数据预报强对流天气，是当下“气象+AI”的重点和难点领域，受到了团队年轻老师的关注，大家跃跃欲试。相比于之前的文献研究、参与项目等研究AI的方法，通过参加竞赛，一是可以提高大家的编程能力，检验大家的学习成果，通过真正动手实践，暴露出理论学习中发现不了的问题，避免了“纸上谈AI”；二是通过参加比赛，能够

更好地磨合团队，让每位队员都能在“气象+AI”的大海中找到自己擅长或感兴趣的领域；三是参与竞争，开拓视野，看看我们与强队之间的差距，了解自身不足。

张悦含：我们三人都是干部学院人工智能教学团队的成员（链接1），想借此机会练练兵，在实践中学习。同时，作为干部学院的一线教师，也希望积累一些授课的素材。

链接1：干部学院参赛团队

团队的3位队员，专业背景涵盖了计算机、气象学、电子信息等方向，平均年龄30岁左右。



赵阳（1988.03）

研究方向为深度学习与课件开发。

张悦含（1994.04）

我是团队里唯一学气象专业的，主要研究方向是降水日变化及形成机理。



王晴旭（1993.01）

我所学专业为计算机技术，主要研究方向是自然语言处理。

采访人：团队此次参赛，采用了哪种机器学习算法，预报结果如何，是否和传统短临预报方法——光

流矢量计算方法的预报结果进行对比，有何优势？

王晴旭：强对流预报，即时空序列预测问题，我们采用了时空序列预测的经典开源算法ConvLSTM（图1）。ConvLSTM是2015年香港科技大学的施行健博士提出的，结合了卷积（Conv）可以提取空间特征的能力以及长短期记忆网络（LSTM）可以处理序列的能力，特别适合时空序列预测任务。很多文献研究表面，相比于光流法预测结果，ConvLSTM拥有更好的效果。另外我们团队采用了本地配置的计算资源，相比于云端高性能计算资源，胜在配置简单，能够更快上手。

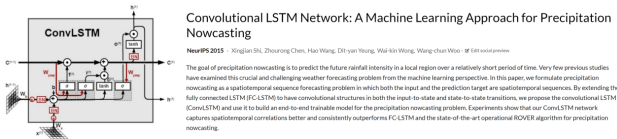


图1 经典时空序列预测算法——卷积长短期记忆网络 ConvLSTM

采访人：刚才提到团队三位队员来自不同专业，平均年龄仅30岁，那么在全球1784支参赛队伍中能够脱颖而出取得优异成绩，具有哪些优势条件？团队在参赛过程中还面临哪些挑战，是否遇到了瓶颈问题？

王晴旭：我们团队充分利用了多学科交叉和合理的团队分工。在我们刚获悉题目的时候，一开始大家都没什么比赛的经验，但是通过一两天的集中讨论，确定了每个人的方向，比如赵阳负责编程，悦含负责研究评分标准和数据处理，我负责算法调研和开发，这样每个人的优势都能发挥出来，进而阶段性地整理每个人的成果并汇总形成方案。另一方面，虽然缺乏参赛经验，但是作为干部学院的教师，平时在教学培训工作中也积累了一些其他的经验，例如调研、编程，以及细致严谨的教学经验，这使得我们在比赛中可以很快地实现初步方案。最后一点，干部学院的领导和团队中其他指导老师或前辈也给予了我们很多帮助和支持，让我们能够在比赛中坚持不懈，始终保持信心和热情。

赵阳：相对于成绩更佳队伍，我们还有很多不足。例如在算法方面，通过这次挑战赛，我发现自身对于AI的很多底层逻辑理解还有欠缺。我们采用的是2015年开源的算法，对于近两年来的新算法研究不足。另外在数据处理、训练技巧、评分检验等方面，也都暴露出很多经验不足的问题。发现不足，补齐短板，这是我们这次比赛最满意的“收获”。

张悦含：和优胜队伍比，我们确实在技术实现上有很大差距，我想这主要是由于我们三个都不是科班出身，并且没有整段的时间全身心投入比赛吧，毕竟

业务工作还是比较繁忙的。

采访人：构建优良的数据集是AI研究的核心。本次气象AI算法挑战赛提供的数据集对您今后的业务工作是否有所借鉴？

赵阳：一个优秀的数据集是发展AI最重要的基础，比如众所周知的ImageNet数据集在十多年前就成为为了图像识别领域飞速发展的奠基石（图2）。本次挑战赛的数据集也给我们的AI教学提供了参考，例如将多源数据进行时空匹配，算是一个很好的思路。对于气象领域来说，我们不缺数据，但是缺乏对数据集的整理，包括质量控制、重点天气现象标注、数据匹配等工作，这算是一个很难但意义重大的工作，对于“气象+AI”的长远发展是必不可少的。

IMAGENET

<p>ImageNet图像识别数据集(2009) 华裔学者李飞飞创办 1500万张人工标注图片 2万多类别</p>	<p>ImageNet挑战赛 (ILSVRC, 2010-2017) 诞生了AlexNet(2012), ResNet(2015)等经典模型 2012年深度学习获首次冠军(Top5误差15.4%) 2015年深度学习超过人类(Top5误差<4%)</p>
--	---

图2 ImageNet图像识别数据集

采访人：团队共参与了几次气象AI算法挑战赛？具体到方法上来说，是否可以以不变应万变，用一个算法就可以参加不同的气象AI大赛？

赵阳：这是我们第一次参加类似的比赛。对于AI的研究者来说，我们最理想的情况当然是找到一个通用算法来满足所有的任务，但是在现阶段是不可能的。受限于AI的发展，计算资源的有限，业界每年都会有新的思路、新的算法出现，AI也在向着（计算速度）更快、（评价分数）更高、（实现效果）更强的方向不断“内卷”。这样也更能激励我们在AI领域不断探索、继续前进！

采访人：通过此次实战，团队对“气象+AI”是否有了新的认识，人工智能在气象领域融合应用的现状如何，实现深度融合应用还面临哪些困难与挑战？

赵阳：自2016年起，先后有40余个国家和地区将推动人工智能发展上升到国家战略高度。人工智能已成为科技创新的关键领域和数字经济时代的重要支柱。目前，人工智能在医疗、制造、自动驾驶等领域的应用已持续深入。“气象+AI”有很多可以结合的点，

比如本次比赛所涉及的强对流天气的短临预报，还有气候预测，卫星云图识别等。此外，AI领域最成熟的图像处理技术，例如图像识别或语义分割，还可以应用在气象服务领域，例如农业气象中对农作物状态的智能识别，还有旅游气象、交通气象等领域。

我们都知道AI是运用了大量统计学的知识，可以通过训练从过往数据中寻找特征，继而运用到新的情况，这和老资历的预报员通过经验分析是类似的，甚至说AI会越来越强、越来越稳定，而不会依赖个人经验的差异。同时传统预报方法还结合了很多大气动力学的知识，如何把气象学的先验知识应用到AI中，是我们需要深入研究的。

采访人：如您所说AI会越来越强、越来越稳定，那么随着人工智能技术的发展与深度应用，对预报员来说，借助外脑作用是否是一种替代关系？

赵阳：当下AI是一个高速发展的领域，目前AI的门槛也是逐年降低，即使是非计算机专业的人员也能轻易上手。在不远的未来，AI会更加接地气，成为像互联网一样的“新基建”，一个很普通的工具。但同时也要看到，AI还存在很多问题，最大的问题就是常说的黑箱，即存在不稳定性。例如数据集中的噪声会

对结果产生不可预见的不良影响，这对于预报业务来说是难以接受的。因此在短时间内，AI还难以承担核心的预报业务工作，但可以作为一个辅助工具提供参考结果，或者参与到一些细分工作中。直到其准确率获得认可，再应用到核心业务中。

采访人：本次挑战赛既是推动AI在气象领域应用实践的一次重要探索，同时也是形成解决预报瓶颈的一种新型众创机制的具体尝试。但此次竞赛在气象领域并不是首次探索。2022年6月11—12日，ECMWF举办了2022年黑客马拉松：可视化气象数据（#VisMetData，链接2），探索如何将天气和气候数据可视化。深圳市气象局2017—2018年连续两年通过与阿里巴巴公司、香港天文台合作，共同组织了“全球AI气象挑战赛”，通过众创机制促进智能临近预报的发展。您认为采用竞赛方式对于AI技术而言具有怎样的特殊意义？

赵阳：AI的优势就在于它是建模无关的，能够高效地、大规模地提取数据中的特征。以我们团队为例，大家都没有观测预报的相关工作经验，却也能够取得一定的成果。这证明了非专业的人员也有一席之地。“气象+AI”并不是要完全替代传统预报方法，而

链接2：2022年黑客马拉松：可视化气象数据

黑客马拉松（Hackathons），又称编程马拉松，是指将程序员或者软件工程师等互联网相关的从业者聚在一起，在特定时间内自由完成比赛题目。通过为技术人员提供一个自我表达和展现创造力的平台，使富有创新性的技术得到开发和应用，它是程序员和软件开发人员向往的业内顶级盛会，也成为业界进行开发创新和融资的契机。

2022年黑客马拉松：可视化气象数据由ECMWF在其总部英国雷丁举办，目的是探索如何将气象数据、天气和气候可视化，以便对用户和更广泛的公众更有用、更易于理解和更有影响力。此次活动注重吸引来自不同背景和多样性的参与者，不仅包含编码人员，还包含了设计师、数据管理员、气象学家、讲故事的人、记者等任何对气象数据和可视化感兴趣的人员，共有9支队伍37人参加。



Hackathon 2022年提出了三个挑战：可视化数据（#VisData）；用数据讲故事（#StorytellingData）；数据处理（#101MemberEnsemble）。除此之外还有一个公开挑战（#OpenHack）。在每个挑战中，ECMWF工作人员都会提出项目供参与者开展工作或帮助激发他们的项目。比赛时间为24小时，每个项目根据四个标准进行评估：原创性、影响力、令人印象深刻和数据的使用。除了目前可从ECMWF获得的所有公开数据之外，参与者还可以获得许多数据。

最终#isitnormal团队成为获胜者，他们的项目关注的问题是“今天的天气与历史正常相比如何？”通过使用来自ECMWF的ERA5再分析的温度数据，以气候条纹为灵感，绘制了可点击的历史月度温度时间序列和创新的小提琴图，以展示当今温度与世界各地不同城市的历史值和平均值相比。该团队还创建了一个地图功能，以展示跨地区和局部的温度变化。

是借助AI的一些优势或数学思路，来解决预报的一些难点。例如图像识别中所用到的卷积可以用来提取特征，例如无监督学习可以让机器自主学习到前所未知的特征，这些都是可以借鉴的思路。

张悦含：在短临预报中，机器学习的效果确实不错。而这种通过比赛集思广益的形式也非常振奋人心。在我看来，是一种双赢机制，对于业务单位来说，低成本地收获了一些新的研究思路和算法；对于参赛选手来讲，不仅开拓了思维，还获取了一系列的“激励”。

采访人：作为干部学院的AI教学团队，从干部学院作为国家级气象管理干部和高层次专业技术人才培训基地出发，在气象业务培训中是否可以借鉴这种新

型众创机制，未来的业务培训将呈现怎样的趋势？

王晴旭：将AI应用到预报中是一个热门的研究方向。在过去两年时间里，干部学院的AI教学团队已经开展了多期研究型天气业务培训班，邀请到了各单位或高校的“气象+AI”领域的专家作为授课老师。全国的预报员或研究人员都踊跃报名参加，收到了非常好的反响。但AI毕竟是一个新兴的、高速发展的学科，来自全国各地的学员其数学和编程水平也都不尽相同，使得当前的AI培训更偏向前沿讲座的性质。我们的团队也在积极探索“气象+AI”的培训方式，争取在未来的预报员培训中看到更多、更先进、更实用的AI内容。

采访人：谢谢团队接受采访，愿团队今后在AI领域取得更多的成果！

专家点评：AI领域独有的竞赛机制赋能气象智能

钟琦

钟琦博士，中国气象局干部学院正研级高工，青年气象英才，学科带头人，研究生导师，新技术培训部副主任。牵头数值预报应用、人工智能气象应用课程体系研究，主持和骨干参加国家重点研发计划、科技部重大科学研究计划、国家自然科学基金重点、面上和青年项目等国家级科研项目10余项，发表高质量论文20余篇，获得软件著作权5项。

【点评内容】

竞赛是促进交流合作和行业发展的常见和有效手段，比如人工智能领域，知名的ImageNet挑战赛在目标识别领域取得了前所未有的突破，同时催生了很多优秀的计算机视觉解决方案，并反向促进了图像数据量和GPU性能的飞速增长，数据、模型和算力三者合力带来了一场席卷全球的深度学习革命。

气象领域的预报竞赛也不是新鲜事，全国气象行业职业技能竞赛已举办了十四届，国际规模最大的天气预报竞赛WXchallenges已举办了十七年。随着大数据、人工智能等科学技术的发展，人工智能与气象的融合优势开始崭露头角，跨学科的竞赛方式也日渐风行。如2019年美国华盛顿大学预报团队应用机器学习建立的多模式集成预报系统获得WXchallenges冠军并打破了赛会纪录。2021年世界气象组织发起了应用人工智能提升次季节-季节预报的竞赛，国内高校、企业和气象业务单位也举办了多期人工智能天气预报大赛，主题涵盖AI助力精准气象和海洋预测，强对流短临预报等。

这种AI竞赛机制对新技术融入和推动气象应用具有显而易见的好处：促进理论、业务与技术开发的融合。竞赛的主题大都是接近真实业务的问题需求或技术瓶颈，具有很强的实战性和前沿性，对实际业务中的关键问题常有最直接的促进作用。提供了众筹的科技创新模式，加速AI技术在气象行业落地，不仅拓展天气预报的技术发展空间，更促进了社会协同和开放创新的气象生态环境形成。

目前的AI竞赛也存在显著局限：第一是数据集准备还不够充分。数据是深度学习的生命，数据决定了机器学习的上限，而算法只是逼近这个上限，这是业内共识。目前国内气象AI竞赛在数据开放和获取上不论是种类还是质量都存在限制，这点在国际赛事上体现相对好些，比如全球顶级的数据科学Kaggle竞赛提供较完备的分级训练数据，比如WXchallenges预报竞赛提供较充分的地面站、探空、卫星、雷达、模式等各种资料。第二是比赛虽然采用了3个月至半年不等的比赛时间，竞赛过程中也多有跨学科协作，答辩时也兼具了统计评分和历史疑难个例，但目前更多地是吸引和测试了AI领域的模型，选手目标单一明确，仅服务于竞赛指标，而对关键预报问题的理解并不深入，在模型和调参都难以达到实际最优。此外，在基于纯数据驱动的预报问题仅依赖深度模型优势尚有市场，但气象预报中的难点问题往往与多尺度非线性的物理本质有关，AI方法要真正比已有业务客观预报方法具有优势，需是考虑了物理融合的智能模型。

2020年7月，干部培训学院首批气象教学团队率先成立了“人工智能气象应用团队”，着手“教研咨”一体的课程建设和开展分层分类的人工智能气象应用培训。团队参加本次竞赛是锻炼队伍的很好机会，未来，教学团队将聚焦卫星、雷达、数值预报等多源资料，进一步深入探索物理与深度学习模型的融合；将竞赛小组的成果转化开发成实际案例，在培训实习中接受全国学员的检验，为学员带回应用到本地业务提供支撑。

