

НЕИЕРАРХИЧЕСКИЙ КЛАСТЕР-АНАЛИЗ В РЕШЕНИИ ЗАДАЧ КЛАССИФИКАЦИИ ДЕТЕЙ НА ОСНОВЕ СОМАТОМЕТРИЧЕСКИХ ПРИЗНАКОВ

А. З. МАРТИРОСЯН

РИВЦ МЗ Армянской ССР, Ереван

Аннотация — Получена объективная классификация 17-летних мальчиков на основе 30 соматометрических признаков с помощью предложенной ранее неперархической модели кластер-анализа. На основе линейной разделимости полученных групп эта классификация сравнивалась с классификацией, полученной из перархической модели кластер-анализа в пространстве варимакс-факторов.

Անոտացիա — Ստացված է 17-ամյա տղաների օբյեկտիվ դասակարգումը 30 սոմատոմետրիկ նշանների հիման վրա՝ նախկինում առաջարկված կլաստեր-անալիզի ոչ հիերարխիական մոդելի միջոցով: Ստացված խմբերի զեմային բաժանելիության հիման վրա այդ դասակարգումը համեմատված է վարիմակս-ֆակտորների տարածության մեջ կլաստեր-անալիզի հիերարխիական մոդելից ստացված դասակարգման հետ:

Abstract—The objective classification of 17-year-old boys on the basis of 30 somatometric signs with the help of the earlier proposed non-hierarchical model of cluster-analysis has been obtained. On the basis of linear divisibility of the obtained groups this classification has been compared with those, obtained from hierarchic model of cluster-analysis in the space of varimax-factors.

Ключевые слова: кластер-анализ, варимакс-фактор, морфологический статус, дискриминальный анализ.

Последние годы характеризуются все более широким проникновением математических методов в биомедицинские исследования, причем антропологические данные в этих работах всегда занимали одно из ведущих мест.

Одной из интересных задач, решаемых с привлечением математических методов, является объективная классификация детей согласно морфологическому статусу. Наиболее приемлемой математической моделью для решения классификационных задач, как известно [5], является кластер-анализ. Необходимо отметить, однако, что несмотря на бурное развитие прикладных работ по кластеру-анализу [2, 6], применяемые модели носят в основном иерархический характер со всеми вытекающими из их природы ограничениями [3, 4].

В настоящей работе для объективной классификации на основе антропометрических признаков используется один из ранее предложенных неиерархических алгоритмов, преимущества которого были проиллюстрированы на искусственно сгенерированных данных [3].

Для апробации предложенной модели проведена классификация 100 мальчиков 17-летнего возраста, учащихся средних школ г. Еревана. Каждый объект представлен точкой в 30-мерном пространстве антропометрических признаков, в основном описы-

вающих его морфологический статус [1]. Количественная оценка работы алгоритма осуществлена на основе достигнутого уровня линейной разделимости классов с помощью результатов пошагового дискриминантного анализа.

В результате работы алгоритма выделено 7 четко различающихся групп численностью 30, 14, 13, 7, 7, 7 и 5 человек, которые обозначены буквами А, В, С, D, E, F, G.

Пошаговый алгоритм отбора наиболее информативных признаков позволил построить линейные дискриминантные функции (ЛДФ) для реклассификации и прогноза по методу скользящего контроля (табл. 1 и 2). Как следует из табл. 1, информативными оказались 9 признаков,

Таблица 1. Линейные дискриминантные функции

Группы	А	В	С	D	Е	Г	Г
№ параметр							
9	3.67	3.52	3.54	3.65	3.55	3.80	3.77
11	-0.29	-0.29	-0.31	-0.37	-0.18	-0.35	-0.29
17	5.73	6.00	5.56	5.81	5.41	5.74	6.12
18	-1.60	-1.55	-1.49	-1.58	-1.46	-1.65	-1.76
25	2.29	2.40	2.16	2.37	2.06	2.28	2.57
28	0.07	0.10	0.03	0.18	0.11	0.07	0.20
29	3.03	3.13	2.89	2.93	2.84	2.98	3.14
33	-0.76	-0.91	-0.53	-0.63	-0.14	-0.65	-0.94
СО	-4788.18	-5205.97	-4445.54	-4853.60	-4461.78	-4952.78	-5315.12

Примечание: 9—рост, 11—длина туловища, 17—длина кисти, 18—плечевой диаметр, 25—обхват грудной клетки, 28—обхват бедра, 29—обхват голени, 33—минимальный лобный диаметр, СО—константа.

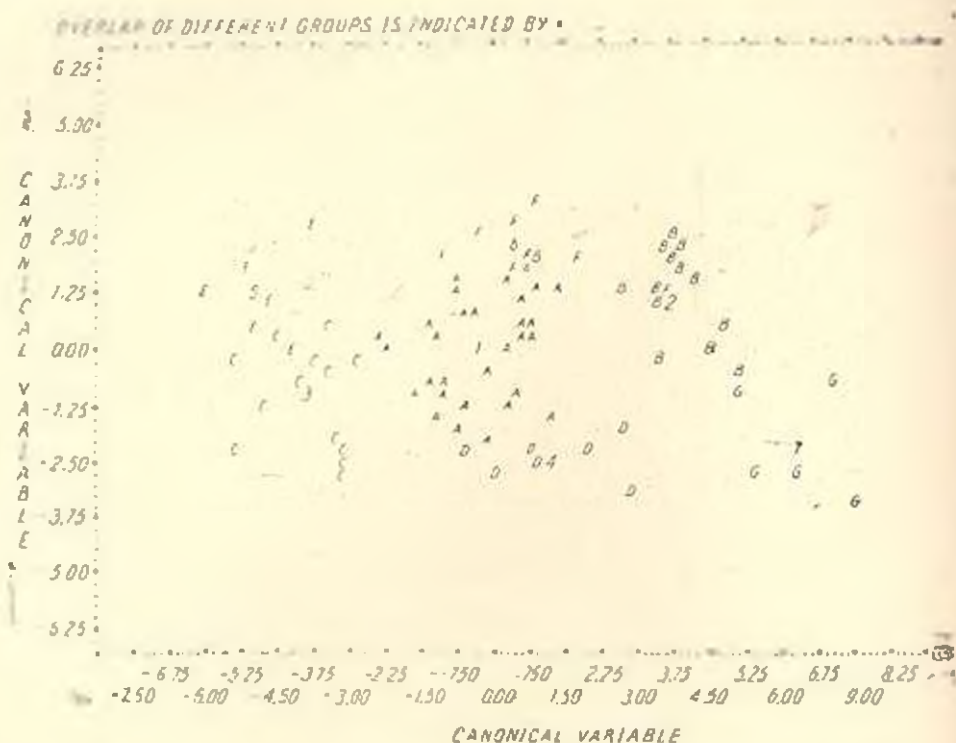
Таблица 2. Реклассификация и прогноз

Группы	Реклассификация								Прогноз							
	%	А	В	С	D	Е	Г		%	А	В	С	D	Е	Г	
А	93.3	28					2	93.0	28						2	
В	92.9		13				1	92.9		13					1	
С	100.0			13				100.0				13				
D	85.7	1			6			85.7	1				6			
Е	100.0					7		71.4				2		5		
Г	100.0						7	85.7	1						6	
Г	100.0							80.0			1					4
Средняя	95.2	29	13	13	6	7	10	5	90.4	30	14	15	6	5	9	4

описывающих основные длиннотные, широтные и обхватные размеры как головы, так и туловища и конечностей. Данные табл. 2 свидетельствуют о высокой описательной и прогностической мощности полученных ЛДФ: в среднем—95,2% реклассификации и 90,4% прогноза (результат, достаточно редко достигаемый в биомедицинских исследованиях). Графическая иллюстрация разделения групп представлена проекцией на плоскость двух первых канонических переменных (рис.).

Предметная интерпретация выделенных групп осуществлена на ос-

нове рисунка и анализа основных одномерных характеристик групп. Среди выявленных 7 групп можно выделить 3 класса детей, характеризующихся определенным уровнем соматического развития. Группы А, D и F представляют собой объекты, характеризующие так называемую норму, группы С и Е—микросоматики, В и G—макросоматики.



Проекция анализируемого множества на плоскость Фишера.

Интересно сравнить полученные результаты с данными иерархического среднесвязывающего кластер-анализа, проведенного на том же материале [1] и показавшего, что стандартный иерархический алгоритм не смог выявить статистически значимую классификацию в анализируемом 26-мерном пространстве соматометрических признаков (4 дополнительных параметра, использованных в нашей работе, имеют достаточно высокую корреляцию с рядом характеристик из 26 показателей). Даже применение этой модели в пространстве варимакс-факторов не привело к получению столь значимых результатов классификации и прогноза, какие достигнуты в настоящем исследовании. Автором достигнут уровень, равный приблизительно 80% (реклассификации и прогноза) [1].

Таким образом, полученные результаты позволяют утверждать, что предлагаемый неиерархический подход кластер-анализа предпочтителен, по меньшей мере, при решении задачи классификации объектов с использованием антропометрических признаков. Вместе с тем этот результат свидетельствует о применимости разработанных подходов не только для случаев ортогональных параметров (что показано

ранее [3]), но и об их адекватности для решения широкого ряда классификационных задач биомедицинских исследований.

ЛИТЕРАТУРА

1. Епископосян Л. М. Биолог. ж. Армения, 39, 9, 1986.
2. Ордуханян А. А., Саркисян А. С., Ландау М. А., Мнацкян Ш. Г., Тер-Закарян Ю. Э. Хим. фарм. ж., 1, 63—66, 1983.
3. Ордуханян А. А., Мартirosян А. Э. Математические методы распознавания образов, 139—141, Дилижан, 1985.
4. Соломон Г. В. сб.: Классификация и кластер, 129—147, М., 1980.
5. Стюпер Э., Брюсер Ч., Джурс П. Машинный анализ связи химической структуры и биологической активности, 67, М., 1982.
6. Gabrielyan E. S., Amroyal E. A., Nalbandyan S. G., Ordubekyan A. A. In Application of New Electronic Instruments in Clinical and Experimental Pharmacology, Jhringen, 102—115, 19-4.

Поступило 15.IV 1986 г

Биолог, ж. Армения, т. 39, № 9, с. 747—752, 1986.

Y2K 612 8:61.007

ОРГАНИЗАЦИЯ ЦИФРОВОГО СПЕКТРАЛЬНОГО АНАЛИЗА В СИСТЕМАХ АВТОМАТИЗАЦИИ ФИЗИОЛОГИЧЕСКОГО ЭКСПЕРИМЕНТА

Т. Г. АРЕШЯН, В. Д. БАРСЕГЯН, Е. Б. ДУНАЛВИЦЕР, Л. С. ЧОБАНИАН

Институт физиологии им. Л. А. Орбели АН Армянской ССР, лаборатория математического моделирования нейронных систем (ММНС), Ереван

Аниотация — Представлен комплекс программных модулей, с помощью которого можно производить обработку широкого круга вызванных биоэлектрических реакций. Приведены блок-схема комплексных программных модулей, а также примеры обработки вызванных реакций нервной системы, иллюстрирующие возможности комплекса.

Անուագրւած — Աւարազգիքն է Ժրագոռայն ճագօշներն կոմայելա: որր Շահապօ-
րաթլուսն է տալիս Ժրակէ լայն քրտնք Հարազգիքն քննաւթիկարական քնակցիա-
ները: ներկայացված է Ժրագոռայն Ժագոռակներ Կ-ժագոռակի քրտնքիման. ինչպիս
նակ Աւարազգիքն Դժմաճագոռի Հարազգիքն քննաւթիկարակն Ժրակէն քրտնքիմանք:

որոնք քրտնքիման Դն Կոմայելա Շահապօրաթլուսները:

Abstract — The complex of program modules has been presented by means of which it is possible to elaborate the evoked bioelectric reactions on a large scale. The block-scheme of complex of the program modules has been given, as well as the examples of elaboration of evoked reactions of the nervous system, illustrating the possibilities of the complex.

Ключевые слова: измерительно-вычислительный комплекс, прогрессивный модуль, частотная характеристика, биоэлектрический сигнал.

В работе рассматриваются вопросы построения универсального программного обеспечения измерительно-вычислительного комплекса (ИВК), состоящего из специализированной (многоканальный анали-