ВОПРОСЫ ОБРАБОТКИ ИНФОРМАЦИИ ПРИ ПРОЕКТИРОВАНИИ СИСТЕМ
(Вычислительные системы)
1977 год Выпуск 69

УДК 51:155.001.57

О МЕТОДЕ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ РЕШАЮЩИХ ПРАВИЛ

Б.Я.Ковалерчук

В работе предлагается одно из возможных уточнений понятия согласованности решающих правил и исходных данных в распознавании образов. Приводится теорема о том,что для некоторых известных решающих правил существуют исходные данные,не согласованные в этом смысле. Предлагается метод согласования. Для некоторых решающих правил и исходных данных экспериментально исследуется возможность использования метода как средства повышения эффективности (уменьшения числа ошибок) распознавания. Развиваемый метод – м е т о д с о п р я ж е н и я – применим как к отдельным решающим правилам, так и к их совокупностям. Данная работа может рассматриваться как развитие и конкретизация работ [1,2] .

I. В настоящее время существует несколько подходов к решению проблемы повышения эффективности решающих правил. Согласно одному подходу решающие правила некоторым образом объединяются (в комитеты функционалов [3], коллективы решающих правил [4]) так, чтобы эффективность совокупности была выше,чем эффективность каждого из решающих правил в отдельности. Сюда же можно отнести работы,в которых предлагается отдавать предпочтение более простому решению задачи с помощью решающего правила [5].

В ряде работ [6,7], исследующих вопросы принятия решения по векторным критериям, а также группового выбора, разрабатываются методы повышения эффективности решающих правил.

В основу многих других подходов к повышению эффективности положены преобразование и сокращение признакового пространства.Предлагаемый метод сопряжения может использоваться как наряду с существующими, так и в случаях, когда какие-либо из них не применимы. Например, для применения методов, излагаемых в некоторых из упомянутых работ, требуется фиксация не менее двух решающих правил. Эти методы в отличие от метода сопряжения не применимы, когда зафиксировано только одно решающее правило.

II. Для простоты изложения ограничимся рассмотрением задачи распознавания только с бинарными признаками и двумя образами. В основу метода сопряжения положен эвристический п р и н ц и п о т н о с и т е л ь н о й у с т о й ч и в о с т и р е ш е н и й, согласно которому предлагается отдавать предпочтение тому решению, которое наиболее устойчиво при переходах к с о п р я ж е н н ы м т а б л и ц а м исходных данных.

Поясним понятие сопряженной таблицы и правила принятия решений на ее основе. В дальнейшем эти понятия будут сформулированы точно. Пусть таблица $\mathcal{T}$ (таблица материала обучения и контроля) – исходных данных имеет вид:

Признаки

Объекты	P_1	P_2	P_3	S
α	1	1	1	1
β	0	0	0	0
c	0	0	1	

Пусть также фиксировано некоторое решающее правило R , с помощью которого требуется принять решение о значении S(c) (незаполненного элемента таблицы $\mathcal{T}$). Признак S будем называть целевым и будем считать, что S(c) = 1 и S(c) = 0 означают соответственно отнесение объекта c к первому или второму образам.

Тогда примером сопряженной таблицы T* (формально таблица T* может быть задана парой $\langle T_1, P_1(c) \rangle$) будет таблица T_1 с двумя выделенными элементами S(c) и $P_1(c)$

Признаки

Объекты	P_1	P_2	P_3	S
α	1	1	1	1
β	0	0	0	0
c	0	0	1	1

В T_1 принята гипотеза, что S(c) = 1 . Остальные элементы T_1 совпадают с элементами $\mathcal{T}$. Процедура принятия решений о значении $S(c)$ на основе сопряженной таблицы T_1 состоит в следующем. Закрывается фактическое значение $P_1(c)$, и предсказывается решающим правилом R значение $P_1(c)$ (в предположении, что $S(c) = 1$, которое сделано в T_1). Если предсказанное значение совпадает с фактическим значением $P_1(c)$, то считается, что предположение $S(c) = 1$ "подтверждено", "согласовано" и принимается решение $S(c) = 1$.

Взяв пару $\langle T_1, S(a) \rangle$ с тем же гипотетическим значением $S(c) = 1$, получим другой пример сопряженной таблицы и процедуры принятия решений на ее основе. Таким варьированием выделенных элементов может быть получен целый ряд сопряженных таблиц и процедур принятия решений на их основе. Каждая из этих процедур в итоге предсказывает значение целевого признака, и в этом смысле его выделенность по отношению к другим признакам сохраняется.

Пусть $\{f_i\}$ – множество функций, сопоставляющих решающему правилу R множество указанных процедур. Будем называть эти функции с п о с о б а м и п р и м е н е н и я р е ш а ю щ е г о п р а в и л а R. Тогда вместо того, чтобы говорить об этих процедурах, можно говорить о способах применения решающего правила R.

В реальных задачах может быть некоторая априорная информация или информация, почерпнутая из материала обучения (в дальнейшем этот вопрос будет рассмотрен подробно), которая ограничивает число допустимых сопряженных таблиц. В рамках этих ограничений для распознавания может быть выбрана л ю б а я допустимая сопряженная таблица или сама таблица $\mathcal{T}$. Этот произвол выбора не играл бы никакой роли, если бы все таблицы давали один и тот же результат распознавания значения $S(c)$. В дальнейшем будет показано (см. теорему стр.136), что существуют такие решающие правила в классе известных в распознавании решающих правил, для которых этот результат меняется. Иначе говоря, существуют решающие правила, для которых способ их применения влияет на результаты распознавания. В этом смысле для них ставится задача согласования как задача обеспечения единственности решения. Смысл принципа относительной устойчивости решения состоит в том, чтобы выбирать наиболее устойчивое решение к "субъективным шумам", вносимым произволом в способе применения решающего правила. При такой трактовке предлагаемый принцип близок к идее "грубого подхода" [1,2].

II. Оставим в силе все сделанные выше предположения. Пусть $\mathcal{T}$ – бинарная таблица $m \times (n+1)$, m – число строк, $n+1$ – число столбцов. Строки интерпретируются как объекты, столбцы – как признаки. Признаки будем обозначать через $P_1, P_2, \ldots, P_n$, S, а объект контроля – через c. Пусть, для определенности, здесь и в дальнейшем $(n+1)$-й признак – это признак S, который будем называть целевым. Также, для определенности, положим здесь и в дальнейшем, что m-й объект – объект контроля c. В таблице $\mathcal{T}$ отсутствует значение $S(c)$, которое требуется предсказать. Будем, как и ранее, считать, что $S(c) = 1$ и $S(c) = 0$ означают соответственно отнесение к первому или второму образам. Будем называть таблицу $\mathcal{T}$ т а б л и ц е й и с х о д н ы х данных, и эту же таблицу $\mathcal{T}$ без m-й строки, соответствующей объекту контроля c, будем называть т а б л и ц е й м а т е р и а л а о б у ч е н и я.

Построим с помощью $\mathcal{T}$ две таблицы: T_1 и T_2. Пусть в таблице T_1 значение $S(c) = 1$, а в T_2 значение $S(c) = 0$, в остальном обе таблицы совпадают с $\mathcal{T}$, т.е. в T_1 сделано предположение, что контрольный объект принадлежит первому образу, а в T_2 – второму образу (см.рис.I). Таблицы T_1 и T_2 будем называть г и п о т е т и ч е с к и м и.

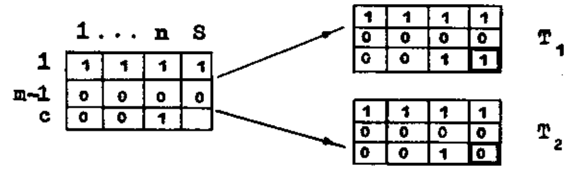


Рис. I

Тройку $\mu = \langle \check{\mu}, i, j \rangle$ будем называть э л е м е н т о м таблицы, если $\check{\mu}$ – значение, стоящее в клетке таблицы с координатами (i, j), где i – номер строки, j – номер столбца. Значение $S(c)$ в T_1 и T_2 будем обозначать через $S^1(c)$, $S^2(c)$ и называть г и п о т е т и ч е с к и м и з н а ч е н и я м и. Тройки, соответствующие гипотетическим значениям $S(c)$, $\langle 1, m, n+1 \rangle$ и $\langle 0, m, n+1 \rangle$, будем называть г и п о т е т и ч е с к и м и э л е м е н т а м и таблиц T_1 и T_2 и обозначать через γ_1 и γ_2.

С о п р я ж е н н о й т а б л и ц е й будем называть каждую таблицу T_i с двумя выделенными элементами μ и γ_i $(i = 1, 2)$ такими, что $\mu \neq \gamma_i$. Элемент μ будем называть с в е р я е м ы м э л е м е н т о м сопряженной таблицы.

Пусть $T_i \setminus \mu$ – бинарная таблица T_i без элемента $\mu = \langle \check{\mu}, 1, j \rangle$. Функцию R, определенную на таблицах вида $T_i \setminus \mu$ со значениями в $\{0, 1, z\}$, будем называть *решающим правилом.* ж) Значения R интерпретируются либо как отсутствие (0) или наличие (I) j-го признака у l-го объекта, либо как отказ принятия решения (z). На содержательном уровне эти понятия были пояснены выше. Отметим, что при $\mu = \langle 1, m, n+1 \rangle$ имеет место $T_i \setminus \mu = \mathcal{T}$, т.е. $\mathcal{T}$ принадлежит области определения R.

Напомним, что $\check{\mu} \in \{0, 1\}$, так как $\mathcal{T}$ – бинарная таблица.

Определим функцию $\mathcal{R}$ такую, что

$$\mathcal{R}(T_i, \mu) = \begin{cases} \text{И, если } R(T_i \setminus \mu) = \check{\mu}, \\ \text{Z, если } R(T_i \setminus \mu) = z, \\ \text{Л, если } R(T_i \setminus \mu) \neq \check{\mu} \text{ и } R(T_i \setminus \mu) \neq z. \end{cases}$$

Поясним смысл функции $\mathcal{R}$. Пусть, для определенности, $T_i = T_1$ и $\mathcal{R}(T_1, \mu) =$ И (истина). Это означает, что принятая в T_1 гипотеза $S(c) = 1$ *согласуется* со значением $\check{\mu}$, т.е. в предположении $S(c) = 1$ с помощью R предсказывается значение j-го признака l-го объекта, совпадающее с его фактическим значением.

Таким образом, значения $\mathcal{R}$ означают соответственно: согласованность по R значений гипотетического и сверяемого элементов – "истина" (И), несогласованность – "ложь" (Л) и отказ от принятия решения по поводу согласованности – Z.

Каждую пару $\langle T_i, \mu \rangle$ такую, что $\gamma_i \neq \mu$ и $\mu \in T_i$, будем называть *сопряженной таблицей* и обозначать через $T^*_{i\mu}$. Здесь выделены элементы γ_i и μ. С помощью $\mathcal{R}$ построим множество функций $R^*_{i\mu}$ следующим образом:

$$S(c) = R^*_{i\mu}(\mathcal{T}) = \begin{cases} S^i(c), \text{ если } \mathcal{R}(T_i, \mu) = \text{И}, \\ z, \text{ если } \mathcal{R}(T_i, \mu) = \text{Z}, \\ \overline{S^i(c)}, \text{ если } \mathcal{R}(T_i, \mu) = \text{Л}, \end{cases}$$

$$i = 1, 2.$$

Через $\overline{S^i(c)}$ обозначено значение $S(c)$, отличное от $S^i(c)$ и z. Смысл каждого из $R^*_{i\mu}$ состоит в том, что если значения гипотетического и сверяемого элементов согласуются по R, то $R^*_{i\mu}$ принимает

ж) Решающие правила рассматриваются нерасчлененными на этапы, в частности этап обучения не выделяется, поскольку это не требуется для данной работы. Синонимами понятия "решающего правила" для нас являются "метод распознавания" и "алгоритм распознавания". Примерами решающих правил являются метод потенциальных функций, алгоритмы вычисления оценок и т.д.

гипотетическое значение $S(c)$ в качестве решения. Итак, $R^*_{i\mu}$ – это упомянутое выше множество процедур принятия решений на основании сопряженных таблиц.

IУ. *Сопряженную таблицу* $T^*_{i\mu}$ будем называть *согласованной с* $R(\mathcal{T})$, если $S^i(c) = R(\mathcal{T})$.

Решающее правило R будем называть *β-согласованным*, если $\forall \mathcal{T}_\mu\ (\mu \in \Pi_i) \mathcal{R}(T_i, \mu) =$ И, где T_i такое, что $R(\mathcal{T}) = S^i(c)$.

Пусть Π_i – множество всех элементов таблицы T_i. Если решающее правило β-согласованно, то это означает, что для любой таблицы исходных данных произвол в выборе для распознавания как самой таблицы $\mathcal{T}$, так и любой согласованной сопряженной таблицы не играет никакой роли, поскольку результат распознавания значения $S(c)$ один и тот же. Решающее правило R будем называть *строго β-несогласованным* с $\mathcal{T}$, если $\exists T_i (R(\mathcal{T}) = S^i(c)) \exists \mu (\gamma_i \neq \mu) \mathcal{R}(T_i, \mu) =$ Л.

Отметим, что просто отрицание β-согласованности означает лишь $(\mathcal{R}(T_i \setminus \mu) = \text{Л}) \vee (\mathcal{R}(T_i \setminus \mu) = z)$.

Рассмотрим в качестве *$\mathcal{L}$-класса решающих правил* совокупность решающих правил, основанных на вычислении оценок [8,9], зависящих от трех параметров δ_1, δ_2 и k следующим образом:

$$R(\mathcal{T}) = S(c) = \begin{cases} 1, \text{ если } \Gamma_1 > \Gamma_2 + \delta_1 \text{ и } \dfrac{\Gamma_1}{\Gamma_1 + \Gamma_2} > \delta_2, \\ 0, \text{ если } \Gamma_2 > \Gamma_1 + \delta_1 \text{ и } \dfrac{\Gamma_2}{\Gamma_2 - \Gamma_1} > \delta_2, \\ z \text{ – в противном случае,} \end{cases}$$

где $\Gamma_1 = \sum_{i=1}^{m_1} C^k_{n-\rho(\alpha_i, c)}$; $\Gamma_2 = \sum_{i=1}^{m_2} C^k_{n-\rho(\beta_i, c)}$; $\delta_1 \geq 0$; $0 \leq \delta_2 < 1$; $n \geq k \geq 1$.

Здесь, как и ранее, n – число признаков (без целевого признака); m_1, m_2 – соответственно число объектов обучения первого и второго образов; α_i, β_i – объекты обучения соответствующих образов и c – объект контроля, рассматриваемые как n-мерные булевы векторы; $\rho(\alpha_i, c), \rho(\beta_i, c)$ – хэммингова расстояния.

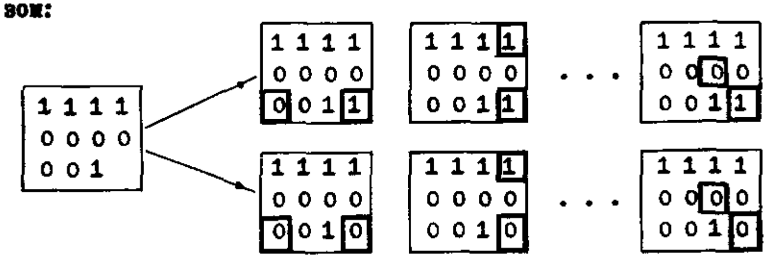
Таблицу $\mathcal{T}$ будем называть *нетривиальной*, если объект c не равен ни одному из объектов обучения α_i, β_i и $\mathcal{T}$ не содержит ни одного столбца со всеми равными значениями элементов столбца. Обозначим нетривиальную таблицу с числом строк, равным m, и с числом столбцов, равным $n+1$, через $\mathcal{T}_{m,n+1}$.

Имеет место следующая

ТЕОРЕМА [13]. **Для каждого R из $\mathcal{L}$-класса и каждого $m (m \geq 3)$ существует $n_0 (n_0 \geq 2)$ такое, что для каждого $n (n \geq n_0)$ существует нетривиальная таблица $\mathcal{T}_{m,n+1}$, строго β-несогласованная с R.**

V. Перейдем к изложению **метода сопряжения** (метода синтеза решающих правил из известных решающих правил на основе упомянутого выше принципа относительной устойчивости решений).

Первый этап – порождение всех сопряженных таблиц для таблицы исходных данных $\mathcal{T}$ – будем изображать следующим образом:



Второй этап – выделение всех допустимых сопряженных таблиц – равносилен фиксации некоторого множества сверяемых элементов Π. Ограничения на множество допустимых сверяемых элементов можно почерпнуть из материала обучения следующим образом. Из таблицы $\mathcal{T}$ исключается m-я строка, соответствующая объекту контроля c. В оставшейся таблице обучения $\mathcal{T}'$ закрывается некоторый элемент таблицы и предсказывается решающим правилом R значение этого элемента. Если оно совпадает с фактическим значением, элемент отмечается. Считается, что этот элемент "закономерно связан" со всей остальной информацией в таблице обучения. Такая процедура проводится для всех элементов таблицы обучения. В результате будут отобраны элементы, на которых успешно проводилось распознавание в указанном выше смысле. Обозначим это множество через Π'. Затем вновь рассматривается таблица $\mathcal{T}$, из нее исключается столбец, соответствующий целевому $(n+1)$-у признаку S. В оставшейся таблице $\mathcal{T}''$ проводится процедура, которая была описана для $\mathcal{T}'$. В результате здесь также будет выделена некоторая совокупность элементов таблицы, "закономерно связанных" в ней с информацией. Обозначим ее через Π''. Рассмотрим также множество

$$\Pi_S' = \{\mu = \langle\check{\mu}, i, j\rangle : j = n+1, \mu\in\Pi'\},\quad \Pi_c'' = \{\mu = \langle\check{\mu}, i, j\rangle : i = m, \mu\in\Pi''\}.$$

Множество Π_S' состоит из элементов $(n+1)$-го столбца, принадлежащих Π', а множество Π_c'' состоит из элементов m-й строки, принадлежащих Π''. Выделение Π_S' и Π_c'' нужно для того, чтобы сформировать множество элементов, "закономерно связанных" с обеими таблицами. В качестве такого множества берется множество $\Pi_S' \cup \Pi_c'' \cup (\Pi' \cap \Pi'')$, т.е. пересечение всех отмеченных элементов таблиц $\mathcal{T}'$ и $\mathcal{T}''$, к которому добавлены отмеченные элементы, принадлежащие либо целевому столбцу, либо контрольной строке. Эти элементы оставим в качестве допустимых сверяемых элементов, отобранных по материалу обучения.

Кроме описанных выше ограничений, на класс допустимых сопряженных таблиц и сверяемых элементов могут накладываться априорные ограничения.

Третий этап – оценка степени согласованности решений. Пусть

$$M_1^+ = |\{\langle T_1, \mu_i\rangle : \mathcal{R}(T_1, \mu_i) = И, \mu_i \in \Pi\}|, \tag{1}$$

$$M_1^- = |\{\langle T_1, \mu_i\rangle : \mathcal{R}(T_1, \mu_i) = Л, \mu_i \in \Pi\}|, \tag{2}$$

$$M_2^+ = |\{\langle T_2, \mu_i\rangle : \mathcal{R}(T_2, \mu_i) = И, \mu_i \in \Pi\}|, \tag{3}$$

$$M_2^- = |\{\langle T_2, \mu_i\rangle : \mathcal{R}(T_2, \mu_i) = Л, \mu_i \in \Pi\}|, \tag{4}$$

т.е. $M_1^+, M_1^-, M_2^+, M_2^-$ равны мощностям соответствующих множеств. Величины (1), (2) характеризуют степени согласованности (несогласованности) гипотезы $S(c) = 1$, а (3) и (4) – степень согласованности (несогласованности) гипотезы $S(c) = 0$. В приведенных терминах принцип относительной устойчивости решений для одного решающего правила R может быть формализован в виде **критерия** предпочтения решений K,

$$S(c) = K(R, \mathcal{T}) = \begin{cases} 1, \text{ если } M_1^+ - M_1^- > M_2^+ - M_2^-, \\ 0, \text{ если } M_1^+ - M_1^- < M_2^+ + M_2^-, \\ z \text{ (отказ), если } M_1^+ - M_1^- = M_2^+ - M_2^-. \end{cases}$$

Таким образом, K для каждого фиксированного R реализует новое **решающее правило** $[k] = K(R)$.

Критерий можно проиллюстрировать на следующем примере. Пусть $M_1^+ = 10$, $M_1^- = 3$, $M_2^- = 4$, $M_2^+ = 6$. Согласно К,за принятие $S(c)=1$ – 7 "голосов" и за принятие $S(c) = 0$ – 2 "голоса".

Таким образом,будет принято $S(c) = 1$ и отвергнуто $S(c)=0$. На этом заканчивается описание метода сопряжения.

УI. Метод сопряжения программно реализован для $\mathcal{L}$-класса при $\delta_1 = 0$ и случая, когда в качестве претендентов на свеpяемые элементы рассматриваются $P_1(c), P_2(c), \ldots, P_n(c)$,т.е. элементы $\mathcal{T}$, соответствующие объекту контроля c. АЛГОЛ-программа включена в качестве блока в программный распознающий комплекс ПРАСК-I [10]. Для того чтобы иметь интегральные характеристики оценки результатов применения критерия и решающих правил, мы воспользовались следующими формулами: $D = N_1 + \frac{N_2}{2}$ и $t = N_1 - N_3$, где N_1 – процент верного распознавания на материале контроля, N_2 – процент отказов и N_3 – процент неверного распознавания также на материале контроля. Формула D согласуется с интуицией при крайних значениях N_1 и N_2 (а) $N_1 = 100\%$, $N_2 = 0$; б) $N_1 = N_2 = 0$). Для фиксированного числа образов D является модификацией формулы η , предложенной в [11, с.59]. При двух образах $D = 100-0{,}02\cdot\eta$.

Пусть D_i и t_i – величины D и t,вычисленные для решений,полученных с помощью решающего правила R_i . И пусть $D_{крит.}$ и $t_{крит.}$ – величины D и t ,вычисленные для решений, полученных с помощью критерия К.В качестве характеристик эффективности критерия по сравнению с решающими правилами,к которым он применялся,будем использовать величины $d = D_{крит.} - D_i$, $\tau = t_{крит.} - t_i$.

Рассмотрим результаты машинных экспериментов.

А. Прогнозирование точек заложения скважин. Требовалось выделить перспективные точки заложения скважин для обнаружения полезных ископаемых в одном из рудных полей*).

Нами рассматривались следующие правила из класса алгоритмов вычисления оценок:

R_1 с параметрами $k = 8$ и $\delta_2 = 0{,}501$;
R_2 с параметрами $k = 1$ и $\delta_2 = 0{,}57$;
R_3 с параметрами $k = 8$ и $\delta_2 = 0{,}55$.

Невысокий процент верного распознавания,полученный самими решающими правилами, к которым применялся критерий, в значительной степени объясняется тем,что с геологической точки зрения целью работы с данным материалом было выяснение возможности использования для дальнейшего прогнозирования рассмотренных признаков и эталонов. Результат эксперимента приведен в табл. I.

*) Исходные данные представлены Т.М.Марипловым.

Таблица I

Характеристики		Решающие правила					
		R_1	$[R_1]$	R_2	$[R_2]$	R_3	$[R_3]$
верное распознавание, %		65	67,5	0	22,5	55	65
неверное распознавание, %		35	22,5	0	12,5	27,5	32,5
отказы, %		0	10	100	65	17,5	2,5
D		65	72,5	50	65	63,75	66,25
t%		30	7,5	0	10	27,5	32,5
Повышение эффективности	τ,%		15		10		5
	d		7,5		15		2,5

Как видно из таблицы,среднее повышение эффективности равно 10%.

Б. Моделирование процессов принятия решений должностными лицами по претензионным и исковым производствам. При решении этой задачи преследовались следующие цели: подтвердить еще на одной реальной задаче работоспособность критерия и показать возможность математического моделирования упомянутых процессов принятия решений.

Будем считать цель достигнутой, если удастся: а) сформулировать задачу принятия решений должностными лицами по претензионным и исковым производствам как задачу распознавания образов; б) провести контрольные эксперименты на ЭВМ, подтверждающие высокую степень согласованности решений, предлагаемых на основе решающих правил из класса алгоритмов вычисления оценок и критерия с решениями, принимаемыми должностными лицами.

Ввиду того, что задача данного пункта является новой, мы подробнее, чем в задаче из области геологии, остановимся на ее нематематической стороне.

Для реализации "а" достаточно исходные данные представить в виде совокупностей признаков (нормативные акты) и совокупностей значений этих признаков (претензии, арбитражные дела).

В качестве неформализованных исходных данных были взяты: "Инструкция о порядке приемки продукции производственно-техниче –

ского назначения и товаров народного потребления по количеству" с учетом "Положения о порядке предъявления и рассмотрения претензий предприятиями, организациями и учреждениями и урегулирования разногласий по договорам" и арбитражные дела по спорам о недостачах продукции *).

В результате "Инструкция" была представлена в виде 81-бинарного признака, а каждое дело из материала контроля – в виде 81-мерного вектора. Дела из материала обучения были представлены в виде 82-мерного вектора, куда включено значение признака-решения (удовлетворить иск, отклонить иск). Возможен также целевой признак, принимающий и другие значения, например частичное удовлетворение иска.

Одной из положительных особенностей рассматриваемой задачи является довольно строгая фиксация признакового пространства, оно определяется нормативными актами, в частности, постановлениями правительства.

Этим данная задача выгодно отличается, например, от геологической задачи из предыдущего пункта, где требуется проведение специальной работы (успех которой зависит от опыта и интуиции специалиста) для нахождения подходящего признакового пространства.

Нами рассматривались решающие правила из класса алгоритмов вычисления оценок с параметрами k и δ_2:

R_4 с параметрами $k = 8$ и $\delta_2 = 0{,}57$;

R_5 с параметрами $k = 20$ и $\delta_2 = 0{,}59$;

R_6 с параметрами $k = 39$ и $\delta_2 = 0{,}506$.

На рассмотрение 40 контрольных дел для принятия решений на основе критерия по всем трем решающим правилам понадобилось около 30 мин на ЭВМ БЭСМ-6, за которое сами R_1, R_2, R_3 применялись около 20 тыс. раз (аналогичные цифры были для задачи из области геологии).

В табл. 2 приведены результаты экспериментов. Как видно из таблицы, эффективность повысилась также на 10%.

VII. Метод сопряжения естественным образом обобщается на случай совокупностей из q решающих правил. Для этого в $M_1^+, M_1^-, M_2^+, M_2^-$ внесем указания на решающие правила, для которых эти величины вычисляются, т.е. $M_1^+(R_1)$, $M_2^+(R_1)$, ..., $M_2^-(R_1)$ и т.д. А в качестве $M_1^+, M_1^-, M_2^+, M_2^-$ будем рассматривать

*) Указанные исходные данные были формализованы Л.Б.Гальпериным и Я.Н.Ковалерчуком с участием автора.

Таблица 2

Характеристики	Решающие правила					
	R_4	$[R_4]$	R_5	$[R_5]$	R_6	$[R_6]$
верное распознавание, %	82,5	90	87,5	97,5	0	12,5
неверное распознавание, %	0	0	0	0	0	0
отказы, %	17,5	10	12,5	2,5	100	87,5
D	91,25	95	93,75	98,75	50	56,25
t,%	82,5	90	87,5	97,5	0	12,5
Повышение эффективности τ,%		7,5		10		12,5
Повышение эффективности d		3,75		5		6,25

$$M_1^+ = \sum_{i=1}^{q} M_1^+(R_i), \quad M_1^- = \sum_{i=1}^{q} M_1^-(R_i),$$

$$M_2^+ = \sum_{i=1}^{q} M_2^+(R_i), \quad M_2^- = \sum_{i=1}^{q} M_1^-(R_i).$$

Тогда вид формул, определяющих критерий $K(R_i, R_j, \mathcal{T})$, останется тем же, что для $K(R_i, \mathcal{T})$. Если зафиксирована пара решающих правил R_i и R_j, то критерий K реализует некоторое решающее правило $[R_i, R_j] = K(R_i, R_j,)$.

В табл. 3 и 4 сведены результаты машинных экспериментов с парами решающих правил для тех же задач из области геологии и юриспруденции, что были рассмотрены выше.

В задаче из области геологии рассматривались пары решающих правил, образованные из тех же R_1, R_2, R_3, по которым проводились эксперименты с каждым в отдельности. Кроме того, рассматривалось R_7 с параметрами $k = 1$ и $\delta_2 = 0{,}51$.

В табл. 3 $\tau_i = t_{крит} - t_i$, $\tau_j = t_{крит} - t_j$, т.е. τ_i, τ_j – это соответственно τ для первого и второго аргументов функции $K(R_i, R_j)$. Как видно из нее, достигнуто повышение эффективности по наиболее эффективным элементам пар решающих правил в среднем на 11,2% и по наименее эффективным – на 31%.

Таблица 3

Характеристики		Решающие правила							
		R_1	$[R_1,R_2]$	R_2	$[R_2,R_3]$	R_3	$[R_1,R_3]$	$[R_3,R_7]$	R_7
верное распознавание, %		65	72,5	0	65	55	72,5	65	37,5
неверное распознавание, %		35	27,5	0	30	27,5	25	32,5	30
отказы, %		0	0	100	5	17,5	2,5	2,5	32,5
D		65	72,5	50	67,5	63,75	73,75	66,75	53,75
t,%		30	45	0	35	27,5	47,5	32,5	7,5
Повышение эффективности	τ_i%		15		35		17,5	5	
	τ_j%		45		7,5		20	25	
	d_i		7,5		17,5		8,75	3	
	d_j		22,5		3,5		10	10	

Таблица 4

Характеристики		Решающие правила					
		R_4	$[R_4,R_5]$	$[R_5]$	R_6	$[R_6,R_8]$	$[R_8]$
верное распознавание, %		82,5	97,5	87,5	0	82,5	75
неверное распознавание, %		0	0	0	0	10	10
отказов, %		17,5	2,5	12,5	100	7,5	15
D		91,25	98,75	93,75	50	86,25	82,5
t,%		82,5	97,5	87,5	0	72,5	65
Повышение эффективности	τ_i, %		15			72,5	
	τ_j,%		10			7,5	
	d_i		7,5			46,25	
	d_j		5			3,75	

В задаче из области юриспруденции рассматривались те же решающие правила R_4, R_5, R_6, R_7 и

R_8 с параметрами $k = 7$ и $\delta_2 = 0.59$,
R_9 с параметрами $k = 1$ и $\delta_2 = 0.506$,
R_{10} с параметрами $k = 1$ и $\delta_2 = 0.508$;
R_{11} с параметрами $k = 39$ и $\delta_2 = 0.65$,
R_{12} с параметрами $k = 3$ и $\delta_2 = 0.65$,
R_{13} с параметрами $k = 3$ и $\delta_2 = 0.59$.

В табл.4 приведены два типичных примера применения метода сопряжения к парам упомянутых решающих правил. В среднем в этой задаче было достигнуто повышение эффективности на 7% по наиболее эффективным решающим правилам рассмотренных пар решающих правил и на 47,7 % по наименее эффективным решающим правилам этих пар.

УШ. 1. Метод сопряжения может быть обобщен на случай t образов ($t>2$). Для этого понадобится: а) ввести не две гипотетические таблицы T_1 и T_2, как ранее, а t гипотетических таблиц $\{T_j\}$ по числу образов, к которым можно отнести контрольный элемент c; б) рассмотреть решающее правило R, относящее контрольный объект к одному из этих t образов.

Пусть $\mu_{ii} \in \Pi$ и

$$M_j^+ = |\{\langle T_j,\mu_{ii}\rangle ; \mathcal{R}(T_j,\mu_{ii}) = И\}|,$$

$$M_j^- = |\{\langle T_j,\mu_{ii}\rangle ; \mathcal{R}(T_j,\mu_{ii}) = Л\}|$$

характеризуют согласованность (несогласованность) гипотезы об отнесении объекта к j-у образу. А $S^j(c)$, аналогично введенному ранее, пусть означает отнесение контрольного объекта к j-образу.

Тогда критерий К запишется в следующем виде:

$$S(c) = K(R,\mathcal{T}) = \begin{cases} S^1(c), \text{ если } \forall q(q \neq 1)\ M_1^+ - M_1^- > M_q^+ - M_q^-, \\ S^2(c), \text{ если } \forall q(q \neq 2)\ M_2^+ - M_2^- > M_q^+ - M_q^-, \\ S^j(c), \text{ если } \forall q(q \neq j)\ M_j^+ - M_j^- > M_q^+ - M_q^-, \\ S^t(c), \text{ если } \forall q(q \neq t)\ M_t^+ - M_t^- > M_q^+ - M_q^-, \\ \quad *, \text{ если не существует } j \text{ такого, что} \\ \quad \forall q(q \neq j)\ M_j^+ - M_j^- > M_q^+ - M_q^-. \end{cases}$$

2. Снимем ограничение о рассмотрении только бинарных признаков $P_1, P_2, \ldots, P_n$. Пусть признак P_i принимает r_i значений (r_i - конечное, заранее заданное число). Таблица исходных данных $\mathcal{T}$ представлена значениями таких признаков. Примером P_i может быть признак "профессия", а значениями r - конкретные профессии: врач, математик, химик и т.д., из заранее заданного конечного списка профессий. Пусть также задано решающее правило, с помощью которого можно предсказывать значения таких признаков на основе исходных дан-

них $\mathcal{T}$. Это обычные в распознавании образов решающие правила.

Тогда, учитывая п.I и сняв в определении $\mathcal{R}$ ограничение $\check{\mu} \in \{0,1\}$, мы видим, что метод сопряжения остается неизменным для введенных в этом пункте изменений в $\mathcal{T}$.

3. Рассмотрим случай, когда таблица исходных данных содержит значения признаков, не удовлетворяющих свойству п.2. Например, пусть признак P_j означает длину объекта и изменяется в интервале от одного до двух метров, объекты, отличающиеся на I см, считаются неразличимыми. Пусть решающее правило таково, что с его помощью можно предсказывать конкретные числовые значения длины [ж)]. Тогда, с учетом п.I и некоторой модификации $\mathcal{R}$, метод сопряжения для п.3 останется неизменным.

Модификация $\mathcal{R}$ состоит в следующем:

$$\mathcal{R}(T_i \backslash \mu) = \begin{cases} \text{И, если } |R(T_i \backslash \mu) - \check{\mu}| \le \varepsilon_j, \\ Z, \text{ если } R(T_i \backslash \mu) = z, \\ \text{Л, если } |R(T_i \backslash \mu) - \check{\mu}| > \varepsilon_j, \end{cases}$$

где ε_j – порог различения по j-у признаку. Например, если $R(T_i \backslash \mu) = 115{,}7$ см и $\check{\mu} = 116{,}4$ см, то в пределах порога различения рассмотренного выше признака, они совпадают и $\mathcal{R}(T_i, \mu) =$ И.

Пусть возможности решающего правила ограничены прогнозированием значений признаков, описанных в п.2. Тогда для использования такого решающего правила в методе сопряжения для таблиц, содержащих признаки, отличные от описанных выше в п.2, понадобится ввести априорные ограничения на множество допустимых сверяемых элементов. Эти ограничения состоят в том, что заранее из числа претендентов в допустимые сверяемые элементы исключаются значения тех признаков, которые отличаются от описанных в п.2.

Может оказаться, что таким путем для конкретной таблицы $\mathcal{T}$ и конкретного R будут исключены из множества П-допустимых сверяемых элементов, все элементы $\mathcal{T}$, соответствующие признакам $P_1, P_2, \ldots, P_n$. При этом заведомо в качестве претендентов на допустимые сверяемые элементы останутся элементы $S(a_1), S(a_2), \ldots, S(a_{m-1})$, где $\{a_i\}$ – объекты обучения. Ввиду того, что множество $\{a_i\}$ непусто, рассмотренные здесь априорные ограничения не могут породить ситуацию, когда единственной допустимой таблицей является сама $\mathcal{T}$, а все сопряженные таблицы недопустимы. Это бы означало, что полностью ликвидирован произвол в способе применения решающего правила R. Таким образом, применение метода сопряжения возможно и в этом случае. Для этого требуется либо некоторое "обеднение" признаков, либо использование в качестве сверяемых элементов только элементов таблицы, соответствующих целевому признаку.

ж) Решающие правила такого типа могут рассматриваться как обобщение обычных в распознавании решающих правил. Их называют методами прогнозирования, эмпирического предсказания, заполнения пробелов и т.д.

4. Несколько замечаний по поводу технической реализуемости метода сопряжения. Не требуется существенных дополнительных затрат машинной памяти для реализации метода сопряжения по сравнению с реализацией самих исходных решающих правил. Требуются значительные дополнительные затраты машинного времени, но, как показывают описанные в работе эксперименты на ЭВМ, по крайней мере для алгоритмов вычисления оценок, метод сопряжения практически реализуем по времени счета. В случаях, когда метод не реализуем в полном объеме, сокращение времени счета может быть получено за счет модификации метода. Если метод сопряжения практически реализуем не для всякого множества сверяемых элементов, а только для любого из его подмножеств мощности λ, то модификация может состоять в том, что подмножество мощности λ выбирается с помощью датчика случайных чисел. Возможны также и другие модификации метода.

Литература

I. ЗАГОРУЙКО Н.Г. Искусственный интеллект и эмпирическое предсказание. Новосибирск, НГУ, I976.

2. HUBER R. Robust Statistic: a review. – "Ann.Math.Stat.", 1972, v.43, p.1044–1067.

3. Метод комитетов в распознавании образов. Отв.ред. И.И.Еремин. Свердловск, УНЦ АН СССР, I974.

4. РАСТРИГИН Л.А., ЭРЕНШТЕЙН Р.Х. Принятия решений коллективом решающих правил в задачах распознавания образов. –"Автоматика и телемеханика", I975, № 9, с. I32–I44.

5. ГАВРИЛКО Б.П., ЗАГОРУЙКО Н.Г., САМОХВАЛОВ К.Ф. Уточнение гипотезы простоты. –В кн.: Вычислительные системы. Вып. 37. Новосибирск, I969, с. 3–9.

6. ГЕРМЕЙЕР Ю.Б. Введение в теорию исследований операций. М., "Наука", I97I.

7. МИРКИН Б.Г. Проблема группового выбора. М., "Наука", I974.

8. ЖУРАВЛЕВ Ю.И., НИКИФОРОВ В.В. Алгоритмы распознавания, основанные на вычислении оценок. –"Кибернетика", I97I, № 3.

9. ЖУРАВЛЕВ Ю.И., КАМИЛОВ М.М., ТУЛЯГАНОВ Ш.Е. Алгоритмы вычисления оценок и их применение. Ташкент, "Фан", 1974.

10. КАМИЛОВ М.М., КИМ А.Н., БУЗУРХАНОВ В. Программный распознающий комплекс, основанный на алгоритмах вычисления оценок (ПРАСК-I). -В кн.: Алгоритмы и программы. Вып. 17. Ташкент, ИК с ВЦ АН УзССР, 1974.

11. БОНГАРД М.М. Проблема узнавания. М., "Наука", 1967.

12. КОВАЛЕРЧУК Б.Я. К вопросу о методе выбора решающих правил. -В кн.: Вычислительные системы. Вып. 61. Новосибирск, 1975, с. 43-50.

13. КОВАЛЕРЧУК Б.Я. О согласовании решающих правил и исходных данных. Отчет ИМ СО АН. Новосибирск, 1976.

Поступила в ред.-изд.отд.
17 сентября 1976 года