А. Н. ДМИТРИЕВ, Ю. И. ЖУРАВЛЕВ, Ф. П. КРЕНДЕЛЕВ
(Новосибирск)

ЛОГИЧЕСКИЙ СПОСОБ ПОСТРОЕНИЯ МНОГОМЕРНЫХ КЛАССИФИКАЦИЙ В ГЕОЛОГИИ

В геологии существует класс задач, в которых по неполному набору признаков, характеризующих объекты данного множества, требуется упорядочить элементы множества по возрастанию или убыванию какой-либо величины (запасы данного месторождения, стоимость алмаза и т. д.).

Пусть задано конечное множество $M = \{\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_k\}$, на элементах которого определены числовые функции $x_1(\alpha), \ldots, x_n(\alpha), x_{n+1}(\alpha)$. Функцию $x_{n+1}(\alpha)$ будем называть целевой и упорядочение объектов производить по возрастанию или убыванию этой величины.

Рассмотрим описание объектов множества M в виде таблицы

	x_1	x_2	x_3	...	x_n	x_{n+1}
$\tilde{\alpha}_1$	x_{11}	x_{12}	x_{13}	...	x_{1n}	x_{1n+1}
$\tilde{\alpha}_2$	x_{21}	x_{22}	x_{23}	...	x_{2n}	x_{2n+1}
.	. . .	. . .	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .	. . .	. . .
	. . .	. . .	. . .	. . .	. . .	. . .
$\tilde{\alpha}_r$	x_{r1}	x_{r2}	x_{r3}	...	x_{rn}	x_{rn+1}

Здесь x_{ij} — значение функции $x_{ij}(\alpha)$ на объекте α_i. Некоторые x_{ij} могут быть неизвестны, тогда на соответствующем месте в таблице ставится прочерк.

Если $n=1$, можно построить линейную классификацию. Так, если элементы $\tilde{\alpha}$ есть алмазы и $x(\tilde{\alpha})$ — вес алмаза в каратах, то нетрудно построить классификацию по стоимости s, так как $s = x_{n+1} = \varphi(x(\tilde{\alpha}))$, и значение φ легко вычислить.

Существуют примеры классификаций по двум, трем и т. д. признакам.

Оказывается, если каждая из функций $x_1, x_2, \ldots, x_n$ принимает лишь конечное число значений, а значения числовой характеристики $x_{n+1}(\tilde{\alpha})$ для многих $\tilde{\alpha}$ неизвестны, то величина $x_{n+1}(\tilde{\alpha})$ может быть экстраполирована следующим образом.

Рассмотрим для простоты случай, когда функции x_i, $i=1, 2, \ldots, n$ принимают только два значения: 1 — «да», 0 — «нет» (и прочерк — «не знаю»), т. е. $x_1, x_2, \ldots, x_n$ являются предикатами.

Выделим все такие признаки x_i, у которых $x_i(\alpha) \equiv \mathrm{const}$ на всех элементах множества M, для которых известна величина x_{n+1}. Такие признаки назовем отождествляющими и удалим из таблицы. В таблице оставим лишь объекты, для которых известна x_{n+1}.

Рассмотрим все тупиковые тесты[1] оставшейся таблицы T. Пусть k — число всех тупиковых тестов, k_i — число тестов, в которые вошел столбец x_i.

Назовем $p_i = \frac{k_i}{k}$ весом предиката x_i. Пусть задан объект β, для которого неизвестно значение x_{n+1}, и все отождест-

[1] Определение тупикового теста см. в работе И. А. Ченгиси, С. В. Яблонского «Логические способы контроля электрических схем». Труды Мат. ин-та им. В. А. Стеклова, т. 51, 1958.

вляющие признаки принимают те же значения, что и на элементах $\tilde{\alpha}$, для которых известно значение x_{n+1}. Тогда величина

$$I(\beta) = \underset{\substack{\text{по всем при-}\\ \text{знакам } i\text{, вхо-}\\ \text{дящим в опи-}\\ \text{сание } \beta \text{ со зна-}\\ \text{чением 1}}}{\sum p(i)} + \frac{1}{2} \underset{\substack{\text{по всем при-}\\ \text{знакам } i\text{, вхо-}\\ \text{дящим в опи-}\\ \text{сание } \beta \text{ со зна-}\\ \text{чением «—»}}}{\sum p(i)}$$

в ряде задач с точностью до мультипликативной константы, одинаковой для всех β, дает значение $x_{n+1}(\beta)$.