

Аломатларни иерархияга асосланган гуруҳлаш асосида объектларнинг умумлашган баҳолари

Ибрагимов М.Ф.¹, Сапаев Ш.О.², Абдулхамидов С.³

Ўзбекистон Миллий Университети

²rich.sapayev@mail.ru

Ўзаро кесишмайдиган K_1, K_2 синфлар билан берилган масала қаралади. Бошланғич E_0 танламада аломатларнинг $X(n)$ тўпламини ўзаро кесишмайдиган $X_1(k_1), \dots, X_\tau(k_\tau), \tau \geq 1, k_1 + \dots + k_\tau \leq n$ тўплам остиларга бўлиш қондаси берилган деб ҳисобланади. Ҳар бир $X_i(k_i)$ бўйича $S_j \in E_0, j = 1, \dots, m$ объектлар тавсифидаги $X_i(k_i)$ аломатлар қийматларини сон ўқидаги қийматларга (умумлашган баҳога) акслантириш учун A_i алгоритмини аниқлаш талаб қилинади (Ю. И. Журавлёвнинг [1] образларни англашга алгебраик ёндошуви терминологиясидаги англаш оператори).

Миқдорий ва номинал аломатлар номерлари тўпламларини мос равишда I ва J орқали белгилайлик.

Умумлашган баҳолар қийматларини кетма-кет ҳисоблаш жараёни иерархияга асосланган агломератив гуруҳлаш алгоритми билан қуйида тавсифланган қоида бўйича амалга оширилади. Иерархияга асосланган гуруҳлашнинг p -қадамида $0 \leq p < n$ аломатларни идентификациялаш учун $\{x_i^p\}_{i \in (I \cup J)}$ белгилашдан фойдаланамиз.

Гуруҳлаш ва умумлашган баҳоларни шакллантириш жараёнида I ва $J, |I| + |J| \leq n$ тўпламлар элементлари ва қуввати ўзгариб боради. Гуруҳларга бирлашадиган аломатларнинг ўлчов шкаласига боғлиқравишда сон ўқида акслантириш учун улар параметрларини ҳисоблашнинг турли усуллардан фойдаланилади.

E_0 танланма объектларининг $x_j^p, j \in I, p \geq 0$ аломат бўйича қийматларининг тартибланган тўпламини, ҳар бири номинал аломат градацияси сифатида қараладиган $[c_1^{jp}, c_2^{jp}], (c_2^{jp}, c_3^{jp}]$ интервалларга бўламиз.

$x_j^p, j \in I$ аломатнинг $[c_1^{jp}, c_2^{jp}], (c_2^{jp}, c_3^{jp}]$ интервалдаги $K_i, i = 1, 2 (|K_i| > 1)$ синфдаги қийматларининг миқдори мос равишда u_i^1, u_i^2 бўлсин, $v-x_j^p$ аломатнинг E_0 танланмадаги ўсиш бўйича тартибланган $r_{j_1}, \dots, r_{j_v}, \dots, r_{j_m}$ кетма-кетлигининг $c_1^{jp} = r_{j_1}, c_2^{jp} = r_{j_v}, c_3^{jp} = r_{j_m}$ кўринишида интервал чегарасини аниқлаб берувчи тартиб номер бўлиб,

$$\left(\frac{\sum_{i=1}^2 u_i^1 (u_i^1 - 1) + u_i^2 (u_i^2 - 1)}{\sum_{i=1}^2 |K_i| (|K_i| - 1)} \right) \left(\frac{\sum_{d=1}^2 \sum_{i=1}^2 u_i^d (|K_{3-i}| - u_{3-i}^d)}{2 |K_1| |K_2|} \right) \rightarrow \max_{c_1^{jp} < c_2^{jp} < c_3^{jp}} (1)$$

мезони бўйича $[c_1^{jp}, c_2^{jp}], (c_2^{jp}, c_3^{jp}]$ интерваллар оралигидаги чегара қийматини баҳолашга имкон беради. Мезоннинг чап қавс ичидаги ифода синф ичидаги ўхшашликни, ўнг-дагиси - синфлараро фарқланишни билдиради. Мезоннинг экстремуми x_j^p аломатнинг w_j^p ($0 \leq w_j^p \leq 1$) вазни сифатида ишлатилади. Агар $w_j^p = 1$ бўлса, x_j^p аломатнинг K_1, K_2 синфлар объектлардаги қийматлари ўзаро кесишмайдиган ҳолатни билдиради.

Фойдаланилган адабиётлар

1. Игнатъев Н.А. "Интеллектуальный анализ данных на базе непараметрических методов классификации и разделения выборок объектов поверхностями". Монография.- Ташкент: Национальный университет Узбекистана им. Мирзо Улугбека, 2010.