

DATA SCIENCE JARAYONIDADA KATTA MA'LUMOTLAR EKOTIMLARI

Umarov Bekzod Azizovich

*Farg'ona davlat universiteti amaliy matematika va informatika
kafedrası o'qituvchisi* Gmail: ubaumarov@mail.ru

Shuhratjonova Gulchiroy Shuhratjon qizi

Farg'ona davlat universiteti talabasi
Gmail: shuhratjonovagulchiroy@gmail.com

Annotatsiya: *Katta ma'lumotlar (Big Data) va Data Science sohalaridagi ekotizmlar, ma'lumotlar yig'ish, saqlash, qayta ishlash va tahlil qilishda ishlatiladigan texnologiyalar, metodologiyalar va jarayonlarning tizimli aloqalarini tashkil etadi. Ushbu maqolada katta ma'lumotlar ekotizmlarining Data Sciencedagi o'rni, ularning texnologiyalari, muammolari va rivojlanish imkoniyatlari yoritiladi. Maqola katta ma'lumotlar tahlili jarayonlaridagi texnologiyalar, metodologiyalar va ularning biznes va ijtimoiy tizimlarga qanday ta'sir ko'rsatishini ko'rib chiqadi.*

Kalit so'zlar: *Katta ma'lumotlar, Data Science, ekotizmlar, texnologiyalar, tahlil, xavfsizlik, ekologiya.*

Аннотация: *Экосистемы в области больших данных (Big Data) и Data Science представляют собой систематические соединения технологий, методологий и процессов, используемых при сборе, хранении, обработке и анализе данных. В этой статье освещается роль экосистем больших данных в науке о данных, их технологии, проблемы и возможности развития. В статье рассматриваются технологии и методологии процессов анализа больших данных и их влияние на бизнес и социальные системы.*

Ключевые слова: *Большие данные, Data Science, экосистемы, технологии, анализ, безопасность, экология.*

Annotation: *Ecosystems in the fields of Big Data (Big Data) and Data Science constitute a systematic connection of technologies, methodologies and processes used in data collection, storage, processing and analysis. This article discusses the role of big data ecosystems in Data Science, their technologies, problems and development opportunities. The article examines the technologies, methodologies in big data analysis processes and their impact on business and social systems.*

Keywords: *Big Data, Data Science, ecosystems, technologies, analysis, security, ecology.*

Kirish:

Katta ma'lumotlar va Data Science sohaları jahon iqtisodiyoti va ilm-fanida tez rivojlanayotgan sohalaridir. Katta ma'lumotlar ekotizmi esa, bu sohada ishlatiladigan texnologik va metodologik vositalarni o'z ichiga oladi. Katta hajmdagi ma'lumotlarning

samarali tahlil qilinishi uchun barcha jarayonlar birlashgan tizim sifatida ishlaydi. Maqolada katta ma'lumotlar ekotizmlari va ularni Data Science kontekstida qo'llanilishining muhim jihatlari tahlil qilinadi.

Data Science jarayoni — bu ma'lumotlarni yig'ish, tahlil qilish, modellash va qaror qabul qilish uchun bir qator usul va texnologiyalarni o'z ichiga olgan ilmiy yondashuvdir. Data Science jarayonida katta ma'lumotlardan (Big Data) foydalangan holda analitik natijalar olinadi va bu natijalar biznes, ilm-fan, tibbiyot, ijtimoiy sohalarda qaror qabul qilishda qo'llaniladi.



1-Rasm : Big Data

Katta ma'lumotlar ekotizmi — bu katta hajmdagi ma'lumotlarni yig'ish, saqlash, qayta ishlash va tahlil qilish jarayonlarini ta'minlaydigan texnologiyalar, vositalar va platformalar yig'indisidir. Bu ekotizmlar quyidagi asosiy komponentlardan iborat:

Ma'lumotlarni yig'ish va to'plash: Ma'lumotlar turli manbalardan, jumladan, sensorlar, ijtimoiy tarmoqlar, IoT (Internet of Things) qurilmalari va onlayn platformalardan yig'iladi. Ma'lumotlar ko'plab formatlarda va turli sohalarda bo'lishi mumkin, masalan, matn, rasm, video va boshqa ko'plab ma'lumotlar turlari.

Saqlash va boshqarish: Katta ma'lumotlarni saqlash va boshqarish uchun maxsus tizimlar ishlab chiqilgan. Masalan, Hadoop, NoSQL, Data Lakes kabi texnologiyalar yordamida ma'lumotlar saqlanadi. Bu platformalar ma'lumotlarning katta hajmini samarali boshqarishni ta'minlaydi.



2-Rasm: Hadoop platformasi

Ma'lumotlarni qayta ishlash va tahlil qilish: Ma'lumotlarni tahlil qilish uchun Apache Spark, MapReduce kabi tizimlar ishlatiladi. Bu texnologiyalar katta ma'lumotlarni tezda qayta ishlash va samarali tahlil qilish imkonini beradi. Ma'lumotlarni ishlashning optimallashtirilgan metodlari yirik tizimlarda samarali ishlashni ta'minlaydi.

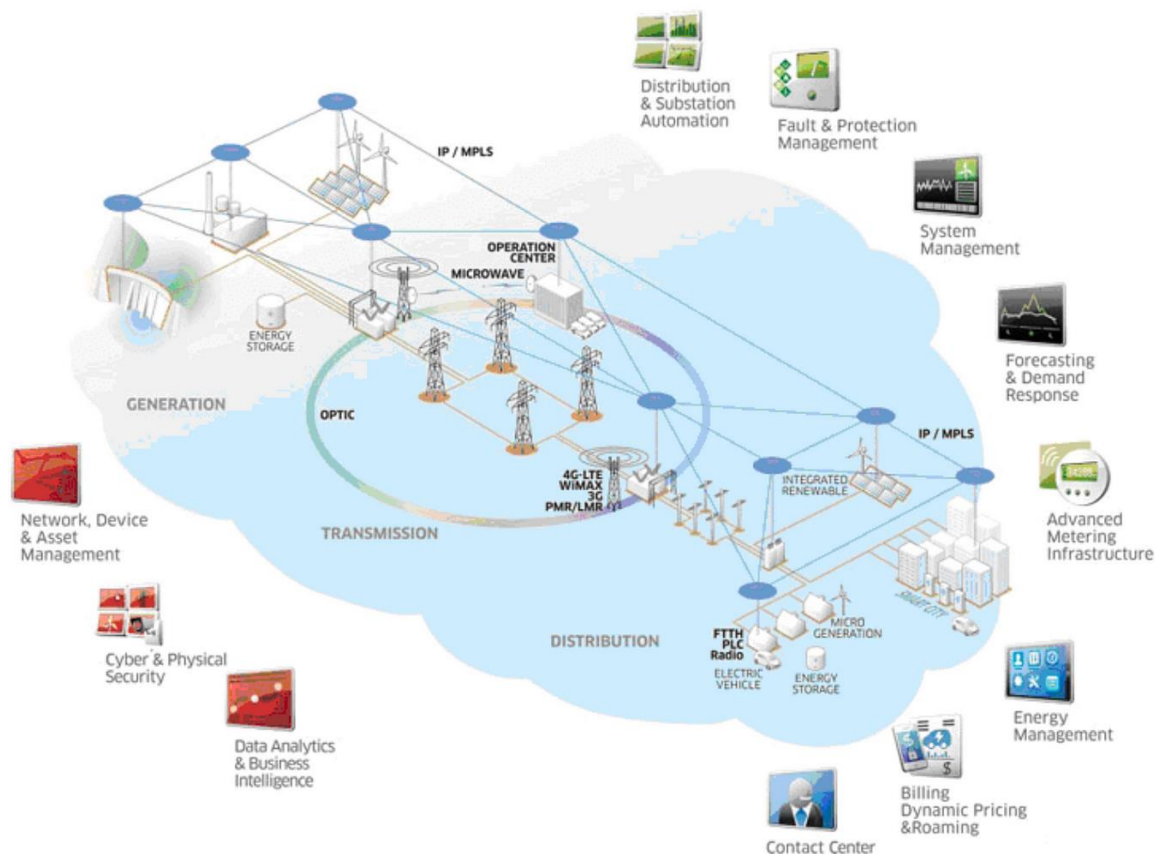
Mashina o'rganish va sun'iy intellekt: Data Science'dagi katta ma'lumotlarni tahlil qilish uchun mashina o'rganish (ML) va sun'iy intellekt (AI) algoritmlari qo'llaniladi. Bu texnologiyalar ma'lumotlardan prognozlar qilish, tasniflash, klasterlash va boshqa analitik natijalar olishga yordam beradi.

Data Science, katta ma'lumotlardan yangi bilimlar olish va qarorlar qabul qilishda qo'llaniladigan ilmiy yondashuvdir. Data Science jarayoni bir nechta bosqichdan iborat:

1. Ma'lumotlarni yig'ish: Data Science jarayonining dastlabki bosqichi ma'lumotlarni yig'ishdir. Bu bosqichda, ijtimoiy tarmoqlardan, sensorlardan, IoT qurilmalardan va boshqa manbalardan katta ma'lumotlar yig'iladi. Yig'ilgan ma'lumotlar to'liq, aniq va ishonchli bo'lishi zarur.

2. Ma'lumotlarni tahlil qilish: Ma'lumotlarni to'plashdan so'ng, ular tahlil qilinadi. Tahlil qilishda mashina o'rganish va sun'iy intellekt yordamida ma'lumotlar tahlil qilinadi. Bu jarayonda ma'lumotlar to'planadi, tasniflanadi, va prognozlar qilish uchun ishlatiladi. Tahlil jarayonlari biznes strategiyalarini shakllantirishda, mijozlarga xizmat ko'rsatishda va boshqa sohalarda muhim rol o'ynaydi.

3. Model yaratish va qaror qabul qilish: Data Science jarayonida ishlab chiqilgan modellardan foydalanib, qarorlar qabul qilish uchun ma'lumotlardan foydalaniladi. Model yaratish bosqichida mashina o'rganish algoritmlari, masalan, regressiya tahlili, tasniflash va klasterlash metodlari qo'llaniladi. Bu modellardan biznesda, tibbiyotda va ijtimoiy sohalarda foydalanish mumkin.



3-Rasm :Aloqalarni uzatish

Katta ma'lumotlar ekotizmlarini yaratishda bir qator muammolar kelib chiqan . Bu muomolar asosan ma'lumotlarning to'liq, aniq va ishonchli bo'lishi katta ahamiyatga ega.Katta hajmdagi ma'lumotlarni yig'ish va qayta ishlashda ma'lumotlar sifatini ta'minlash kerak. Xatoliklar, noaniqliklar va yetishmayotgan ma'lumotlar tahlilning to'g'riligiga ta'sir qilishi mumkin. Xavfsizlik va maxfiylikda katta ma'lumotlarni saqlashda va ishlatishda xavfsizlik muammolari yuzaga keladi. Shaxsiy ma'lumotlarning himoyasi va ularga ruxsatsiz kirishni oldini olish zarur. Maxfiylikka oid qonunlar, masalan, GDPR, ma'lumotlarga ruxsatsiz kirishni oldini olishga qaratilgan choralardir.

Texnologik infratuzilmalarda ham bir nechata muomolar kelib chiqqan .Ulardatta ma'lumotlar bilan ishlash uchun yuqori quvvatli texnologik infratuzilma kerak. Bulutli hisoblash tizimlari va yuqori samarali hisoblash quvvatlari katta ma'lumotlar bilan ishlashda muhim rol o'ynaydi. Bulutli tizimlar katta ma'lumotlarni samarali saqlash va qayta ishlash imkoniyatlarini taqdim etadi.

Data Science jarayonida katta ma'lumotlar ekologik ekotizmlar	
Ekologik monitoring	Resurslarni boshqarish

Katta ma'lumotlar yordamida atrof-muhitni kuzatish, havo sifatini o'lchash, suv resurslarini boshqarish va iqlim o'zgarishlarini prognozlash mumkin. Ekologik monitoring katta ma'lumotlar tahlili yordamida amalga oshiriladi va bu, masalan, iqlim o'zgarishlarini oldindan aniqlashda foydalidir.	Katta ma'lumotlar yordamida tabiat resurslarini samarali boshqarish imkoniyatlari yaratiladi. Energiyani tejash, chiqindilarni kamaytirish va barqaror rivojlanishni ta'minlash uchun katta ma'lumotlar asosida qarorlar qabul qilinadi.
---	---

Natija: Big Data va Data Science sohalaridagi ekotizmlar, ma'lumotlar yig'ish, saqlash, qayta ishlash va tahlil qilishda ishlatiladigan texnologiyalar, metodologiyalar va jarayonlarning tizimli aloqalarini tashkil etadi.biz shuni aytib otishim joizki, katta ma'lumotlar tahlili jarayonlaridagi texnologiyalar, metodologiyalar va ularning biznes va ijtimoiy tizimlarga qanday ta'sir ko'rsatishini katta ma'lumot ekotizmlari Data Science sohasida ko'

Xulosa:

Katta ma'lumot ekotizmlari Data Science sohasidagi muvaffaqiyatli ishlash uchun zarur bo'lgan texnologik, metodologik va ijtimoiy tizimlarni tashkil etadi. Ushbu ekotizmlar yordamida katta ma'lumotlar tahlil qilinadi va samarali qarorlar qabul qilish imkoniyatlari yaratiladi. Katta ma'lumotlar va Data Science sohaları o'rtasidagi o'zaro aloqalar, texnologik rivojlanish va ijtimoiy ta'sirlar kelajakda yanada muhimroq bo'ladi. Ekologik va barqaror rivojlanish sohalarida katta ma'lumotlar ekotizmlari yordamida yangi yechimlar ishlab chiqilishi kutilmoqda.

FOYDALANILGAN ADABIYOTLAR:

1. **Big Data: Principles and Paradigms** Rajendra Akerkar, 2014.
2. **Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking** Foster Provost, Tom Fawcett, 2013.
3. **Data Science from Scratch: First Principles with Python** Joel Grus, 2015.
4. **Big Data and the Internet of Things: Enterprise Information Architecture for a New Age** Mike Barlow, 2015.
5. **Big Data: A Revolution That Will Transform How We Live, Work, and Think** Viktor Mayer-Schönberger, Kenneth Cukier, 2013.
6. **Data Mining: Concepts and Techniques** Jiawei Han, Micheline Kamber, Jian Pei, 2011.