

Tojimamatov Israiljon Nurmamatovich

Farg'ona davlat universiteti

Gmail: israiltojimatov@gmail.com

Umarov Bekzod Azizovich

*Farg'ona davlat universiteti amaliy matematika
va informatika kafedrasio 'qituvchisi,*

Gmail: ubaumarov@mail.ru

Tursunaliyeva Odinoxon Erkinboy qizi

Farg'ona davlat universiteti

Gmail: tursunaliyevaodinaxon69@gmail.com

Annotsiya: *Data Mining (ma'lumotlarni qazib olish) sohasida turli metodlar va bosqichlar mavjud bo'lib, ular katta hajmdagi ma'lumotlardan foydali bilimlarni aniqlashga yordam beradi. Ushbu maqolada, data mining metodlari, uning asosiy bosqichlari va amaliy qo'llanilishi haqida so'z boradi. Maqolada, data mining jarayonining boshlanishidan to yakuniga qadar bo'lgan bosqichlar, shuningdek, turli metodlarning ishlash prinsiplariga e'tibor qaratilgan. Bular orasida tasniflash, regressiya, klasterlash, assotsiativ qoidalar va neyron tarmoqlar kabi metodlar muhokama qilinadi. Har bir metodning o'ziga xos xususiyatlari va qo'llanilish sohalari batafsil ko'rib chiqiladi.*

Kalit so'zlar: *Data Mining, ma'lumotlarni qazib olish, tasniflash, regressiya, klasterlash, assotsiativ qoidalar, neyron tarmoqlar, ma'lumotlar tahlili, bosqichlari.*

Annotation: *In the field of Data Mining (data mining), there are various methods and stages that help to identify useful knowledge from large amounts of data. This article discusses data mining techniques, its main stages, and practical applications. The article focuses on the stages from the beginning to the end of the data mining process, as well as the principles of operation of various techniques. These include discussion of methods such as classification, regression, clustering, associative rules, and neural networks. The specific characteristics and areas of application of each method are examined in detail.*

Keywords: *Data Mining, data mining, classification, regression, clustering, associative rules, neural networks, data analysis, stages.*

Аннотация: *В области интеллектуального анализа данных (интеллектуального анализа данных) существуют различные методы и этапы, которые могут помочь вам извлечь полезные знания из больших объемов данных. В данной статье будут рассмотрены методы Data Mining, его основные этапы и практическое применение. В статье основное внимание уделяется этапам процесса data mining от начала до завершения, а также принципам работы различных методов. Среди них*

обсуждаются такие методы, как классификация, регрессия, кластеризация, ассоциативные правила и нейронные сети. Будут подробно рассмотрены особенности и области применения каждого метода.

Ключевые слова: Интеллектуальный анализ данных, интеллектуальный анализ данных, классификация, регрессия, кластеризация, ассоциативные правила, нейронные сети, анализ данных, этапы.

Kirish: Data Mining — bu katta hajmdagi ma'lumotlardan foydali bilimlarni olish jarayonidir. Ma'lumotlar bazalarida yoki boshqa saqlash tizimlarida mavjud bo'lgan katta hajmdagi ma'lumotlar ichida qandaydir yashirin bilim yoki munosabatlar mavjud bo'lishi mumkin. Ushbu jarayonni amalga oshirishda, turli metodlar yordamida ma'lumotlarni tahlil qilish, qimmatli ma'lumotlar olish va ularni amaliyotda qo'llash mumkin. Data Mining sohasida rivojlanish, asosan, ma'lumotlar bazalari va yuqori tezlikda ishlov beruvchi texnologiyalarni yaratish bilan bog'liq. Har bir metodning o'ziga xos xususiyatlari bor va ular ma'lum bir tahlil turini yoki sohani yoritishda foydalidir. Data mining jarayoni ko'plab sohalarda, jumladan, biznes, tibbiyot, marketing va ilmiy izlanishlarda keng qo'llaniladi.

Data Mining bosqichlari. Data mining jarayoni quyidagi bosqichlardan iborat:

Ma'lumotlarni to'plash va tayyorlash. Data mining jarayonining birinchi bosqichi ma'lumotlarni yig'ish va tayyorlashdir. Ma'lumotlarni to'plash jarayonida, kerakli ma'lumotlar bazasidan yoki boshqa manbalardan olinadi. So'ngra, ma'lumotlar tayyorlanadi: tozalanadi, normallashtiriladi va kerakli formatda bo'lishi ta'minlanadi. Bu bosqichda, ma'lumotlarni tozalash, ularning xatoliklarini aniqlash va ularni tuzatish ishlari amalga oshiriladi.



Ma'lumotlarni tahlil qilish (Exploratory Data Analysis, EDA). Ma'lumotlarni tahlil qilish (EDA) bosqichi, ma'lumotlarning umumiy ko'rinishini tahlil qilish va ularning o'zaro bog'lanishlarini o'rganishdan iborat. Bu bosqichda, statistik tahlil usullari yordamida ma'lumotlar o'rganiladi va muhim xususiyatlar aniqlanadi. Shuningdek, ma'lumotlarni vizualizatsiya qilish orqali ularning tuzilishini yaxshiroq tushunish mumkin.

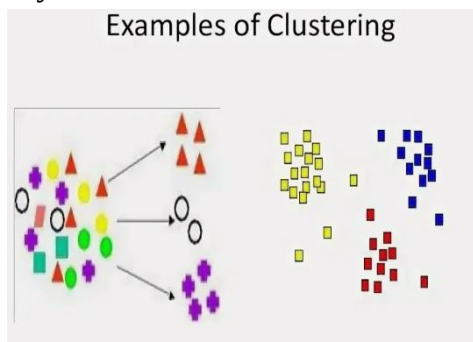
Model qurish. Model qurish jarayoni, ma'lumotlarga asoslangan modelni yaratishni o'z ichiga oladi. Bu bosqichda tasniflash, regressiya, klasterlash va boshqa metodlardan foydalanib, turli maqsadlar uchun modellar quriladi. Modelni qurish jarayonida, avvalgi bosqichlarda aniqlangan xususiyatlar va munosabatlar asosida yangi model yaratilib, sinovdan o'tkaziladi.

Modelni baholash va tasdiqlash. Yaratilgan modelni baholash va tasdiqlash bosqichi, modelning samaradorligini tekshirish uchun o'tkaziladi. Baholashda, odatda, aniqlik, sezgirlik, xatoliklar darajasi va boshqa mezonlar ishlatiladi. Shuningdek, modelni tekshirish uchun trening va test ma'lumotlari yordamida sinovlar o'tkaziladi.

Modelni qo'llash va chiqarish. Modelni qo'llash va chiqarish bosqichida, yaratilgan model real dunyo ma'lumotlari bilan ishlaydi va amaliyotga tatbiq etiladi. Bu bosqichda, modeldan olinadigan natijalar va bilimlar qaror qabul qilish jarayonida foydalaniladi. Masalan, biznesda prognoz qilish, tibbiyotda kasalliklarni aniqlash yoki marketingda mahsulotni taklif etishda modeldan foydalanish mumkin. Data Mining Metodlari. Data mining metodlari turli maqsadlar uchun ishlatiladi. Quyida eng keng tarqalgan metodlar va ularning ishlash prinsiplarini ko'rib chiqamiz.

Tasniflash (Classification). Ushbu uslub o'zining kelib chiqishini mashinani o'rganishda topadi. U ma'lumotlar to'plamidagi elementlar yoki o'zgaruvchilarni oldindan belgilangan guruhlar yoki sinflarga tasniflaydi. U boshqa usullar qatorida ma'lumotlarni qazib olishda chiziqli dasturlash, statistika, qarorlar daraxtlari va sun'iy neyron tarmog'idan foydalanadi. Tasniflash ma'lumotlar to'plamidagi narsalarni turli sinflarga tasniflash qobiliyatiga ega bo'ladigan tarzda modellashirilishi mumkin bo'lgan dasturiy ta'minotni ishlab chiqish uchun ishlatiladi. Masalan, biz undan suhbatda qatnashgan barcha nomzodlarni ikki guruhga ajratish uchun foydalanishimiz mumkin -birinchi guruh tanlangan nomzodlar ro'yxati, ikkinchisi esa rad etilgan nomzodlar ro'yxati. Ushbu tasniflash ishini bajarish uchun ma'lumotlarni qidirish dasturidan foydalanish mumkin.

Regressiya (Regression). Regressiya usuli, biron bir o'zgaruvchining boshqa o'zgaruvchilarga qanday ta'sir qilishini aniqlash uchun ishlatiladi. Misol uchun, sotuvlarni prognoz qilishda narx va talab o'rtasidagi bog'liqlikni tahlil qilishda regressiya metodidan foydalaniladi.



Klasterlash. (Clustering). Ma'lumotlar qazib olishning yana bir metodologiyasi klasterlashdir. Bu bir xil xususiyatlarga ega bo'lgan mazmunli ob'ekt klasterlarini yaratadi. Odamlar ko'pincha uni tasniflash bilan chalkashtirib yuborishadi, lekin agar ular ushbu ma'lumotlarni qidirish metodologiyasi yoki usullari qanday ishlashini to'g'ri tushunsalar, ularda hech qanday muammo bo'lmaydi.

Ob'ektlarni oldindan belgilangan sinflarga joylashtiradigan tasniflashdan farqli o'laroq, klasterlash ob'ektlarni o'zi tomonidan belgilangan sinflarga joylashtiradi. Keling, bir misol keltiraylik. Kutubxona turli mavzulardagi kitoblar bilan to'la. Endi muammo o'sha kitoblarni o'quvchilarga ma'lum bir mavzu bo'yicha kitoblarni topishda hech qanday muammo tug'dirmaydigan tarzda tartibga solishdir. Biz o'xshashliklari bo'lgan kitoblarni bitta javonda saqlash va keyin bu javonlarga mazmunli nom berish uchun klasterlashdan foydalanishimiz mumkin. Muayyan mavzu bo'yicha kitob izlayotgan o'quvchilar to'g'ridan-to'g'ri o'sha javonga borishlari mumkin. Ulardan kitoblarini topish uchun butun kutubxonani kezish talab qilinmaydi. Klaster tahlili bir-biriga o'xshash ma'lumotlarni aniqlaydi. Bu ma'lumotlar

TA'LIM, TARBIYA VA INNOVATSIYALAR

I son, Dekabr

o'rtasidagi o'xshashlik va farqlarni aniqlaydi. U segmentatsiya sifatida tanilgan va ma'lumotlar bazasida sodir bo'layotgan voqealar haqida tushuncha beradi.

Assotsiativ qoidalar (Association rules). Assotsiativ qoidalar, ma'lumotlar orasidagi bog'lanishlarni aniqlashda ishlatiladi. Bunda, masalan, ma'lum bir mahsulotni sotib olgan mijozlar boshqa mahsulotni sotib olish ehtimolini prognoz qilish mumkin. Bu metod retail (savdo) sohasida keng qo'llaniladi.

Neyron tarmoqlar (Neural networks). Neyron tarmoqlari inson miya faoliyatiga o'xshash ishlaydi. Ushbu metod murakkab ma'lumotlar asosida prognoz qilishda samarali hisoblanadi. Masalan, tasvirni tanish, nutqni aniqlash yoki tibbiyotda kasalliklarni prognoz qilishda neyron tarmoqlari ishlatiladi.

Natija: Data Mining metodlari va bosqichlari har bir soha va ilmiy tadqiqotlar uchun katta ahamiyatga ega. Ushbu jarayon ma'lumotlardan foydali bilimlarni olish, yangi tendensiyalarni aniqlash va qarorlar qabul qilishda qo'llaniladi. Ma'lumotlarni to'plash, tayyorlash, tahlil qilish, model qurish, baholash va amaliyotga tatbiq etish kabi bosqichlar orqali ma'lumotlar o'z ichidan qimmatli ma'lumotlar chiqariladi.

Data Mining metodlari, jumladan tasniflash, regressiya, klasterlash, assotsiativ qoidalar va neyron tarmoqlari, har biri o'ziga xos xususiyatlari va qo'llanilish sohalariga ega.

Xulosa. Data Mining metodlari va bosqichlari har bir tashkilot va ilmiy jamoa uchun katta ahamiyatga ega. Ma'lumotlar asosida qarorlar qabul qilish, ish faoliyatini yaxshilash va yangi bilimlarni kashf etish imkoniyatlarini taqdim etadi. To'g'ri metodlar va bosqichlardan foydalanish orqali tashkilotlar va ilmiy tadqiqotlar yuqori sifatli va samarali qarorlar qabul qilishlari mumkin.

FOYDALANILGAN ADABIYOTLAR:

1. Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Elsevier.
2. Witten, I. H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (2nd ed.). Morgan Kaufmann.
3. Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
4. <https://www.sites.google.com/site/upravlenieznaniami/tehnologii-upravlenia-znaniami/data-mining> (24-mart2023 yil)
5. Berry, M. J. A., & Linoff, G. S. (2004). *Data Mining Techniques: For Marketing, Sales, and Customer Support*. Wiley.