

YIRIK XAJMDAGI AXBOROTLARGA ISHLOV BERISHDA TIMSOLLARNI ANIQLASH ALGORITMI

Yuldashev Sherzod Ilimjonovich

Toshkent amaliy fanlar universiteti informatika kafedrası o'qituvchisi

yuldashev-1405@mail.ru

<https://doi.org/10.5281/zenodo.13366014>

Annotatsiya: Mazkur ishda pretsedent axborotlar bilan bog'liqlikni tiklashni kombinatorik nazariyasi doirasida bajarilgan bo'lib, bunda eng yaqin qo'shnilar algoritmgiga asoslangan tanib olish algoritmlari tadqiq qilingan.

Kalit so'zlar: Big Data, Data Centre, BioData, Data Sciences, DataMining, Web Mining, k yaqin qo'shnilar algoritmlari.

1 KIRISH

Ma'lumki, ma'lumotlar bazalarida axborotlar tobora ortib borayotganligi, yirik xajmdagi axborotlardan foydalanishga va ularga ishlov berishga imkoniyat bermogda hamda yanada takomillashtirish zarurati tug'ilmoqda. O'z navbatida yirik hajmdagi ma'lumotlarni soqlovchi kontenerlarni Big Data, Data Centre, BioData deb ataluvchi terminlar ham vujudga kelgan. Bu kabi yirik xajmdagi ma'lumotlarni intellektual tahlil qiluvchi Data Sciences, DataMining, Web Mining kabi fan sohalari ham shakllanib ulgurgan. Mazkur Data Sciences, DataMining, Web Mining fan yo'nalishlarining ma'lumotlarni intellektual tahlil qilishning fundamental usullaridan biri bu timsollarni aniqlash algoritmlaridir. Yirik hajmdagi ma'lumotlar bazalaridan obyektlar o'xshashliklarini sinflashtirish orqali dastlabki ma'lumotlarni tahlil qilishni takomillashtirishga zarurat keskin ortib bormogda [1].

2 TADQIQOT METODOLOGIYASI

Yirik xajmdagi ma'lumotlarni intellektual tahlil qilish masalalarini hal qilishda timsollarni aniqlash usullarini yangi modifikatsiyalangan variantlarini ishlab chiqishga to'g'ri keladi. Buning uchun mavjud algoritmlarni parametrlarini sozlash yoki yangi parametrlarni kiritish yordamida algoritmlar modellari boyitiladi va bu jarayon quyidagi bosqichlardan iborat bo'ladi: belgilar vaznlarini aniqlash orqali faol belgilar qism to'plamini ajratib olish; obyektlar vaznlarini joriy qilish orqali ularni faollik darajalaridagi guruhlariga ajratish; obyektlarni filtrlash orqali tayanch o'quv tanlanmani ajratib olish.

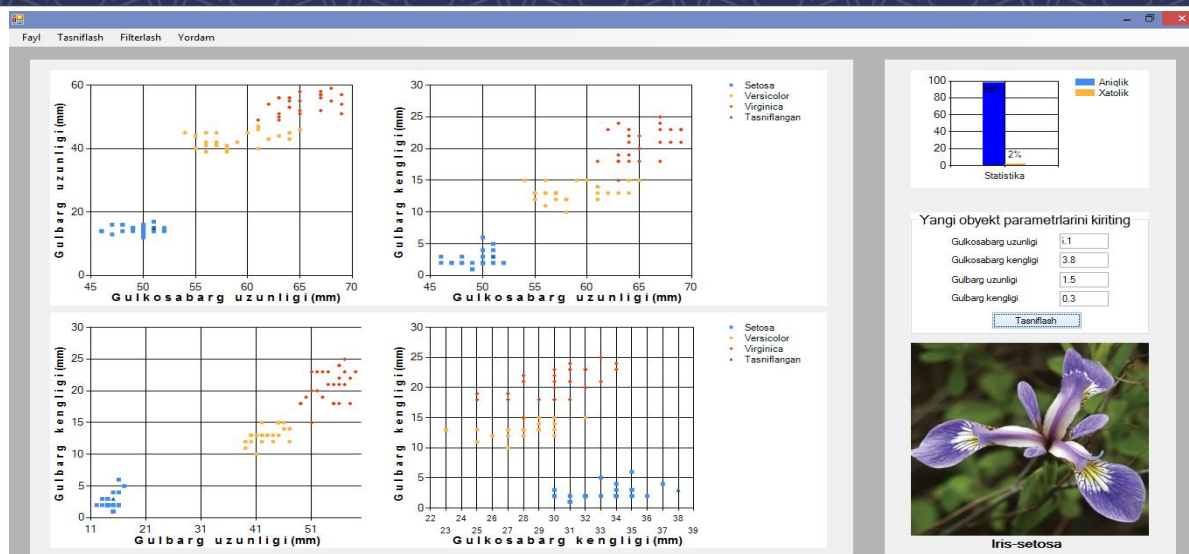
3 MASALANI QO'YILISHI

Ushbu ishda sinflashtirishni metrik algoritmlaridan biri bo'lgan yaqin qo'shnilardan foydalangan holda tanlanmadagi obyektlarni filtrlash masalasi ko'rib chiqiladi. Tayanch obyektlarni tanlash uchun kompaktilik profillarini tabaqalarga ajratish orqali tahlil qilishga asoslangan umumlashgan imkoniyatga ega kombinator baholashlardan foydalaniladi. Bu jarayonni

amalga oshirish uchun quyidagicha masalalarni yechish talab etiladi: kompaktilik profillarini tabaqalashtirishga asoslangan tayanch obyektlarni tanlash algoritmini ishlab chiqish. Tayanch obyektlarni tanlashda evristik usullardan foydalangan holda model masalalarda sinov tajribalarini o'tkazish.

4 TAHLIL VA NATIJALAR

Bosh to'plamdan tayanch obyektlarni qism to'plamini ajratib olishni ikki yondoshuvi mavjud: tayanch obyektlar metrikalarini sintez qilish; to'plamni qism to'plamlarga ajratish. Ushbu ishda ikkinchi yondoshuvga to'xtash ko'zda tutilgan. Bu masalani yechish uchun k yaqin qo'shnilar algoritmlaridan foydalanish maqsadga muvofiqdir. Yuqorida aytilganidek, har bir sinf o'zining tabaqalashgan o'quv tanlanmasiga ega. Ularning markazida tayanch qism to'plam mavjud bo'lgan holda har bir tabaqa o'z chegarasi ega bo'ladi. S obyektlar, Γ javoblar to'plami hamda S -ni Γ -ga akslantirishni amalga oshiruvchi A -algoritmlar to'plamlari berilgan bo'lsin. O'z navbatida $S = \langle S_\ell, S_q \rangle$, bu yerda $S_\ell = (s_i, \Gamma_i)_{i=1}^\ell$ -o'quv tanlanma, $S_q = (s_j)_{j=1}^q$ -nazorat tanlanma obyektlaridan tashkil topgan [1]. S_ℓ o'quv tanlanmada A algoritmini o'rtacha xatoligini hisoblovchi funksional berilgan: $\varphi(A, S_\ell) = \frac{1}{\ell} \sum_{i=1}^\ell \rho_i(S_\ell, S_q)$, ($\ell \neq q$), bu yerda $\rho_i(S_\ell, S_q)$ -funksiya s obyektni aniqlashda A algoritmnining xatolik qiymatini xarakterlovchi miqdor. Dastlabki o'quv tanlanmada obyektlari soni M , birinchi sinf obyektlari soni m_1 , ikkinchi sinf obyektlari soni esa m_2 bo'lsa, u holda o'quv tanlanmaning mumkin bo'lgan variantlari har bir sinf uchun bittadan pretsedentlar guruhi bo'yicha $M = \bigcup_{i=1}^\ell m_i$ ga teng o'rinli bo'ladi. Har bir s_{ij} obyekt belgilanishga ega, bu yerda j -o'quv tanlanmadagi obyektning raqami bo'lsa, i -sinfni raqami, ($i = \overline{1, \ell}$). Har bir sinf o'z o'quv tanlanmasidagi obyektlar masofaga nisbatan o'sish tartibida tartiblanadi.



1-rasm. Nazorat obyektlarini sinflashtirish jarayoni

Ko'rinib turganidek sinflarni kompaktlini ta'minlash orqali 100 % aniqlikka erishildi.

XULOSA VA TAKLIFLAR

Yuqorida keltirilgan nazariya asosida ikki moduldan tashkil topgan dasturiy majmua ishlab chiqilgan bo'lib, birinchisi o'quv tanlanmadagi obyektlarni

ADABIYOTLAR

- [1] Камилов М.М., Худайбердиев М.Х., Хамроев А.Ш. О подходе к решению задачи формирования активных признаков для распознавания одного рода объектов растительного мира. Журнал "Проблемы вычислительной и прикладной математики". Ташкент, 2016, №3. С.44-49.
- [2] Воронцов К.В. комбинаторные оценки качества обучения по прецедентам // Докл. РАН.-2004. -Т.394, №2. -С.175-178.
- [3] Вапник В.Н. Восстановление зависимостей по эмпирическим данным. -М.: Наука, 1979.
- [4] Воронцов К.В. комбинаторный подход к оценке качества обучаемых алгоритмов // Математические вопросы кибернетики. -2004.-Т 477, № 4. -С. 231-238.
- [5] Загуройко Н.Г. Прикладные методы анализа данных и знаний. -Новосибирск: ИМ СО РАН, 1999.

vizuallashtirishga, ikkinchisi esa obyektlarni o'z sinfiga yaqinligini aniqlashga xizmat qiladi. Quyidagi rasmda qism to'plamlarini yuqorida keltirilgan nazariya asosida olingan natijalar keltirilgan.

- Advances in Relation and Hybrid methods.-Kluwer Academic Publishers, 2000. -p.308.
- [6] Беркинов Х., Беркинова А., Султонов Б., Холдоров Х. Иктисодий масалаларда корреляцион-регрессион тахлил моделлари тадбири. //Ўқувуслублий қўлланма. Тошкент. «IQTISOD-MOLIYA». 2008. 108 б.
- [7] Воронцов К.В. Оптимизационные методы линейной и монотонной коррекции в алгебраическом подходе к проблеме распознавания// ЖВМ и МФ.-2000. -Т.40, №1. -С.166-176.
- [8] Гореева Н.М., Демидова Л.Н., Клозогуб Л.М. и другие. Эконометрика: учебное пособие. //М.: Эксмо, 2008. -224 с.
- [9] Mazurov V.I.D., Hachay M.Yu. The committee machine constructions //News of the Ural University. Series Mathematic-mechanic. 1999. Issue 2 (14). P. 77-109.
- [10] <https://xb.uz/uz/interactive/rates/2018-04-03>.