

DATA MINING KLASSIFIKATSIYA VA KLASTERIZATSIYA

Muxtorov Doston Naim o'g'li

Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti, Assistent

E-mail: dmuxtorov062@gmail.com

SAFAROV ULUG'BEK ANORBOY O'G'LI

Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti, Jizzax filiali talabasi

E-mail: safarovulugbek@gmail.com

G'ANIYEV SOBIR QODIR O'G'LI

Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti, Jizzax filiali talabasi

E-mail: ganiyevsobir@gmail.com

Annotatsiya: *Klassifikatsiya(Tasniflash). Klassifikatsiya eng oddiy va shu bilan birga eng tez-tez hal qilinadigan Data Mining vazifasidir. Tasniflash vazifalarining keng tarqalganligi sababli ushbu kontseptsiyaning mohiyatini aniq tushunish kerak. Bu yerda ba'zi ta'riflar keltirib o'tilgan*

Kalit so'zlar:*Klassifikatsiya, Tasniflagich, O'quv to'plami*

Kirish

Klassifikatsiya - o'rganilayotgan predmetlar, hodisalar, jarayonlarning jinsi, turlari, turlari bo'yicha, ularni o'rganish qulayligi uchun har qanday muhim belgilar bo'yicha tizimli ravishda taqsimlash; asl tushunchalarni guruhlash va ularni o'xshashlik darajasini aks ettiruvchi ma'lum bir tartibda joylashtirish.

Klassifikatsiya - bu ba'zi bir printsipga muvofiq buyurtma qilingan ob'yektlar to'plami bo'lib, ular o'xshash tasniflash xususiyatlariga ega (bir yoki bir nechta xususiyatlar), ushbu ob'yektlar o'rtasidagi o'xshashlik yoki farqni aniqlash uchun tanlangan.

Tasniflash quyidagi qoidalarga rioya qilishni talab qiladi:

har bir bo'linish aktida faqat bitta bazani qo'llash kerak;

bo'linish mutanosib bo'lishi kerak, ya'ni. aniq kontseptsiyalarning umumiy hajmi

bo'linadigan umumiy kontsepsiya hajmiga teng bo'lishi kerak;

bo'linma a'zolari o'zaro mutlaqo bo'lishi kerak, ularning hajmi bir-biriga zid bo'lmasligi kerak;

bo'linish ketma-ket bo'lishi kerak.

tashqi xususiyatga ko'ra tuzilgan va ob'yektlar (jarayonlar, hodisalar) to'plamini kerakli tartibda berishga xizmat qiladigan yordamchi (sun'iy) tasnif;

ob'yektlar va hodisalarning ichki hamjamiyatini tavsiflovchi muhim xususiyatlarga ko'ra tuzilgan tabiiy tasnif. Bu ilmiy izlanishlarning natijasi va muhim vositasidir, chunki tasniflangan ob'yektlarning naqshlarini o'rganish natijalarini taklif qiladi va birlashtiradi. Tanlangan xususiyatlarga, ularning kombinatsiyasiga va tushunchalarni bo'lish tartibiga qarab tasniflash quyidagicha bo'lishi mumkin:

oddiy - umumiy tushunchani faqat barcha turlar ochilgunga qadar va faqat bir marta bo'lish. Bunday tasniflashning misoli dixotomiya bo'lib, unda faqat ikkita tushuncha bo'linish a'zolari bo'lib, ularning har biri boshqasiga ziddir (ya'ni, printsipga rioya qilinadi: "A va A emas");

murakkab - bitta kontsepsiyani turli asoslar bo'yicha ajratish va bunday sodda bo'linishlarning yaxlitligini birlashtirish. Bunday tasniflashga misol kimyoviy elementlarning davriy jadvali.

Tasniflash deganda biz predmetlarni (kuzatuvlar, voqealar) oldindan ma'lum sinflardan biriga tayinlashni anglatadi.

Klassifikatsiya - bu ma'lum bir guruhning xususiyatlarini aniqlash to'g'risida xulosa chiqarish imkonini beradigan naqsh. Shunday qilib, tasniflash uchun, u yoki bu hodisa yoki ob'yekt tegishli bo'lgan guruhni tavsiflovchi belgilar bo'lishi kerak (odatda, ba'zi qoidalar allaqachon tasniflangan hodisalarni tahlil qilish asosida tuzilgan).

Klassifikatsiya boshqariladigan yoki boshqariladigan ta'lim deb ham yuritiladigan, nazorat qilinadigan o'quv strategiyasini anglatadi.

Klassifikatsiya vazifasi odatda doimiy va / yoki kategoriyali o'zgaruvchilar namunasiga asoslangan kategoriyaga bog'liq o'zgaruvchini (ya'ni, toifaga bog'liq bo'lgan o'zgaruvchini) bashorat qilish deb nomlanadi.

Masalan, firmaning mijozlaridan qaysi biri ma'lum bir mahsulotni potentsial xaridor ekanligini va kim bo'lmaganligini, kompaniyaning xizmatlaridan kim foydalanishini va kim xohlamasligini va hokazolarni taxmin qilishingiz mumkin. Muammoning bu turi ikkilik tasniflash muammolariga tegishli bo'lib, bunda bog'liq o'zgaruvchi faqat ikkita qiymatni olishi mumkin (masalan, ha yoki yo'q, 0 yoki 1). Boshqa klassifikatsiya opsiyasi, agar bog'liq o'zgaruvchi oldindan belgilangan sinflar to'plamidan qiymatlarni olishi mumkin bo'lsa, paydo bo'ladi. Masalan, mijoz qaysi markadagi avtomobilni sotib olishni xohlashini oldindan aytib berish kerak bo'lganda. Ushbu holatlarda ko'plab o'zgaruvchilar bog'liq bo'lgan o'zgaruvchi uchun ko'rib chiqiladi. Klassifikatsiya bir o'lchovli (bitta atribut) va ko'p o'lchovli (ikki yoki undan ortiq atribut) bo'lishi mumkin.

Ko'p o'lchovli klassifikatsiya biologlar tomonidan organizmlarni tasniflash uchun diskriminatsiya masalalarini hal qilish uchun ishlab chiqilgan. Ushbu yo'nalishga bag'ishlangan dastlabki ishlardan biri R. Fisher (1930) ning ishi bo'lib, unda organizmlar fizik parametrlarini o'lchash natijalariga qarab kichik turlarga bo'lingan. Biologiya ko'p o'lchovli tasnif usullarini ishlab chiqish uchun eng mashhur va qulay muhit bo'lib kelgan va shunday bo'lib qolmoqda.

Klassifikatsiya muammosini oddiy misol yordamida ko'rib chiqamiz. Aytaylik, sizning yoshingiz va oylik daromadingiz to'g'risida ma'lumotlarga ega bo'lgan sayyohlik agentligining mijozlari ma'lumotlar bazasi mavjud. Reklama materiallarining ikki turi mavjud: qimmatroq va qulay dam olish va arzonroq, yoshlar ta'tili. Shunga ko'ra, mijozlarning ikkita klassi aniqlanadi: 1-sinf va 2-sinf. Ma'lumotlar bazasi 5.1.1-jadvalda keltirilgan.

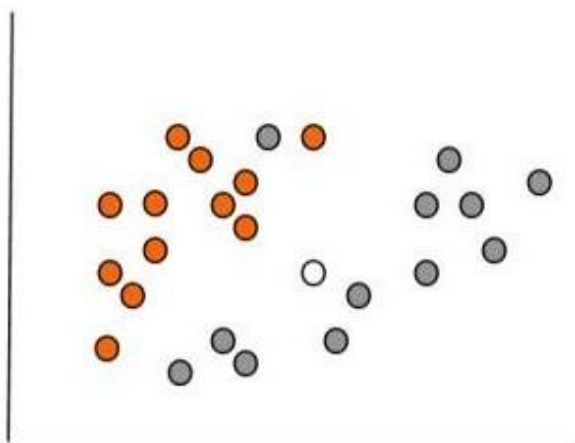
Mijoz ID	Yoshi	Foyda	Sinf
1	18	25	1
2	22	100	1
3	30	70	1
4	32	120	1
5	24	15	2

6	25	22	1
7	32	50	2
8	19	45	2
9	22	75	1
10	40	90	2

5.1.1-jadval. Sayyohlik agentligining mijozlar bazasi

Vazifa. Yangi mijoz qaysi sinfga tegishli ekanligini va qaysi turdagi reklama materiallarini yuborishi kerakligini aniqlang.

Aniqlik uchun, keling, bizning ma'lumotlar bazamizni 1-sinfga (apelsin yorlig'i) va 2-sinfga (kul rang yorliq) tegishli ob'yektlar to'plami sifatida ikki o'lchovli o'lchovda (yosh va daromad) taqdim etamiz. Shaklda 5.1.1-da ikkita sinfdagi ob'yektlar ko'rsatilgan.



Shakl: 5.1.1. 2D ichida bir nechta ma'lumotlar bazasi ob'yektlari

Bizning muammoni hal qilish oq yorlig'i bilan rasmda ko'rsatilgan yangi mijoz qaysi sinfga tegishli ekanligini aniqlash bo'ladi.

Klassifikatsiya jarayoni. Klassifikatsiya jarayonining maqsadi predmetli atributlarni kirish sifatida ishlatadigan va bog'liq bo'lgan atributning qiymatini oladigan modelni yaratishdir. Tasniflash jarayoni muayyan mezon bo'yicha ob'yektlar to'plamini sinflarga bo'lishdan iborat.

Tasniflagich(Klassifikator) - bu atributlar vektoriga tegishli predmetning qaysi sinfga tegishli ekanligini aniqlaydigan muayyan ob'yekt.

Klassifikatsiyani matematik usullardan foydalangan holda amalga oshirish uchun

tasniflashning matematik apparati yordamida boshqarilishi mumkin bo'lgan ob'yektning rasmiy tavsifi bo'lishi kerak. Bizning holatda, bunday tavsif ma'lumotlar bazasi. Har bir ob'yekt (ma'lumotlar bazasi yozuvi) ob'yektning ba'zi mulki haqida ma'lumotni olib yuradi. Dastlabki ma'lumotlar to'plami (yoki ma'lumotlar namunasi) ikkita to'plamga bo'lingan: o'quv va sinov.

O'quv to'plami - bu modelni tayyorlash (qurish) uchun ishlatiladigan ma'lumotlarni o'z ichiga olgan to'plam. Ushbu to'plam misollar kirish va chiqish (maqsad) qiymatlarini o'z ichiga oladi. Chiqish qiymatlari modelni o'qitish uchun mo'ljallangan. Sinov to'plamida, shuningdek, misollar kirish va chiqish qiymatlari mavjud. Bu erda, chiqish qiymatlari modelning sog'lig'ini sinash uchun ishlatiladi.

Foydalanilgan adabiyotlar:

1. Javlon, Xolmatov, and Salimov Jamshid Obid o'g'li. "Классификатор движения рук с использованием биомиметического распознавания образов с помощью сверточных нейронных сетей с методом динамического порога для извлечения движения с использованием датчиков EF." *Journal of new century innovations* 19.6 (2022): 352-357.
2. Amanbayevna, Anarova Shahzoda, and Muxtorov Doston Naim o'g'li. "GEOMETRIC MODELING AND VISUALIZATION OF SELF-SIMILAR STRUCTURES BASED ON FRACTAL THEORY." *JOURNAL OF INNOVATIONS IN SCIENTIFIC AND EDUCATIONAL RESEARCH* 5.4 (2022): 187-188.
3. Akbar, Qudratov, and Muxtorov Doston. "BIR QATLAMLI SUN'IY NEYRON TO'RLARI." *International Journal of Contemporary Scientific and Technical Research* (2023): 214-219.
4. Sherali, Qo'chqorov, and Muxtorov Doston. "BERILGANLARNI KLAUSTERLASH MASALASI." *International Journal of Contemporary Scientific and Technical Research* (2023): 268-273.
5. Javohir, Mamirov, and Muxtorov Doston. "SUN'IY NEYRON TARMOQLARINI O'RGATISH ALGORITMLARI." *International Journal of Contemporary Scientific and Technical Research* (2023): 125-130.

6. Анарова, Шахзода, and Достон Мухторов. "ТИББИЙ ТУЗИЛИШЛИ МУРАККАБ ОБЪЕКТЛАРНИНГ ФРАКТАЛ ЎЛЧОВЛАРНИ АНИҚЛАШ." *International Journal of Contemporary Scientific and Technical Research* (2023): 196-200.
7. Doston, Muxtorov, and Shermuhammedov Abdulatif. "Sun'iy intellekt texnologiyalari va ularni sohalarda qo'llanilishi." *Educational Research in Universal Sciences* 1.5 (2022): 144-147.
8. Muxtorov, Doston, and Zarina Burxonova. "ОСНОВНЫЕ ПРОБЛЕМЫ ИНФОРМАТИКИ." *Молодые ученые* 1.1 (2023): 54-56.