УДК 681.327.12

Белозерский Л.А.
Донецкий государственный институт искусственного интеллекта

НЕОПРЕДЕЛЕННОСТЬ РЕШЕНИЙ ВЕРОЯТНОСТНЫХ СИСТЕМ РАСПОЗНАВАНИЯ

Как известно, решение о принадлежности распознаваемых объектов в вероятностных системах основывается на критерии минимума среднего условного апостериорного риска:

$$g(\vec{X}_{изм}) = \arg\min_{1\le i\le m} \sum_{j=1}^{m} П_{ij} \cdot P(\Omega_j / \vec{X}_{изм}) \;, \tag{1}$$

где: $\overline{R}_i(\vec{X}_{изм}) = \sum_{i=1}^{m} П_{ij} \cdot P(\Omega_j / \vec{X}_{изм})$ - упомянутый риск;

$П_{ij}$ - плата за ошибку i -го решения при истинности j -го;

$P(\Omega_j / \vec{X}_{изм})$ - апостериорная вероятности класса Ω_j;

$\vec{X}_{изм}$ - измеренное значение вектора признаков распознаваемого объекта;

m - число классов.

Апостериорная вероятность каждого из распознаваемых классов $P(\Omega_j / \vec{X}_{изм})$ является известной функцией их априорного описания:

$$P(\Omega_j / \vec{X}_{изм}) = F\{P(\Omega_j), f(\vec{X}_{изм} / \Omega_j), \vec{X}\}.$$

Соответственно приведенному решающему правилу (1) легко определяется граница между классами в виде значений аргумента $\vec{X}^{Гр}_{изм}$, при которых происходит скачок функции $g(\vec{X}_{изм})$. Известно, что каждой такой точке отвечает равенство апостериорных рисков отнесений объекта к двум соседствующим классам (Ω_i, Ω_k), приводящее к критерию минимума среднего риска [1]. Для определения границы можно записать:

$$\sum_{j=1}^{m} П_{ij} \cdot P(\Omega_j) \cdot f(\vec{X}^{Гр}_{изм} / \Omega_j) - \sum_{j=1}^{m} П_{kj} \cdot P(\Omega_j) \cdot f(\vec{X}^{Гр}_{изм} / \Omega_j) = 0 \tag{2}$$

Интересуясь природой аргументов описания классов, понимаем, что их источником могут являться экспериментальные данные, а, следовательно, как они, так и определяемая ими зависимость $P(\Omega_j / \vec{X}_{изм})$ представляют собой статистические оценки со всеми вытекающими отсюда последствиями. Поэтому для корректности дальнейшего изложения перейдем к соответствующим обозначениям, подчеркивающим статистический характер составляющих:

$$\widetilde{P}(\Omega_j / \vec{X}_{изм}) = F\{\widetilde{P}(\Omega_j), \widetilde{f}(\vec{X}_{изм} / \Omega_j), \vec{X}\}$$

Для большей ясности перепишем и средний апостериорный риск аналогичным образом:

$$\widetilde{\overline{R}}_i(\vec{X}_{изм}) = \sum_{j=1}^{m} П_{ij} \cdot \widetilde{P}(\Omega_j / \vec{X}_{изм})$$

Отсюда непосредственно следует статистический характер оценки среднего риска, дисперсия которой определяется как:

$$D_{Ri}(\vec{X}_{изм}) = \sum_{j=1}^{m} \Pi_{ij}^2 \cdot D\left[\widetilde{P}(\Omega_j / \vec{X}_{изм})\right]$$

Таким образом, согласно (2), граница между примыкающими классами в каждой точке $\vec{X}_{изм}^{Гр}$ носит статистический характер. С достаточным основанием можно предположить подчинение рассматриваемой разности рисков нормальному закону распределения с математическим ожиданием, равным нулю, и дисперсией, представляющей собой сумму дисперсий рисков отнесений объекта к каждому из соседних классов:

$$D_{\Delta R}(\vec{X}_{изм}) = D_{Ri}(\vec{X}_{изм}) + D_{Rk}(\vec{X}_{изм})$$

Полученная характеристика разброса разности рисков уже сама является оценкой неопределенности принимаемых решений. Если же использовать классическое определение неопределенности, то соответствующая энтропия (H) окажется равной [2]:

$$H(\Delta R) = \log\left[\frac{\sqrt{2\pi e D_{\Delta R}(\vec{X}_{изм})}}{\delta R}\right], \tag{3}$$

где: δR – степень точности определений состояний системы распознавания на границе соседних классов.

Как следует из выполненного рассмотрения, уменьшение дисперсии оценок компонент среднего условного апостериорного риска ведет к уменьшению неопределенности решений вероятностной системы распознавания.

Остается связать дисперсию апостериорной вероятности с особенностями обработки априорной информации при формировании описания вероятностной системы распознавания. В конечном итоге, интересующей является дисперсия $D\left[\widetilde{P}(\Omega_j / \vec{X}_{изм})\right]$, компоненты которой определяются формулой Байеса. Тогда, линеаризуя соответствующую зависимость в точках статистических оценок $\widetilde{P}(\Omega_j)$, $\widetilde{f}(\vec{X}_{изм} / \Omega_j)$, получаем:

$$D[P(\Omega_j / \vec{X}_{изм})] = \sum_{i=1}^{m}\left(\frac{d[P(\Omega_j / \vec{X}_{изм})]}{dP(\Omega_i)}\right)^2_{P(\Omega_i)} \cdot D[P(\Omega_i)] +$$

$$+ \sum_{i=1}^{m}\left(\frac{d[P(\Omega_j / \vec{X}_{изм})]}{df(\vec{X}_{изм} / \Omega_i)}\right)^2_{f(\vec{X}_{изм} / \Omega_i)} \cdot D[f(\vec{X}_{изм} / \Omega_i)]$$

Производные при $i = j$ здесь равны:

$$\frac{d[\Omega_j / \vec{X}_{изм}]}{dP(\Omega_i)\big|_{i=j}} = \frac{f(\vec{X}_{изм} / \Omega_j) \cdot \sum_{k=1}^{m} P(\Omega_k) \cdot f(\vec{X}_{изм} / \Omega_k) - P(\Omega_j) \cdot f^2(\vec{X}_{изм} / \Omega_j)}{\left[\sum_{k=1}^{m} P(\Omega_k) \cdot f(\vec{X}_{изм} / \Omega_k)\right]^2};$$

$$\frac{d[\Omega_j / \vec{X}_{изм}]}{df(\vec{X}_{изм} / \Omega_i)\Big|_{i=j}} = \frac{P(\Omega_j) \cdot \sum_{k=1}^{m} P(\Omega_k) \cdot f(\vec{X}_{изм} / \Omega_k) - P^2(\Omega_j) \cdot f(\vec{X}_{изм} / \Omega_j)}{\left[\sum_{k=1}^{m} P(\Omega_k) \cdot f(\vec{X}_{изм} / \Omega_k)\right]^2}.$$

Точно так же легко получаются выражения производных для $i \neq j$.

Определение дисперсий $D[P(\Omega_i)]$ и $D[f(\vec{X}_{изм}]$ непосредственно связано с обработкой статистического материала. В результате для дисперсии априорных вероятностей [3] имеем:

$$D[P(\Omega_i)] = \frac{1}{N_i}\{P(\Omega_i) \cdot [1 - P(\Omega_i)]\}$$

где: N_i - число испытаний для оценки априорных вероятностей в классе Ω_i.

Точность определения плотности распределения вероятности каждого из независимых признаков распознавания является функцией количества измерений, так как для нее имеем:

$$D[f(X_{изм} / \Omega_i] = \frac{1}{n_i} \cdot \frac{1}{\Delta_X^2}\left\{f(X_{изм} + \frac{\Delta_X}{2} / \Omega_i) \cdot \Delta_X \cdot \left[1 - f(X_{изм} + \frac{\Delta_X}{2} / \Omega_i) \cdot \Delta_X\right]\right\}$$

где: n_i - число измерений признака для оценки плотностей вероятностей описания класса Ω_i.

Итоговая неопределенность решений (3) вероятностной системы распознавания теперь описывается приведенной цепочкой преобразований исходных априорных данных, позволяющей проводить корректные оценки влияния соответствующих количественных характеристик на изменение информационного обеспечения выбора конкретных схем и разработки систем распознавания в классах, не требующих обучения, обучающихся и самообучающихся.

ЛИТЕРАТУРА

1. Ту Дж., Гонсалес Р. Принципы распознавания образов. Перев. с англ. Под. ред. Ю.И. Журавлева. - М.: Мир, 1978. - 411 с.
2. Вентцель Е.С. Теория вероятностей. - М.: Физматгиз, 1964. - 564 с.
3. Лившиц Н.А., Пугачев В.Н. Вероятностный анализ систем автоматического управления. - М.: Сов. Радио, 1963. - 895 с.