

粗糙集神经网络在昆虫总科阶元分类学上的应用

杜瑞卿¹ 褚学英¹ 王庆林¹ 赵秋红² 庞发虎¹

(1. 南阳师范学院 生命科学系, 河南 南阳 473061; 2. 南阳师范学院 数学系, 河南 南阳 473061)

摘要 为研究粗糙集和神经网络相结合方法的实践性,以及昆虫的数学形态特征在总科阶元上作为分类特征的可行性、可靠性和重要性。从总科角度对鳞翅目(Lepidoptera)和鞘翅目(Coleoptera)5个总科23种虫体图像中提取的昆虫面积、周长等11项数学形态特征进行粗糙集神经网络分析。结果表明:在总科阶元上,11项特征的可靠性大小为,面积、亮斑数>周长、横轴长、形状参数、圆形性、似圆度、偏心率>纵轴长、叶状性、球状性,与赵汗青等的统计分析结果中属性特征的重要性大多数一致;神经网络模式识别结果与传统分类结果完全一致。应用粗糙集理论进行昆虫数学形态特征分类结果准确;在昆虫总科阶元分类上粗糙集神经网络较统计学方法具有优势。

关键词 昆虫分类;粗糙集;神经网络;数学形态特征

中图分类号 Q 964

文章编号 1007-4333(2007)01-0033-06

文献标识码 A

Application of a rough-set neural network to superfamily level in insect taxonomy

Du Ruiqing¹, Chu Xueying¹, Wang Qinglin¹, Zhao Qihong², Pang Fahu¹

(1. Department of Biology, Nanyang Normal University, Nanyang 473061, China;

2. Department of Mathematics, Nanyang Normal University, Nanyang 473061, China)

Abstract To study the feasibility, reliability and importance of taxonomy character at superfamily level in insect morphological features by a combination approach of the rough-set theory and the neural network we analyzed 11 morphological features (MMFs), such as area, perimeter, and so on, from the images of 23 species of insects, in five superfamilies of Lepidoptera and Coleoptera, Noctuoidea, Bombycoidea, Papilionoidea, Scarabaeoidea and Chrysomeloidea. The results were compared with those of the statistical analysis reported by Zhao Hanqing. The rank of reliability MMFs in the identification of insect superfamilies is from high to low; the area and hot-holenumbe are grater than perimeter; the X-length, form parameter, circularity, roundness-likelihood, eccentricity greater then Y-length, lobation, sphericity. The results are not completely identical with those from Zhao Hanqing, but the most characteristics are identical. The results of pattern recognition by neural network are completely identical with those of traditional classifications. This theory applied to insect taxonomy is better compared with the statistical analysis method and has a great significance at superfamily level when used with rough-set neural network.

Key words insect taxonomy; theory of rough set; neural network; morphological features (MMF)

20世纪中叶以来,生物分类学领域出现了一些新的理论学派,最著名的有数值分类学派、支序分类学派和进化分类学派^[1-2]。数值分类自20世纪70年代介绍到我国后,首先被应用到昆虫分类中,但发展缓慢^[3-4]。数值分类方法主要有聚类分析、主成分分析、统计分析等。粗糙集理论是一门研究和处理不

精确性的新兴学科^[5],和神经网络在生物分类上的应用刚刚开始,特别是粗糙集理论和神经网络相结合的应用未见报道。笔者用粗糙集模糊聚类分析法在目阶元上对昆虫的数学形态特征进行了分析^[6],并与赵汗青等^[7]的统计分析法进行比较。本着发展数值分类在昆虫分类上的应用,本研究采用粗糙

收稿日期:2006-03-08

基金项目:南阳师范学院青年科学研究资助项目(nytc2004k01)

作者简介:杜瑞卿,讲师,主要从事生物数学生物工程专业研究,E-mail:duruiqing8@163.com

集理论和神经网络相结合的方法在科阶元上对昆虫的数学形态特征进一步分析,并与赵汗青等^[8]的统计分析法进行比较,以此来说明粗糙集理论和神经网络相结合作为昆虫数值分类的一种新方法较统计分析方法更为科学简便,且结果准确。

1 粗糙集理论和神经网络及其应用

1.1 粗糙集理论方法简述

为了便于理解,结合本研究中的23种昆虫对象(表1),只对涉及到的知识进行文字性简述。

1) 等价关系的划分。将要研究的对象23种昆虫构成的集合 U 称为论域。11项数学形态特征属性(指标)似圆度、偏心率等构成等价关系族 R , 每个特征为1个等价关系。 $U = \{e_1, e_2, \dots, e_{23}\}$, $R = \{A, P, L_X, L_Y, F, B, S, C, R, E, H\}$ 。若 $P \subseteq R$ 且 P 则 P (P 中所有等价关系的交集)也是一个等价关系,称为 P 的不可区分关系,记为 $\text{ind}(P)$,利用等价关系和等价关系族对 U 划分成若干子集。例如,可利用 B 划分,也可利用 $P = \{B, S, C, R\}$ 划分,记为 U/B 或 U/P 。

表1 23种昆虫的11项数学形态特征提取值

Table 1 Eleven mathematical morphological characters extracted from 23 species of insect

科	昆虫名称	面积 $A/$ 10^{-4}cm^2	周长 $P/$ 10^{-2}cm	横轴长 $L_X/$ 10^{-2}cm	纵轴长 $L_Y/$ 10^{-2}cm	形状参 数 F	叶状性 B	球状性 S	圆形性 C	似圆度 R	偏心率 E	亮斑数 H
夜蛾 总科	小地老虎	20 365	932.9	273.5	117.8	3.315	0.089	0.156	5.900	0.335	0.431	10.12
	棉铃虫	10 394	768.8	188.2	86.0	4.686	0.085	0.151	4.909	0.372	0.441	26.17
	白点雍夜蛾	13 841	717.8	228.4	96.0	2.980	0.076	0.139	5.874	0.338	0.409	1.84
蚕蛾 总科	大蚕蛾	72 155	1 685.0	485.1	124.7	3.142	0.048	0.084	7.750	0.442	0.254	5.20
	丁目大蚕蛾	82 879	1 821.0	503.4	147.8	3.201	0.069	0.087	9.366	0.426	0.279	3.40
	黄目大蚕蛾	103 368	2 143.0	561.0	156.4	3.545	0.041	0.072	9.243	0.420	0.270	53.17
	猫目大蚕蛾	149 687	2 486.0	640.9	178.4	3.297	0.050	0.086	9.758	0.464	0.261	10.63
凤蝶 总科	菜粉蝶	37 948	1 206.0	298.0	92.7	2.901	0.086	0.145	6.923	0.551	0.316	7.73
	黄粉蝶	38 479	1 105.0	305.9	105.1	2.538	0.110	0.189	6.886	0.526	0.344	1.22
	山楂粉蝶	54 812	494.0	375.5	117.5	3.278	0.064	0.109	8.275	0.495	0.264	1.39
	尖钩粉蝶	60 053	1 344.0	377.2	125.4	2.404	0.114	0.198	8.521	0.542	0.321	1.76
	黄环蛱蝶	42 783	1 317.0	404.3	102.2	3.348	0.070	0.133	6.818	0.334	0.228	43.55
金龟甲 总科	独角仙	31 773	1 231.0	169.4	250.4	3.930	0.314	0.325	5.710	1.432	1.493	5.35
	棕色鳃金龟	5 331	303.0	60.7	113.6	1.373	0.420	0.424	5.554	1.844	1.875	1.63
	华北大黑鳃金龟	5 252	299.6	60.7	113.0	1.362	0.422	0.421	5.564	1.828	1.870	1.57
	中华弧绿丽金龟	4 350	266.8	59.3	98.0	1.306	0.436	0.492	5.163	1.571	1.652	6.74
	铜绿丽金龟	4 606	275.8	61.3	101.0	1.319	0.445	0.514	5.342	1.558	1.648	2.77
	白星花金龟	6 091	320.1	69.3	114.5	1.341	0.453	0.516	5.517	1.614	1.653	12.04
叶甲 总科	黄斑星天牛	8 011.5	488.4	60.8	167.3	2.389	0.349	0.222	5.011	2.743	2.750	16.20
	松幽天牛	2 402.9	234.4	32.0	91.2	1.860	0.363	0.226	4.139	2.991	2.853	1.37
	榆绿天牛	1 871.9	220.0	27.3	88.7	2.067	0.409	0.223	3.682	3.214	3.265	1.00
	绿翅契天牛	2 700.8	262.0	32.9	98.9	2.032	0.411	0.249	3.919	3.176	3.013	1.35
	漆伪叶甲	2 190.3	206.2	36.6	79.7	1.549	0.369	0.307	4.098	2.120	2.185	1.00

2) 上近似集与下近似集的确定。利用某个等价关系例如 B (或某个等价关系族, 如 P) 对 U 进行划分后形成若干子集, 如果子集元素肯定属于某总科例如夜蛾总科 m_1 , 就称 m_1 的 B (或 P) 下近似集, 如果子集元素可能属于某总科例如夜蛾总科 m_1 , 就称 m_1 的 B (或 P) 上近似集。

3) 知识约简^[9-10]。利用不同的等价关系族, 如 $P_1 = \{S, C, R\}$, $P_2 = \{S, C, R, H\}$, $P_3 = \{B, S, C, R, H\}$, 分别对 U 进行划分, 如果 P_2 划分结果与 P_3 划分结果完全相同, 就称 B 是不必要属性; 如果 P_1 划分结果与 P_2 划分结果不相同, 就称 H 是必要属性; 如果 P_2 中每个属性都是必要的, 就称 P_2 是一个约简。令 5 总科为决策属性 $D = \{m_1, m_2, m_3, m_4, m_5\}$, 如果 P_1 划分的子集能够分别属于 D 的子集, 就称 H 为 D 相对不必要属性, 记为 $POS_{ind}(P_1)^{(ind(D))} = POS_{ind}(P_2 - (H))^{(ind(D))}$, 否则就称为相对必要属性。如果 P_1 中每个属性都是 D 的相对必要属性, 就称 P_1 为 D 的一个相对约简。 D 可能有多个相对约简, 所有相对约简共有的属性称为核。以上约简的求解过程是通过在每个对象的每个属性赋予 1 个信息值, 形成信息表, 然后进行布尔函数运算^[9]。

1.2 人工神经网络方法简述

人工神经网络的类型很多, 目前应用广泛的前馈网络是 BP 网络。3 层的 BP 网络包括输入层、中间层 (隐层) 和输出层。本研究中 BP 网络的训练过程如下:

1) 赋予中间层 (隐层) 和输出层权值矩阵和偏置值的初值。本研究中间层权值矩阵和偏置值构成矩阵 $W(5 \times 5)$, 输出层权值矩阵为 $H(5 \times 5)$ 。

2) 设置参数 (学习速度) 和 (误差要求), 并赋予一定的值 $\eta = 0.1$, $\epsilon = 0.01$ 。给定激活函数 $f(x) = 1 / (1 + \exp(-x))$ 。

3) 输入样本特征向量为 x_k (k 为训练样本数), 给定样本目标输出向量为 dk (dk, x_k 均为 5×1 矩阵):

$$\begin{aligned} dk(m_1) &= (0.9 \quad 0.6 \quad 0.6 \quad 0.6 \quad 0.6) \\ dk(m_2) &= (0.6 \quad 0.9 \quad 0.6 \quad 0.6 \quad 0.6) \\ dk(m_3) &= (0.6 \quad 0.6 \quad 0.9 \quad 0.6 \quad 0.6) \\ dk(m_4) &= (0.6 \quad 0.6 \quad 0.6 \quad 0.9 \quad 0.6) \\ dk(m_5) &= (0.6 \quad 0.6 \quad 0.6 \quad 0.6 \quad 0.9) \end{aligned}$$

4) 按如下方法运算: $uk = W \cdot x_k, vk = f(uk)$,

$$yk = f(H \cdot vk), ek = |yk - dk|^2 / 2; uk = (uk_1, uk_2, uk_3, uk_4, uk_5), xk = (xk_1, xk_2, xk_3, xk_4, xk_5), yk = (yk_1, yk_2, yk_3, yk_4, yk_5)。H = (h_1, h_2, h_3, h_4, h_5), dk = (dk_1, dk_2, dk_3, dk_4, dk_5)。$$

如 $ek > \epsilon$, 则对 W 和 H 的反馈性修正:

$$\begin{aligned} gH_1 &= yk_1 \cdot (1 - yk_1) \cdot (dk_1 - yk_1) \\ gH_2 &= yk_2 \cdot (1 - yk_2) \cdot (dk_2 - yk_2) \\ gH_3 &= yk_3 \cdot (1 - yk_3) \cdot (dk_3 - yk_3) \\ gH_4 &= yk_4 \cdot (1 - yk_4) \cdot (dk_4 - yk_4) \\ gH_5 &= yk_5 \cdot (1 - yk_5) \cdot (dk_5 - yk_5) \\ gH &= (gH_1, gH_2, gH_3, gH_4, gH_5) \end{aligned}$$

$$\begin{aligned} gw_1 &= h_1 \cdot gH \cdot f'(uk_1) \\ gw_2 &= h_2 \cdot gH \cdot f'(uk_2) \\ gw_3 &= h_3 \cdot gH \cdot f'(uk_3) \\ gw_4 &= h_4 \cdot gH \cdot f'(uk_4) \\ gw_5 &= h_5 \cdot gH \cdot f'(uk_5) \end{aligned}$$

$$gw = (gw_1, gw_2, gw_3, gw_4, gw_5)$$

$$H = H + \eta \cdot gw \cdot vk$$

$$H(n+1) = H(n) + \eta \cdot gw \cdot vk$$

$$w = w + \eta \cdot gw \cdot xk$$

$$w(n+1) = w(n) + \eta \cdot gw \cdot xk$$

其中 n 为修正次数。

5) 根据 1) ~ 4) 的数学模型及运算过程, 用 MATLAB7.0 软件编程运算, 并且从 1) ~ 4) 反复训练, 直到 W 和 H 收敛。然后把待检测样本输入修正好的网络进行运算, 由 $ek = |yk - dk|^2 / 2$ 值的差异判断能否将样本正确分类, 其中 ek 为目标输出向量值与实际输出向量值差向量模长平方的 1/2。

2 昆虫总科阶元分类

2.1 粗糙集理论在昆虫总科阶元分类上的应用

根据粗糙集理论, 建立知识表达系统: $U = \{e_1, e_2, \dots, e_{23}\}$ 为 23 种昆虫对象集, $A = R \cup D, R = \{A, P, L_x, L_y, F, B, S, C, R, E, H\}$ 为 11 项数学形态特征属性集, 称为条件属性; 决策属性 $D = \{m_1, m_2, m_3, m_4, m_5\}$, 其中 m_1, m_2, \dots, m_5 分别为夜蛾总科、蚕蛾总科、凤蝶总科、金龟甲总科和叶甲总科。

根据表 1 中各项数学形态特征进行布尔变量, 对每个对象的每个属性赋予 1 个信息值。主要根据

数字的集中分布特征进行分类。根据 $S = (U, A, V, f)$ 知识表达系统, 利用表 1 形成信息决策表 2。

$$\begin{aligned}
 V_F &= \begin{cases} 1 & F < 1.5 \\ 2 & 1.5 < F < 2.5 \\ 3 & 2.5 < F < 3.5 \\ 4 & 3.5 < F \end{cases} &
 V_B &= \begin{cases} 1 & B < 0.15 \\ 2 & 0.15 < B < 0.25 \\ 3 & 0.25 < B < 0.35 \\ 4 & 0.35 < B \end{cases} \\
 V_S &= \begin{cases} 1 & S < 0.2 \\ 2 & 0.2 < S < 0.25 \\ 3 & 0.25 < S < 0.35 \\ 4 & 0.35 < S \end{cases} &
 V_C &= \begin{cases} 1 & C < 4.5 \\ 2 & 4.5 < C < 5.5 \\ 3 & 5.5 < C < 6.5 \\ 4 & 6.5 < C \end{cases} \\
 V_E &= \begin{cases} 1 & E < 1 \\ 2 & 1 < E < 2 \\ 3 & 2 < E < 3 \\ 4 & 3 < E \end{cases} &
 V_H &= \begin{cases} 1 & H < 1 \\ 2 & 1 < H < 2 \\ 3 & 2 < H < 6 \\ 4 & 6 < H \end{cases} \\
 V_R &= \begin{cases} 1 & R < 1 \\ 2 & 1 < R < 2 \\ 3 & 2 < R < 3 \\ 4 & 3 < R \end{cases} &
 V_A &= \begin{cases} 1 & A < 10\,000 \\ 2 & 10\,000 < A < 50\,000 \\ 3 & 50\,000 < A < 100\,000 \\ 4 & 100\,000 < A \end{cases} \\
 V_P &= \begin{cases} 1 & P < 300 \\ 2 & 300 < P < 500 \\ 3 & 500 < P < 1\,000 \\ 4 & 1\,000 < P \end{cases} &
 V_{L_Y} &= \begin{cases} 1 & L_Y < 120 \\ 2 & 120 < L_Y < 150 \\ 3 & 150 < L_Y < 200 \\ 4 & 200 < L_Y \end{cases}
 \end{aligned}$$

表 2 23 种昆虫的粗糙集知识表达系统信息决策表

Table 2 Decision of the classification to each insect species with the knowledge system of rough set analysis

对象集	条件属性											决策属性
	面积 A	周长 P	横轴长 L_x	纵轴长 L_y	形状参数 F	叶状性 B	球状性 S	圆形性 C	似圆度 R	偏心率 E	亮斑数 H	
e_1	2	3	2	1	4	1	1	3	1	1	4	m_1
e_2	2	3	2	1	4	1	1	2	1	1	4	m_1
e_3	2	3	2	1	3	1	1	3	1	1	1	m_1
e_4	3	4	3	2	3	1	1	4	1	1	3	m_2
e_5	3	4	4	2	5	1	1	4	1	1	3	m_2
e_6	4	4	4	3	4	1	1	4	1	1	4	m_2
e_7	4	4	4	3	3	1	1	4	1	1	4	m_2
e_8	2	4	2	1	3	1	1	4	1	1	4	m_3
e_9	2	4	3	1	3	1	1	4	1	1	1	m_3
e_{10}	3	4	3	1	3	1	1	4	1	1	1	m_3
e_{11}	3	4	3	2	2	1	1	4	1	1	1	m_3
e_{12}	2	4	3	1	3	1	1	4	1	1	4	m_3
e_{13}	2	4	2	4	4	3	3	3	2	2	3	m_4
e_{14}	1	2	1	1	1	4	4	3	2	2	1	m_4
e_{15}	1	1	1	1	1	4	4	3	2	2	1	m_4
e_{16}	1	1	1	1	1	4	4	2	2	2	4	m_4
e_{17}	1	1	1	1	1	4	4	2	2	2	2	m_4
e_{18}	1	2	1	1	1	4	4	3	2	2	4	m_4
e_{19}	1	2	1	3	2	3	1	2	3	3	4	m_5
e_{20}	1	1	1	1	2	4	1	1	3	3	1	m_5
e_{21}	1	1	1	1	2	4	1	1	4	4	1	m_5
e_{22}	1	1	1	1	2	4	1	1	4	4	1	m_5
e_{23}	1	1	1	1	2	4	1	1	3	3	1	m_5

1) 利用区分函数进行约简。

利用区分函数对表 2 中的数据进行约简运算。运算结果为： $= A \ P \ X_L \ F \ C \ R \ H$ 。由于 $U/R = U/E$ 与 E 是完全等价关系，所以属性 E 的作用与 R 同。这个知识表达系统只有 1 个约简 $\{A, P, L_X, F, C, R, H\}$ ，同时也是核。 L_Y, B, S 属性是不必要的，其意义就是对 U 对象按等价关系族 R 分类时， L_Y, B, S 等价关系是不必要的，按 $A \ P \ X_L \ F \ C \ R \ H$ 分类即可。即有

$$U/\{A, P, L_X, F, C, R, H\} =$$

$$U/\{A, P, L_X, L_Y, F, B, S, C, R, E, H\}$$

2) 求知识表达系统相对约简和相对核。

令 $P = \{A, P, L_X, F, C, R, H\}$ 。由于 $POS_{ind(P)}^{(ind(D))} = POS_{ind(P - \{C, R, H\})}^{(ind(D))}$ ，所以 C, R, H 属性为 D 的相对不必要属性， $\{A, P, L_X, F\}$ 为 D 的 1 个相对约简， $POS_{ind(P)}^{(ind(D))}$ 简记为 $POS_p(D)$

同样可知， $\{L_X, C, R, H\}$ 和 $\{A, C, R, H\}$ 都是 D 的 1 个相对约简， D 有很多相对约简，但每个约简都必须含有 H 或 A ，这说明 H 和 A 的重要性。

$POS_{ind(P)}^{(ind(D))} = POS_{ind(P - \{C, R, H\})}^{(ind(D))}$ 就是按照相对约简的属性知识划分对象 U 形成的子集分别属于 D 的子集。

2.2 人工神经网络在昆虫总科阶元分类上的应用

通过粗糙集理论的运算，对昆虫正确分类的 11 项特征已简化为 4 项 ($\{A, P, L_X, F\}$ 或 $\{L_X, C, R, H\}$ 或 $\{A, C, R, H\}$) 相对约简核，本研究采用 $\{A, C, R, H\}$ ，这大大减少了人工神经网络的运算，增大了正确可分性。这里分别从 5 总科中各取 2 个样本组成 10 个样本数的训练样本，输入网络进行反复训练，最后得到已收敛的权值矩阵：

$$H =$$

$$\begin{bmatrix} 1.0071 & -0.0845 & -0.0839 & -0.0841 & -0.0823 \\ -0.0853 & 1.0072 & -0.0852 & -0.0839 & -0.0837 \\ -0.0081 & 0.0068 & 1.0970 & 0.0739 & -0.0580 \\ 0.0083 & 0.0079 & -0.0084 & 1.0930 & 0.0079 \\ 0.0091 & 0.0082 & 0.0079 & 0.0067 & 1.0899 \end{bmatrix}$$

$$W =$$

$$\begin{bmatrix} 0.0013 & 0.0072 & 0.0850 & 1.9710 & 1.9974 \\ 0.0930 & 0.0014 & 1.1969 & 0.0047 & 1.9974 \\ 0.0783 & 0.0037 & 1.2010 & 0.0039 & 1.9996 \\ 0.0038 & 0.0041 & 1.1989 & 0.0950 & 1.9996 \\ 0.0071 & 0.0990 & 1.1999 & 0.0063 & 1.9999 \end{bmatrix}$$

将剩余的 13 种昆虫作为待检测样本分别输入已训练好的网络进行运算，结果见表 3。通过神经网络的运算可以看出， ek 值在 5 总科上是不同的，能够正确分类。

表 3 神经网络模型对待测样本的检验结果

Table 3 The results of neural network in testing pending samples

对象集	条件属性				ek 值	决策属性
	面积 A	圆形性 C	似圆度 R	亮斑数 H		
e_2	2	2	1	4	0.1547	m_1
e_5	3	4	1	3	0.1709	m_2
e_6	4	4	1	4	0.1712	m_2
e_9	2	4	1	1	0.1912	m_3
e_{11}	3	4	1	1	0.1908	m_3
e_{12}	2	4	1	1	0.1913	m_3
e_{14}	1	3	2	1	0.1237	m_4
e_{16}	1	2	2	4	0.1237	m_4
e_{17}	1	2	2	2	0.1228	m_4
e_{18}	1	3	2	4	0.1230	m_4
e_{19}	1	2	3	4	0.1045	m_5
e_{20}	1	1	3	1	0.1098	m_5
e_{21}	1	1	4	1	0.1093	m_5

注：目标输出向量值与实际输出向量值差向量模长平方的 $1/2$ ， $ek = |yk - dk|^2/2$ 。

3 结果与讨论

通过粗糙集理论的分析与运算，11 个属性的重要性依次为：面积、亮斑数 > 周长、横轴长、形状参数、圆形性、似圆度、偏心率 > 纵轴长、叶状性、球状性。面积、亮斑数是最重要的不可缺少的属性，纵轴长、叶状性、球状性为不必要的属性，即具有必要属性和相对属性就可以分类和划分到决策属性。

赵汗青等^[8]用统计学对夜蛾总科、蚕蛾总科、凤蝶总科、金龟甲总科、叶甲总科 5 总科的 11 项数学形态特征属性(指标)进行假设检验，从数学形态特征两两显著性的差别确定其重要性，而没有把 11 项属性综合分析。其属性特征的重要性大小为：圆形性 > 面积、周长、横轴、球状性、似圆度、偏心率 > 纵轴长、叶状性 > 形状参数、亮斑数。

本粗糙集理论的分析结果与赵汗青等的结果不

完全一致,但大多数属性特征的重要性是一致的。笔者认为,粗糙集理论分析过程较为客观,所分11项数学形态特征为:必要性指标、相对不必要指标和没必要指标,结果精细、明确。

神经网络能在粗糙集滤过简化的指标(特征属性)的基础上进行很好的模式识别,粗糙集的滤过使得神经网络运算大为简化,结果可靠。研究也表明,对昆虫进行分类时,应根据所要分类的类别确定所选分类指标,并非指标越多越好。运用粗糙集理论与人工神经网络进行分类能取得较为理想的结果,与统计学方法相比有其特有的优势,值得进一步深入研究和推广。

参 考 文 献

- [1] 郑乐怡. 系统发育与分类[M]. 郑乐怡编著. 动物分类原理与方法. 北京:高等教育出版社,1987:130-177
- [2] 王心丽. 浅谈昆虫分类学的研究方向[J]. 昆虫知识, 2001,38(6):421
- [3] 喻泓,肖曙光,李海涛. 数值分类在我国昆虫分类工作的应用[J]. 河南林业科技,2004,24(4):33-34
- [4] 赵铁桥. 数值分类的基本原理和研究程式[J]. 昆虫知识,1998,25(5):314-318
- [5] 刘清. Rough 集理论:现状与前景[J]. 计算机科学, 1997,24(4):1-5
- [6] 杜瑞卿,张征天,刘光亮,等. 粗糙集模糊聚类分析法在昆虫目阶元分类学上的应用[J]. 昆虫学报,2006,49(1):106-111
- [7] 赵汗青,沈佐锐,于新文. 数学形态学在昆虫分类学上的应用研究: . 在目级阶元上的应用研究[J]. 昆虫学报, 2003,46(1):45-50
- [8] 赵汗青,沈佐锐,于新文. 数学形态学在昆虫分类学上的应用研究: . 在总科级阶元上的应用研究[J]. 昆虫学报, 2003,46(2):201-208
- [9] 苗夺谦,胡桂荣. 知识约简的一种启发式算法[J]. 计算机研究与发展,1999,36(6):681-684
- [10] 张文修,吴伟志. 粗糙集理论与方法[M]. 北京:科学出版社,2001:22-23