

<https://doi.org/10.5281/zenodo.17330891>

Qurbonboyeva Shaxrizoda Shoxruxbek qizi

Amaliy matematika va intellektual

texnologiyalar fakulteti talabasi

Annotatsiya: *Bugungi raqamli davrda har soniya ma'lumotlar juda katta hajmda ortib bormoqda. Har bir tashkilot o'z ishi faoliyati davomida juda katta ma'lumotlar to'plashadi biroq ularning aksariyat qismidan foydalanishmaydi. Bunday ma'lumotlar qorong'i ma'lumotlar yani Dark Data deb ataladi. Suniy intellekt(AI) texnologiyalari bizga bu ma'lumotlarni qayta ishlab tahlil qilish va ulardan foydalanish imkonini beradi. Qorong'i ma'lumotlarning turlari, suniy intellekt yordamida tahlil qilish usullari va bir qator ishlatilish sohalarini ko'rib chiqamiz.*

Kalit so'zlar: *Dark Data, Suniy intellekt, qorong'i malumotlar, dasturlar, zamonaviy texnologiyalar.*

Abstract: *In today's digital era, data is growing at a huge rate every second. Every organization collects a huge amount of data during its business activities, but most of it is not used. Such data is called dark data, or Dark Data. Artificial intelligence (AI) technologies allow us to process, analyze and use this data. We will consider the types of dark data, methods of analysis using artificial intelligence, and a number of areas of use.*

Keywords: *Dark Data, Artificial Intelligence, dark data, programs, modern technologies.*

XXI asr raqamli texnologiyalar davrida ma'lumotlar kun sayin ortib bormoqda. Turli xil tashkilotlar, kompaniyalar, universitetlar yokida turli xil raqamli qurilmalardagi ilovalar malumotlar yaratishadi. Ularning katta qismi ishlatilmasdan qolib ketadi. Bunday malumotlar qorong'i malumotlar(Dark Data) deb ataladi. IBM(2015) tarifiga ko'ra "Qorong'i ma'lumotlar-bu ko'p hollarda e'tibordan chetda qoladigan va foydalanilmaydigan, biroq katta qiymatga ega bo'lishi mumkin bolgan malumotlar majmui" deb atalgan. Yani tashkilot yoki shaxs tomonidan yig'ilib, ammo tahlil qilinmasdan qolib ketadigan malumotlarga aytiladi. Ularni saqlab qo'yishdan maqsad esa dastlab foydasiz ko'ringan malumotlar kelajakda yangi texnologiyalar paydo bo'lganda yoki yashirin tendensiyalarda qiymatli resursga aylanishi mumkin. Tadqiqotlarga ko'ra tashkilotlardagi ma'lumotlarning 80% gacha qismi qorong'i ma'lumotlar hisoblanadi.

Tadqiqotchilar fikriga ko'ra, qorong'i ma'lumotlarni ularning tuzilishi, manbasi va amaliy ahamiyatiga ko'ra bir nechta turlarga ajratish mumkin. Dark Data malumotlari turlariga arxivlangan ma'lumotlar, mijozlar bilan aloqa yozuvlari, tuzilmagan

ma'lumotlar, qisman tuzilgan ma'lumotlar, transaksion va operatsion yozuvlar, jarayonlarda qoladigan ma'lumotlar va sensor IoT ma'lumotlar kiradi.

1. *Arxivlangan ma'lumotlar* bu eski tizimda saqlangan fayllar yani hujjat arxivlari va skanerlangan hujjatlar kiradi.

2. *Mijozlar bilan aloqa yozuvlari* bu Call-center suhbatlari kiradi.

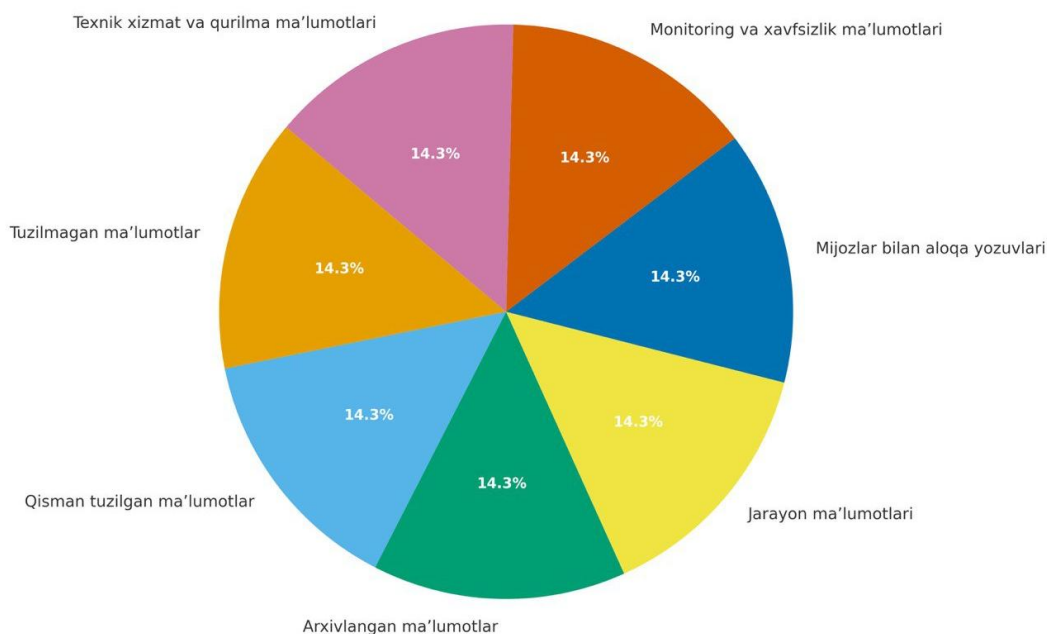
3. *Tuzilmagan ma'lumotlar* bu mat hujjatlari, elektron pochta, chat yozishmalari, suratlar, audio va video fayllari kiradi.

4. *Qisman tuzilgan ma'lumotlar* bu XML, JSON, log-fayllar va internet orqali yig'iladigan veb-sahifa yozuvlari kiradi.

5. *Transaksion va operatsion yozuvlar* bu bank operatsiya yozuvlari, onlayn xaridlar haqidagi ma'lumotlar kiradi.

6. *Jarayonlarda qoladigan ma'lumotlar* bu server loglari, tizim faoliyat jurnallari va qurilma faoliyati haqidagi yozuvlar kiradi.

7. *Sensor va IoT ma'lumotlari* bu smart qurilmalaridagi yig'ilgan ma'lumotlar



Подпись

kiradi. Asosan bu ma'lumotlar juda katta hajmda bo'ladi, ammo kamdan-kam ishlatiladi.

Qorong'i ma'lumotlar (Dark Data) ko'pincha tuzilmagan yoki qisman tuzilganligi sababli an'anaviy ETL va analitik jarayonlarga mos kelmaydi ularni foydali ma'lumotga aylantirish uchun AI yordamida tahlil qilish kerak bo'ladi. Suniy intellekt yordamida tahlil qilish usullariga quyidagilar kiradi:

1. *Ma'lumotlarni yig'ish va tayyorlash* - Dark Data turli formatlarda bo'ladi ularni AI modellariga kiritish uchun tozalash, strukturaga keltirish va tegishli formatga aylantirish zarur.

2. *Tabiiy tilni qayta ishlash(NLP)*-elektron xatlar, hisobotlarni NLP yordamida asosiy mavzularini aniqlash va yashirin bilimlarini topish mumkin.

3. *Kompyuter ko'rishi*-skaner qilingan hujjatlar, fotosuratlarini va videolarni tahlil qilish.

4. *Mashinaviy o'rganish*-log fayllaridan noodatiy xatti-harakatlarni aniqlash, sensor ma'lumotlari asosida prognoz qilish.

5. *Big Data va AI integratsiyasi*-qorong'i ma'lumotlar juda katta hajmda bo'lgani uchun Hadoop yoki Spark kabi platformalardan foydalanish mumkin.

6. *Amaliyotda qo'llanishi*-tibbiyotda, biznesda va xavfsizlik sohaslarida keng foydalaniladi.

Umumiy qilib aytganda, AI qorong'i ma'lumotlarni yoritib, ulardan qimmatli ma'lumotlar chiqarishga yordam beradi.

Qorong'i malumotlar faqat saqlanib qolmasdan, balki turli sohalarda ham muhim o'ringa ega. Ularning qo'llanilish sohasi keng bo'lib, quyidagilar kiradi.

1. *Biznes va marketing*: ijtimoiy tarmoq sharhlari va call-markaz yozuvlarini qayta ishlash orqali xizmat va mahsulotni yaxshilash

2. *Sog'liqni saqlash*: elektron tibbiy yozuvlarda qolib ketgan ma'lumotlarni qayta ishlashda qo'llaniladi.

3. *Moliyaviy xizmatlar va sug'urta*: kredit risklarini baholashda qo'shimcha yashirin ma'lumotlardan foydalanish.

4. *Sanoat va ishlab chiqarish*: sensor va IoT qurilmalaridan kelayotgan, ammo ishlatilmagan ma'lumotlarni qayta tahlil qilish.

5. *Davlat boshqaruvi va xavfsizlik*: jinoyatchilik va xavfsizlikda yashirin yozuvlardan foydalanish.

6. *Energetika*: energiya yo'qotishlarini aniqlash va barqaror ta'minotni rejalashtirish.

Qorong'u malumotlarning afzalliklari: Qaror qabul qilishda foydaliligi, tarixiy ma'lumotlarning qiymati va yangi bilimlar manbai kiradi. Kamchiliklariga esa: Ma'lumotlarni saqlash xarajati, xavfsizlik xatarlari va huquqiy muammolar kiradi.

Xulosa qilib aytganda, Hozirgi davrda qorong'i ma'lumotlar bazasi kun sayin ortib bormoqda. Bu esa zamonaviy axborot texnologiyalar rivojlanishida muhim o'rin kasb etadi. Ularni suniy intellekt orqali tahlil qilinib kerakli sohalarga yo'naltirilsa katta yutuqlarga erishish mumkin. Masalan tibbiyot sohasini olib qarasak, hozirgi kunda bu soha jadal rivojlanib bormoqda. Buning sabablaridan biri esa qorong'i ma'lumotlar hisoblanadi. Qorong'i ma'lumotlar orqali ular kasalliklar tarixini o'rganishlari mumkin hisoblanadi. Dark Data faqatgina tibbiyotda emas boshqa sohalarda ham foydalaniladi. Shuning uchun ham qorong'i ma'lumotlardan foydalanishda to'g'ri usul orqali foydalanilsa kelajak ravnaqi uchun katta yutuqlar olib kelishi mumkin. Suniy intellekt yordamida biz qulay va tez ma'lumotlarni o'zimiz istagan shaklda topib foydalanishimiz mumkin. Lekin undan bazi insonlar o'z manfaati yo'lida yoki maxfiy ma'lumotlarni tarqatish uchun foydalanishlari ham mumkin. Qorong'i ma'lumotlar muhofaza qilish-axborot xavfsizligining eng muhit yo'nalishi hisoblanadi. Lekin ba'zi vaqtlari ularni muhofaza qilish

zaiflashib qoladi. Shu sababli, ularni saqlashda qattiq muhofaza qilish kerak bo'ladi.

FOYDALANILGAN ADABIYOTLAR:

1. Гантц Ж., Реинсел Д. "The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East". IDC iView, IDC, 2012.
2. IBM Corporation. "The Four V's of Big Data". IBM Big Data & Analytics Hub, 2015.
3. Мэпп Б. "Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results". Wiley, 2016.
4. Лани Д. "3D Data Management: Controlling Data Volume, Velocity, and Variety". META Group, 2001.
5. Рассел С., Норвиг П. "Artificial Intelligence: A Modern Approach". 4-нашр. Pearson, 2020.
6. Провост Ф., Фоусетт Т. "Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking". O'Reilly Media, 2013.
7. Катал А., Вазид М., Гоудар Р. Х. "Big Data: Issues, Challenges, Tools and Good Practices". 2013 Sixth International Conference on Contemporary Computing (IC3), IEEE, 2013.
8. Уссус А., Бенжеллун Ф., Лахсен А. А., Белфких С. "Big Data technologies: A survey". Journal of King Saud University - Computer and Information Sciences, 30(4), 431–448, 2018.
9. Мэника Дж., Чуи М., Браун Б., Бугин Дж., Доббс Р., Роксбург Ч., Хунг Б. А. "Big Data: The Next Frontier for Innovation, Competition, and Productivity". McKinsey Global Institute, 2011.
10. ISO/IEC 20546:2019. "Information technology — Big data — Overview and vocabulary". International Organization for Standardization, 2019