

TURLI TOIFADAGI ALOMATLARNI NOCHIZIQLI O'GIRISHLARINI QO'LLANGAN HOLDA FAZO O'LCHAMINI QISQARTIRISH

Ergasheva Shohsanam Elmurod qizi

*Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti, Toshkent,
O'zbekiston, e-mail: ergasheva.shohsanam1998@gmail.com*

Annotatsiya. Ob'ektlarning binar alomatlardagi tavsiflari bilan unifikatsiya qilish uchun turli toifadagi alomatlariga dastlabki ishlov berish masalasi qaraladi. Dastlabki ishlov berish vositasi sifatida, o'zaro kesishmaydigan sinflarga tegishlilik funksiyasi ishlatiladi. Fazo o'lchamini qisqartirish uchun binar alomatlarning aniqlanuvchi to'plam ostidagi bo'yicha umumlashgan baho orqali amalga oshiriladi. Anglash aniqligi bo'yicha qiyosiy taqqoslashni amalga oshirish uchun NN, KNN, SVM usullaridan foydalaniladi.

Kalit so'zlari: tegishlilik funksiyasi, umumlashgan baho, nochiziqli o'girish

Turli toifadagi (miqdoriy va sifat) ma'lumotlarni tahlil qilishning umumiy qabul qilingan metodologiyasi mavjud emas. Miqdoriy (xom va latent) ma'lumotlarni tahlil qilish natijalarini tushuntirishda ko'pincha "ko'rsatkich me'yordan kattaroq (kamroq)" munosabatlari qo'llaniladi. Qoida tariqasida, normal qiymat yotadigan interval ko'rsatiladi. Tibbiy amaliyotda har bir insonning o'ziga xos "sog'liqni saqlash standarti" bor, degan muqobil nuqtai nazar mavjud, bu salomatlik holatini baholash uchun rasman qabul qilingan ko'rsatkichlar bilan mutlaqo mos kelmaydi. Ushbu nuqtai nazarni asoslash uchun odamlarning alohida guruhlar uchun normal qiymatlar yotadigan bir nechta intervallar mavjudligi haqidagi gipotezani sinab ko'rish mantiqa to'g'ri keladi. Masalan, ba'zi sport turlari bo'yicha jahon elitasining vakillari. Interval chegaralarini hisoblash uchun maxsus mezonlarni ishlab chiqish va ulardan foydalanish zarurati mavjud.

Ikki sinfli anglash muammosida ob'ektlarni tavsiflash uchun yangi fazoni tanlashda, alomat qiymatlarini o'zaro kesishmaydigan intervallarga bo'lish uchun ikkita mezon taklif qilindi. Bu ikki xususiyatlar mezonlar orasidagi farqni ko'rsatadi:

- alomat qiymatlarining tartibi;
- har bir sinf vakillarini alomatlarni qiymatlarini uchrashishlar chastotasi.

Birinchi xususiyatdan foydalanib, alomatning farqi (ixchamligi) o'lchovi va sinflar orasidagi chegara (eshik) aniqlanadi. Ikkinchi xususiyat ob'ektlarning sinfga tegishlilik funksiyasi qiymatlarini hisoblash uchun ishlatiladi.

Miqdoriy alomatning dastlabgi qiymatlarini sinflarga tegishlilik funksiyasi qiymati bilan almashtirish nochiziqli transformatsiya [2] sifatida qaraladi. Nominal alomatlar uchun gradatsiyalar o'rniga tartib shkalasida ob'ektlarning tavsifidan

foydalanish yoki gradatsiyalar sonini ikkitagacha kamaytirish mumkin bo'ldi. Dastlabki " xom " alomatlarining qiymatlaridan uni ikkita gradatsiya orqali taqdim etishga o'tish jarayonining tavsifianglash algoritmlarining metamodelini yaratish uchun elementar klassifikatorlar ansamblini shakllantirish stacking texnologiyasidan foydalangan[3].

Sinflar juftligi uchun latent ko'rsatkichlarni (alomatlarini) sintez qilish orqali turli toifadagi alomatlarining informativ majmualarini tanlash jarayonini amalga oshirish taklif etiladi. Yashirin (latent) xarakteristikalar asosida sinflar ob'ektlari orasidagi farqlarni aniqlash uchun maxsus o'lchov ko'rib chiqiladi. Sinflar soni 2 dan ortiq bo'lsa, informativ alomatlarini tanlashning kutilayotgan natijalari:

- barcha sinflar uchun xom alomatlar to'plami;
- har bir sinf juftligi uchun farqli xom alomatlar to'plami.

Ushbu ishda latent alomatlardan foydalanish sinflar juftligi uchun ob'ektlarni umumlashtirilgan baholash sifatida ularning qiymatlarini hisoblash bilan bog'liq. O'zaro kesishmaydigan sinflar ob'ektlarining alomatlari bir-biriga qarama-qarshilikni belgilaydi (kasal - sog'lom; kuchli - zaif; kambag'al - boy). Xom alomatlar to'plamiga asoslangan umumlashgan bahoni hisoblash ularni oldindan qayta ishlash bilan yoki ularsiz amalga oshirilishi mumkin. Dastlabki ishlov berishdan avval fazo o'lchami ob'ektlar tavsifidagi xom alomatlarining qiymatlari $\{1,2\}$ da ko'rsatiladi.

Ob'ektlarni umumlashgan bahoga asoslangan latent alomatlardan foydalanish istiqbollari ikki sinfli masalalarda fazo o'lchamini qisqartirish uchun boshlang'ich alomatlarini guruhlash (to'plamlarini shakllantirish) mezonlarini ishlab chiqish bilan bog'liq. Mezonlarni qo'llash natijalariga ko'ra, o'zaro kesishmaydigan bo'sh bo'lmagan alomatlar to'plami shakllantiriladi. Tadqiqot muammolaridan biri bu alomatlar to'plamining mavjudligini isbotlash, ob'ektlarni ikki sinfga bo'lishning aniqligi barcha alomatlar to'plami bo'yicha hisoblangan umumlashgan bahosi yuqori bo'lganini hisoblashdir. Mavjudlikni isbotlash klassik usullar yordamida hal qilinishi mumkin bo'lmagan katta o'lchamli muammolar uchun ham tegishli.

O'zaro kesishmaydigan l ta K_1, \dots, K_l , vakilli sinflarni o'z ichiga oluvchi $E_0 = \{S_1, \dots, S_m\}$ ob'ektlar to'plami berilgan deb hisoblanadi. Sinf indekslarining $\{1, \dots, l\}$ qiymatlari tartibli shkaladagi tasniflash alomatining ruxsat etilgan ketma-ketlik to'plamiga tegishli deb hisoblanadi. E_0 dagi ob'ektlarning tavsiflari $X(n) = (x_1, \dots, x_n)$ turli toifadagi alomatlar to'plami bilan ifodalanadi, ulardan σ nominal, $n - \sigma$ miqdoriy alomatlar.

Quyidagi amallar aniqlangan bo'lsin:

- $X(n)$ to'plamni (K_i, K_j) sinflar juftligi bo'yicha o'zaro kesishmaydigan $X^{ij}(n_1), \dots, X^{ij}(n_t)$ to'plamostisiga ajratish, $i < j$, $t \geq 1$, $\forall c \in \{1, \dots, t\} n_c > 1$, $n_1 + \dots + n_t = \theta \leq n$ va ular asosida $Y^{ij}(t) = (y_1^{ij}, \dots, y_t^{ij})$ latent alomatlar termasini hisoblash;

- $K_i \cup K_j$ da $y_c^{ij} \in Y^{ij}(t)$ va $F(i, j) = \max_c \delta(y_c^{ij})$ bo'yicha $\delta(y_c^{ij})$, $c \in \{1, \dots, t\}$ - kompaktlik o'lchovi qiymatlarini hisoblash.

Talab etiladi:

-har bir $i \in \{1, \dots, l-1\}$ uchun $\Omega_i = \{X^{ij}(d) | j > i, j = i+1, \dots, l, d = \arg \max_t \delta(y_t^{ij})\}$

to'plamni ajratib olish;

- $A = \{\alpha_{ij}\}_{i,j \in \{1, \dots, l\}}$ matritsani hosil qilish, bunda $\alpha_{ij} = \begin{cases} 0, & i \geq j, \\ F(i, j), & i < j; \end{cases}$

- $\alpha_{i,i+1} \wedge \alpha_{i,i+2} = \dots = \alpha_{i,l-1} \wedge \alpha_{i,l} = \text{true}$ shartni qanoatlantiruvchi Ω_i , $i \in \{1, \dots, l-2\}$ to'plam ostilari ajratib olinsin, bu yerda $H \in \{\leq, \geq\}$.

Fazo o'lchamini qisqartirish uchun yuqorida keltirilgan amallar ketma-ket bajariladi, ya'ni sinflar juftligi bo'yicha $F(i, j)$ kompaktlik o'lchovining maksimal qiymatiga muvofiq $Y^{ij}(t) = (y_1^{ij}, \dots, y_t^{ij})$ latent alomatlar termasi quriladi, bu yerda t ($t < n$) fazoning o'lchami.

Foydalanilgan adabiyotlar

1. Н.А. Игнатъев, М.А. Рахимова. Формирование и анализ наборов информативных признаков объектов по парам классов
2. Ignatiev N. A. On Nonlinear Transformations of Features Based on the Functions of Objects Belonging to Classes // Pattern Recognition and Image Analysis. 2021. V. 31. № 2. P. 197–204.
3. Joseph R. (2019, April 23). Ensemble methods: bagging, boosting and stacking. Understanding the key concepts of ensemble learning. <https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205>