

利用专家神经网络试报广西前汛期暴雨

杨望月 那基坤 曾 奋

(广西气象台, 南宁 530022)

提 要

介绍了专家神经网络的基本方法及用此方法作广西前汛期暴雨预报的经验, 以探索人工智能用于灾害性天气预报的新方法。

关键词: 专家神经网络 暴雨预报 BP网络 能量函数

引 言

广西前汛期(5—6月)暴雨是广西重大灾害性天气之一, 尤其是全区性暴雨, 降水特点是强度大、范围广、来得快、分布不均匀, 24小时雨量常可达100—200mm, 最大达600mm左右, 一日中往往有10个以上站点同时出现日雨量 ≥ 50 mm的暴雨, 因此常常造成洪涝灾害。但是从预报角度看, 此类灾害性天气仍属小概率事件, 其气候概率仅为8%左右, 预报难度较大。近几年经过试验, 使用人工智能专家系统作预报, 效果较为理想。但是该系统的预报效果不仅依赖于所选择的因子, 而且还取决于预报规则的编写, 即系统的功能好坏取决于编写者的知识与能力。由于专家系统非常依赖于专家的经验及知识的人工编写, 这在很大程度上给预报人员, 尤其是缺乏经验的新预报员带来困难。尽管以前研制的智能专家系统也具有机器学习功能, 可以由机器通过对历史个例学习, 编出预报规则, 但是水平很低。近来, 国际、国内有关专家研制的神经网络方法^[1]却在系统的自适应或自学习方面具有很大的优越性。为此, 我

们在北京大学张承福副教授、北京气象学院王耀生教授等有关专家的指导和帮助下, 尝试利用专家神经网络作前汛期暴雨预报, 为其它灾害性的天气预报探索新方法。

1 专家神经网络的预报方法

1.1 基本方法

我们采用的是多层前传网络的误差反传法(BP网络法)。即采用正向传播信息, 反向传播误差, 对多层网络逐层训练算法, 通过训练找出最佳的独立网络群, 然后用投票表决作预报。

1.2 样本的选取及格式

试验采用1981—1990年广西前汛期暴雨专家系统预报因子作为样本。

1.2.1 预报因子的获取

广西前汛期暴雨专家系统预报因子共有39个, 其中9个为反映天气形势(如高空槽、切变线等)的因子, 其余的因子为指标形式, 绝大多数采用自动识别的方法获取, 与省级气象台站实时业务系统连接, 因此不需人工输入因子。

考虑到历史资料中有些因子缺得较多,

因此在专家神经网络试验中仅选取了 34 个因子。

1.2.2 样本格式

把每天的预报因子列成一行,作为一个样本,其资料格式如下:

因子最大值(指历史上出现的)

因子最小值(同上)

序号(日期)	因子值	实况
如: 810501	5211341.....	R
810502	3122131.....	N
.....		

R 为有雨, N 为无雨。1981—1990 年共 610 个样本。

1.3 网络的结构

由输入层、隐层、输出层组成(图略)。

1.4 网络的训练

采用 BP 网络 S 型曲线作用函数 $f(x) = \frac{1}{1+e^{-x}}$, $f(x)$ 在 -1 到 1 之间取值(图略), 把预报实况的历史资料分为相应的几段(如 0, 1), 进行训练。

训练的方法是采取梯度下降算法, 即 BP 网络算法。引入能量函数的概念:

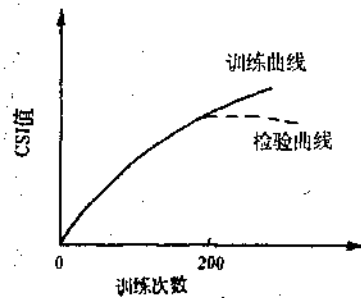
$$E = \frac{1}{2} \sum_{i=1}^M (t_i^{(a)} - o_i^{(a)})^2$$

其中, $o_i^{(a)}$ 表示输出层的实际输出值, $t_i^{(a)}$ 则是目标输出。E 实际上代表误差的大小, 学习算法就是使 E 取极小值, 用改变权重的分配函数达到这一目标。因此程序中用能量函数进行调整, 其原则是使分类集中, 折衷最优。从图象来看实质上是高维点聚图。运行时能量曲线表示较直观。

隐单元的选取, 经过试验采用 6 个较为合适, 因为用 6 个隐单元训练时, 训练 200 次

左右即可找出最高的成功指数(CSI)的最佳点, 如附图所示训练曲线(CSI 值)与检验曲线拟合的最高点。在试验时我们也取过 8 个隐单元, 但训练的时间很长, 而且不易收敛, 效果没有 6 个隐单元好。

训练时各隐单元权重及阈值的初始值是随机的, 根据给定期望的成功指数值, 机器可以自动进行训练, 也可以在专家指导人工干预下训练。



附图

1.5 最佳网络投票

经过对历史资料的学习训练(学习样本数至少 100—200 个), 选出成功界限指数 $[CSI = \text{正确次数} / (\text{正确次数} + \text{漏报次数} + \text{错报次数})]$ 相对最高的若干个专家网络(类似于规则), 然后进行综合投票, 作出预报。

1.6 历史验证情况

广西前汛期暴雨预报利用 1981—1990 年 5、6 月份的 10 年逐日资料(600 多个样本)进行试验, 前 6 年用于学习训练, 后 4 年作检验, 对训练出的 6 个最佳专家网络进行综合投票作为预报结果, 其 CSI 值为 67% 左右, 高于专家系统预报水平。(专家系统预报 1986—1990 年 CSI 平均值为 50%)。

2 1993年试报结果及与专家系统平行试验对比

1993年前汛期期间的广西大暴雨预报,用神经网络试报及专家系统预报的结果见附表。

附表 1993年两种方法的预报结果

	神经网络预报			专家系统预报			
	有	无	合计	有	无	合计	
实况	有	4	5	9	6	3	9
	无	6	43	49	3	46	49
	合计	10	48	58	9	49	58

由附表可知,专家神经网络预报的CSI值 $=4/(4+6+5) \approx 0.27$,专家系统预报的CSI值 $=6/(6+3+3) = 0.50$ 。神经网络试报的CSI值低于专家系统预报及历史验证水平。但所使用的单个专家网络的CSI值分别为36%、38%、42%、37%、41%,而综合投票预报结果没有达到预期愿望。尤其6月份效果较差,有5场大暴雨过程未报出。

3 讨论

为什么会出现学习检验的CSI值比较高(达60%以上),而1993年试报的CSI值显著下降(仅27%左右)?我们认为这与所用资料的年代有关,因为1993年试报时用的是1981—1986年资料学习训练的网络。试验结果表明,历史样本对未来预报的能力将随着时间的推移而明显下降,这表明蕴含于历史数据中的预报信息会随着时间的推移变得愈来愈模糊,这一点在动力方法或统计方法的预报中也已显现。

然而神经网络的预报稳定性问题(也可叫做可预报性问题)是一个值得进一步研究的方面,它不但有具体应用的意义,也有着一定的理论价值,对进一步探索气象演变的规律性可能会有所裨益,对此,我们正在作进一步的试验。

参考文献

- 1 张承福,神经网络系统.力学进展,1988,18(2).

An Attempt of Guangxi Early Rainy Season Rainstorm Forecast with Expert Neural Network

Yang Wangyue Na Jikun Zeng Fen

(Guangxi Meteorological Observatory, Nanning 530022)

Abstract

A method of expert neural network and the trail rainstorm forecasting with the method during Guangxi early rainy season is introduced, in order to explore a new method of disastrous weather forecast with artificial intelligence.

Key Words: expert neural network rainstorm forecast BP network energy function