



STATISTIK DASTURLASH TILLARI (R, PYTHON) VA ULARNING IT SOHASIDAGI QO'LLANILISHI

Tashbayeva Nafisa Nurmuxamedovna

Mustaqillik izlanuvchi

Anotatsiya: Ushbu tezis statistik dasturlash tillari bo'yicha R va Python tillarining nazariy asoslari, farqlari, afzalliklari va IT sohasidagi qo'llanilishini ilmiy jihatdan tahlil qiladi. R va Pythonning ma'lumotlar tahlili, mashina o'rganishi, katta ma'lumotlar va statistik modellashtirishdagi roli, shuningdek, O'zbekistondagi ta'lim, ilmiy-tadqiqot va amaliy loyihalardagi qo'llanilishi batafsil yoritiladi. Ilmiy tadqiqotlar, amaliy misollar va kod fragmentlari asosida tillarning samaradorligi baholanadi, bu esa data science va IT mutaxassislariga qimmatli ma'lumotlar taqdim etadi.

Kalit so'zlar: statistik dasturlash, R tili, Python tili, ma'lumotlar tahlili, IT sohasi, mashina o'rganishi, katta ma'lumotlar, data science, O'zbekiston, vizualizatsiya, modellashtirish.

Kirish

Zamonaviy IT sohasida statistik dasturlash tillari ma'lumotlar tahlili, qaror qabul qilish va murakkab jarayonlarni modellashtirishning asosiy vositalari sifatida muhim ahamiyatga ega. R va Python tillari bu sohada yetakchi o'rinni egallaydi, chunki ular statistik hisob-kitoblar, ma'lumotlar vizualizatsiyasi, mashina o'rganishi va katta ma'lumotlar bilan ishlashni samarali ta'minlaydi. R tili statistik va grafik tahlillar uchun maxsus ishlab chiqilgan bo'lsa, Python umumiy maqsadli til sifatida IT infratuzilmasiga oson integratsiya qilinadi va keng qo'llaniladi. Ushbu tillarning qo'llanilishi iqtisodiyot, tibbiyot, ta'lim, sanoat va biznes sohalarda yuqori samaradorlikni ta'minlaydi, masalan, katta ma'lumotlar (big data) tahlilida yoki sun'iy intellekt loyihalarida. O'zbekistonda ham ushbu tillar ta'lim muassasalarida, ilmiy-tadqiqot ishlarida va raqamli loyihalarda keng



tarqalmoqda, bu milliy raqamli iqtisodiyotni rivojlantirishga va huquqiy savodxonlikni oshirishga hissa qo'shadi. Maqola ushbu tillarning nazariy asoslari, amaliy qo'llanilishi, farqlari, afzalliklari va O'zbekistondagi holatini ilmiy yondashuv asosida kengroq yoritadi, shuningdek, ilmiy tadqiqotlar va amaliy misollar orqali ularning samaradorligini baholaydi.

Statistik dasturlash tillari bo'yicha R va Python tillari IT sohasida ma'lumotlar tahlili, statistik modellash va murakkab algoritmlarni ishlab chiqishning asosiy vositalari sifatida markaziy o'rin egallaydi, chunki ular vaqt qatorlari tahlili, regressiya modellari, ma'lumotlar vizualizatsiyasi va mashina o'rganishini samarali ta'minlaydi. R tili, 1990-yillarda Ross Ihaka va Robert Gentleman tomonidan ishlab chiqilgan, statistik hisob-kitoblar va grafikalar uchun maxsus mo'ljallangan ochiq kodli til bo'lib, uning CRAN (Comprehensive R Archive Network) kutubxonasida 19 mingdan ortiq paket mavjud, masalan, ggplot2 grafikalar uchun va dplyr ma'lumotlar manipulyatsiyasi uchun ishlatiladi, bu esa ilmiy tadqiqotlarda va data science loyihalarida yuqori aniqlikdagi vizualizatsiyalarni ta'minlaydi. R tilining nazariy asosi stoxastik jarayonlar, ehtimollik nazariyasi va Bayyes statistikasiga asoslanib, u katta ma'lumotlar to'plamlarida statistik testlar o'tkazishda yuqori samaradorlik ko'rsatadi, masalan, lineer regressiya uchun `lm()` funksiyasi orqali model yaratish mumkin: `data(mtcars); model <- lm(mpg ~ wt, data = mtcars); summary(model)`, bu esa modelning parametrlari va statistik ahamiyatini baholashga yordam beradi. R ning vektorlashtirilgan operatsiyalari va funksional dasturlash paradigmasi ma'lumotlar tahlilini tezlashtiradi, masalan, `apply()` oilasi funksiyalari orqali massivlar ustida iteratsiyalar o'tkazish, bu IT sohasida ma'lumotlar qayta ishlash jarayonlarini optimallashtiradi. Python tili esa, 1991-yilda Guido van Rossum tomonidan yaratilgan umumiy maqsadli til bo'lib, uning PyPI (Python Package Index) kutubxonasida 300 mingdan ortiq paket mavjud, masalan, Pandas ma'lumotlar manipulyatsiyasi, NumPy ilmiy hisob-kitoblar, Scikit-learn mashina o'rganishi va Matplotlib vizualizatsiya uchun qo'llaniladi. Pythonning sintaksisi ingliz tiliga yaqin bo'lib, o'qilishi oson va indentatsiya orqali strukturalangan, bu boshlang'ich dasturchilar



uchun qulay, masalan, o'rtacha hisoblash uchun `import numpy as np; data = np.array([1, 2, 3, 4, 5]); mean_value = np.mean(data)`, bu esa statistik ko'rsatkichlarni tez hisoblashga imkon beradi. Pythonning obyektga yo'naltirilgan va funksional dasturlash elementlari IT loyihalarida murakkab strukturalarni yaratishga yordam beradi, masalan, class lar orqali ma'lumotlar modellarini ta'riflash. IT sohasidagi farqlari shundaki, R akademik va ilmiy tadqiqotlarda statistik tahlil va vizualizatsiya uchun afzalroq, masalan, genomika tadqiqotlari yoki mijozlar xatti-harakati tahlilida, chunki u ggplot2 orqali murakkab grafikalar yaratishga imkon beradi: `library(ggplot2); ggplot(mtcars, aes(x=wt, y=mpg)) + geom_point() + geom_smooth(method="lm")`, bu esa ma'lumotlar munosabatlarini vizual ravishda ko'rsatadi. Python esa keng qo'llanilishi tufayli veb-illovalar, mashina o'rganishi, katta ma'lumotlar bilan ishlash va avtomatlashtirishda ustun, masalan, TensorFlow orqali chuqur o'rganish modellari yaratish yoki Hadoop va Spark bilan integratsiya qilish orqali big data loyihalarida qo'llash mumkin, bu esa sanoat darajasidagi tizimlarda samaradorlikni oshiradi. R tilining afzalliklari orasida statistik paketlarning boyligi va vizualizatsiya imkoniyatlari bor, masalan, Shiny paketi orqali interaktiv veb-illovalar yaratish, bu esa IT sohasida ma'lumotlar hisobotlarini taqdim etishda samarali va foydalanuvchilar uchun qulay interfeyslar yaratishga yordam beradi. Biroq, R ning kamchiliklari hisoblash tezligining sekinligi va katta ma'lumotlar bilan ishlashda qiyinchiliklar, chunki u bir yadroli operatsiyalarga asoslangan, bu holatda Pythonning NumPy va Pandas kutubxonalari afzalroq, chunki ular vektorlashtirilgan va optimallashtirilgan operatsiyalarni ta'minlaydi va veb-ishlab chiqishda requests kutubxonasi orqali ma'lumotlarni yig'ishga yordam beradi, masalan, `import requests; response = requests.get('https://api.example.com/data'); data = response.json()`. IT sohasidagi qo'llanilishida R tibbiyot va moliya sohalarda statistik modellar qurish uchun ishlatiladi, masalan, Poisson regressiyasi orqali bemorlar oqimini prognozlash, bu resurslarni optimal taqsimlashga yordam beradi va `glm()` funksiyasi orqali amalga oshiriladi: `model <- glm(count ~ predictor, family = poisson(link = "log"), data = dataset)`. Python esa sanoat va biznesda ta'minot zanjirini optimallashtirish uchun talabni



prognozlashda qo'llaniladi, masalan, Scikit-learn orqali regressiya modellari yaratish:

```
from sklearn.linear_model import LinearRegression; model = LinearRegression();  
model.fit(X, y); predictions = model.predict(new_data),
```

bu inventar xarajatlarini kamaytiradi va qaror qabul qilishni tezlashtiradi. Ilmiy tadqiqotlarda ko'rsatilishicha, Pythonning mashina o'rganishi bilan integratsiyasi (masalan, Keras va TensorFlow) R ga nisbatan 20-30% yuqori aniqlik beradi, chunki u chuqur neyron tarmoqlarini samarali boshqaradi, ammo R kichik ma'lumotlar to'plamlarida statistik tahlilda ustun, masalan, bootstrap usullari orqali noaniqlikni baholashda. O'zbekistonda R va Python tillari ta'lim va ilmiy-tadqiqot ishlarida keng qo'llanilmoqda, masalan, maktab va oliy ta'lim muassasalarida Python o'qitilishi raqamli savodxonlikni oshirishga yordam beradi, bu Yangi O'zbekiston strategiyasining bir qismi bo'lib, Pythonning sodda sintaksisi va turtle kutubxonasi orqali chizmalar va animatsiyalar yaratish imkoniyati bolalar uchun qiziqarli darslar tashkil etadi. Python O'zbekistonda veb-ishlab chiqish, sun'iy intellekt va ma'lumotlar tahlilida qo'llaniladi, masalan, Telegram botlari yaratish yoki ma'lumotlar bazalari bilan ishlashda, uning o'zbek tilidagi o'quv qo'llanmalar va uslubiy materiallari boshlang'ich dasturchilarga qulay, shuningdek, Pythonning katta kutubxonalari (masalan, NLTK orqali tabiiy tilni qayta ishlash) o'zbek tilining raqamlashtirilishida qo'llanilmoqda. R tili O'zbekiston statistika tizimida ma'lumotlar tahlili uchun ishlatiladi, masalan, iqtisodiy ko'rsatkichlarni prognozlashda va normal taqsimlanish parametrlarini baholashda, bu milliy statistika islohotlarida muhim rol o'ynaydi va R ning paketlari orqali murakkab modellar qurishga imkon beradi. O'zbekistonning ilmiy jurnallarida Python va R ning qo'llanilishi muhokama qilinib, ularning ta'lim sohasidagi integratsiyasi dolzarb muammo sifatida ko'rib chiqilmoqda, masalan, Python darsliklari va uslubiy qo'llanmalar tarqatilmoqda, shuningdek, kompyuter tilshunosligida Python orqali o'zbek tilidagi matnlar bilan ishlash tadqiqotlari olib borilmoqda. Ushbu tillarning gibrid qo'llanilishi, masalan, Python va R ni birgalikda ishlatish (reticulate paketi orqali), IT loyihalarida samaradorlikni oshiradi, masalan, R ning statistik tahlilini Pythonning veb-interfeysiga integratsiya qilish, ammo O'zbekistonda malakali mutaxassislar yetishmasligi va o'zbek



tilidagi adabiyotlarning cheklanganligi muammosi mavjud, bu esa qo'shimcha ta'lim dasturlarini talab qiladi. Pythonning O'zbekistondagi qo'llanilishiga misol sifatida, ta'lim jarayonida turtle kutubxonasi orqali grafik modellashtirish va animatsiyalar yaratish, shuningdek, machine learning loyihalarida Pandas va Scikit-learn orqali ma'lumotlar tahlili ko'rsatilishi mumkin, bu bolalar va talabalar uchun dasturlashni qiziqarli qiladi. R ning O'zbekistondagi ilmiy tadqiqotlardagi roli esa, masalan, statistik metodlar orqali iqtisodiy modellarni qurishda namoyon bo'ladi, bu milliy statistika tizimini rivojlantirishga hissa qo'shadi. Ushbu tillarni rivojlantirish orqali ma'lumotlar tahlili jarayonini takomillashtirish mumkin, ammo ma'lumotlar sifatini nazorat qilish, yangi kutubxonalarni integratsiya qilish va O'zbekistonda o'zbek tilidagi resurslarni ko'paytirish zarur, masalan, Pythonning o'zbek tilidagi darsliklari va R paketlari uchun lokalizatsiya.

Xulosa qilib aytganda statistik dasturlash tillari R va Python IT sohasida ma'lumotlar tahlili, statistik modellashtirish va mashina o'rganishining muhim vositalari bo'lib, R vizualizatsiya va ilmiy tahlilda, Python esa keng integratsiya va amaliy loyihalarda ustunlik qiladi. O'zbekistonda ushbu tillarning qo'llanilishi ta'lim, ilmiy tadqiqotlar va raqamli transformatsiyani rivojlantirishga hissa qo'shadi, biroq malaka oshirish, o'zbek tilidagi adabiyotlar va resurslarni ko'paytirish zarur. Ushbu tillarni gibril qo'llash orqali IT loyihalarining samaradorligini yanada oshirish mumkin, bu esa O'zbekistonning raqamli iqtisodiyotiga yordam beradi.

Foydalanilgan adabiyotlar ro'yxati:

1. Python dasturida kompyuter tilshunosligiga doir ba'zi masalalarning ... - ResearchGate, 2023.
2. PYTHON DASTURLASH TILIDA TABIIY TILNI QAYTA ISHLASH (NLP) TIZIMLARI - Academia.edu, 2023.
3. Python Dasturlash Tili o'quv Qo'llanma - Studylib, 2023.
4. Normal Taqsimlanish Parametrlarini Baxolashning Statistik Metodlari - Scribd, 2023.
5. Python Uzbekistan | Shaxsiy blog | Farrux Elomonov - Telegram, 2023.