

doi:10.11676/qxxb2019.070

## AI 技术可以将地面气温预报精度提高多少？

### ——记 AI Challenger 2018 全球天气预报挑战赛\*

嵇磊<sup>1</sup> 王在文<sup>1</sup> 陈敏<sup>1</sup> 范水勇<sup>1</sup> 王迎春<sup>1,2</sup> 沈志远<sup>3</sup>

Ji Lei<sup>1</sup> WANG Zaiwen<sup>1</sup> CHEN Min<sup>1</sup> FAN Shuiyong<sup>1</sup> WANG Yingchun<sup>1,2</sup> SHEN Zhiyuan<sup>3</sup>

1. 北京城市气象研究院, 北京, 100089

2. 北京市气象局, 北京, 100089

3. 创新工场人工智能工程院, 北京, 100080

1. *Institute of Urban Meteorology (IUM), Beijing 100089, China*

2. *Beijing meteorological Service, Beijing 100089, China*

3. *Sinovation Ventures AI Institute, Beijing 100080, China*

2019-04-26 收稿, 2019-08-16 改回.

**摘要:** 2018 年 8 月, 北京城市气象研究院(IUM)与创新工场等公司联合举办了“天气预报”竞赛(WFC)——这是一项面向全球的人工智能(AI)挑战赛, 旨在通过发挥 AI 技术的优势以提高天气预报水平。全球有超过 1000 支队伍参加本次 WFC 竞赛, 约 250 支队伍完成了实时天气预报赛程。最终, 决赛排名前 5 的队伍获得了奖励。竞赛结果表明: 多 AI 模型集合方法显著提高了地面 2-m 气温、2-m 相对湿度和 10-m 风速的预报水平。与 IUM 在业务中应用的相似集合预报方法相比, 基于时间序列分析、梯度提升树、深度概率预测等 AI 模型构建的集合预报方法, 显著提升了地面 2-m 气温预报准确度, 前 2 名队伍在决赛期间平均分别达到 24.2%和 17.0%。同时, 合理的数据处理技术和 AI 模型集合框架对预报效果的提升具有重要的作用。

**关键字:** 人工智能 (AI), 天气预报, 地面气象要素, AI 集合气象模型

嵇磊, 王在文, 陈敏, 范水勇, 王迎春, 沈志远. 2019. AI 技术可以将地面气温预报精度提高多少? ——记 AI Challenger 2018 全球天气预报挑战赛. 气象学报, 77(5):

Ji Lei, Wang Zaiwen, Chen Min, Fan Shuiyong, Wang Yingchun, Shen Zhiyuan. 2019. How much can AI techniques improve surface air temperature forecast? —A report from AI Challenger 2018 Global Weather Forecast Contest. *Acta Meteorologica Sinica*, 77(5): XX-XX.

#### 1. 人工智能技术在气象领域的研究应用现状

人工智能(Artificial Intelligence, AI)一词最早出现在 1956 年的达特茅斯科学会议上, 它涵盖自然科学、社会科学和技术科学的交叉领域。2006 年, 机器学习领域泰斗 Geoffrey Hinton 教授首次提出深度学习的概念, 极大的提高了神经网络学习的性能, 进而掀起了人工

\*资助课题: 国家重点研发计划专项(2018YFC1506801), 国家自然科学基金项目(41505117), 中央级公益性科研院所基本科研业务费专项(IUMKY201904)。

作者简介: 嵇磊, 高级工程师, 主要从事大数据, 水资源, 云和降水物理观测与模式研究。E-mail: lji@ium.cn  
通信作者: 王在文, 高级工程师, 主要从事大数据, 数值模式释用预报研究。E-mail: zwwang@ium.cn

智能研究的新一轮浪潮 (Hinton, 2006)。近年来, 人工智能技术在很多领域得到了快速发展和广泛应用。鉴于此, 基于大数据和机器学习技术, 如何应用人工智能建立气象预报模型, 提高气象预报准确性, 正成为广大气象工作者及人工智能专家们共同关心的热点科学问题之一。

目前, 国内外众多气象组织正纷纷与人工智能研究机构和公司开展合作, 探索人工智能在气象领域应用的解决方案。美国 AccuWeather 公司与 Google 公司开展合作, 利用云计算和人工智能技术, 初步实现了 0-90 天内逐分钟、逐小时和逐日的精细预报 (<https://qz.com/535345/ibm-is-going-to-change-how-we-forecast-the-weather-with-watson/>)。英国气象局与美国亚马逊公司合作研发数据存储、云计算技术, 同时与微软公司合作发展人工智能技术 (<http://www.odtms.org/2017/07/machine-learning-in-weather-forecasting/>)。Earth Risk 公司基于 ECMWF 的数值预报, 通过多智能模型集合框架研发出 TempRisk Apollo 方法, 以提供更加可靠的气温概率预报 (EarthRisk Technologies, 2013)。中国国家气象中心与清华大学合作, 采用分布式深度学习框架、时空记忆深度循环网络算法提升了雷达回波外推预报准确率 (Wang, 2018), 相比交叉相关法平均提升了 40% (毕宝贵, 2017)。与此同时, AI 技术已初步用于冰雹和雷暴等强对流天气的预报 (Zhou et al., 2019)。北京市气象局利用机器学习方法开展了地面温度 (戴翼等, 2019)、雷暴大风 (杨璐等, 2018) 和强对流天气 (郭瀚阳等, 2019) 预报。深圳市气象局与阿里平台合作开展了基于雷达回波观测的短时临近降水预报试验 (Yao and Li, 2017)。

人工智能在天气预报业务中已有初步应用, 涵盖了雷暴大风 (Lagerquist et al, 2018)、降水强度 (Mattioli et al., 2018) 等短时临近预报, 极地暴风雪 (Burrows et al., 2018)、低能见度 (Kneringer et al., 2018) 等极端灾害性天气预警。同时, 人工智能在天气、气候分析中也发挥着积极的作用 (Collins et al., 2018; Kunkel et al., 2018)。

以上成果表明, 基于多源观测和全球数值预报的 AI 方法, 在提升不同时间和空间尺度气象要素预报准确性方面都展现出一定效果。近年来, 随着高性能计算和网络技术的进步, 我国高分辨率区域气象数值模式预报和高时、空气象监测网建设正在逐步完善。因此, 如何充分发挥两者的优势, 以达到精细化气象业务预报需求, 是亟需思考的问题。鉴于此, 作为面向全球人工智能人才的开源平台, 由创新工场 (<http://www.chuangxin.com/>) 于 2017 年创建的 AI Challenger 全球挑战赛 ([https://challenger.ai/news/ai\\_challenger](https://challenger.ai/news/ai_challenger)), 为解决当前天气预报所面临的问题提供了契机。

## 2. AI Challenger 2018 全球挑战赛

2018 年 8 月 29 日, 第二届 AI Challenger 全球挑战赛在北京拉开序幕。该挑战赛共包含 10 个竞赛项目 (天气预报, 无人驾驶视觉感知, 农作物病害检测等), 拟通过 AI 技术解决行业难点问题, “用 AI 挑战真实世界的问题”。

立足中国首都北京气象预报业务实际需求, 北京城市气象研究院 (IUM) 作为协办方, 策划并组织了“天气预报”竞赛项目 (<https://challenger.ai/competition/wf2018>), 旨在面向全球智力资源, 跨学科探寻提高未来精细化气象业务预报的新思路、新方法。“天气预报”竞赛项目吸引了来自全球 1000 多支队伍报名; 选手覆盖中国, 美国, 日本, 俄罗斯等 19 个国家和地区; 参赛机构包括中国科学院大学, 清华大学, 北京大学, 美国斯坦福大学等国内外众多高等学府, 以及多家知名科技公司。最终, 约有 250 支队伍完成了两周一次的双周赛和持续一周的决赛赛程。

## 3. “天气预报”竞赛项目

### 3.1 赛题设置

“天气预报”竞赛是 2018 年 AI Challenger 全球挑战赛的实验项目之一。该项目要求选

77 手基于 IUM 提供的气象站点的“观测”和“睿图”<sup>1</sup> (范水勇等, 2013)数据集, 建立科学有效的  
78 AI 模型, 用于预报气象站点未来 36 h 整点的 2 m 温度( $t_2$ )、2 m 相对湿度( $rh_2$ )和 10 m 风速( $w_{10}$ )。

79 “观测”和“睿图”数据集包含了北京 10 个地面自动气象站点的逐小时气象要素时间序列  
80 值, 范围自 2015 年 3 月 1 日至 2018 年 11 月 3 日。数据集中的气象站点经纬度和站号信息  
81 对参赛队伍做了隐藏。“观测”数据集包含各气象站点的 2 m 温度、2 m 相对湿度、10 m 风  
82 速和地面气压等 9 个气象要素;“睿图”数据集包含 29 个气象要素(地面和不同气压层的温、  
83 湿、风、压等)。两组数据集共同包含 1188 天训练样本和 89 天验证样本, 用于参赛选手搭  
84 建和调试 AI 模型。在此基础上, 实时更新的测试集数据则分别用于双周赛和持续一周的决  
85 赛。本次竞赛提供的数据量, 仅占气象行业实际天气预报业务可用数据总量的约 1%。参赛  
86 队伍需要基于提供的有限数据, 建立 AI 模型, 用于预报北京 10 个气象站点 36 h 内逐时的  
87  $t_2$ 、 $rh_2$  和  $w_{10}$ 。

88 由于实际天气预报具有极高的时效性要求, 竞赛采取实时天气预报的方式, 由 IUM 在  
89 竞赛日更新发布数据集, 并要求参赛队伍在数据集公布后 6 h 内提交预报结果。对比其它竞  
90 赛项目, “天气预报”竞赛具有更高的挑战性。实时天气预报的竞赛方式, 使得“观测”和“睿  
91 图”数据集需要实时更新, 无法提前准备; 预报结果需要在相对很短的时间内提交, 以保证  
92 实用性; 决赛需要连续 7 天提交预报结果, 以证明 AI 模型的健壮性。

### 93 3.2 竞赛结果

94 决赛前 5 名队伍 (AI01—AI05) 被选择进入最终的答辩环节。AI01—AI05 的最终成绩  
95 与目前 IUM 释用效果最好的相似集合预报方法(AnEn, 王在文等, 2019)进行对比, 以评估各  
96 队 AI 建模的性能。

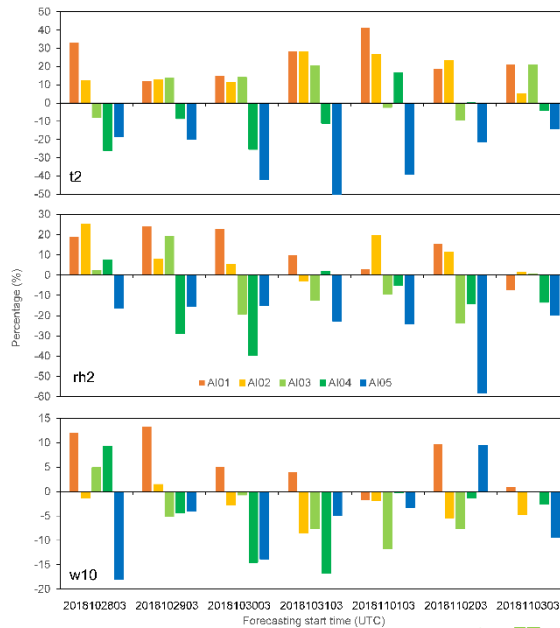
97 评判标准如下: 以 AnEn 预报为基准, 首先计算其与观测值的均方根误差( $RMSE_{AnEn}$ ),  
98 再分别计算各支队伍 (AI01—AI05) 的预报结果与观测值之间的均方根误差( $RMSE_{Team}$ )。  
99 最后, 各支队伍的 RMSE 成绩相对于  $RMSE_{AnEn}$  的误差减小比例  $RMSE_p$  可计算为

$$RMSE_p = (RMSE_{AnEn} - RMSE_{Team}) / RMSE_{AnEn} \times 100\% \quad (1)$$

100 该方程表明,  $RMSE_p$  越大, 则 AI 方法相对于 AnEn 方法的预报效果越好。由于 AnEn  
101 方法是目前 IUM 最好的预报方法, 因此真正优秀的 AI 方法必须要取得比 AnEn 方法更准确  
102 的预报结果。

103 图 1 给出决赛阶段上述 5 支参赛队逐日的 36 h 预报整体评估结果。5 支队伍大多采用了  
104 多 AI 模型集合的方法, 但在具体模型选择以及数据处理技巧方面, 有一定区别。主要采用  
105 的 AI 模型有: 时间序列分析 (Prophet), 梯度提升树 (GBM), 深度概率预测 (Seq2Seq),  
106 双向长短记忆神经网络 (Bi-LSTM), 循环神经网络 (RNN), 人工神经网络 (ANN) 等。  
107 由于每天“睿图”集提供未来 36 h 预报, 因此相邻两天的预报, 在预报时段上会有 12 h 重  
108 叠, 部分参赛队选取重叠时段平均值作为最后结果, 并对时间特征做 sin/cos 编码。缺失值  
109 主要采用线性插值法, 多站点均值填充法, 长时段 (37 h) 缺测值剔除法等。

<sup>1</sup> “睿图”由北京城市气象研究院的数值预报系统“RMAPS”音译而来。RMAPS 全称为 the Rapid-refresh Multi-scale Analysis and Prediction System。



110

111 图 1 2018 年 10 月 28 日至 11 月 3 日决赛期间，前 5 名参赛队（AI01—AI05）预报的 2 m 温  
112 度、2 m 相对湿度和 10 m 风速的RMSE<sub>p</sub>对比

113 Fig.1 Percentage change of RMSE (RMSE<sub>p</sub>) for t<sub>2</sub>, rh<sub>2</sub>, w<sub>10</sub> from the forecasts made by the five  
114 teams (AI01–AI05), during the week-long final contest period of 28 October to 3 November 2018.

115 整体来看（表 1），AI01（浙江大学 yuanpengli 代表队）和 AI02（西南交通大学 CCIT007  
116 代表队）的预报技巧相比 AnEn 方法有明显提升，且逐日预报性能较为稳定，AI03 的预报  
117 技巧略逊一筹，而 AI04-AI05 的预报效果总体低于 AnEn 方法。以上结果表明：合理构建  
118 AI 模型集合框架，对 2 m 温度，2 m 相对湿度和 10 m 风速的预报准确性提升具有至关重要的  
119 的作用。这为今后应用 AI 技术提高天气预报业务水平提供了有价值的参考。以 2 m 温度预  
120 报为例，AI01—AI03 优于 AnEn 方法的预报效果，其中 AI01 和 AI02 的优势更加明显，较  
121 AnEn 方法分别提高 24.2%和 17.0%。

122 表 1 决赛阶段 AI01—AI05 队伍对 2 m 温度，2 m 相对湿度和 10 m 风速预报的 RMSE<sub>p</sub>  
123 周平均成绩

124 Table 1. Weekly mean RMSE<sub>p</sub> values for t<sub>2</sub>, rh<sub>2</sub>, w<sub>10</sub> from forecasts by AI01–AI05

RMSE <sub>p</sub>	AI01	AI02	AI03	AI04	AI05
t <sub>2</sub>	24.2	17.0	7.0	-8.3	-29.9
rh <sub>2</sub>	12.4	9.7	-6.1	-13.2	-24.6
w <sub>10</sub>	6.2	-3.3	-4.0	-4.4	-6.3

125

#### 126 4 展望

127 本次“天气预报”竞赛项目共吸引全球超过 1000 支队伍参加。竞赛选手来自全球多个  
128 国家和地区的众多高等学府和高科技公司。在仅提供极为有限的气象数据集的条件下，本次  
129 竞赛构建的部分 AI 集合气象模型，由于采用了较合理的智能模型组合和数据处理技巧，对  
130 2 m 温度、2 m 相对湿度和 10 m 风速等气象要素展示出较好的预报能力。其中 AI01 和 AI02  
131 参赛队伍对 2 m 温度的预报准确率较 AnEn 方法分别提高 24.2%和 17.0%。

132 本次竞赛也存在一些不足。例如：降水预报效果并没有进行评估。然而，竞赛结果仍然  
133 表明，AI 技术在气象领域拥有重大的潜力等待被发掘。未来，北京城市气象研究院也将持

134 续探索 AI 技术在气象领域的应用能力，并针对社会影响广泛的天气类型，切实提高其精细化  
135 预报水平。

136 AI 技术可以渗透到基于大数据的众多学科，目前，除了应用广泛的医疗，交通，教育  
137 等领域，其未来在气象，水文，地质等自然科学领域也将会有巨大的机遇。

138 **致谢：**感谢创新工场 CEO 李开复先生，CTO 王咏刚先生，吴卓浩副总裁和东静女士，以  
139 及为本次“天气预报”竞赛项目圆满举办贡献力量所有团队成员们。感谢全球所有热情投  
140 身于本次“天气预报”竞赛项目的广大竞赛团队的大力支持与智力奉献。

#### 141 参考文献

142 毕宝贵. 2017. 智能天气预报进展及思考. 2017 年全国气象台长会议技术报告. 银川, 2017,  
143 10.12. Bi B G. 2017. Progresses and thoughts on weather forecasting using artificial  
144 intelligence technology. Proc. National Conference of Weather Forecast Center Directors,  
145 Yinchuan, China, 12 October 2017 (in Chinese)

146 戴翼, 何娜, 付宗钰等. 2019. 北京智能网格温度客观预报方法(BJTM)及预报效果检验. 干  
147 旱气象, 37(2): 339-344. Dai Y, He N, Fu Z Y, et al. 2019. Beijing intelligent grid  
148 temperature objective prediction method (BJTM) and verification of forecast result.  
149 *Chinese J Arid Meteor*, **37(2)**: 339-344 (in Chinese)

150 范水勇, 王洪利, 陈敏等. 2013. 雷达反射率资料的三维变分同化研究. 气象学报, 71(3):  
151 527-537. Fan S Y, Wang H L, Chen M, et al. 2013. Study of the data assimilation of radar  
152 reflectivity with the WRF 3D-Var. *Acta Meteor Sinica*, **71(3)**: 527-537 (in Chinese)

153 郭瀚阳, 陈明轩, 韩雷等. 2019. 基于深度学习的强对流高分辨率临近预报试验. 气象学报,  
154 77(4): 715-727. Guo H Y, Chen M X, Han L, et al. 2019. High resolution nowcasting  
155 experiment of severe convection based on deep learning. *Acta Meteor Sinica*,  
156 **77(4)**:715-727(in Chinese)

157 王在文, 陈敏, L D Monache 等. 2019. 相似集合预报方法在北京区域地面气温和风速预报中  
158 的应用. 气象学报, 77(5): xxx-xxx. Wang Z W, Chen M, Monache L D, et al. 2019.  
159 Application of analog ensemble method to surface temperature and wind speed prediction  
160 in Beijing area. *Acta Meteor Sinica*, **77(5)**: xxx-xxx (in Chinese)

161 杨璐, 韩丰, 陈明轩等. 2018. 基于支持向量机的雷暴大风识别方法. 应用气象学报, 29(6):  
162 680-689. Yang L, Han F, Chen M X, et al. 2018. Thunderstorm gale identification method  
163 based on support vector machine. *J Appl Meteor Sci*, **29(6)**: 680-689 (in Chinese)

164 Burrows W R, Mooney C J. 2018. Automated products for forecasting arctic blizzard conditions.  
165 J36.4 in Proc. Annual Meeting of the Amer. Meteor. Soc., Austin, Texas, 6-11 January  
166 2018. <http://ams.confex.com/ams/98Annual/webprogram/Paper336043.html>

167 Collins W, Prabhat M, Racah E, et al. 2018. Deep learning for detecting extreme weather and  
168 climate patterns. TJ7.1 in Proc. Annual Meeting of the Amer. Meteor. Soc., Austin. Texas,  
169 6-11 January 2018. <http://ams.confex.com/ams/98Annual/webprogram/Paper328029.html>

170 EarthRisk Technologies. 2013. TempRisk Apollo White Paper.  
171 [http://www.earthrisktech.com/resources/reports/white\\_papers/TempRiskApollo\\_WhitePa](http://www.earthrisktech.com/resources/reports/white_papers/TempRiskApollo_WhitePaper_Oct2013.pdf)  
172 [per\\_Oct2013.pdf](http://www.earthrisktech.com/resources/reports/white_papers/TempRiskApollo_WhitePaper_Oct2013.pdf).

173 Hinton G E, Osindero S, Teh Y. 2006, A fast learning algorithm for deep belief nets. *Neural*  
174 *Computation*, **18**, 1527-1554(请核实是否有期号，若有请填上)

175 Kneringer P, Dietz S J, Mayr G J, et al. 2018. An ordered hurdle model for probabilistic  
176 low-visibility nowcasting to support decisions at airports. J36.6 in Proc. Annual Meeting  
177 of the Amer. Meteor. Soc., Austin, Texas, 6-11 January 2018.

- 178 <http://ams.confex.com/ams/98Annual/webprogram/Paper325064.html>. Accessed on 16  
179 [August 2019](#)
- 180 Kunkel K E, Biard J C, Racah E. 2018. Automated detection of fronts using a deep learning  
181 algorithm. TJ7.4 in Proc. Annual Meeting of the Amer. Meteor. Soc., Austin. Texas, 6-11  
182 January 2018. <http://ams.confex.com/ams/98Annual/webprogram/Paper333480.html>
- 183 Lagerquist R, McGovern A, Richman M B, et al. 2018. Using machine learning to forecast severe  
184 thunderstorm winds on a CONUS-Wide grid. 3.1 in Proc. Annual Meeting of the Amer.  
185 Meteor. Soc., Austin. Texas, 6-11 January 2018.  
186 <http://ams.confex.com/ams/98Annual/webprogram/Paper335039.html>
- 187 Mattioli C J, Veillette M S, Iskenderian H. 2018. Dual application of convolutional neural  
188 networks: Forecasts of radar precipitation intensity and offshore radar-like mosaics. 695  
189 in Proc. Annual Meeting of the Amer. Meteor. Soc., Austin. Texas, 6-11 January 2018.  
190 <http://ams.confex.com/ams/98Annual/webprogram/Paper323735.html>
- 191 Wang Y, Long M, Wang J, et al. 2018. PredRNN: Recurrent neural networks for predictive leaning  
192 using spatiotemporal LSTMs. Proc. 31st Conference on Neural Information Processing  
193 Systems (NIPS 2017), Long Beach, CA, USA, 4-9 December 2017.  
194 <http://papers.nips.cc/paper/6689-predrnn-recurrent-neural-networks-for-predictive-learning-using-spatiotemporal-lstms>  
195
- 196 Yao Y C, Li Z J. 2017. Short-term precipitation forecasting based on radar reflectivity images.  
197 Proc. International Conference on Information and Knowledge Management, Singapore,  
198 6-10 November 2017.  
199 [https://github.com/yaoyichen/CIKM-Cup-2017/blob/master/CIKM\\_AnalytiCup\\_2017\\_Team\\_Marmot.pdf](https://github.com/yaoyichen/CIKM-Cup-2017/blob/master/CIKM_AnalytiCup_2017_Team_Marmot.pdf)  
200
- 201 Zhou K H, Zheng Y G, Li B, et al. 2019. Forecasting different types of convective weather: A deep  
202 learning approach. *J Meteor Res*, **33** (期号), XXX—XXX