

DATA SCIENCE

Sayfullayeva Marjona

Mirzo Ulug'bek nomidagi O'zbekiston Milliy universiteti

Tarix fakulteti 2-kurs talabasi

Annotatsiya: Ushbu maqolada Data science nima ekanligi, katta hajmdagi ma'lumotlar bilan ishlash tamoyillari, muammolari va rivojlanish istiqbollari muhokama qilinadi.

АННОТАЦИЯ: В данной статье рассматривается, что такое наука о данных, принципы, проблемы и перспективы развития работы с большими объемами данных.

Abstract: This article discusses what Data science is, the principles, problems, and development prospects of working with large volumes of data.

Kalit so'zlar: ma'lumotlar, Data science, IT, axborot, ishlov berish, tahlil qilish, ma'lumotlar, tizimlashtirish.

Bugungi davrda raqamli texnologiyalar har soniyada rivojlanayotgan zamonda katta hajmdagi ma'lumotlar ishlab chiqarilishi biznes uchun ham qiyin, ham yanada katta imkoniyatlarni ochib beradi. Ushbu ma'lumotlardan kerakli bo'lgan fikrlarni olish turli sohalarda muvaffaqiyatga erishishning muhim manbai sanaladi. Bu yerda asosan dinamik va fanlararo soha bo'lgan data science muhim ahamiyatga egadir. Ushbu maqolada biz data science dunyosiga kirib boramiz va uning axborot texnologiyalari (IT) sohasidagi ahamiyatini o'rganamiz.

Data science ¹ yani ma'lumotlar fani fanlararo akademik soha bo'lib, foydalanadigan statistika, ilmiy hisoblash, ilmiy usullar, jarayonlar, algoritmlar va tizimlar, tuzilgan va tuzilmagan ma'lumotlardan bilim va tushunchalarni olish yoki ekstropolyatsiya qilish sanaladi.² Data scienceni tadqiqot paradigmasi, tadqiqot usuli, ish jarayoni va kasb sifatida ta'riflash mumkin.³ Bundan tashqari ma'lumotlar fani (Data Science) statistika, ma'lumotlarni tahlil qilish, informatika va ular bilan bog'liq usullarni birlashtirish tushunchasi bo'lib, ma'lumotlar bilan haqiqiy hodisalarni tushunish va tahlil qilish imkonini yaratib beradi.⁴ Shunday qilib, u kompyuter fanlari, statistika, axborot fanlari, matematika, sun'iy intellekt (AI), ma'lumotlarni vizualizatsiya qilish, axborotni vizualizatsiya qilish, ma'lumotlarni sonifikatsiya qilish,

¹ Donoho, David(2017). "50 Years of Data Science. Journal of Computational and Graphical Statistics. 26 (4):745-766.

² Dhar, V (2013). Data science and prediction. Communications of the ACM. 56(12): 64-73.

³ Mike, Koby, Hazzan, Orit (20 January 2023). "What is Data Science?". Communications of the ACM. 66(2) 12-13.

⁴ Hayashi, Chikio (1 January 1998). "What is Data Science? Fundamental Concepts and a Heuristic Example". In Hayashi, Chikio; Yajima, Keiji; Bock, Hans-Hermann; Ohsumi, Noboru; Tanaka, Yutaka; Baba, Yasumasa(eds). Data Science, Classification, and Related Methods. Studies in Classification, Data Analysis, And Knowledge Organization. Springer Japan.pp. 40-51.

ma'lumotlar integratsiyasi, murakkab tizimlar, aloqa va biznes ko'nikmalarini o'z ichiga oladi.^{5 6} Data Science bir necha bosqichlardan iboratdir. Ularga ma'lumotlarni yutish, ma'lumotlarni saqlash va ma'lumotlarni qayta ishlash, ma'lumotlarni tahlil qilish va muloqot kabilar kiradi.

Data science sezilarli ta'sir ko'rsatadigan bir nechta asosiy sohalar:

Data science mashinani o'qitish algoritmlari va AI modellarini ishlab chiqish va amalga oshirishda hal qiluvchi rol o'ynaydi. Axborot fanlari texnikasidan foydalangan holda, IT mutaxassisleri samaradorlikni oshiradigan, jarayonlarni avtomatlashtiradigan va foydalanuvchi tajribasini yaxshilaydigan aqlli tizimlarni yaratishi mumkin

Data Science va BI bir birini istisno qilmaydi. Raqamli tushunadigan tashkilotchilar o'z ma'lumotlarini to'liq tushunish va qiymatini olish uchun foydalanadilar. Data science ma'lumotlar muhandisligi va ma'lumotlarni boshqarish amaliyotlarini o'z ichiga oladi. Bu IT mutaxassislariga ma'lumotlarni yig'ish, integratsiya qilish, saqlash va olish jarayonlarini soddalashtirishga yordam beradi. Ma'lumotlar fani shuningdek, ma'lumotlarning to'g'ri, ishonchli va me'yoriy standartlarga muvofiqligini ta'minlab, ma'lumotlar sifatini baholashni osonlashtiradi.

Umuman olganda, sun'iy intellekt yaratishning 70–80% vaqti ma'lumotlarni tahlil qilishdan va qolgan 20–30% esa bevosita dasturlashdan iborat. Shuning uchun bu sohaga kirgan insondan matematika, statistika va ehtimollar nazariyasidan minimal bilimlarga ega bo'lish talab qilinadi. Soha mutaxassisligida Data Science uchun qulay dasturlash tillari mavjud.

•**R:** Bevosita ma'lumotlar bilan ishlash, turli matematik va statistik hisob-kitob va tahlil uchun ishlatiladi. Agar siz matematikada kuchli bo'lsangiz va bevosita ma'lumotlar tahlili bilan shug'ullanishni istasangiz, sizning tanlovingiz R dasturlash tili.

•**Python:** Yuqori darajadagi, umumiy maqsadli dasturlash tili. Python oddiy dasturlardan tortib, murakkab hisob-kitoblar va tizimlar yaratish uchun ham ishlatiladi.

Agar sizning maqsadingiz matematik hisob-kitoblardan o'tib, sun'iy intellekt asosida ishlaydigan dasturlar yaratish bo'lsa, Pythondan afzalroq tilni topa olmaysiz. Python o'rganish uchun oson, dasturlash uchun qulay va muhimi Data Sciencga oid tayyor funksiyalar, kutubxonalar va frameworklarga boy tildir.

Bugungi kunda Google, Facebook, Tesla, IBM, Amazon kabi kompaniyalar ham

⁵ "About Data Science". Data Science Association. Archived from the original on 10 August 2020. Retrieved 3 April 2020.

⁶ 1. Introduction: What is Data Science? . Doing Data Science (Book). O'Reilly. Retrieved 3 April 2020.

o'z sun'iy intellekt dasturlarini yaratishda Pythondan foydalanib kelishadi.⁷

Raqamli hayotimizni himoya qilish. Kibertahdidlarning keng tarqalgan xususiyatini hisobga olgan holda, jismoniy shaxslar va tashkilotlar uchun kiberxavfsizlikka ustuvor ahamiyat berish juda muhimdir. Ushbu raqamli asrda o'zimizni himoya qilish uchun ba'zi muhim qadamlar:

Kuchli va noyob parollar: Kuchli, murakkab parollar yarating va ularni turli platformalarda qayta ishlatishdan saqlaning. Har bir onlayn hisob uchun noyob parollarni xavfsiz saqlash va yaratish uchun parol menejeridan foydalanishni o'ylab ko'ring.

Ko'p faktorli autentifikatsiya: iloji boricha ko'p faktorli autentifikatsiyani (MFA) yoqing, chunki u mobil qurilmangizga yuborilgan parol va noyob kod kabi bir nechta tekshirish usullarini talab qilish orqali qo'shimcha xavfsizlik darajasini ta'minlaydi.⁸

Dasturiy ta'minotni muntazam yangilash: Qurilmalaringiz va dasturiy ta'minotningizni so'nggi xavfsizlik yamoqlari bilan yangilab turing. Dasturiy ta'minot yangilanishlari ko'pincha ma'lum zaifliklarni bartaraf etadigan muhim xavfsizlik tuzatishlarini o'z ichiga oladi.

Xavfsiz Wi-Fi tarmoqlari: xavfsiz, shifrlangan Wi-Fi tarmoqlaridan foydalaning va tegishli xavfsizlik choralarisiz umumiy tarmoqlarga ulanishdan saqlaning. Ommaviy Wi-Fi tarmoqlari xakerlar uchun maxfiy ma'lumotlarni tutib olish uchun asosiy maqsad hisoblanadi.

Fishingdan ehtiyot bo'ling: shaxsiy ma'lumotlarni so'ragan yoki zudlik bilan chora ko'rishga undaydigan shubhali elektron pochta xabarlarini, xabarlar yoki telefon qo'ng'iroqlaridan ehtiyot bo'ling. Shubhali havolalarni bosishdan saqlaning va har qanday nozik ma'lumotlarni almashishdan oldin jo'natuvchining qonuniyligini tekshiring.⁹

Ma'lumot bering va xabardorlikni oshiring: so'nggi kiber tahdidlar haqida xabardor bo'ling va o'zingizni va boshqalarni kiberxavfsizlik bo'yicha eng yaxshi amaliyotlar bo'yicha o'rgating. Tashkilotingizda, oilangiz va do'stlaringiz orasida kiberxavfsizlikdan xabardorlik madaniyatini targ'ib qiling¹⁰.

⁷ Pythonda dasturlash asoslari: <https://python.sariq.dev>

⁸ Veretennikov, A. V. Ma'lumotlar: bugungi kunda katta ma'lumotlarni tahlil qilish. Yosh olim. - 2017. – №32 (166).

⁹ Ma'lumotlar: katta ma'lumotlar qanday tartibga solingan va hozirda qayerda ishlatiladi // URL: <https://tproger.ru/articles/big-data-explanation-megafon>

¹⁰ Ma'lumotlar texnologiyalari: marketingda Datadan qanday foydalanish // URL: <https://www.uplab.ru/blog/big-data-technologies/>

Data biz uchun ishlab chiqarishni rejalashtirish, ta'lim, sog'liqni saqlash va boshqa sohalarda yangi ufqlarni ochadi. Agar ularning rivojlanishi davom etsa, Data texnologiyalari ishlab chiqarish omili sifatida axborotni mutlaqo yangi sifat darajasiga ko'tarishi mumkin. Axborot nafaqat mehnat va kapitalga tenglashadi, balki zamonaviy iqtisodiyotning eng muhim resursiga aylanadi.¹¹

Umuman olganda Data Science hozirgi kunda jadal sur'atlar bilan rivojlanib kelmoqda. Aynan shu soha bo'yicha mutahassislar ham asta sekinlik bilan ko'paymoqda. Data Science yoki ma'lumotlar fani IT sanoatining keng sohasi bo'lib, uning ishlash printsiptini bitta maqolada tasvirlash juda qiyin. Chunki u keng tarmoqli soha bo'lib, uning ichiga qancha kirsak shuncha ko'p ma'lumotlarga va juda ko'p chigalliklarga duch kelamiz. Shunga qaramasdan biz harakatdan to'xtamasdan oldinga qarab borishimiz lozim. O'zimiz uchun yangi bo'lgan qirralarini ochishimiz kerak. Bu sohani qancha o'rgansak o'zimiz uchun ham kelajak avlod uchun ham foydali bo'ladi.

¹¹ Andrey Naydich. Ma'lumotlar: muammo, texnologiya, bozor // Kompyuter matbuoti // URL: <https://compress.ru/article.aspx?id=22725>