

感知器算法在台风风暴潮灾情等级评估中的应用^{*}

叶 雯, 刘美南, 陈晓宏

(中山大学水资源与环境系, 广东 广州 510275)

摘 要: 风暴潮灾害每年都给我国沿海地区造成巨大的经济损失。在前人研究的基础上, 引入风暴潮灾情等级的模式识别思想, 运用感知器方法的理论及其算法, 并以广东省湛江地区为例, 建立了该地区的台风风暴潮灾情等级判别函数, 经还原检验和判别预测获得满意的结果, 证明该方法用于台风风暴潮灾情等级判别是可行的。这对于实现风暴潮灾情等级的快速评估, 为抢险救灾的人力、财力和物力分配提供依据。

关键词: 模式识别; 台风风暴潮; 灾情等级

中图分类号: P426.62 **文献标识码:** A **文章编号:** 0529-6579 (2004) 02-0117-04

风暴潮灾害是发生频率最高, 造成的经济损失最严重, 危害最大的海洋自然灾害^[1]。我国拥有漫长的海岸线, 沿海众多区域对风暴潮灾害尤其敏感和脆弱, 而这些地区是我国人口、财产密度最高、社会经济最发达的地区, 因此, 风暴潮造成的损失非常严重。如何在灾情发生后, 对灾情进行快速评估, 确定灾情等级, 为救灾和减灾的决策与实施提供第一手材料, 一直以来是人们热切关注与急待解决的问题之一。然而, 由于受到自然环境、社会经济等诸多因素的影响, 风暴潮灾情常表现出各种不确定或未知的信息, 给灾情等级评估工作带来很大困难。现有的风暴潮灾情等级判别方法多是公式加经验的方法, 带有一定的主观经验性。

实际上, 风暴潮灾情等级评估从根本上来说是一个模式识别问题, 相继有学者进行了模式识别理论应用于自然灾害灾情等级评估的研究工作^[2-4]。本文在前人研究的基础上, 引入模式识别技术中的感知器算法, 建立台风风暴潮灾情等级判别模型, 并结合实例说明其应用。

1 方法简介

1.1 判别函数

模式识别 (pattern recognition) 是 20 世纪 60 年代初迅速发展起来的一门新兴学科, 主要功能是通过设计合适的分类器来判别各个模式所属的类别。本文所采用的分类器是线性判别函数。即对于训练模式样本集 $\{X\} = (Z_1, Z_2, \dots, Z_i, \dots, Z_n), Z_i =$

$(x_{i1}, x_{i2}, \dots, x_{in})$; M 类线性可分模式集为 $\Omega = \{w_1, w_2, \dots, w_M\}$, 存在这样一个直线方程:

$$d(z) = a_1 x_1 + a_2 x_2 + \dots + a_n x_n + a_{n+1} = A_0^T Z + a_{n+1}$$

作为模式类别划分的依据。 $d(Z)$ 即为判别函数 (discriminant function)。

1.2 判别函数的感知器算法

感知器一词是人们对一种分类学习机模型的称呼, 最早的感知器模型是美国学者 Fank Rosenblatt 在 1957 年提出的。感知器训练算法 (the perceptron approach) 是指通过对训练模式样本集的学习, 从而得到判别函数的系数解, 产生线性可分的模式判别函数^[5]。该算法的优点是不需要对各类模式的统计性质作任何假设, 属于确定性方法。

1.3 赏罚概念

赏罚过程 (reward and punishment procedure) 贯穿于感知器算法的始终, 是感知器算法的一个主要内容。其基本概念是对正确分类的模式则“赏” (本文中用“不罚”), 即权向量 A 不变; 对错误分类的模式则“罚”, 即将权向量 A 加上一个正比于模式样本 Z 的分量。用全部模式样本训练过一轮后, 只要有一个样本是判别错误的, 则需要进行第二轮迭代, 继续这一赏罚过程, 直到全部分类正确为止。

1.4 感知器的迭代算法

本文采用的模式划分方法为不含 IR (不确定性

^{*} 收稿日期: 2003-05-16

基金项目: 国家自然科学基金资助项目 (49971019)

作者简介: 叶雯 (1978 年生), 女, 硕士研究生; 通讯联系人: 刘美南; E-mail: lmn@zsu.edu.cn

区域) $\frac{w_i}{w_j}$ 的两分法。即

$$d_i - d_j = \begin{cases} > 0 & \text{则 } Z \in w_i \\ \leq 0 & \text{则 } Z \in w_j \end{cases} \quad (1)$$
$$i, j = 1, 2, \dots, M, i \neq j$$

该划分方法保证了不确定性区域的不存在, 并且将 M 类模式一一划分只需要 M 个判别函数。

对 M 类模式, 假设在训练过程中的第 K 次迭代时, 一个属于 w_i 类的模式样本 Z 被送入分类器中, 计算其 M 个判别函数:

$$d_j(k) = A_j(k)Z, \quad j = 1, 2, \dots, M \quad (2)$$

的结果。在迭代过程中按赏罚概念对权向量进行修正。

(1) 如果 $d_i(k) > d_j(k), j = 1, 2, \dots, M, j \neq i$ 的条件成立, 则权向量不变。即 $A_j(k+1) = A_j(k), j = 1, 2, \dots, M$ 。

(2) 如果其中第 l 个权向量使 $d_i(k) \leq d_l(k)$, 则相应的权向量应作如下调整:

$$A_i(k+1) = A_i(k) + cZ \quad (3)$$

$$A_l(k+1) = A_l(k) - cZ \quad (4)$$

由上述两种情况可知, 感知器迭代算法为:

$$\begin{cases} A_j(k+1) = A_j(k) & \text{若 } d_j(k) > d_i(k) \\ A_i(k+1) = A_i(k) + cZ & \text{若 } d_i(k) > d_l(k) \\ A_l(k+1) = A_l(k) - cZ & \text{若 } d_i(k) \leq d_l(k) \end{cases}$$
$$j = 1, 2, \dots, M, j \neq i, j \neq l \quad (5)$$

式中, c 为一正常数, 称为惩罚因子。通常 c 值都取 1。只要确定了权向量的初始值 $A_i(1)$ (可视情况任意选取) 和 c , 即可由全部训练模式样本进行迭代计算。训练过一轮后, 只要有一个样本是判别错误的, 就要进行第二轮迭代, 即用全部模式再训练一次, 建立新的 A 。如此不断反复, 直到用全部模式进行训练都能获得正确的分类结果时, 结束迭代, 此时的 A 即为算法的结果。由 A 及 (2) 式就可以得到模式判别函数。

1.5 风暴潮灾情的判别

本文将风暴潮灾情按其严重程度划分为 I、II、III 灾度等级, 分别对应小灾、中灾、大灾。每一个等级为一个模式类, 显然每一模式类都包含了一些灾度相同的风暴潮灾害个例, 从而可以将分属于各模式类的风暴潮灾害个例看成一个模式样本集。同时, 选取死亡人数 (D)、受灾面积 (A)、直接经济损失 (L) 3 项指标作为训练模式样本元素, 组成训练模式样本集, 由感知器算法推求风暴潮灾情等级判别函数。

2 应用实例

2.1 湛江地区台风风暴潮灾情等级判别函数的建立

为了更好地说明感知器算法应用于台风风暴潮灾情等级判别的全过程, 本文收集了 1976—2002 年期间对湛江地区造成灾害的 30 多场台风风暴潮的灾情统计资料。由于这期间湛江地区经济发展较快, 物价波动较大, 以当年的灾害直接经济损失数据难以准确反映风暴潮直接经济损失增长趋势, 不同年份之间的经济损失不具有可比性。因此我们对原始数据进行了可比性处理, 全部统一到 1975 年的物价水平。同时为了消除不同量纲之间的差异, 加快算法的收敛速度, 对原始数据进行了如下的归一化处理, 将原始数据转换到 $[0, 1]$ 区间之内:

$$D_i = \frac{D_i - \min(D)}{\max(D) - \min(D)}$$
$$A_i = \frac{A_i - \min(A)}{\max(A) - \min(A)}$$
$$L_i = \frac{L_i - \min(L)}{\max(L) - \min(L)} \quad (6)$$

在建立判别函数之前, 为了将本文利用判别函数所得到的分类结果与其他方法的结果进行比较, 我们参考李桂忱等^[9]划分暴雨洪涝灾害灾情等级的灾度分析法, 即将各个训练模式样本的 3 个样本元素 D 、 A 、 L 按式 (6) 归一化后的灾情指数之和作为该样本的灾度值, 把按照灾度值划分的灾情等级作为各样本建立判别函数时所属的原始灾情等级 (见表 1)。

将风暴潮灾情的 3 个等级: 小灾、中灾和大灾分别为模式类 w_1, w_2, w_3 , 以表 1 中属于小灾等级的 1~19 号风暴潮、属于中灾等级的 22~26 号风暴潮及属于大灾等级的 29~31 号风暴潮, 共 27 个组成训练模式样本集, 其余 5 个样本作为测试样本。由感知器算法推求风暴潮灾情等级判别函数。

权向量初值 $A_i(1) (i = 1, 2, \dots, M)$ 和赏罚因子 c 的选择没有固定的规律, 一般可视迭代计算情况任意选定。本文取 $A_1(1) = A_2(1) = A_3(1) = 0, c = 1$ 。将各训练模式样本依次一一送入分类器, 按式 (5) 编制程序进行迭代计算。当所有训练模式样本都已经分类正确, 则迭代结束, 得到相应的判别函数为:

$$d_1 = -5.21D - 23.74A - 18.68L - 27$$
$$d_2 = -3.04D - 10.68A - 4.60L - 33 \quad (7)$$
$$d_3 = 4.41D - 0.77A + 11.93L - 41$$

式 (7) 就是本文建立的湛江地区台风风暴潮灾情等级评估模型。对于该地区发生的一场风暴潮灾

害, 只要我们收集到其造成的死亡人数, 受灾面积及直接经济损失的数据, 按照式 (6) 进行归一化处理, 将处理后的数据代入式 (7) 计算出各 d_i 值, 该次风暴潮灾情等级即可归入判别函数值 d_i 最大的所属的灾度中。如此, 我们就可以快速评估风暴潮灾害的灾情等级。

表 1 湛江地区部分台风风暴潮灾情分类特征值
Tab 1 Eigenvalues of part of typhoon storm surges in Zhanjiang District

序号	台风编号	D	A	L	灾情等级
1	0220	0.000	0.114	0.002	小灾
2	0214	0.000	0.142	0.016	小灾
3	0107	0.000	0.113	0.016	小灾
4	9910	0.000	0.051	0.011	小灾
5	9710	0.000	0.035	0.002	小灾
6	9506	0.000	0.0119	0.008	小灾
7	9318	0.000	0.022	0.002	小灾
8	9316	0.047	0.085	0.043	小灾
9	9302	0.000	0.095	0.003	小灾
10	9108	0.000	0.063	0.008	小灾
11	9004	0.004	0.344	0.011	小灾
12	8908	0.004	0.084	0.009	小灾
13	8903	0.023	0.059	0.023	小灾
14	8805	0.012	0.021	0.000	小灾
15	8702	0.004	0.000	0.000	小灾
16	8609	0.043	0.251	0.016	小灾
17	8607	0.004	0.049	0.001	小灾
18	8517	0.035	0.397	0.020	小灾
19	8311	0.133	0.067	0.018	小灾
20	8303	0.000	0.079	0.002	小灾
21	8217	0.039	0.193	0.071	小灾
22	9516	0.000	0.476	0.026	中灾
23	9515	0.000	0.646	0.038	中灾
24	9403	0.145	0.091	0.334	中灾
25	9309	0.000	0.582	0.002	中灾
26	8616	0.121	0.382	0.067	中灾
27	85 未编号台风	0.301	0.393	0.032	中灾
28	8410	0.070	0.505	0.025	中灾
29	0103	0.000	0.563	0.222	大灾
30	9615	1.000	1.000	1.000	大灾
31	8007	1.000	0.150	0.109	大灾
32	7619	0.219	0.617	0.118	大灾

为了检验该判别函数, 随机选取 5 个训练模式样本作为还原检验样本, 5 个非训练模式样本进行灾情等级判别预测, 结果如表 2。由表中可以看出, 分类的结果与灾度分析法所得到的结果完全一致, 符合率达到 100%, 说明本文建立的判别函数具有很高的可靠性。

2.2 风暴潮灾情等级分类结果合理性检验

对上述分类结果进行合理性检验分析, 得出如下结论:

(1) 上述分类结果符合宝塔状分布规律, 即大灾应该为少数, 中、小灾占多数。对 32 场风暴潮灾害划分等级的结果是: 大灾 4 次, 占总数的 12.5%; 中灾 7 次, 占总数的 21.9%; 小灾 21 次, 占总数的 65.6%。划分结果表明, 符合宝塔状分布规律。

(2) 符合历史事实, 即对少数灾害, 尤其是对全市造成很大影响的风暴潮灾害, 应该划分成大灾。被划分为大灾的 4 场风暴潮灾害, 无论是造成的死亡人数、受灾面积等, 还是造成的直接经济损失, 均为本文所采用的样本系列中的最大值, 如 9615 号强台风造成的经济损失达 129.39 亿元 (当年价) 之巨, 占了当年 GDP 的 40.98%, 是湛江市建国以来遭受的最严重的经济损失; 而 8007 号强台风更是造成了南渡站 5.90m 的最大风暴潮增水, 名列世界第三。因此, 把这 4 次风暴潮灾害划分为大灾是符合历史事实和客观情理的。

2.3 风暴潮灾情等级判别函数的适用性

风暴潮灾情等级判别研究的主要成果是一种普遍适用的判别理论和方法的提出, 而本文得出的判别函数只是该理论与方法的一个具体结果, 它是一个地区致灾因子、孕灾环境与当地自然条件的综合产物, 只能运用于当地, 而对于不同地区、不同情况可以得出不同的判别函数。只要各地区存在能够反映本地区风暴潮灾情特征的足够数量的样本资料, 就可以按本文的理论和方法确定适用于本地区的判别函数。

3 结 论

以上的介绍和应用实例研究表明, 模式识别技术的感知器算法为风暴潮灾情等级评估提出了一种全新的方法, 本文所建立的风暴潮灾情等级判别函数能够较准确而且快速地对灾情进行评估, 具有一定的实用性。

模式识别理论应用于自然灾害灾情评估, 尤其是风暴潮灾情等级评估的研究工作尚处于探索阶段, 不论在理论和实践上都还有许多问题需要进一步研究。如训练模式样本集的线性可分性的判别问题, 因为如果训练模式样本集不是线性可分的, 会造成感知器算法的不收敛。但我们相信, 模式识别技术在自然灾害灾情等级评估中将具有广阔的应用前景。

表 2 风暴潮灾情等级判别还原检验及判别预测¹⁾

Tah 2 Test and forecast of evaluation of stom surge disaster degree

样本范围	台风编号	D	A	L	模式判别函数值			样本属类	判别是 否正确
					d_1	d_2	d_3		
训练 模式 样本	8311	0.133	0.067	0.018	-29.6198	-34.2027	-40.2503	I	✓
	8217	0.039	0.193	0.071	-33.1113	-35.5064	-40.1296	I	✓
	9309	0.000	0.582	0.002	-40.8540	-39.2250	-41.4243	II	✓
	8616	0.121	0.382	0.067	-37.9507	-37.7558	-39.9612	II	✓
	8007	1.000	0.150	0.109	-37.8071	-38.1434	-35.4051	III	✓
非训练 模式 样本	8303	0.000	0.079	0.002	-25.9128	-33.8529	-41.0370	I	✓
	8217	0.039	0.193	0.071	-33.1113	-35.5064	-40.1296	I	✓
	85 未编 号台风	0.301	0.393	0.032	-38.4958	-38.2595	-39.5934	II	✓
	8410	0.070	0.505	0.025	-39.8204	-38.7212	-40.7819	II	✓
	7619	0.219	0.617	0.118	-44.9928	-40.7981	-39.1016	III	✓

1) I 小灾, II 中灾, III大灾, ✓分类正确, ×分类错误

参考文献:

[1] LE K T. Marine disaster in China and the strategy for their mitigation // ELSSBH M I, et al. Land-based and marine Hazards[M] . Kluwer Academic Publishers, 1996: 256— 277.

[2] 任鲁川. 灾害损失等级划分的模糊灾度判别法[J] . 自然灾害学报, 1996, 5(3): 13— 17.

[3] 魏一鸣, 万庆, 周成虎. 基于神经网络的自然灾害灾情评估模型研究[J] . 自然灾害学报, 1997, 6(2): 1— 6.

[4] 杨思全, 陈亚宁. 基于模糊模式识别理论的灾害损失等级划分[J] . 自然灾害学报, 1999, 8(2): 56— 60.

[5] 陈晓宏. 分层型水库水温水质分布模拟预测研究[D] . 武汉: 武汉水利电力学院, 1991.

[6] 李桂忱, 胡朝霞. 暴雨洪涝灾害灾情级别及其空间分布[J] . 中国减灾, 1997, 7(2): 47— 50.

The Application of the Perception Approach in the Evaluation of Typhoon Storm Surge Disaster Degree

YE Wen, LIU Mei nan, CHEN Xiao hong

(Department of Water Resources and Environment, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: Storm surge disasters result in great economic loss in coastal areas of China every year. The idea of pattern recognition is introduced into storm surge disaster degree evaluation for the first time. Using the theory and method of the perceptron approach and the loss data of Zhanjiang District of Guangdong Province, a discriminant function of typhoon storm surge disaster degree is established with satisfactory results. It is proved that the perceptron approach is feasible in storm surge disaster degree evaluation. It is significant in realizing rapid evaluation of storm surge disaster degree and distributing the manpower and material resources reasonably when providing disaster relief.

Key words: pattern recognition; typhoon storm surge; evaluation of disaster degree