

DATA MINING STANDARTLARI

Onarqulov Maqsadjon Karimberdiyevich

Farg'ona davlat universiteti amaliy matematika va informatika kafedrası dotsenti

maqsadjon@gmail.com

Abdusoliyeva Umida Akramjon qizi

Farg'ona davlat universiteti 3-kurs talabasi

abdusoliyevaumidaxon@gmail.com

ANNOTATSIYA

Ushbu tezis data mining jarayonini standartlashtirishning ahamiyati va afzalliklariga bag'ishlangan. Data mining — katta hajmdagi ma'lumotlardan foydali bilimlarni ajratib olish jarayoni bo'lib, uni standartlashtirish turli sohalarda samaradorlikni oshiradi. Tezisda CRISP-DM, PMML kabi keng qo'llaniladigan standartlar va ularning qo'llanilish sohalari (bank-moliya, sog'liqni saqlash, marketing, davlat boshqaruvi) batafsil yoritilgan. Shuningdek, standartlarni joriy etishda yuzaga keladigan qiyinchiliklar, jumladan, texnologik infratuzilma va mutaxassislar yetishmasligi muammolari ko'rib chiqiladi. Annotatsiyada, data miningning sun'iy intellekt, avtomatlashtirish va bulutli texnologiyalar bilan uyg'unlashuvi kelajakdagi tendensiyalar sifatida ta'kidlangan. Ushbu standartlarning joriy etilishi ma'lumotlar bilan ishlashda yuqori samaradorlik va ishonchlilikni ta'minlaydi.

Kalit so'zlar: Data mining, Standartlari, CRISP-DM, PMML, ISO/IEC standartlari, ma'lumotlarni qazib olish, Predictive analytics, big data, sun'iy intellekt, avtomatlashtirish, bulutli texnologiyalar, ma'lumot xavfsizligi, samaradorlik, ma'lumotlarni qayta ishlash, innovatsiyalar .

АННОТАЦИЯ

Данная тезис посвящена важности и преимуществам стандартизации процесса data mining. Data mining — это процесс извлечения полезных знаний из больших объемов данных, а его стандартизация повышает эффективность в различных сферах. В тезисе подробно рассматриваются такие широко применяемые стандарты, как CRISP-DM, PMML, и области их применения (банковско-финансовая сфера, здравоохранение, маркетинг, государственное управление). Также обсуждаются трудности внедрения стандартов, включая проблемы с технологической инфраструктурой и нехваткой специалистов. В аннотации подчеркивается, что интеграция data mining с искусственным интеллектом, автоматизацией и облачными технологиями является тенденцией будущего. Внедрение этих стандартов обеспечивает высокую эффективность и надежность при работе с данными.

Ключевые слова: Data mining, Стандарты, CRISP-DM, PMML, Стандарты ISO/IEC, Извлечение данных, Предиктивная аналитика, Большие данные, Искусственный интеллект, Автоматизация, Облачные технологии, Информационная безопасность, Эффективность, Обработка данных, Инновации

ANNOTATION

This thesis is dedicated to the importance and advantages of standardizing the data mining process. Data mining is the process of extracting useful knowledge from large datasets, and its standardization enhances efficiency across various fields. The thesis elaborates on widely used standards such as CRISP-DM and PMML, as well as their application areas (banking and finance, healthcare, marketing, public administration). Challenges in implementing these standards, including issues with technological infrastructure and a lack of specialists, are also discussed. The annotation highlights the integration of data mining with artificial intelligence, automation, and cloud technologies as future trends. The adoption of these standards ensures high efficiency and reliability in data handling.

Keywords: Data mining, Standards, CRISP-DM, PMML, ISO/IEC standards, Data extraction, Predictive analytics, Big data, Artificial intelligence, Automation, Cloud technologies, Data security, Efficiency, Data processing, Innovatsion

Hozirgi davrda ma'lumotlarni qayta ishlash va ulardan samarali foydalanish korxonalar va tashkilotlar uchun muhim ahamiyatga ega. Data mining (ma'lumotlarni qazib olish) bu jarayonda asosiy vositalardan biri bo'lib, katta hajmdagi ma'lumotlardan foydali bilimlarni ajratib olish imkonini beradi. Ushbu jarayonni standartlashtirish uni samaradorroq va bir xil usulda bajarishga yordam beradi.

1. Data Mining standartlarining ahamiyati

Data mining jarayonida qo'llaniladigan metodlar va vositalarni standartlashtirish quyidagi maqsadlarga xizmat qiladi:

- Uyg'unlik: Turli vositalar va algoritmlar bir-biriga mos kelishini ta'minlash.
- Samaradorlik: Jarayonlarni takroriy va tez amalga oshirish imkonini yaratish.
- Ishonchlilik: Ma'lumotlarni qayta ishlash natijalarining sifati va aniqligini oshirish.
- Tashkilotlararo hamkorlik: Turli tashkilotlar va mutaxassislar o'rtasida ma'lumot almashinuvini osonlashtirish.

2. Data Mining standartlari turlari:

Data mining uchun qo'llaniladigan asosiy standartlar:

- CRISP-DM (Cross Industry Standard Process for Data Mining): Ma'lumotlarni qazib olishning asosiy bosqichlarini belgilaydi (biznesni tushunish, ma'lumotlarni tushunish, ma'lumotlarni tayyorlash, modellashtirish, baholash, va ishlatish).
- PMML (Predictive Model Markup Language): Predictive analytics modellarini turli dasturlash tillarida qo'llash va ulash uchun yaratilgan XML formati.
- ISO/IEC standartlari: Ma'lumot xavfsizligi, sifatni nazorat qilish va ma'lumotlarni boshqarish bilan bog'liq standartlar.
- DMG (Data Mining Group) tavsiyalari: Ochiq formatda predictive analytics va big

data texnologiyalarini qo'llash bo'yicha tavsiyalar.

3. Standartlarning qo'llanilishi

Bank-moliya sektori: Firibgarlikni aniqlash va mijozlarni segmentatsiya qilish.
Sog'liqni saqlash: Kasalliklarni aniqlash va davolashni optimallashtirish.
Marketing: Xaridor xatti-harakatlarini tahlil qilish va maqsadli reklama.
Davlat boshqaruvi: Jinoyatlarning oldini olish va transport tizimlarini rejalashtirish.

4. Standartlarni joriy etishning qiyinchiliklari

Texnologik infrastrukturani moslashtirishga bo'lgan ehtiyoj.
Standartlar bo'yicha bilim va malakaga ega mutaxassislarning yetishmasligi.
Katta hajmdagi ma'lumotlar xavfsizligini ta'minlash.

5. Yangi tendensiyalar

AI bilan integratsiya: Data mining va sun'iy intellekt usullarining birlashishi.
Avtomatlashtirish: Jarayonlarning yanada avtomatlashtirilishi.
Bulutli texnologiyalar: Ma'lumotlarni qayta ishlashni bulutda amalga oshirish.

Xulosa:

Data mining standartlari ma'lumotlar bilan ishlashni tartibga solish va optimallashtirish imkonini beradi. Ularni qo'llash korxonalar va tashkilotlarga samaradorlik, ishonchlik, va innovatsiyalarni tatbiq etishda yordam beradi. Shu bilan birga, standartlarni amaliyotga joriy etish uchun texnologik imkoniyatlar va bilim darajasini oshirish zarur.

FOYDALANILGAN ADABIYOTLAR:

1. Onarkulov, M. K. (2024). ИНТЕГРАЦИЯ НЕЙРОННЫХ СЕТЕЙ В РАСПОЗНАВАНИИ РЕЧИ И РАЗРАБОТКА НОВЫХ МЕТОДОВ. ОБРАЗОВАНИЕ НАУКА И ИННОВАЦИОННЫЕ ИДЕИ В МИРЕ, 47(7), 193-197.
2. Onarkulov, M. K., угли Юсупов, М. А., & угли Умиржонов, Л. А. (2023). ПРИМИНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В РАСПОЗНАВАНИИ РЕЧИ. Educational Research in Universal Sciences, 2(3), 1206-1210.
3. Umarjon o'g, Y. Y. L. (2024). PROBLEMS AND MODERN TRENDS IN COMPUTING ENGINEERING. Multidisciplinary and Multidimensional Journal, 3(6), 17-21.
4. Onarkulov, M., & Nabijonov, A. (2024). DB2: KATTA HAJMDAGI MA'LUMOTLARNI SAMARALI BOSHQARISH. Инновационные исследования в науке, 3(5), 99-104.
5. Onarkulov, M., & G'oyibova, G. (2024). SQL SERVER ILOVASIDA UNIVERSAL MODELNI QO'LLASH VA KONVERTATSIYA ETISHNING SAMARALI YO'LLARI. Академические исследования в современной науке, 3(18), 147-152.
6. Karimberdiyevich, M. O. (2024). RBF TURLARINING O 'QITISH ALGORITMI VA XOR MASALASI. worldly knowledge conferens, 8(1), 176-180.
7. Karimberdiyevich, M. O. (2024). GAUSS FUNKSIYASI. worldly knowledge conferens, 8(1), 239-244.
8. Karimberdiyevich, O. M. (2024). SQL TILIDA PROTSEDURA VA FUNKSIYALARDAN FOYDALANISHNING AHAMIYATI. worldly knowledge conferens, 8(1), 145-148.
9. Karimberdiyevich, O. M. (2024). BIR QATLAMLI NEYRON TO'RLARI VA ULARNI YARATISH USULLARI. IQRO INDEXING, 9(2), 104-108.

10. Onarkulov, M., & Omonaliyeva, E. (2024). QARORLAR DARAXTI VA UNI KIRITISH ALGORITIMI. *Science and innovation in the education system*, 3(6), 66-73.
11. Onarkulov, M., & Isaqova, S. (2024). NEYROCHIPLAR, MAXSUS MATRITSALI KUCHAYTIRGICHLAR VA NEYROEMULYATORLAR. *Science and innovation in the education system*, 3(6), 52-58.
12. Karimberdiyevich, M. O. (2024). EKSPERT TIZIMLARI YARATISH VA ULARNING MUAMMOLARI. ILM-FAN YANGILIKLARI KONFERENSIYASI, 2(1), 123-126.
13. Onarkulov, M., & Meliboyeva, A. (2024). HEMMING NEYRON TO'RLAR VA ULARNING ARXITEKTURASI. *Current approaches and new research in modern sciences*, 3(4), 177-181.
14. Onarkulov, M., & Satinova, G. (2024). NEYRON TO 'RLARIDA FAOLLASHTIRISH FUNKSIYALARI. *Theoretical aspects in the formation of pedagogical sciences*, 3(8), 26-30.10:20