

Международный научно-практический форум
«НАУКА И БИЗНЕС»

ТЕЗИСЫ ДОКЛАДОВ ФОРУМА

1 июля 2016 года
Днепр,
Украина

БЫСТРОЕ ОБУЧЕНИЕ НЕЧЕТКОГО КЛАССИФИКАТОРА ПО МЕТОДУ ГЛАВНЫХ КОНКУРЕНТОВ

д.т.н. Штовба С. Д., Галушак А. В.

Винницкий национальный технический университет

В нечетких классификаторах принятие решений происходит по лингвистическим правилам <Если – то>, которые содержат нечеткие термы "низкий", "средний", "высокий" и т.п. [1]. Каждое правило описывает зону факторного пространства, в пределах которой объекты принадлежат одному классу. Границы зон нечеткие, поэтому один и тот же объект может одновременно принадлежать нескольким классам, но с разной степенью.

Для повышения безошибочности нечеткий классификатор обучают по экспериментальным данным. Ниже рассматривается параметрическое обучение, во время которого семантика правил не изменяется, а модифицируются функции принадлежности нечетких термов и весовые коэффициенты правил [2–3]. Обучение в этом случае сводится к решению задачи оптимизации с непрерывными управляемыми переменными.

В последние годы исследовались различные приемы повышения качества обучения нечетких классификаторов, один из которых состоит в выборе функции ошибки – целевой функции соответствующей задачи оптимизации. Минимизация этой функции на обучающей выборке должно происходить быстро.

Полученная при таком обучении нечеткая модель должна обеспечивать низкую (в идеале – минимальную) частоту ошибочной классификации на тестовой выборке. В [4] исследованы 3 функции ошибки: частота ошибок; квадратичная невязка между двумя нечеткими множествами – желаемыми и реальными результатами классификации; квадратичная невязка между нечеткими желаемыми и реальными результатами классификации с дополнительным штрафом за ошибочное решение. Целью настоящей работы является исследования новых функций ошибок на основе разницы принадлежностей нечеткого вывода только к главным конкурентам. В случае правильной классификации главным конкурентом принятого решения является класс со второй по величине степенью принадлежности. Эту разницу следует максимизировать, чтобы отдалить от класса, с которым легче всего перепутать правильное решение. В случае неправильной классификации ошибочно принятое решение является главным конкурентом правильного класса, поэтому при обучении разницу между степенями принадлежности к этим классам следует уменьшать.

1. Нечеткий классификатор

Введем такие обозначения: $\mathbf{X} = (x_1, x_2, \dots, x_n)$ – вектор информативных признаков объекта классификации (вектор входных атрибутов); l_1, l_2, \dots, l_m – классы решений; y – результат классификации.

Нечеткий классификатор представляет собой отображение

$\mathbf{X} = (x_1, x_2, \dots, x_n) \rightarrow y \in \{l_1, l_2, \dots, l_m\}$ на основе базы нечетких правил. Основываясь на [3] базу правил нечеткого классификатора запишем так:

Если $(x_1 = \tilde{a}_{1j}$ и $x_2 = \tilde{a}_{2j}$ и ... и $x_n = \tilde{a}_{nj}$ с весом w_j), то $y = d_j, j = \overline{1, N}$, (1)

где N – количество правил;

$d_j \in \{l_1, l_2, \dots, l_m\}$ – значение консеквента j -го правила;

$w_j \in [0, 1]$ – весовой коэффициент, задающий достоверность j -го правила;

\tilde{a}_{ij} – нечеткий терм, оценивающий атрибут x_i в j -ом правиле.

Классификация текущего объекта, заданного вектором $\mathbf{X}^* = (x_1^*, x_2^*, \dots, x_n^*)$, осуществляется следующим образом. Сначала рассчитывается степень выполнения j -го правила из базы (1):

$$\mu_j(\mathbf{X}^*) = w_j \cdot (\mu_j(x_1^*) \wedge \mu_j(x_2^*) \wedge \dots \wedge \mu_j(x_n^*)), \quad j = \overline{1, N}, \quad (2)$$

где $\mu_j(x_i^*)$ – степень принадлежности значения x_i^* нечеткому терму \tilde{a}_{ij} ;

\wedge – t -норма, которую реализуют операцией минимума или умножением.

Степень принадлежности входного вектора \mathbf{X}^* классам l_1, l_2, \dots, l_m рассчитывается так:

$$\mu_{l_s}(y^*) = \underset{\forall j: d_j=l_s}{\text{agg}} \left(\mu_j(\mathbf{X}^*) \right), \quad s = \overline{1, m}. \quad (3)$$

Принятие решения осуществляется по двум схемам. По схеме с единственным правилом-победителем решением выбирается консеквент правила с максимальной степенью выполнения [2]. По схеме голосования решением выбирается класс с максимальной суммарной принадлежностью [5].

По схеме с единственным правилом-победителем агрегирование (3) осуществляется следующим образом:

$$\mu_{l_s}(y^*) = \max_{\substack{\forall j: d_j=l_s, \\ j=\overline{1, N}}} \left(\mu_j(\mathbf{X}^*) \right). \quad (4)$$

По схеме голосования агрегирование (3) реализуется так:

$$\mu_{l_s}(y^*) = \sum_{\substack{\forall j: d_j=l_s, \\ j=\overline{1, N}}} \mu_j(\mathbf{X}^*). \quad (5)$$

Если в (5) максимальная степень принадлежности больше 1, тогда осуществим нормирование:

$$\mu_{l_s}(y^*) = \frac{\mu_{l_s}(y^*)}{\max_{s=1, m} (\mu_{l_s}(y^*))}.$$

Нечетким решением задачи классификации будет нечеткое множество

$$\tilde{y}^* = \left(\frac{\mu_{l_1}(y^*)}{l_1}, \frac{\mu_{l_2}(y^*)}{l_2}, \dots, \frac{\mu_{l_m}(y^*)}{l_m} \right). \quad (6)$$

Результатом выберем класс с максимальной степенью принадлежности:

$$y^* = \arg \max_{\{l_1, l_2, \dots, l_m\}} (\mu_{l_s}(y^*))$$

2. Обучение нечеткого классификатора

Обучающую выборку из M пар “входы – выход” запишем так:

$$(\mathbf{X}_r, y_r), \quad r = \overline{1, M}, \quad (7)$$

где $\mathbf{X}_r = (x_{r1}, x_{r2}, \dots, x_{rm})$ - входные атрибуты r -го объекта;

$y_r \in \{l_1, l_2, \dots, l_m\}$ - класс r -го объекта.

Введем следующие обозначения:

\mathbf{p} – вектор параметров функций принадлежности термов из базы (1);

\mathbf{w} – вектор весовых коэффициентов правил из базы (1);

$F(\mathbf{K}, \mathbf{X}_r) \in \{l_1, l_2, \dots, l_m\}$ – результат классификации по базе правил (1) с параметрами $\mathbf{K} = (\mathbf{P}, \mathbf{W})$ для входного вектора \mathbf{X}_r из r -ой строки выборки (7).

Обучение нечеткого классификатора состоит в нахождении вектора \mathbf{K} , минимизирующего частоту ошибок классификации на тестовой выборке. При этом для настройки параметров используется только обучающая выборка (7). Обучение осуществим по пяти критериям.

Критерий 1 – частота ошибок классификации [5].

Критерий 2 – квадратичная невязка между двумя нечеткими множествами – желаемыми и реальными результатами классификации [2].

Критерий 3 – квадратичная невязка между нечеткими желаемыми и реальными результатами классификации с дополнительным штрафом за ошибочно принятое решение [4].

Критерий 4 – расстояние между главными конкурентами с штрафом за ошибочное решение. Это новый критерий обучения. Он учитывает разницу принадлежностей нечеткого вывода только к главным конкурентам. По алгоритму логического вывода решением выбирается класс с максимальной степенью принадлежности. Обозначим этот класс-победитель как *win* и присвоим ему первый ранг. В случае правильной классификации главным конкурентом принятого решения является *vicewin* - класс со вторым рангом, то есть класс со второй по величине степенью

принадлежности (рис. 1а). Чем больше разница между степенями принадлежности к классам *win* и *vicewin*, тем больше уверенность в логическом выводе.

Обозначим через $smax$ - операцию нахождения второго по величине элемента множества. Тогда, для *r*-го объекта выборки

$$(7) \quad \mu_{win}(\mathbf{X}_r) = \max_{s=1, m}(\mu_{l_s}(\mathbf{X}_r)) \quad \text{и} \quad \mu_{vicewin}(\mathbf{X}_r) = smax_{s=1, m}(\mu_{l_s}(\mathbf{X}_r)).$$

Соответственно, разница между главными конкурентами составляет:

$$\mu_{win}(\mathbf{X}_r) - \mu_{vicewin}(\mathbf{X}_r).$$

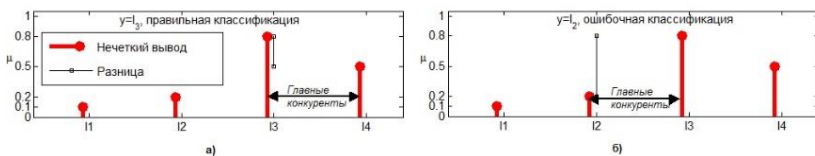


Рис. 1. Главные конкуренты:
а) правильная классификация; б) ошибочная классификация

При неправильной классификации ошибочно принятое решение будет главным конкурентом правильного класса (рис. 1б). Следовательно, желательно уменьшить разницу между степенями принадлежности к ошибочному решению и к правильному классу. Разницу между главными конкурентами в этом случае запишем так:

$$\mu_{win}(\mathbf{X}_r) - \mu_{y_r}(\mathbf{X}_r).$$

В критерии обучения будем учитывать относительные показатели, разделив разницу на степень принадлежности класса победителя. При правильной классификации относительная

разница равна $D_r^1 = \frac{\mu_{win}(\mathbf{X}_r) - \mu_{vicewin}(\mathbf{X}_r)}{\mu_{win}(\mathbf{X}_r)}$, а при

неправильной – $D_r^0 = \frac{\mu_{win}(\mathbf{X}_r) - \mu_{y_r}(\mathbf{X}_r)}{\mu_{win}(\mathbf{X}_r)}$. Кроме того, аналогично

критерию 3, при ошибочной классификации взвесим разницу штрафным коэффициентом. Математически критерий обучения запишем следующим образом:

$$Crit_4 = p \cdot \sum_{r=1, M} \sum_{y_r \neq F(\mathbf{K}, \mathbf{X}_r)} D_r^0(\mathbf{K}) - \sum_{r=1, M} \sum_{y_r = F(\mathbf{K}, \mathbf{X}_r)} D_r^1(\mathbf{K}),$$

где $p \geq 1$ – штрафной коэффициент.

Критерий 5 – квадратичное расстояние между главными конкурентами с штрафом за ошибочное решение. Этот критерий является модификацией предыдущего. Отличие заключается в использовании не абсолютных расстояний, и их квадратов:

$$Crit_5 = p \cdot \sum_{r=1, M} \sum_{y_r \neq F(\mathbf{K}, \mathbf{X}_r)} D_r^0(\mathbf{K})^2 - \sum_{r=1, M} \sum_{y_r = F(\mathbf{K}, \mathbf{X}_r)} D_r^1(\mathbf{K})^2.$$

3. Компьютерные эксперименты

Эксперименты проведены на задаче распознавания итальянских вин из UCI Machine Learning Repository. Использовался классификатор из 5 нечетких правил. Он учитывал 3 характеристики вина: flavanoids, color intensity и proline. Обучающая выборка сформирована из 100 строк данных, а

тестовая из 78. Эксперименты проведены для нечетких классификаторов как с единственным правилом-победителем, так и с голосующими правилами. Для каждой функции ошибки проведено 1000 экспериментов по обучению нечеткой базы знаний на основе квазиньютоновского алгоритма из разных начальных точек. После обучения каждый классификатор проверен на тестовой выборке.

Эксперименты показали существенное преимущество новых критериев обучения. Эти критерии обеспечивают лучшую сходимость обучения – частота ошибок на тестовой выборке нечетких классификаторов существенно ниже, чем при обучении с использованием традиционных критериев. Среди новых критериев обучения небольшое преимущество имеет критерий в форме квадратичного расстояния между главными конкурентами со штрафом за ошибочное решение.

Литература

1. Kuncheva L.I. Fuzzy classifier design / Studies in Fuzziness and Soft Computing. Vol. 49. – Berlin – Heidelberg: Springer-Verlag, 2000. – 314 p.
2. Rotshtein A. Design and tuning of fuzzy rule-based system for medical diagnosis. In “Fuzzy and Neuro-Fuzzy Systems in Medicine” (Eds.: Teodorescu N. H., Kandel A., and Jain L.C.). Boca-Raton: CRC-Press, 1998. - P. 243–289.
3. Штовба С.Д. Проектирование нечетких систем средствами MATLAB. М.: Горячая линия – Телеком, 2007. – 288 с.
4. Штовба С.Д., Панкевич О.Д., Нагорна А.В. Анализ критериев обучения нечеткого классификатора // Автоматика и вычислительная техника. - 2015. - №3. - С. 5-16.
5. Ishibuchi H., Nakashima T., Morisawa T. Voting in fuzzy rule-based systems for pattern classification problems // Fuzzy Sets and Systems. – 1999. – Vol. 103, №2. – P. 223–238.