PYTHON DASTURLASH TILIDA TABIIY TILNI QAYTA ISHLASH

Xursandov Hamidillo Fayzillo o‘g‘li – Sharof Rashidov nomidagi Samarqand davlat universiteti Urgut filiali assistenti
e-mail: hamidilloxursandov@gmail.com
Ahtamqulov Muhriddin Ahtamqul o‘g‘li – Sharof Rashidov nomidagi Samarqand davlat universiteti Urgut filiali assistenti
Hamzayev Jaloliddin Jo‘rabek o‘g‘li - Sharof Rashidov nomidagi Samarqand davlat universiteti Urgut filiali 2-kurs talabasi

Annotatsiya: Globalizatsiya va raqamli axborotlashtirish jarayonida o‘zbek tilining rivojida axborot texnologiyalari o‘ta muhim o‘ringa ega. Kompyuter lingvistikasining tabiiy tilni qayta ishlash (Natural Language Processing – NLP) jarayoni va tabiiy tilning kompyuter tushunadigan til (formal) shakli yaratilishi natijasida tilga oid masalalar (tahrir, tahlil, tarjima, elektron matnni ovozlashtirish, og‘zaki nutqni elektron matnga aylantirish, robot bilan muloqot qilish, yirik matnni ixcham matnga aylantirish kabi) bo‘yicha kompyuter dasturlari yaratilmoqda.

Kalit so‘zlar: tabiiy tilni qayta ishlash, NLP, NLTK, axborot texnologiyalarining lingvistik dasturlari, lingvistik tizimlar.

Jahon tilshunosligida XX asrning 50-yillaridan til va matn muammolari bilan bog‘liq masalalar kompyuter til va matn muammolari bilan bog‘liq masalalar kompyuter avtomatik tushunish, mashina tarjimasi, matnni referatlash (gipermatnni asosiy mazmunni o‘zida saqlagan minimatnga keltirish, ya’ni matn hajmini qisqartirish, uni qisqa bayon holiga aylantirish), tasniflash (mavzu, uslub va janr jihatidan o‘zaro yaqin matnlarni guruhlash), matnni tahrir va tahlil demakdir.

Kompyuter lingvistikasi, aynan, NLP har bir tilning axborot texnologiyalari makonida o‘z o‘rnini topishiga xizmat qiladi. Bugungi kunda o‘zbek tilining qo‘llanilish doirasini kengaytirish, elektron lug‘atlar, tarjimon dasturlari, lingvistik dasturlari va tizimlarini yaratish, o‘zbek tili Milliy korpusi va boshqa turdagi til korpuslarini yaratish ustuvor vazifalaridan hisoblanadi. Buning uchun o‘zbek tilini qayta ishlash jarayonlari ustida ilmiy va amaliy izlanishlar olib borilmoqda va, albatta, muayyan natijalarga ham erishilmoqda. Jumladan,

Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universitetida o‘zbek tili Milliy korpusi va ta’limiy korpus, o‘zbek nutq sintezatori, yaratildi. Turkiy tillarning elektron platformasini yaratish ustida izlanishlar boshlangan. Alisher Navoiy

va Zahiriddin Muhammad Bobur ijod mahsullari, shuningdek, yana bir qancha adiblar kitoblarining mobil ilovalari yaratilib, foydalanuvchilar taqdim etildi[1].

O‘zbekiston Respublikasi mustaqillik odimlarini dadil qo‘yayotgan hozirgi davrda axborotlashgan jamiyat qurish masalasi mamlakatimiz uchun naqadar katta ahamiyat kasb etayotgani hech kimga sir emasdir. Internet hayotimizning bir bo‘lagiga aylanib ulgurdi, biz uning xizmatlaridan har kuni foydalanishga odatlandik. Respublikamizda o‘qitish texnologiyalarini zamonaviylashtirishni jadallashtirish rivojlangan iqtisodiyotli mamlakatlarga qaraganda yanada dolzarb ahamiyatga ega hisoblanadi. Chunki hozirgi kunda milliy ta'lim tizimining salohiyati iqtisodiy rivojlanishning yanada yuqori pog‘onasiga ko‘tarilishga amaliy imkoniyat ta'minlovchi asosiy ijtimoiy resurs sifatida gavdalanadi. Dasturlash tillari hozirgi vaqda rivojlanib bormoqda. Ana shu tillardan biri bu- Python dasturlash tilidir. Pythonning rasmiy sayti - www.python.org hisoblanadi. Python hozirgi kunda mashhur dasturlash tillaridan biridir. Bu til Guido van Rossum tomonidan yaratilgan va 1991-yilda chiqarilgan.

Python - bu o‘rganishga oson, shu bilan birgalikda imkoniyatlari yuqor i bo‘lgan oz sonlik zamonaviy dasturlash tillari qatoriga kiradi. Python yuqori darajadagi ma'lumotlar strukturasi va oddiy lekin samarador obyektga yo‘naltirilgan dasturlash uslublarini taqdim etadi[2].

Stack Overflow saytining 2019-yildagi dasturchilar o‘rtasida dasturlash tillari bo‘yicha olib borilgan so‘rovnomasida, eng qulay va ko‘p foydalaniladigan dasturlash tillari ro‘yxatida Python JavaScriptdan so‘ng ikkinchi o‘rinni egallagan. Shu bilan bir qatorda dunyoning Twitter, Pinterest, HP, Symantec, Instagram va Groupon kabi yirik korxonalar aynan Python dasturlash tilidan foydalanib kelmoqda. YouTube, DropBox, Google va Quora kabi dunyoning mashhur online platformalarining dasturiy ta'minoti ham aynan mana shu python dasturlash tilida yozilganligi ushbu dasturlash tiliga bo‘lgan talabning yuqori ekanligini anglatadi. Python nafaqat web sohasida balki sun'iy intellekt va robotexnika sohasida ham yuqori talabga ega tillardan biri hisoblanadi.

Pythonni hattoki katta proyektlarda ishlatish mumkin. Bularga misol qilib Google, Instagram, YouTube ni misol qilib ko‘rsatsak bo‘ladi. Bu loyihalar aynan Python dasturlash tilida yozilgan. Chunki, uni chegarasi yo‘q, imkoniyati juda yuqori.

Shuningdek, u sodda va universalligi bilan dasturlash tillari orasida eng yaxshisi hisoblanadi.

Python dasturlash tili sintaksisi o‘zi kabi juda sodda:

Satr oxiri instruksiyaning oxiri hisoblanadi (nuqta vergul shart emas). Pythonda sintaksis juda sodda tuzilishga ega. Quyida "Salom dunyo" gapini ekranga chiqaruvchi kodni ko‘rsatib o‘tamiz:

print ("Hello world")

Har bir qator boshidagi bo‘sh joy (oTCTyn) muhim ahamiyatga ega hisoblanadi. Kiritilgan amallar bo‘sh joylarning kattaligiga qarab bloklarga birlashadi. Bo‘sh joy istalgancha bo‘lishi mumkin asosiysi bitta kiritilgan blok chegarasida bo‘sh joy bir xil bo‘lishi kerak. Noto‘g‘ri qo‘yilgan bo‘sh joylar

xatolik yuz berishiga olib kelishi ham mumkin. Bitta probel bilan bo‘sh joy hosil qilish yaxshi qaror emas uni o‘rniga to‘rtta probel yoki Tab belgisini ishlatishimiz kerak.

Ma'lum bo‘lganidek, NLP – bu raqamli texnologiyalar uchun matnni qulay va foydali usulda tahlil qilish, tushunish va inson tilidan ma'no olish jarayonidir. Tabiiy tilni qayta ishlash natijasida (NLPdan foydalanib), aynan Python dasturlash tilining yuqorida imkoniyatlari ko‘rsatib o‘tilgan uning kutubxonalari yordamida, avtomatik referatlash, kompyuter tarjimasi, nomlangan obyektni aniqlash, nutq sintezatorini yaratish, tuyg‘ularni tahlil qilish, nutqni tanish va matnlar segmentatsiyasi, so‘z turkumlarini teglash, matnni tahlil qilish: tokenizatsiya, stemming, lemmatizatsiya, parsing kabi vazifalarni bajarish uchun kompyuter dastuarlari va tizimlarini yaratish mumkin.

Foydalanilgan adabiyotlar

1. AbjalovaM.A. Toshkent: ToshDO‘TAU, 23.04.2021. – B.12-17. http://compling.navoiy-uni.uz/index.php/conferences/article/view/12/10

2. Abjalova M. Tahrir va tahlil dasturlarining lingvistik modullari. [Matn]: monografiya / M.A. Abjalova. - Toshkent: Nodirabegim, 2020. – 176 b.

3. Narzullayev Anvar. Pythonda dasturlash asoslari. — T.: “Akademnashr”, 2021. - B. 6.