Sinflashning metrik algoritmlari

Reja:

1.Yaqin qo’shni usuli va uni umumlashtirish.

2.k-yaqin qo’shni algaritmi.

3.Farzin darchasi usuli.

4.Yadro funksiyasi.

Ko’pgina amaliy masalalarda ob’ektlarniing o’xshashlik darajasini o’lchash xususiyatlari sinflarini shakllantirishdan ancha osondir.Misol uchun, ba’zi xususiyatlarini ixtiro qilish va xususiyatlar sinflarini solishtirishdan ko’ra yuzlarning fotosuratlari,imzolar,vaqt seriyalari yoki oqsillarning birlamchi tuzilmalari kabi murakkab ob’ektlarni qandaydir “hizalama bilan qoplash” orqali to’g’ridan to’g’ri bir-biri bilan solishtirish tabiiyroqdir.Agar ob’ektlarning o’xshashligi o’lchovi juda muvaffaqiyatli kiritilgan bolsa,unda,qoida tariqasida,o’xshash ob’ektlar ko’pincha o’xshash javoblarga mos keladi.Sinflash muammolarida bu sinflar ixcham lokalizatsiya qilingan kichik to’plamlarni tashkil qilishini anglatadi.Bu taxmin odatda ixchamlik gipotezasi deb ataladi,”o’xshashlik” kontseptsiyasini rasmiylashtirish uchun X ob’ektlar fazosiga masofa funksiyasi kiritiladi.Ob’ektlarning o’xshashligini taxlil qilishga asoslangan o’rganish usullari masofa bo’lsa ham metric deb ataladi.Funksiya metrikaning barcha aksiomalarini (xususan,aksioma uchburchagini )qanoatlintirmaydi.Ingliz tilidagi adabiyotlarda o’xshashlikka asoslangan ta’lim yoki masofaviy ta’lim atamalari qo’llaniladi.Sinflash metrikasi (o’xshashlikka asoslangan sinflash )ob’ektlar orasidagi o’xshashlik ballarini hisoblashga asoslangan sinflash algoritmidir.Eng oddiy metrik sinflash eng yaqin qo’shnilar usuli bo’lib unda sinflanayotgan ob’ekt unga o’xshash ob’ektlarning aksariyati tegishli bo’lgan sinfga tegishlidir.O’xshashlik tushinchasinin rasmiylashtirish uchun ob’ektlar orasidagi masofa funksiyasi kiritiladi.Qoida tariqasida, bu funksiya metric bo’lishi uchun qat’iy talab yoq; xususan, uchburchak tengsizligi bo’lishi mumkin.Sinflashning metric algoritmlari quyidagilarni o’z ichiga oladi:

Eng yaqin qo’shnilar usuli

Potensial funksiya usuli

Radial asosli funksiya usuli

Farzen oynasi usuli

Standartlarni maydalash usuli

Baholarni hisoblash algoritmi

Sinflash metrikalariga qo’shimcha ravishda,mashinani o’rganish ob’ektlarni o’xshashligi haqidagi malumotlardan foydalanadigan,ammo boshqa muammolarni hal qilish uchun ishlatiladigan keng usullar sinfiga ega:

Klasterlash

Parametrik bo’lmagan regressiya

Ko’p o’lchovli masshtablash

Vizualizatsiya,xususan,koxonen xaritasi

Kontaktlik gipotezasi

Sinflash metrikalari ixchamlik gipotezasiga tayanadi,unga ko’ra o’xshash ob’ektlar har xil ob’ektlarga qaraganda bir xil sinfda bo’lish ehtimoli ko’proq.Bu shuni anglatadiki,sinflar orasidagi chegara juda oddiy shaklga ega va sinflar ob’ekt fazosida ixcham maxalliylashtirilgan xududlarni tashkil qiladi.E’tibot bering, matematik tahlilda chegaralangan yopiq to’plamlar ixcham deyiladi.Konpaktlik gipotezasi bunday emas, bu tushuncha bilan hech qanday aloqasi yoq va bu so’zning kundalik manosida tushiniladi.

Xususiyatsiz tan olish

Metric algoritmlarda sinflanga ob’ektlar xususiyatlar to’plami bilan emas,balki to’g’tidan to’g’ri o’quv namunasidagi boshqa ob’ektlrga masofalar vektori bilan sinflash mumkin.

Belgisiz tan olish

Masalan ,matnlarning oxshashligi,kimyoviy formulalar ,aminakislotalar ketma ketligi va boshqalar.Xususiyat sinflariga o’tishdan ko’ra, to’g’ridan to’g’ri o’lchash ancha oson

Metrik tanlash muammosi

Amaliy sinflash sinflash masalalarida,yaxshi masofa funksiyasin oldindan ma’lum bo’lganda, bunday ideal holatlar kamdan kam uchraydi.Agar ob’ektlar raqamli vektorlar bilan tasvirlangan bo’lsa,ko’pincha Evklidn metrikasi qo’llaniladi .Bu tanlov,qoida tariqasida, oqlanmaydi ,bu shunchaki aqlga kelgan birinchi narsa.Shuni esda tutish kerakki, barcha xuxusiyatlarni bir xil miqqiyosda o’lchash kerak,yaxshisi normallashtiriladi.Aks holda,eng yuqori raqamli qiymatlarga ega xususiyat metrikada ustunlik qiladi va qolgan xususiyatlar,aslida,hisobga olinmaydi. Biroq normalizatsiya juda shubhali evristik,chunki savol qolmoqda:”barcha xususiyatlar bir xil darajada axamiyatga ega va taxminan bir xil og’irlikda hisobga olinishi kerakmi?”.

Pretsedentlardan mulohoza yuritish

CBR paradigmasi,vaziyatga asoslangan fikrlash,sun’iy intellect sohalaridan biri sifatida paydo bo’ldi.Ekspert tizimlarida nafaqat ob’ektlarni tasniflash, balki foydalanuvchiga taklif qilingan tasniflash haqida tushuntiri berish muhimdir.Eng yaqin qo’shni usulida bunday tushintirishlar juda asosli ko’rinadi:”Ob’ekt y sinfga tayinlangan,chunki o’quv namunasidagi \* yaqin ob’ekt bir sinfga tegishli edi”.Ushbu “pretsedent” mantiq ko’plab fan sohalari (tibbiyot,geologiya,huquq)mutaxasislari tomonidan yaxshi tushiniladi.