Третья система

Из приведенных графиков видно, что синяя шкала закрытых и выполненных заявок в наибольшей степени преобладает над красной в третьем графике системы ITM (третья система).

Это значит, что SLA системы ITM наиболее эффективно справляется со своей задачей, поскольку в нем отсутствует большое число посредников по сравнению, например, с системой BMS. За счет этого техническая поддержка работает наиболее слаженно и эффективно.

В любой компании соглашение об уровне услуг (далее SLA) играет очень важную роль. Без SLA служба технической поддержки не сможет вовремя обрабатывать заявки, решать срочные проблемы с неработоспособной техникой, заказывать сырьё для ее эксплуатации и т.д. SLA требуется индивидуально подбирать с учетом возможностей и потребностей каждого предприятия, для чего используются специализированные информационные системы, приобретаемые компанией у сторонних фирм либо создаваемые самостоятельно для оптимизации бизнес-процессов.

Были проанализированы измененные и оптимизированные SLA для систем регистрации заявок, проведено их сравнение для выбора наиболее эффективного.

В результате проведенного анализа выяснилось, что SLA с минимальным количеством посредников позволяет наиболее эффективно работать без перегрузки службы технической поддержки.

1. «Введение в реальный ITSM» / Роб Ингланд; Пер. с англ. — М.: Лайвбук, 2010. — 132 с. ISBN 978-5-904584-05-4
2. «Овладевая ITIL» / Роб Ингланд; Пер. с англ. — М.: Лайвбук, 2011. — 200 с. ISBN 978-5-904584-13-9
3. ИТ-Сервис-менеджмент. Введение. М. 2003 ISBN 90-77212-15-9

Полещук О.М.

Кластеризация выходных данных нечеткой линейной регрессионной модели

*Московский государственный технический университет им. Н.Э.Баумана
(Россия, Москва)*

doi: 10.18411/sr-10-06-2018-05

idsp: 000001:sr-10-06-2018-05

Аннотация

В статье на основе взвешенных отрезков разработана модель кластеризации выходных данных нечеткой линейной регрессионной модели. Входные и выходные данные, а также коэффициенты модели являются нечеткими числами. Кластеризация

выходных данных осуществляется в рамках лингвистической шкалы, которая может быть использована экспертами для оценки параметров сложных объектов.

Ключевые слова: нечеткая регрессионная модель, взвешенный отрезок, кластеризация.

Введение

Согласно [1] нечетким множеством $\tilde{A}$ называется множество пар вида $\{(x, \mu_{\tilde{A}}(x)): x \in X\}$, где $\mu_{\tilde{A}}(x): X \rightarrow [0,1]$ - функции принадлежности $\tilde{A}$.

Нечетким числом называется нечеткое множество $\{(x, \mu_{\tilde{A}}(x)): x \in R\}$, где $\mu_{\tilde{A}}(x): R \rightarrow [0,1]$

T -числом (нечетким числом T -типа) называется нечеткое число $\tilde{A}$ с функцией принадлежности

$$\mu_{\tilde{A}}(x) = \begin{cases} 1 + \frac{x-a_1}{a_L}, & a_1 - a_L \leq x < a_1, a_L > 0 \\ 1 - \frac{x-a_2}{a_R}, & a_2 < x \leq a_2 + a_R, a_R > 0 \\ 1, & a_1 \leq x \leq a_2 \\ 0, & x < a_1 - a_L \text{ или } x > a_2 + a_R. \end{cases}$$

Символически T -число записывается в виде $\tilde{A} \equiv (a_1, a_2, a_L, a_R)$. Нормальное треугольное число является частным случаем T -числа при $a_1 = a_2$ и символически записывается в виде $\tilde{A} \equiv (a_1, a_L, a_R)$. При $a_L = a_R$ треугольное число называется симметричным, а при $a_L \neq a_R$ несимметричным. T -число называется неотрицательным при условии $a_1 - a_L \geq 0$.

В [2] дано определение взвешенного отрезка $[A_1, A_2]$ для T -числа

$\tilde{A} \equiv (a_1, a_2, a_L, a_R)$, где

$$A_1 = a_1 - \frac{1}{6}a_L, \quad A_2 = a_2 + \frac{1}{6}a_R.$$

Для двух T -чисел $\tilde{A}, \tilde{B}$ с взвешенными отрезками $[A_1, A_2], [B_1, B_2]$ в [3] определена мера близости:

$$f(\tilde{A}, \tilde{B}) = \sqrt{(A_1 - B_1)^2 + (A_2 - B_2)^2}.$$

Нечеткая линейная регрессионная модель

Пусть $\tilde{Y} = \begin{pmatrix} \tilde{Y}_1 \\ \dots \\ \tilde{Y}_n \end{pmatrix}$, $\tilde{Y}_i \equiv (y_1^i, y_2^i, y_L^i, y_R^i)$, $y_1^i - y_L^i \geq 0$, $i = \overline{1, n}$ - выходные

неотрицательные T -числа, $\tilde{X}_j = \begin{pmatrix} \tilde{X}_j^1 \\ \dots \\ \tilde{X}_j^n \end{pmatrix}$, $\tilde{X}_j^i \equiv (x_1^{ji}, x_2^{ji}, x_L^{ji}, x_R^{ji})$, $x_1^{ji} - x_L^{ji} \geq 0$, $j = \overline{1, m}$, $i = \overline{1, n}$

- входные неотрицательные T -числа.

Зависимость между входными и выходными данными будем находить в виде:

$$\tilde{Y} = \tilde{a}_0 + \tilde{a}_1 \tilde{X}_1 + \dots + \tilde{a}_m \tilde{X}_m,$$

$a_j \equiv (b^j, b_L^j, b_R^j)$, $j = \overline{0, m}$ - неизвестные коэффициенты регрессионной модели, которые определяются в виде треугольных чисел.

Определим взвешенные отрезки $\left[y_1^i - \frac{1}{6} y_L^i, y_2^i + \frac{1}{6} y_R^i \right]$, $i = \overline{1, n}$ для наблюдаемых выходных данных $\tilde{Y}_i$.

Обозначим через $\left[\theta_{\tilde{a}_j \tilde{X}_j^i}^1(b^j, b_L^j, b_R^j), \theta_{\tilde{a}_j \tilde{X}_j^i}^2(b^j, b_L^j, b_R^j) \right]$ взвешенный отрезок произведения чисел $\tilde{a}_j$ и $\tilde{X}_j^i$, $j = \overline{1, m}$, $i = \overline{1, n}$.

Рассмотрим неотрицательное T -число $\tilde{X} \equiv (x_1, x_2, x_L, x_R)$, $x_1 - x_L \geq 0$ и треугольное число $\tilde{a} \equiv (b, b_L, b_R)$.

В [3] доказано, что границы взвешенного отрезка $\left[\theta_{\tilde{a} \tilde{X}}^1, \theta_{\tilde{a} \tilde{X}}^2 \right]$ произведения нечетких чисел $\tilde{a}$ и $\tilde{X}$ имеют вид:

$$\begin{aligned} \theta_{\tilde{a} \tilde{X}}^1 &= b \left(x_q + (-1)^q \frac{1}{6} x_{M_q} \right) - b_L \left(\frac{1}{6} x_q + (-1)^q \frac{1}{12} x_{M_q} \right), \\ \theta_{\tilde{a} \tilde{X}}^2 &= b \left(x_r + (-1)^r \frac{1}{6} x_{M_r} \right) + b_R \left(\frac{1}{6} x_r + (-1)^r \frac{1}{12} x_{M_r} \right) \quad q = \begin{cases} 1, b - b_L \geq 0 \\ 2, b + b_R < 0 \end{cases}, M_q = \begin{cases} L, q = 1 \\ R, q = 2 \end{cases}, \\ r &= \begin{cases} 2, b - b_L \geq 0 \\ 1, b + b_R < 0 \end{cases}, M_r = \begin{cases} L, r = 1 \\ R, r = 2 \end{cases}. \end{aligned}$$

Согласно этому утверждению, если $\tilde{a}_j \equiv (b^j, b_L^j, b_R^j)$, $j = \overline{1, m}$ неотрицательное нечеткое число ($b^j + b_R^j \geq 0$), то

$$\begin{aligned} \theta_{\tilde{a}_j \tilde{X}_j^i}^1(b^j, b_L^j, b_R^j) &= b^j \left(x_1^{ji} - \frac{1}{6} x_L^{ji} \right) - b_L^j \left(\frac{1}{6} x_1^{ji} - \frac{1}{12} x_L^{ji} \right), \\ \theta_{\tilde{a}_j \tilde{X}_j^i}^2(b^j, b_L^j, b_R^j) &= b^j \left(x_2^{ji} + \frac{1}{6} x_R^{ji} \right) + b_R^j \left(\frac{1}{6} x_2^{ji} + \frac{1}{12} x_R^{ji} \right). \end{aligned}$$

Если $\tilde{a}_j \equiv (b^j, b_L^j, b_R^j)$, $j = \overline{1, m}$ отрицательное нечеткое число ($b^j + b_R^j < 0$), то

$$\begin{aligned} \theta_{\tilde{a}_j \tilde{X}_j^i}^1(b^j, b_L^j, b_R^j) &= b^j \left(x_2^{ji} + \frac{1}{6} x_R^{ji} \right) - b_L^j \left(\frac{1}{6} x_2^{ji} + \frac{1}{12} x_R^{ji} \right), \\ \theta_{\tilde{a}_j \tilde{X}_j^i}^2(b^j, b_L^j, b_R^j) &= b^j \left(x_1^{ji} - \frac{1}{6} x_L^{ji} \right) + b_R^j \left(\frac{1}{6} x_1^{ji} - \frac{1}{12} x_L^{ji} \right). \end{aligned}$$

Определим взвешенные отрезки

$$\left[b^0 - \frac{1}{6} b_L^0 + \sum_{j=1}^m \theta_{\tilde{a}_j \tilde{X}_j^i}^1(b^j, b_L^j, b_R^j), b^0 + \frac{1}{6} b_R^0 + \sum_{j=1}^m \theta_{\tilde{a}_j \tilde{X}_j^i}^2(b^j, b_L^j, b_R^j) \right], i = \overline{1, n}$$

для модельных выходных данных $\hat{Y}_i = \tilde{a}_0 + \tilde{a}_1 \tilde{X}_1^i + \dots + \tilde{a}_m \tilde{X}_m^i$.

Рассмотрим функционал $F = \sum_{i=1}^n f^2(\hat{Y}_i, \tilde{Y}_i)$, который характеризует меру близости между исходными и модельными выходными данными. Легко показать, что

$$F = \sum_{i=1}^n \left[b^0 - \frac{1}{6} b_L^0 - y_1^i + \frac{1}{6} y_L^i + \sum_{j=1}^m \theta_{\tilde{a}_j \tilde{x}_j^i}^1 (b^j, b_L^j, b_R^j) \right]^2 + \\ + \sum_{i=1}^n \left[b^0 + \frac{1}{6} b_R^0 - y_2^i - \frac{1}{6} y_R^i + \sum_{j=1}^m \theta_{\tilde{a}_j \tilde{x}_j^i}^2 (b^j, b_L^j, b_R^j) \right]^2.$$

Оптимизационная задача ставится следующим образом:

$$F(b^j, b_L^j, b_R^j) = \sum_{i=1}^n f^2(\hat{Y}_i, \tilde{Y}_i) \rightarrow \min, \\ b_L^j \geq 0, b_R^j \geq 0, j = \overline{0, m}.$$

Так как $\theta_{\tilde{a}_j \tilde{x}_j^i}^1(b^j, b_L^j, b_R^j)$ и $\theta_{\tilde{a}_j \tilde{x}_j^i}^2(b^j, b_L^j, b_R^j)$ являются кусочно-линейными функциями в области $b_L^j \geq 0, b_R^j \geq 0, j = \overline{0, m}$, то F является кусочно-дифференцируемой функцией, и решения оптимизационной задачи находятся с помощью известных методов [4].

Кластеризация выходных модельных данных.

Пусть исходные выходные данные $\tilde{Y}_i \equiv (y_1^i, y_2^i, y_L^i, y_R^i)$, $i = \overline{1, n}$ являются формализациями $\tilde{Y}_k \equiv (y_1^k, y_2^k, y_L^k, y_R^k)$, $k = \overline{1, p}$ лингвистических значений $Y_k, k = \overline{1, p}$ характеристики Y . После определения модельных выходных данных $\hat{Y}_i, i = \overline{1, n}$ возникают задача их кластеризации или определение принадлежности к одному из кластеров $Y_k, k = \overline{1, p}$. Кластеризация осуществляется путем идентификации нечетких чисел $\hat{Y}_i, i = \overline{1, n}$ с одним из нечетких чисел $\tilde{Y}_k, k = \overline{1, p}$.

Обозначим через $[A_1^i, A_2^i], i = \overline{1, n}$ взвешенные отрезки $[b^0 - lb_L^0 + \sum_{j=1}^m \theta_{\tilde{a}_j \tilde{x}_j^i}^1(b^j, b_L^j, b_R^j), b^0 + rb_R^0 + \sum_{j=1}^m \theta_{\tilde{a}_j \tilde{x}_j^i}^2(b^j, b_L^j, b_R^j)], i = \overline{1, n}$ выходных модельных данных $\hat{Y}_i, i = \overline{1, n}$, а через $[B_1^k, B_2^k], k = \overline{1, p}$ взвешенные отрезки $[y_1^k - ly_L^k + ry_R^k], k = \overline{1, p}$ формализаций $\tilde{Y}_k \equiv (y_1^k, y_2^k, y_L^k, y_R^k)$, $k = \overline{1, p}$ лингвистических значений $Y_k, k = \overline{1, p}$ характеристики Y .

Пусть $f^2(\hat{Y}_i, \tilde{Y}_k) = (A_1^i - B_1^k)^2 + (A_2^i - B_2^k)^2, i = \overline{1, n}, k = \overline{1, p}$.

Выходное модельное значение $\hat{Y}_i$ принадлежит кластеру с лингвистическим значением Y_s , если

$$f^2(\hat{Y}_i, \tilde{Y}_s) = \min_k f^2(\hat{Y}_i, \tilde{Y}_k), k = \overline{1, p}.$$

Заключение

Методы нечеткого регрессионного анализа значительно расширили возможности классического регрессионного анализа, позволив анализировать и прогнозировать информацию в условиях неопределенности неслучайного характера.

Вместе с тем, появилась проблема распознавания и кластеризации выходных модельных данных, представленных нечеткими множествами, которые чаще всего являются формализациями слов естественного языка. В статье разработана модель кластеризации выходных данных нечеткой линейной регрессионной модели на основе представления полученной информации в виде агрегирующих отрезков. Подобный подход позволяет распознавать и прогнозировать информацию ряда областей с активным влиянием человека-эксперта.

1. L.A.Zadeh The Concept of a linguistic variable and its application to approximate reasoning // Information Sciences. - 1975. - Vol. 8. - Pp. 199-249.
2. O.M. Poleshuk, E.G. Komarov New defuzzification method based on weighted intervals // Annual Conference of the North American Fuzzy Information Processing Society, NAFIPS'2008. – 2008. - Pp. 4531223.
3. O. Poleshchuk, E. Komarov Hybrid fuzzy least-squares regression model for qualitative characteristics // Advances in Intelligent and Soft Computing. -2010. - Vol. 68. -Pp. 187-196.
4. T .F. Coleman, Y. Li A reflective newton method for minimizing a quadratic function subject to bounds on some of the variables // SIAM J. Optim. – 1996. - Vol. 6. - Pp. 1040-1058.