

11-§. Ma'lumotlarni intellectual qidirish (data mining).

Reja:

- 1. Ma'lumotlarni intellectual qidirish (Data Mining)*
- 2. Ma'lumotlarni qidirish usullari*

Ma'lumotni intellektual qidirish odatda tavsiflovchi va bashorat qiluvchi kuchni yaratish uchun ma'lumotlarni yig'ishning to'rtta usulidan foydalanadi: regressiya, assotsiatsiya qoidalarini topish, tasniflash va klasterlash. Ma'lumotni intellektual qidirish - bu tendentsiyalar va naqshlarni aniqlash uchun katta miqdordagi ma'lumotlarni tahlil qilish jarayoni.

Intellektual qidiruv axborotni qayerda saqlanishidan qat'i nazar, iloji boricha tezroq topish va unga kirish jarayonini anglatadi. Aqlli qidiruv foydalanuvchiga moslashtirilgan va aniq qidiruv natijalarini taqdim etish uchun sun'iy intellekt (AI) va mashinani o'rganish (ML) kabi kognitiv texnologiyalar yordamida osonlashtiriladi. Intellektual qidiruv ma'lumotlar turiga qarab izlashga yordam beradi. Bu xodimlar va iste'molchilarga ma'lumotlarga tez va oson kirish imkonini beradi. Yakuniy foydalanuvchilar kontentni tashkilot ichida yoki tashqarisida, har qanday formatda, hujjat boshqaruv tizimlarida, ma'lumotlar bazasida yoki veb-sahifada bo'lgan joydan olish uchun aqlli qidiruvdan foydalanishlari mumkin. Odamlar kundalik hayotida katta miqdordagi ma'lumotlardan bilan qilinadi. Minglab mijozlari bo'lgan korxona qancha ma'lumotlarga ega ekanligini va unga bo'lgan talabni tasavvur qiling. Agar biron bir shaklda tashkil qilinmasa, mijozlar o'zlari qidirayotgan narsani topish uchun veb-saytdagi astronomik miqdordagi tarkibni elakdan o'tkazishlari kerak.

Data Mining - istiqbolli modellashtirish usullari, bir vaqtning o'zida ham miqdoriy, ham sifatni tahlil qilishga imkon beruvchi ko'p faktorli hisoblash modellarini olish uchun ma'lumotlar, shu jumladan o'rganilayotgan tizim to'g'risida to'liq bo'lmagan ma'lumotlar asosida yangiyangi ma'lumotlar ishlab chiquchi Big Data va ma'lumotlar bazasining intellectual texnologiyasidir. Data

Mining - inson faoliyatining turli sohalarida qarorlar qabul qilish uchun zarur bo'lgan ma'lumotlarda ilgari noma'lum, ahamiyatsiz, lekin amaliy foydali hamda foydalanish mumkin bo'lgan bilimlarni, ma'lumotlarni aniqlash usullari, to'plamlari hamda texnologiyalar majmuasidir. Data Miningni ma'lumotlarni qidirish usullari bilan hal qilinadigan vazifalari odatda 2 turga bo'linadi:

1. tavsiflash (ing. descriptive)

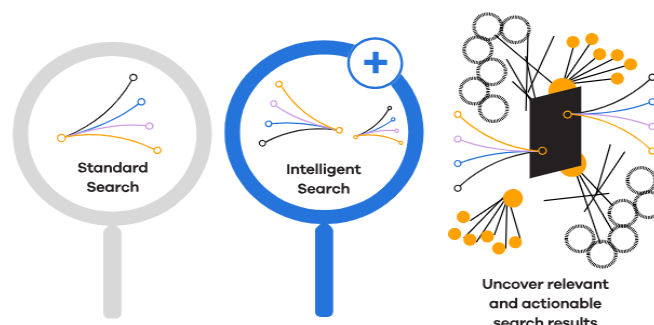
2. bashoratlash (ing. predictive) Data Miningni tavsiflash vazifalari: assotsiatsiya qoidalarini yoki pattern (namunalarni) qidirish, ob'ektlarni guruhlash, klaster tahlil qilish, regressiya modelini yaratish bilan bog'liqdir.

Data Miningni bashoratlash vazifalari: ob'ektlarning tasnifini (oldindan belgilangan sinflar uchun) aniqlash, regressiya tahlili, vaqt qatorlarini tahlil qilish bilan bog'liq.

Data Mining imkoniyatlaridan foydalanilganda quyidagilarni ko'rsatadi:

1. Korrelyatsion tahlil asosida nikoh ma'lumotlarini aniqlashga erishildi.
2. Qarorlar daraxti asosida oilaviy munosabatlar belgilanib olindi.
3. Sun'iy neyron tarmoqlari usuli yordamida oilaviy munosabatlar ajralishlar modeli yaratib olindi.

Natijalarni tahlil qilish shuni ko'rsatdiki, Data Mining ijtimoiy hodisalarni o'rganish imkoniyatlarini tubdan ajratishga, ijtimoiy hodisalarning yangi hisoblash modellarini yaratishga imkon beradi. Ma'lumotlarni intellektual qidirish, shuningdek, ma'lumotlarda bilim kashfiyoti (KDD) sifatida ham tanilgan, katta ma'lumotlar to'plamlaridan naqsh va boshqa qimmatli ma'lumotlarni ochish jarayonidir.



3.1-rasm.Intellektual qidiruv axborot

Biz axborotni haddan tashqari yuklash haqiqatida yashayapmiz, ya'ni aqlli qidiruv foydalanuvchilar va biznes hayotini osonlashtiradigan zarur vositadir.

Intellektual qidiruv foydalanuvchining maqsadlarini tushunish va ushbu maqsadlarning xulosasiga asoslangan natijalarni taqdim etish uchun mo'ljallangan. Bu shuni anglatadiki, aqlli qidiruv to'g'ridan-to'g'ri so'rovlar uchun mosliklarni taqdim etishdan tashqariga chiqadi. U semantik vektor qidiruvining ba'zi elementlaridan, shuningdek, qidiruvni amalga oshirishda foydalanuvchi qo'yadigan maqsadning rasmini yaratish uchun ML texnologiyasining kuchidan foydalanadi.

Foydalanuvchi qidiruv satriga so'zini kiritishi mumkin. Aqlga keladigan savollar: Ehtimol, ular taomning tarixini bilishni xohlashadi. Agar leksik qidiruv usullariga (ma'lumotlar bazasida aniq mosliklarni qidiradigan qidiruv usuli) tayangan bo'lsa, natijalar xilma-xil bo'lar edi va, ehtimol, foydalanuvchi xohlagan narsa bo'lmaydi.

Bu yerda aqlli qidiruv o'ynaydi. Qidiruvning asl maqsadini chuqur tushunish uchun mijoz tomonidan taqdim etilgan turli xil signallarni talab qiladi. Ehtimol, bu ham retsept qidiruvi bo'ladi. Korxona ma'lumotlarini qidirish tizimlari yaqinda paydo bo'lgan hodisa emas. Ular jamoatchilik internetdan foydalanishni boshlashdan ancha oldin paydo bo'lgan.

Ko'p foydalanuvchili asosiy kompyuter tizimlari joriy etilganda, eng muhim afzalliklardan biri bu qidiruv so'rovlarini hujjatlarning katta omborlarida joylashgan matn satrlariga aniq moslashtirish orqali ma'lumotni topishni osonlashtirgan.

Buning ortidan korporativ intranetlar, hamma joyda keng tarqalgan ish stoli kompyuterlari va qidiruv uchun tijorat korxonalari yechimlari davri keldi. Tez orada AltaVista va Google kabi veb-qidiruv tizimlari paydo bo'ldi, ular bepul va hamma uchun ochiq edi. Ushbu qidiruv tizimlari foydalanuvchilarning qidiruv

tizimlaridan kutganlarini o'zgartirib, ma'lumot olish va kontentni topishda inqilob qildi. Shuningdek, ular korporativ qidiruv platformalarini o'zgartirdilar.

Korxonalar nafaqat ma'lumotlar hajmida, balki qidiruv vositalari o'rganishi kerak bo'lgan turli xil tarkiblarda ham tez o'sishga duch keldi. Natijalarni olish tezligi kognitiv qidiruv algoritmlarining asosiy ishlash ko'rsatkichiga aylandi. Zamonaviy ma'lumotlarning hajmi juda katta.

Bugungi kunning aqlli qidiruv echimlari katta ma'lumotlar ish yuklarining talablarini osonlikcha bajara oladigan arxitekturaga asoslangan bo'lishi kerak. Shu sababli, keng qamrovli API asosidagi integratsiyaga ega bulutga asoslangan infratuzilmalar, shuningdek, avtomatlashtirish korporativ qidiruv yechimlari uchun eng yaxshi tanlovdir. Ular bizning tobora kattaroq ma'lumotlarga asoslangan dunyomizda kerakli o'lchov va tezlikni ta'minlaydi.

Aqlli qidiruv qidiruv so'rovi ortidagi maqsad va niyatni qanday tushunishi bilan bog'liq uchta asosiy jarayon ML va tabiiy tilni qayta ishlash (NLP) va semantik vektor qidiruvidir. NLP va ML ikkalasi ham bashorat qilish qobiliyatini faollashtirish va har bir foydalanuvchi uchun moslashtirilgan qidiruv natijalarini taqdim etish uchun ishlaydi. Boshqa tomondan, semantik vektor qidiruvi foydalanuvchilarning xatti-harakatlaridan bilim to'playdi va umumiy vektor maydonida mahsulot va so'rovlarni kodlaydi. Ushbu usullar turli xil ma'lumotlar manbalarini bog'laydi va ularni oson topish va so'rov qilish uchun indekslaydi. O'rnatilgan algoritmlar so'ralganda to'g'ri javob beradi. Bu qadamlar quyidagilarni o'z ichiga oladi:

- Turli ma'lumotlar manbalarini, masalan, CRM'larni, hujjat boshqaruv tizimlarini yoki bulut almashish echimlarini ulash
- Manbadan qat'iy nazar qidiruv natijalarini tartiblash imkonini beruvchi yagona qidiruv indeksi orqali tarkibni indekslash
- Hujjatlarni identifikatsiyalash va tasniflash uchun ma'lumot olish orqali ularni aqlli tushunish

Ushbu uchta texnologiya foydalanuvchi oladigan aqlli qidiruv tajribasini yaratish uchun tandemda ishlaydi.

Nazorat savollari	1.Ma'lumotlarni intellektual qidirish jarayonlarini izohlang.
	2.Ma'lumotlarni qidirish usullari jarayonlarini tahlil qiling.



Mavzuni mustahkamlash uchun savollar.

1.Quyidagi javoblardan qaysi birida kalitli maydonning vazifasi to'g'ri ifodalangan?

- a) Yozuvni bir qiymatli aniqlash.
- b) Jadval ustuni sarlavhasini aniqlash.
- c) Ma'lumotlarni kiritishning to'g'riligini tekshirish uchun cheklash o'rnatish.
- d) Jadval ob'ektini bir qiymatli aniqlash.

2.Relyastion baza tushunchasini kiritgan qaysi olim hisoblanadi?

- a) Kodd
- b) Shennon
- c) Atanasov
- d) Stivenson

3.Predmet sohasi – bu?

- a) boshqarishni tashkil qilish va uni avtomatlashtirish maqsadida o'rganilishi lozim bo'lgan ob'ektiv dunyoning ma'lum bir qismidir.

- b) MB strukturasi
- c) Jadvallar orasidagi bog'lanish
- d) Ma'lumotlar sxemasi