

16-§. Ma'lumotlarni qidirish vazifalari va tasniflash

Reja:

- 1. Ma'lumotlarni qidirish vazifalari va tasniflash.*
- 2. Ma'lumotlarni qidirish muommalarining yechimlari.*

Zamonaviy kompyuter atamasi Data Mining "axborot qazib olish" yoki "ma'lumotlar qazib olish" deb tarjima qilingan. Ko'pincha Data Mining bilan bir qatorda Knowledge Discovery ("bilimlarni kashf etish") va Data Warehouse ("ma'lumotlar ombori") atamalari topiladi. Data Miningning ajralmas qismi bo'lgan ushbu atamalarning paydo bo'lishi ma'lumotlarni qayta ishlash va saqlash vositalari va usullarini ishlab chiqishning yangi bosqichi bilan bog'liq. Shunday qilib, Data Mining maqsadi katta (juda katta) hajmdagi ma'lumotlarda yashirin qoidalar va naqshlarni aniqlashdir.

Gap shundaki, inson ongining o'zi turli xil ma'lumotlarning ulkan massivlarini idrok etishga moslashmagan. O'rtacha, bir kishi, ba'zi shaxslar bundan mustasno, hatto kichik namunalarda ham ikki yoki uchtdan ortiq munosabatlarni qo'lga kirita olmaydi. Ammo uzoq vaqt davomida ma'lumotlarni tahlil qilish uchun asosiy vosita rolini o'ynagan an'anaviy statistika ham real hayotdagi muammolarni hal qilishda muvaffaqiyatsizlikka uchraydi. U ko'pincha xayoliy qiymatlar bo'lgan namunaning o'rtacha xarakteristikalarini bilan ishlaydi (mijozning o'rtacha to'lov qobiliyati, agar xavf funksiyasi yoki yo'qotish funksiyasiga qarab, mijozning to'lov qobiliyati va niyatlarini bashorat qilish kerak bo'lganda; o'rtacha signal intensivligi, siz xarakteristikalarini va signal cho'qqilari fon manfaatdor esa, va hokazo. d.).

Shuning uchun matematik statistika usullari asosan oldindan tuzilgan gipotezalarni sinab ko'rish uchun foydalidir, gipotezani aniqlash esa ba'zan ancha murakkab va ko'p vaqt talab qiladigan vazifadir. Zamonaviy Data Mining texnologiyalari heterojen ko'p o'lchovli ma'lumotlarning har qanday

bo'laklariga xos bo'lgan naqshlarni (naqshlarni) avtomatik ravishda izlash uchun ma'lumotlarni qayta ishlaydi. Onlayn tahliliy ma'lumotlarni qayta ishlashdan (OLAP) farqli o'laroq, Data Mining-da gipotezalarni shakllantirish va noodatiy (kutilmagan) naqshlarni aniqlash yuki odamdan kompyuterga o'tkaziladi.

Ma'lumotni qazib olish - bu bitta emas, balki ko'p sonli turli xil bilimlarni kashf qilish usullarining kombinatsiyasi. Usulni tanlash ko'pincha mavjud ma'lumotlar turiga va qanday ma'lumotlarni olishga harakat qilayotganingizga bog'liq. Bu erda, masalan, ba'zi usullar:

- assotsiatsiya (birlashtirish)
- tasniflash
- klasterlash
- vaqt seriyalarini tahlil qilish
- prognozlash
- neyron tarmoqlar va boshqalar.

Ma'lumotlarni qazib olish — bu ma'lumotlarni tavsiflashdan ko'ra uni modellashtirish va kashf etishga qaratilgan ma'lumotlarni tahlil qilishning o'ziga xos usuli. Biznes razvedka yig'ishga asoslangan ma'lumotlar tahlilini qamrab oladi. Statistik ma'noda, ba'zida ma'lumotlar tahlilini tavsiflovchi statistika, kashfiyot ma'lumotlarini tahlil qilish va statistik gipotezani tekshirishga ajratadilar. Tashkilot ma'lumotlarini tahlil qilish ma'lumotlarning yangi xususiyatlarini aniqlash bilan bog'liq bo'lsa, statistik gipotezalarni tekshirish mavjud farazlarni tasdiqlash yoki rad etish bilan bog'liqdir. Bashoratli tahlil bashorat qilish yoki tasniflash uchun statistik yoki tizimli modellarni qo'llashga qaratilgan bo'lsa, matn tahlili tuzilmagan ma'lumotlarga tegishli matn manbalaridan ma'lumotlarni ajratib olish va tasniflash uchun statistik, lingvistik va tizimli usullarni qo'llaydi. Bu ma'lumotlarni tahlil qilishning barcha turlaridir.

Ma'lumotlar integratsiyasi ma'lumotlar tahlilining asosi hisoblanadi va ma'lumotlar tahlilining o'zi ma'lumotlarni vizualizatsiya qilish va ma'lumotlarni

tarqatish bilan chambarchas bog'liq. "Ma'lumotlar tahlili" atamasi ba'zan ma'lumotlarni modellashtirish uchun sinonim sifatida ishlatiladi.

Ma'lumotlarni qidirish - ixtiyoriy ma'lumotlar bazasida (massiv, to'plam, ma'lumotlar bazasi) berilgan xususiyatga ega bo'lgan bir yoki bir nechta elementlarni aniqlash usullarini o'rganadigan informatika bo'limi. Bu xususiyat turli yo'llar bilan belgilanishi mumkin: berilgan naqsh bilan moslik sifatida, boshqa elementlarga nisbatan ma'lum sifatning mavjudligi (masalan, barcha elementlar orasida maksimalni qidirish) yoki murakkab mezon sifatida (masalan, shaxsiy ma'lumotlar bazasida potentsial mijozlarni qidirish).

Kerakli elementlarning yo'qligini aniqlash ham haqiqiy qidiruv natijasi hisoblanadi. Ma'lumotlarni qidirishdan farqli o'laroq, axborotni qidirishning mazmuni ob'ektlarni: tuzilmagan ma'lumotlarni saqlashda mavjud bo'lgan faktlar, hujjatlar, tasvirlarni topishdir.

Axborot olish uchun muhim ahamiyatga ega - so'rov va qidiruv natijalari qanchalik mos kelishi. Ma'lumotlarni qidirish va ma'lumot qidirishni amalga oshirish uchun qidiruv tizimi qo'llaniladi, bu qidiruvning asosiy bosqichlarini amalga oshiradigan dasturiy ta'minot va lingvistik vositalar to'plamidir: ma'lumotni dastlabki tahlil qilish, qidiruv so'rovini shakllantirish, qidiruvning o'zi va taqdimoti. uning natijalari.

Ma'lumotlarni qidirish vazifalari ma'lumotlarni tahlil qilish muammolari bilan chambarchas bog'liq. Masalan, ba'zi bir ko'rsatkichning optimal qiymatini ta'minlaydigan o'zgaruvchan qiymatlarni qidirish ma'lum bir qiymatlar to'plamini, ba'zan esa butun to'plamni tahlil qilishni o'z ichiga oladi. Shu bilan birga, ma'lumotlarni tahlil qilish qidiruvdan tashqariga chiqadi va asboblari va usullarning ancha katta arsenaliga ega. Intellektual ma'lumotlarni tahlil qilish texnologiyalari odatda Data Mining deb nomlangan fan doirasida ko'rib chiqiladi. Mining inglizchada "kon qazish" degan ma'noni anglatadi va Data Mining iborasi katta hajmdagi ma'lumotlardan foydali bilimlarni olish uchun metaforadir.

Bilim yangi, ilgari noma'lum bo'lishi kerak. Foydalanuvchiga allaqachon ma'lum bo'lgan bilimlarni kashf qilish uchun sarflangan harakatlar o'zini oqlamaydi. Shuning uchun, bu yangi, ilgari noma'lum bo'lgan bilimdir. Bilim ahamiyatsiz bo'lishi kerak. Tahlil natijalari yashirin bilimlarni tashkil etuvchi ma'lumotlarda aniq bo'lmagan, kutilmagan naqshlarni aks ettirishi kerak. Oddiyroq usullar bilan (masalan, vizual ko'rish) olinishi mumkin bo'lgan natijalar kuchli Data Mining usullaridan foydalanishni oqlamaydi.

Bilim amaliy jihatdan foydali bo'lishi kerak. Topilgan bilimlar, shu jumladan yangi ma'lumotlarga nisbatan, etarlicha yuqori darajadagi ishonchlilik bilan qo'llanilishi kerak. Foydaliligi shundaki, bu bilim uni qo'llashda qandaydir foyda keltirishi mumkin. Bilim inson tushunchasi uchun ochiq bo'lishi kerak. Topilgan naqshlar mantiqan tushuntirilishi kerak, aks holda ular tasodifiy bo'lish ehtimoli mavjud. Bundan tashqari, kashf etilgan bilimlar inson tushunadigan shaklda taqdim etilishi kerak.

Data Mining-da modellar olingan bilimlarni ifodalash uchun ishlatiladi. Modellarning turlari ularni yaratish usullariga bog'liq. Eng keng tarqalganlari: qoidalar, qarorlar daraxtlari, klasterlar va matematik funktsiyalar. Data Mining ko'lami cheklanmagan - Data Mining har qanday ma'lumot mavjud bo'lgan joyda kerak bo'ladi. Ko'pgina bunday korxonalar tajribasi shuni ko'rsatadiki, Data Mining-dan foydalanish rentabelligi 1000% ga yetishi mumkin. Masalan, 350 dan 750 ming dollargacha bo'lgan dastlabki xarajatlardan 10-70 baravar yuqori iqtisodiy samara haqida xabarlar mavjud. 20 million dollarlik loyiha haqida ma'lumot berilgan, u bor-yo'g'i 4 oyda o'zini oqladi. Yana bir misol - yiliga 700 000 dollar tejash. Buyuk Britaniyadagi supermarketlar tarmog'ida Data Mining joriy etish orqali. Ma'lumotlarni qidirish menejerlar va tahlilchilar uchun ularning kundalik faoliyatida katta ahamiyatga ega. Ishbilarmonlar Data Mining usullari yordamida sezilarli raqobatdosh ustunliklarga ega bo'lishlari mumkinligini tushundilar.

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

Ma'lumotlarni qidirish vazifalarining tasnifi. DataMining usullari tahlilchi duch keladigan ko'plab muammolarni hal qilishga imkon beradi. Ulardan asosiylari: tasniflash, regressiya, assotsiatsiya qoidalarini izlash va klasterlash. Quyida ma'lumotlarni tahlil qilishning asosiy vazifalarining qisqacha tavsifi keltirilgan.

- Tasniflash vazifasi ob'ektning xususiyatlariga ko'ra sinfini aniqlashga qisqartiriladi. Shuni ta'kidlash kerakki, bu masalada ob'ektni belgilash mumkin bo'lgan sinflar to'plami oldindan ma'lum.
- Regressiya vazifasi, xuddi tasniflash vazifasi kabi, ob'ektning ma'lum xususiyatlaridan kelib chiqib, uning ayrim parametrlarining qiymatini aniqlash imkonini beradi. Tasniflash masalasidan farqli o'laroq, parametrning qiymati cheklangan sinflar to'plami emas, balki haqiqiy sonlar to'plamidir.
- Assotsiatsiya vazifasi. Assotsiatsiya qoidalarini qidirishda maqsad ob'ektlar yoki hodisalar o'rtasidagi tez-tez bog'liqliklarni (yoki assotsiatsiyalarni) topishdir. Topilgan bog'liqliklar qoidalar ko'rinishida taqdim etilgan va tahlil qilingan ma'lumotlarning mohiyatini yaxshiroq tushunish uchun ham, voqealar sodir bo'lishini bashorat qilish uchun ham ishlatilishi mumkin.
- Klasterlashning vazifasi mustaqil guruhlar (klasterlar) va ularning xususiyatlarini tahlil qilinayotgan ma'lumotlarning butun majmuasida izlashdan iborat. Ushbu muammoni hal qilish ma'lumotlarni yaxshiroq tushunishga yordam beradi. Bundan tashqari, bir hil ob'ektlarni guruhlash ularning sonini kamaytirishga va natijada tahlilni osonlashtirishga imkon beradi.

- Ketma-ket naqshlar - vaqt bilan bog'liq bo'lgan hodisalar o'rtasida qonuniyatlarni o'rnatish, ya'ni. Agar X voqea sodir bo'lsa, Y hodisasi ma'lum vaqtdan keyin sodir bo'lishiga bog'liqlikni aniqlash.

Og'ishlarni tahlil qilish - eng xarakterli bo'lmagan naqshlarni aniqlash. Sanab o'tilgan vazifalar maqsadiga ko'ra tavsiflovchi va bashoratli bo'linadi.

Tasniflash – o'rganilayotgan predmetlar, hodisalar, jarayonlarni jinsi, turi, tipi bo'yicha, ularni o'rganish qulayligi uchun har qanday muhim belgilar bo'yicha tizimli taqsimlash; ushbu o'xshashlik darajasini aks ettiruvchi dastlabki tushunchalarni guruhlash va ularni belgilangan tartibda joylashishi.

Tasniflash – o'xshash sinflashtirish belgilariga (bir yoki bir nechta belgilar) ega ba'zi bir tamoilga muvofiq ob'ektlar to'plami bo'lib, ushbu ob'ektlar o'rtasidagi o'xshashlik yoki farqni aniqlash uchun tanlangan. Tasniflash quyidagi qoidalarni o'z ichiga oladi:

- bo'linmaning har bir aktida faqat bir asosni qo'llash zarur;
- bo'linma mutanosib bo'lishi kerak, ya'ni turga oid tushunchalarning hajmi bo'linuvchi turga oid tushunchaga teng bo'lishi kerak;
- bo'linma hadlari o'zaro bir-birini chiqarib tashlashi kerak, ularning hajmi bir-biriga zid bo'lmasligi kerak;
- bo'linma ketma-ket bo'lishi kerak.

Shu bilan farqlanadiki:

- tashqi belgi bo'yicha amalga oshiriluvchi va predmetlar (jarayonlar, hodisalar) to'plamini kerakli tartibda berish uchun xizmat qiluvchi yordamchi (sun'iy) tasniflash;
- predmetlar va hodisalarni ichki umumiylikni tavsiflovchi muhim belgilar bo'yicha amalga oshiriluvchi tabiiy tasniflash.

Tanlangan belgilarga, ularning birlashishi va bo'lish tushunchalariga bog'liq holda tasniflash quyidagicha bo'lishi mumkin:

- oddiy – turga oid tushunchani faqatgina belgi bo'yicha va barcha turlarni ochishdan avval faqat bir marta bo'lish. Bunday tasniflashga misol bo'lib

dixotomiya bo'lib, unda faqat ikkita tushuncha bo'linma hadlari bo'lib, har biri boshqasiga ziddir (ya'ni tamoilga rioya qilinadi: "A va A emas");

- murakkab – bir tushunchani turlicha asoslar bo'yicha bo'lish va bunday bo'lishlarni bir butun holda sintez qilish qo'llaniladi.

Bunday tasniflashga misol bo'lib, kimyoviy elementlarni davriy sistemasi bo'lib hisoblanadi. Tasniflash deganda ob'ektlarni (kuzatishlarni, hodisalarni) ma'lum bir sinflardan biriga tayinlash tushuniladi.

Tasniflash – bu ma'lum bir guruhning xususiyatlarini aniqlash to'g'risidagi nisbiy xulosa chiqarishga imkon beruvchi qonuniyat. Shunday qilib, tasniflashni o'tkazish uchun u yoki boshqa hodisa yoki ob'ektga tegishli bo'lgan guruhni tavsiflovchi belgilar mavjud bo'lishi zarur (odatda, bunda allaqachon sinflashtirilgan hodisalarning tahlili natijasida ba'zi qoidalar tuziladi). Tasniflash nazorat qilinuvchi yoki boshqariluvchi o'qitish deb ham nomlanuvchi o'qituvchili o'qitish (supervised learning) strategiyasini anglatadi. Tasniflash masalalari odatda uzluksiz va/yoki kategoriyali o'zgaruvchilar asosida kategoriyali bog'liq o'zgaruvchini avvaldan aytish (ya'ni, kategoriya bo'lib hisoblangan bog'liq o'zgaruvchi) deb nomlanadi. Masalan, oldindan aytish mumkinki, kim firmaning mijozlaridan bo'lsa ma'lum bir mahsulotning potensial xaridori bo'lib hisoblanadi, kimdir esa – yo'q, kimdir firmaning xizmatidan foydalanadi, kimdir esa – yo'q, va hokazo. Tasniflashning boshqa turi, agar bog'liq o'zgaruvchi oldindan belgilangan sinflarni ba'zi to'plamlaridan qiymatni qabul qilishi mumkin bo'lsa, yuzaga kelishi mumkin. Masalan, mijoz avtomobilni qanday markalisini xohlashini oldindan aytib berish mumkin bo'lgan paytda. Bunday hollarda bog'liq o'zgaruvchi uchun sinflar to'plami ko'rib chiqiladi.

Tasniflash bir o'lchamli (bir belgisi bo'yicha) va ko'p o'lchamli (ikki yoki undan ko'proq belgilar bo'yicha).

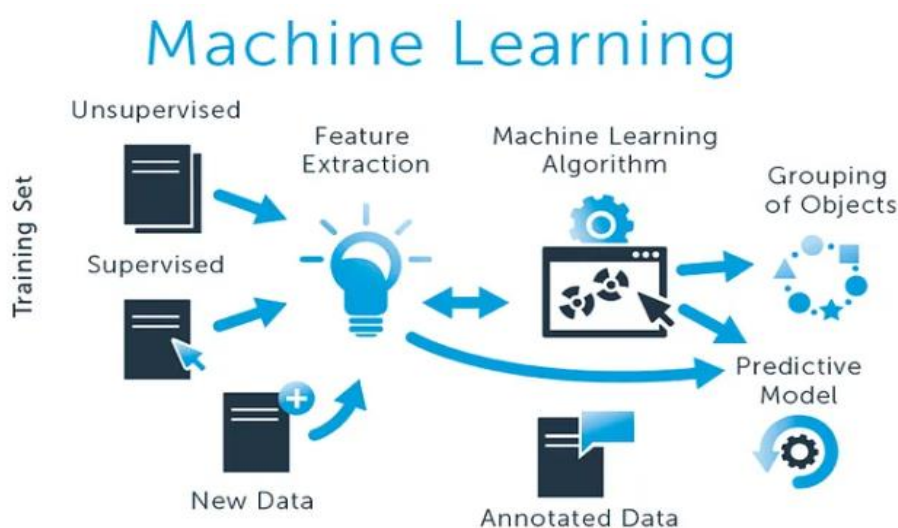
Tasniflash jarayoni. Tasniflash jarayonining maqsadi shundan iboratki, kirish parametrlari sifatida bashorat qilinuvchi atributlardan foydalaniladigan va

bog'liq atribut qiymati olinadigan modelni qurish. Tasniflash jarayoni ob'ektlar to'plamini ma'lum mezon bo'yicha sinflarga bo'lishdan iborat.

Ta'riflash vazifalari tahlil qilinayotgan ma'lumotlarni tushunishni yaxshilashga qaratilgan. Bunday modellardagi asosiy nuqta - natijalarning inson idroki uchun qulayligi va shaffofligi. Ushbu turdagi muammolar klasterlash va assotsiatsiya qoidalarini qidirishni o'z ichiga oladi.

Bashoratli masalalarni yechish ikki bosqichga bo'linadi. Birinchi bosqichda ma'lum natijalarga ega bo'lgan ma'lumotlar to'plamiga asoslangan model quriladi. Ikkinchi bosqichda u yangi ma'lumotlar to'plamiga asoslangan natijalarni bashorat qilish uchun ishlatiladi. Bunday holda, albatta, qurilgan modellarning iloji boricha aniq ishlashi talab qilinadi. Ushbu turdagi muammolar tasniflash va regressiya muammolarini o'z ichiga oladi. Bu, shuningdek, assotsiatsiya qoidalarini topish muammosini o'z ichiga oladi, agar uni hal qilish natijalaridan ma'lum voqealar sodir bo'lishini taxmin qilish uchun foydalanish mumkin bo'lsa.

Muammolarni hal qilish usullariga ko'ra ular nazorat ostida o'qitish (o'qituvchi bilan o'qitish) va nazoratsiz ta'lim (o'qituvchisiz o'rganish) ga bo'linadi. Bu nom ingliz tilidagi adabiyotlarda tez-tez ishlatiladigan va barcha Data Mining texnologiyalariga ishora qiluvchi Machine Learning (mashinalarni o'rganish) atamasidan kelib chiqqan.



3.9-rasm.Mashinali o'qitishda ma'lumotlarni tahlil qilish.

Nazorat ostida o'qitishda ma'lumotlarni tahlil qilish muammosi bir necha bosqichda hal qilinadi. Birinchidan, har qanday Data Mining algoritmidan foydalanib, tahlil qilingan ma'lumotlarning modeli - klassifikator quriladi. Keyin klassifikator o'qitiladi. Boshqacha aytganda, uning ish sifati tekshiriladi va agar u qoniqarsiz bo'lsa, tasniflagich qo'shimcha ravishda o'qitiladi. Bu talab qilinadigan sifat darajasiga erishilgunga qadar yoki tanlangan algoritim ma'lumotlar bilan to'g'ri ishlamasligi yoki ma'lumotlarning o'zi aniqlanishi mumkin bo'lgan tuzilishga ega emasligi aniq bo'lgunga qadar davom etadi. Ushbu turdagi muammolar tasniflash va regressiya muammolarini o'z ichiga oladi.

Data Mining-da tasniflash vazifasi tahlil qilinadigan ob'ektning parametrlaridan birining qiymatini boshqa parametrlarning qiymatlari asosida aniqlash vazifasi sifatida qaraladi. Aniqlanayotgan parametr ko'pincha bog'liq o'zgaruvchi deb ataladi va uni aniqlashda ishtirok etuvchi parametrlar mustaqil o'zgaruvchilar deb ataladi. Ko'rib chiqilgan misollarda mustaqil o'zgaruvchilar quyidagilar edi:

- ish haqi, yoshi, bolalar soni va boshqalar;
- ma'lum so'zlarning chastotasi;
- matritsa piksel rang qiymatlari.

Ushbu misollardagi bog'liq o'zgaruvchilar:

- mijozning kreditga layoqatliligi (ushbu o'zgaruvchining mumkin bo'lgan qiymatlari "ha" va "yo'q");
- xabar turi (ushbu o'zgaruvchining mumkin bo'lgan qiymatlari "spam" va "pochta");
- rasm raqami (ushbu o'zgaruvchining mumkin bo'lgan qiymatlari 0, 1,..., 9).

Tasniflash va regressiya muammosi ikki bosqichda hal qilinadi.

Birinchisi - mashg'ulot to'plami. U mustaqil va qaram o'zgaruvchilarning qiymatlari ma'lum bo'lgan ob'ektlarni o'z ichiga oladi. Yuqorida tavsiflangan misollarda bunday o'quv namunalari bo'lishi mumkin:

- ilgari turli miqdorlarda kredit olgan mijozlar to'g'risidagi ma'lumotlar va ularning qaytarilishi to'g'risidagi ma'lumotlar;
- qo'lda spam yoki elektron pochta sifatida tasniflangan xabarlar;
- raqamli tasvirlarning ilgari tan olingan matritsalar.

Assotsiatsiya qoidalarini izlash vazifasi - Assotsiatsiya qoidalarini qidirish eng mashhur Data Mining ilovalaridan biridir. Muammoning mohiyati bunday to'plamlarning katta to'plamida tez-tez uchraydigan ob'ektlar to'plamini aniqlashdir. Bu muammo tasniflash muammosining alohida holatidir. Dastlab, supermarketlardagi xaridorlarning xatti-harakatlaridagi tendentsiyalarni tahlil qilishda hal qilindi. Ular tomonidan amalga oshirilgan, xaridorlar aravaga (savatga) joylashtirgan xaridlari haqidagi ma'lumotlar tahlil qilindi. Bu ikkinchi tez-tez uchraydigan nomga sabab bo'ldi - bozor savatlarini tahlil qilish (Basket Analysis).

Assotsiatsiya qoidalarini izlash vazifasi nafaqat savdo sohasida ham dolzarbdir. Misol uchun, xizmat ko'rsatish sohasida mijozlar jami xizmatlardan foydalanishni afzal ko'rishlari qiziqishdir. Ushbu ma'lumotni olish uchun muammo bitta mijoz tomonidan ma'lum vaqt (oy, yil) uchun foydalanilgan xizmatlar to'g'risidagi ma'lumotlarga nisbatan hal qilinadi. Bu, masalan, mijozga taqdim etilayotgan xizmatlarni qanday qilib eng yaxshi paketlashni aniqlashga yordam beradi. Tibbiyotda bemorlarda kuzatilgan alomatlar va kasalliklarni tahlil qilish mumkin. Bunday holda, kasallik va alomatlarning qaysi kombinatsiyasi ko'pincha sodir bo'lishini bilish kelajakda to'g'ri tashxis qo'yishga yordam beradi.

Nazorat savollari	
1.	<ul style="list-style-type: none"> Ma'lumotlarni qidirish vazifalari va tasniflash jarayonlarini izohlang.
2.	<ul style="list-style-type: none"> Ma'lumotlarni qidirish muommalarining yechimlarini izohlang.



Mavzuni mustahkamlash uchun savollar.

1. Ma'lumotlarni qidirishga tegishli bo'lmagan vazifasini aniqlang

- Avtomatik yoki yarim avtomatik vositalar yordamida katta hajmdagi ma'lumotlarni mazmunlilarini aniqlash uchun tadqiq qilish va tahlil qilish
- Ma'lumotlarni qidiruv kalit so'zi bo'yicha topish
- Ma'lumotdagi foydali, ehtimoli kutilganlarini topish
- Ma'lumotlardan noaniq, ilgari noma'lum va potentsial foydali ma'lumotlarni olish

2. Amalda eng kam qo'llaniladigan mashinali o'rganish algoritmlari quyidagilarga asoslangan:

- Mustahkamlash va o'rganish algoritmlari
- Nazoratsiz o'qitish algoritmlari
- Nazorat ostidagi o'rganish algoritmlari
- Konvolyutsion neyron tarmoqlar

3. K-o'rtacha algoritmi qanday muammoni hal qilish uchun mo'ljallangan:

- Klasterlash
- Tasniflar
- Bashoratlar

d) o'lchamlarni qisqartirish