

BIG DATA TA'LIM JARAYONIDA GOOGLE COLABORATORY PLATFORMASIDAN FOYDALANISH METODIKASI VA SAMARADORLIGI

Xamrayev Xurshidbek Komiljon o'g'li

University of Business and Science stajor – o'qituvchisi

Aliyev Azamjon Ayubjon o'g'li

University of Business and Science stajor – o'qituvchisi

Annotatsiya Mazkur maqolada oliy ta'lim tizimida Katta hajmli ma'lumotlar (Big Data) bo'yicha mutaxassislar tayyorlashning dolzarb muammosi — resurs talabchan hisoblash infratuzilmasi yetishmovchiligi ko'rib chiqilgan. Ushbu muammoni hal etish vositasi sifatida bepul, bulutga asoslangan Google Colaboratory (Colab) platformasidan foydalanish metodikasi taklif etilgan. Tadqiqotda Apache Spark kabi Big Data texnologiyalarini Colab muhitida sozlash, ma'lumotlarni qayta ishlash va tahlil qilish bo'yicha amaliy mashg'ulotlarni o'tkazishning bosqichma-bosqich yondashuvi ishlab chiqilgan.

Kalit so'zlar: Big Data, Google Colaboratory, ta'lim metodikasi, Apache Spark, PySpark, ma'lumotlar tahlili, bulutli hisoblashlar, raqamli savodxonlik, pedagogik eksperiment, Python, ta'limda innovatsiyalar, resurs tejamkor ta'lim.

METHODOLOGY AND EFFECTIVENESS OF USING THE GOOGLE COLABORATORY PLATFORM IN THE BIG DATA EDUCATIONAL PROCES

Xamrayev Xurshidbek Komiljon o'g'li

University of Business and Science teacher

Aliyev Azamjon Ayubjon o'g'li

University of Business and Science teacher

Abstract This article addresses a pressing issue in higher education for Big Data specialists: the lack of resource-intensive computing infrastructure. As a solution, a methodology using the free, cloud-based Google Colaboratory (Colab) platform is proposed. The study develops a step-by-step approach for configuring Big Data technologies like Apache Spark within the Colab environment and conducting practical sessions on data processing and analysis.

Keywords: Big Data, Google Colaboratory, teaching methodology, Apache Spark, PySpark, data analysis, cloud computing, digital literacy, pedagogical experiment, Python, educational innovation, resource-efficient education.

Kirish (Introduction)

“Sanoat 4.0” va raqamli iqtisodiyot sharoitida Katta hajmli ma’lumotlar (Big Data) global miqyosda eng qimmatli aktivlardan biriga aylandi. Biznes, fan, tibbiyot va davlat boshqaruvida to’g’ri qarorlar qabul qilish uchun ulkan ma’lumotlar oqimini yig’ish, qayta ishlash va tahlil qilish qobiliyati strategik ahamiyat kasb etmoqda. Bu esa, o’z navbatida, mehnat bozorida Big Data texnologiyalari (Hadoop, Spark) bilan ishlay oladigan malakali mutaxassislariga bo’lgan talabni keskin oshirdi.

Ushbu tadqiqotning maqsadi — bepul, resurslarga oson kirish imkonini beruvchi Google Colaboratory platformasi yordamida Big Data texnologiyalarini o’qitishning samarali metodikasini ishlab chiqish va uning natijalarini baholashdan iborat.

Big Data’ni o’qitish muammolari ko‘plab ilmiy ishlarda ko‘tarilgan. Asosiy qiyinchilik sifatida “resurslar tafovuti” (resource gap) keltiriladi, ya’ni talabalarga real loyihalardagidek katta hajmdagi ma’lumotlar va taqsimlangan hisoblash muhitini taqdim etishning imkonsizligi ta’kidlanadi (Fox & Patterson, 2012). Ushbu muammoni yechish uchun virtualizatsiya va bulutli texnologiyalar (Amazon

Web Services, Microsoft Azure) taklif etilgan, biroq ular ham ko‘pincha tijoriy hisoblanadi va ta‘lim muassasalari uchun ma‘lum bir xarajatlarni talab qiladi.

Google Colaboratory (Colab) esa bu borada inqilobiy yechim taklif etadi. U Jupyter Notebook formatiga asoslangan bo‘lib, brauzer orqali Python kodini yozish va ishga tushirish imkonini beradi. Eng muhimi, Colab bepul GPU va TPU resurslarini taqdim etadi va murakkab sozlashlarni talab qilmaydi (Carneiro et al., 2018). Dastlab asosan mashinaviy o‘qitish va neyron tarmoqlar sohasida qo‘llanilgan Colab, to‘g‘ri yondashuv orqali Big Data texnologiyalarini o‘qitish uchun ham kuchli vositaga aylanishi mumkin. Biroq, Apache Spark kabi taqsimlangan hisoblash tizimlarini Colabning yagona mashina (single-node) muhitida o‘rnatish va undan ta‘lim metodikasi sifatida foydalanish bo‘yicha tizimli tadqiqotlar yetarli emas.

Tadqiqotning maqsadi — Google Colab platformasi asosida Big Data texnologiyalarini o‘qitish metodikasini ishlab chiqish va uning samaradorligini pedagogik eksperiment orqali tekshirishdan iborat.

Taklif etilayotgan metodika quyidagi bosqichlarni o‘z ichiga oladi:

1. **Muhitni sozlash:** Talabalarga Colab muhitida Apache Sparkni ishga tushirish o‘rgatiladi. Buning uchun pyspark kutubxonasi pip yordamida o‘rnatiladi va Spark sessiyasi yaratiladi. Bu jarayon talabalardan murakkab tizim administratorligi bilimlarini talab qilmaydi va bir necha qator kod bilan amalga oshiriladi.
2. **Ma‘lumotlar bilan ishlash:** Amaliy mashg‘ulotlar uchun ochiq, katta hajmli ma‘lumotlar to‘plamlari (masalan, Kaggle platformasidagi aviaparvozlar yoki onlayn-do‘kon tranzaksiyalari ma‘lumotlari) tanlanadi. Ma‘lumotlar Google Drive orqali Colabga yuklanadi yoki to‘g‘ridan-to‘g‘ri internetdan o‘qiladi.

Metodikaning samaradorligini baholash uchun Toshkent axborot texnologiyalari universitetining "Kompyuter injiniringi" yo‘nalishi bo‘yicha tahsil olayotgan 60 nafar 3-kurs talabasi ishtirokida eksperiment o‘tkazildi. Talabalar tasodifiy ravishda ikki guruhga bo‘lindi:

- **Nazorat guruhi (30 nafar talaba):** "Big Data" fanini an‘anaviy usulda o‘rgandi (nazariy ma‘ruzalar, shaxsiy kompyuterga o‘rnatilgan kichik hajmli ma‘lumotlar bilan ishlaydigan standart kutubxonalar).
- **Eksperimental guruh (30 nafar talaba):** Xuddi shu fanni yuqorida taklif etilgan Google Colab asosidagi metodika bo‘yicha o‘rgandi.

Jadval 1. Eksperiment natijalarining qiyosiy tahlili

Ko‘rsatkich	Nazorat guruhi	Eksperimental guruh	Farq
Loyihani bajarish vaqti (o‘rtacha, soat)	15.2	9.1	40.1% tezroq
Loyihadagi aniqlik (o‘rtacha, %)	70.4%	88.2%	25.3% yuqori
Test natijalarining o‘lishi	+22%	+45%	23% punkt ko‘p

Natijalardan ko‘rinib turibdiki, eksperimental guruh talabalari amaliy loyihani ancha tezroq va sifatliroq bajargan. Buning asosiy sababi, ular nazariyani darhol Colab muhitida real, katta hajmli ma‘lumotlar bilan sinab ko‘rish imkoniga ega bo‘lganligidir. An‘anaviy guruh esa infratuzilma cheklovlari tufayli asosan kichik ma‘lumotlar bilan ishlashga majbur bo‘lgan.

So‘rovnoma natijalari ham bu xulosalarni tasdiqladi. Eksperimental guruh talabalarining 95%i Colabdan foydalanish o‘quv jarayonini “qiziqarli va tushunarli” qilganini ta’kidlagan. Ularning eng ko‘p qayd etgan afzalliklari: “dasturiy ta’minotni o‘rnatish va sozlash bilan bog‘liq muammolarning yo‘qligi”

va “istalgan joydan, faqat brauzer yordamida amaliy mashg‘ulotlarni bajarish imkoniyati”.

Olingan natijalar Google Colab platformasining nafaqat qulay vosita, balki Big Data ta’limidagi “resurslar tafovuti”ni bartaraf etuvchi kuchli pedagogik instrument ekanligini isbotlaydi. Ushbu metodika ta’lim jarayonini demokratizatsiya qiladi, ya’ni barcha talabalarga, ularning shaxsiy kompyuterlarining quvvatidan qat’i nazar, zamonaviy texnologiyalar bilan ishlash uchun teng imkoniyat yaratadi. Biroq, Colabning ba’zi cheklovlarini ham hisobga olish lozim: sessiya vaqtining chegaralanganligi va uning haqiqiy ko‘p tugunli (multi-node) klaster emasligi. Shunga qaramay, ta’lim maqsadlari uchun, ya’ni Big Data’ning fundamental konsepsiyalarini va dasturlash modellarini o‘rgatish uchun uning imkoniyatlari to‘liq yetarli.

Tadqiqot natijasida Katta hajmli ma’lumotlarni o‘qitishda Google Colaboratory platformasidan foydalanishga asoslangan samarali va resurs tejamkor metodika ishlab chiqildi va uning amaliy afzalliklari isbotlandi. Ushbu yondashuv universitetlarga qimmatbaho infratuzilmaga sarmoya kiritmasdan, talabalarga zamonaviy va mehnat bozorida talab yuqori bo‘lgan amaliy ko‘nikmalarni berish imkonini yaratadi.

Foydalanilgan adabiyotlar (References)

1. Carneiro, T., Medeiros Ribeiro, M. V., & Cunha, A. (2018). *Google Colaboratory: A case study in using a free cloud-based platform for deep learning education*. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18), 1014.
2. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107-113.
3. Fox, A., & Patterson, D. (2012). *Crossing the software education chasm*. *Communications of the ACM*, 55(5), 44-49.

4. Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity, and Variety*. META Group Research Note.