

ISSN 2181-922X

LANGUAGE & CULTURE

O'ZBEKISTON

TIL VA MADANIYAT

**KOMPYUTER
LINGVISTIKASI**

2023 Vol. 4 (6)

www.compling.tsuull.uz

UZBEKISTAN

ISSN 2181-922X

O‘ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER LINGVISTIKASI

2023 Vol. 4 (6)

compling.tsuull.uz

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti

Bosh muharrir:

Botir Elov

Bosh muharrir o'rinbosari:

Shahlo Hamroyeva

Mas'ul kotib:

Oqila Abdullayeva

Tahrir kengashi

Shuhrat Sirojiddinov (O'zbekiston), Eshref Adali (Turkiya), [Viktor Zaxarov] (Rossiya), Vladimir Benko (Slovakiya), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Murat O'rxun (Turkiya), Suyun Karimov (O'zbekiston), Abduvali Qarshiyev (O'zbekiston), Muxammadjon Musayev (O'zbekiston), Kamoliddin Shukurov (O'zbekiston), O'tkir Hamdamov (O'zbekiston), Tal'at Zuparov (O'zbekiston), Bahodir Mo'minov (O'zbekiston), Faxriddin Nurullayev (O'zbekiston), Zulxumor Xolmanova (O'zbekiston), Muqaddas Abdurahmonova (O'zbekiston), Habibulla Madatov (O'zbekiston), Azizaxon Raxmanova (O'zbekiston), Ruhillo Alayev (O'zbekiston), Rasuljon Atamuratov (O'zbekiston), Malika Abdullayeva (O'zbekiston), Mannon Ochilov (O'zbekiston), Xolisa Axmedova (O'zbekiston), Zilola Xusainova (O'zbekiston).

Jurnal haqida ma'lumot

“O'zbekiston: til va madaniyat. Kompyuter lingvistikasi” seriyasi – Oliy attestatsiya komissiyasi ilmiy nashrlar ro'yxatidagi “O'zbekiston: til va madaniyat” akademik jurnalining ilovasi hisoblanib, unda professor-o'qituvchilar, doktorantlar, stajor-tadqiqotchilar, mustaqil izlanuvchilar, magistrantlarning kompyuter lingvistikasi, jumladan, tabiiy tilga ishlov berish (NLP), o'zbek tilining formal grammatikasi, korpus lingvistikasi, mashina tarjimasi, nutqni qayta ishlash tizimlari, intellektual tizimlar, kompyuter leksikografiyasi hamda lingvistik ontologiyalar kabi sohalarga oid tadqiqotlari nashr qilinadi.

Jurnal ilovasi bir yilda to'rt marta chop etiladi.

O'zbek, turk, rus va ingliz tillarida yozilgan maqolalar qabul qilinadi.

Jurnalda kitoblarga yozilgan taqrizlar, adabiyotlar sharhi, konferensiyalar hisobotlari va tadqiqot loyihalari natijalari ham e'lon qilinadi.

Mualliflar fikri tahririyat nuqtayi nazaridan farq qilishi mumkin.

“O'zbekiston: til va madaniyat. Kompyuter lingvistikasi” seriyasi 2023-yildan chiqa boshlagan.

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti. O'zbekiston, Toshkent, Yakkasaroy tumani, Yusuf Xos Hojib ko'chasi, 103-uy.

E-mail: kompling@navoiy-uni.uz

Website: compling.tsuull.uz

Alisher Navoi'i Tashkent State University of the Uzbek Language and Literature

Chief editor:	Botir Elov
Deputy editor-in-chief:	Shahlo Hamroyeva
Responsible secretary:	Oqila Abdullayeva

Editorial board

Shukhrat Sirojiddinov (Uzbekiston), Eshref Adali (Turkiye), [Viktor Zakharov] (Russia), Vladimir Benko (Slovakia), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Murat Orhun (Turkey), Suyun Karimov (Uzbekistan), Abduvali Karshiyev (Uzbekistan), Mukhammadjon Musayev (Uzbekistan), Kamoliddin Shukurov (Uzbekistan), O'tkir Hamdamov (Uzbekistan), Tal'at Zuparov (Uzbekistan), Bahadir Mo'minov (Uzbekistan), Fakhridin Nurullayev (Uzbekistan), Zulkhumor Kholmanova (Uzbekistan), Muqaddas Abdurakhmonova (Uzbekistan), Habibulla Madatov (Uzbekistan), Azizakhan Raxmanova (Uzbekiston), Ruhillo Alayev (Uzbekistan), Rasuljon Atamuratov (Uzbekistan), Malika Abdullayeva (Uzbekistan), Mannon Ochilov (Uzbekistan), Kholisa Akhmedova (Uzbekistan), Zilola Khusainova (Uzbekistan).

Information about the magazine

"Uzbekistan: language and culture. "Computer Linguistics" series is an appendix of the academic journal "Uzbekistan: Language and Culture" in the list of scientific publications of the Higher Attestation Commission, in which computer linguistics, including natural language processing (NLP) of professors-teachers, doctoral students, intern-researchers, independent researchers, master's students, researches related to formal grammar of the Uzbek language, corpus linguistics, machine translation, speech processing systems, intelligent systems, computer lexicography and linguistic ontologies are published.

The magazine supplement is published four times a year.

Articles written in Uzbek, Turkish, Russian and English languages are accepted.

The journal also publishes book reviews, literature reviews, conference reports, and research project results.

The opinion of the authors may differ from the editorial point of view.

"Uzbekistan: language and culture. "Computer Linguistics" series has been published since 2023.

Tashkent State University of Uzbek Language and Literature named after Alisher Navoi. Yusuf Khos Hajib street, 103, Yakkasaray district, Tashkent, Uzbekistan.

E-mail: kompling@navoiy-uni.uz

Website: kompling.tsuull.uz

MUNDARIJA

Mastura Primova

Til korpuslarida matnlarni annotatsiyalash: afzallik va kamchiliklari.....6

Nilufar Muradova

Clarin tizimidagi og'zaki korpuslar xususida.....19

Noila Matyakubova

Iboralarni moslashtirish (phrase alignment)da otli va
fe'lli so'z birikmalar mosligi.....28

Ruxsora Muftillayeva

Dialektal korpuslarning umumiy tavsifi: tajriba va tahlil.....38

Sabura Xudayarova

Jahon tilshunosligida tabiiy tilni modellashtirish nazariyasi va
amaliyoti.....49

Jahongir Berdiyev

Tensorflow kutubxonasining imkoniyatlari.....63

CONTENT

Mastura Primova

Advantages and disadvantages of corpus annotation.....17

Nilufar Muradova

Specifically oral corpuses in the clarin system.....27

Noila Matyakubova

Aligning noun and verb phrases in phrase alignment36

Ruxsora Muftillayeva

General description of dialectal corpses: experiment and analysis.....48

Sabura Xudayarova

Theory and practice of natural language modeling
in world linguistics.....62

Jahongir Berdiyev

Tensorflow library capabilities.....72

JAHON TILSHUNOSLIGIDA TABIIY TILNI MODELLASHTIRISH NAZARIYASI VA AMALIYOTI

Sabura Xudayarova¹

Annotatsiya. Texnika paydo bo'lganidan beri bir necha o'n yillar o'tdi, bu esa kompyuter texnologiyalari shakllanishi va tez rivojlanishini hayotga olib keldi. Tilshunoslikning kompyuter kabi noan'anaviy sohasining rivojlanishi lingvistik bilim va jarayonlarni modellashtirish zaruriyatini keltirib chiqardi. Ushbu maqolada jahon tilshunosligida tabiiy tilni qayta ishlash hamda lingvistik modellashtirish masalalarida olib borilgan tadqiqotlarning farqli va o'xshash jihatlarini tahlil qilish orqali to'xtalib o'tilgan.

Kalit so'zlar: *Model, modellashtirish, deduktivlik, anilitik, sintetik, meta-til, N-gramm, ELM dasturi, Open AT, Neyron.*

Tabiiy til va uning turli xil modellari mavjud. Til faoliyatining har xil turlarida ishlash esa, ko'pincha ushbu modellar o'rtasidagi munosabatlar noaniq bo'lib qolishiga olib keldi. Shuning uchun modellashtirishdagi umumiy xususiyatlarni o'rganish ehtiyoji tug'ildi. Bu modellarni solishtirish imkonini beradigan lingvo-falsafiy asosdir. Ularning o'zaro asosiy o'xshashliklari va farqlarini aniqlab berish vazifasi dolzarb masalalardan biridir, chunki bu sohada o'xshashlik va farqlar mavjud.

Tilshunoslikda model - bu tilshunos tomonidan haqiqiy yoki sun'iy ravishda yaratilgan. Uning xatti-harakatlarini takrorlaydigan, taqlid qiladigan aqliy qurilma (odatda soddalashtirilgan shakl) lingvistik maqsadlar uchun asl nusxaning xatti-harakatidir.

Tilni modellashtirish texnik modellashtirishdan farqli ravishda, o'ziga xos xususiyatlarga ega. L.N.Murzinning fikricha, obykti bevosita kuzatuvga berilmagan har qanday fan obyektini modellashtirishga majbur bo'ladi. Modellashtirish umumilmiy metod bo'lib, deduktivlik, tafakkur eksperimentidan foydalanish va modelni ideal obyekt sifatida talqin qilish kabilar bilan xarakterlanadi.

¹Xudayarova Sabura Shuxrat qizi -Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti, Kompyuter lingvistikasi yo'nalishi 2-kurs magistranti.

E-pochta: xudayorovasabura@navoiy-uni.uz

ORCID: 0000-0001-7102-0824

Lingvistik modellashtirish berilgan semiotik elementlar o'rtasidagi munosabatlar tizimi sifatidagi til bilan yoki verbal kommunikatsiya jarayoni sifatidagi til bilan ish ko'radi.

Demak, modellashtirish struktur tilshunoslik, psixolingvistika, dinamik va kommunikativ lingvistikalarning obyekti sifatida namoyon bo'ladi. Til modeli haqida fikr yuritar ekan, I.I.Revzin shunday yozadi: "Model quyidagicha quriladi: muayyan fan tomonidan to'plangan xilma-xil tushunchalar yig'indisidan birlamchi deb qabul qilinuvchi ayrimlari tanlab olinadi. Birlamchilar o'rtasidagi munosabatlar aniqlanadi va ular pastulat sifatida qabul qilinadi. Qolgan barcha da'volar qat'iy deduktiv ravishda terminlarda namoyon bo'ladi. Ular birlamchi tushunchalar orqali aniqlanadi. Shu ma'noda model sistema sifatidagi tilning qismi emas, balqi qandaydir konstrukt, gipotetik ilmiy qurilma sifatida qabul qilinadi" [Tulupova, Pavlenko].

Lingvistik model tushunchasi tarkibiy tilshunoslikda paydo bo'lgan. K.L.Buler, Z.Z.Harris, Ch.Hokket, ammo XX asrning 60-70-yillarida ilmiy foydalanishga kiradi.

Matematik tilshunoslikning paydo bo'lishi va g'oyalarning tilshunoslikka kirib borishi, turli kibernetika usullari shakllanib bordi. Maqolada mualliflar tamonidan modellashtirish haqida fikr yuritilar ekan uning quydagi turlari e'tirof etiladi.

"Tilni bilishning qaysi tomoni mavzu ekanligiga qarab modellashtirish, nutq faoliyati modellari til modellariga bo'linadi to'g'ri va to'g'ri farqlash qobiliyatini taqlid qiluvchi grammatik to'g'rilik tildagi noto'g'ri va funktsional, o'zaro bog'liqlik qobiliyatini taqlid qilish nutqning mazmuni (tarkib rejasi) uning shakli (ifoda rejasi) bilan.

Modelning "kirish" va "chiqish" dagi ma'lumot turiga qarab grammatik to'g'rilik tan oluvchi va generativga bo'linadi.

Tan olish modeli: Masalan, K. Aydukevichning "kategorik grammatikasiga kirish" da tabiiy tilda yoki uning ichida matnning bir qismini oladi sun'iy tilda mavhum vakillik va "chiqish"da javob beradi, berilgan segment grammatik jihatdan to'g'ri yoki g'ayritabiiy bo'ladimi.

Generativ model: Masalan, N. Xomskiyning "Generativ grammatika" nasaziyasiga ko'ra bu taniydiganga teskari. Birinchisini tanqidiy yengish orqali Xomskiyning "Generativ grammatika" versiyasi modelni yaratishga olib keldi. Generativ semantika (J. Lakoff) modellar bilan ko'p o'xshashliklarga ega gapirish yoki sintez qilish. Nutq faoliyatining qaysi jihati modellashtirilganiga

qarab – tinglash yoki gapirish-funksional modellar quyidagilarga bo'linadi:

1. Analitik
2. Sintetik.

Ba'zi tillarning to'liq analitik modeli kirish fikrlardan tushuniladiki, bunda matnning bir qismini oladi. Odatda, bayonotdan kam emas maxsus semantik meta-til (ya'ni uning talqini). To'liq sintetik to'liq analitik bilan teskari bo'lgan ba'zi tillarning modelidir.

Yuqoridagi fikirlardan tushuniladiki, modellashtirish tili - bu izchil qoidalar to'plami bilan belgilangan tuzilmadagi ma'lumot, bilim yoki tizimlarni ifodalash uchun ishlatilishi mumkin bo'lgan har qanday sun'iy til. Qoidalar tarkibiy qismlarning ma'nosini izohlash uchun ishlatiladi. Modellashtirish tili grafik yoki matnli bo'lishi mumkin. Grafik modellashtirish tillari cheklovlarini ifodalash uchun tushunchalar va belgilarni bog'laydigan chiziqlarni, shuningdek munosabatlarni va boshqa turli xil grafik izohlarni ifodalovchi nomlangan belgilar bilan diagramma usullaridan foydalanadi. Matnni modellashtirish tillari odatda standartlashtirilgan kalit so'zlardan, so'ngra parametrlar bilan kompyuter tomonidan talqin qilinadigan ifodalarni yaratish uchun foydalanadi. Dasturlash, hisoblash tilshunosligining eng nufuzli maktablaridan biri N.Xomskiy yaratgan yo'nalish - uning lingvo-falsafiy asoslaridir.

Soha doirasidagi bilimlar bu muallif tomonidan ham juda chuqur o'rganilgan. Ushbu yo'nalishda uning izdoshlari ham Raqobatbardosh yondashuvlarga asoslanib, daraja yoki tabaqalanish nazariyasi hamda, lingvistik modellar yuzasidan tadqiqotlar olib bordilar.

Tilshunoslikdagi bu ikki asosiy yo'nalishning o'ziga xos belgisitik modellashtirish [Шалыпина, 2007]. Ushbu monografiyada yuqoridagi tadqiqotlar yuzasidan fikr yuritilar ekan, muallif monografiya tadqiqotlarning nazariy jihatlarini ochib berishga qaratilganligini ta'kidlaydi. Til faoliyatining u yoki bu turi doirasida tilni rivojlantirish lingvistik aloqa predmetining harakati sifatida talqin qilinishi mumkin. Qanchalik ikki o'lchovli va undan ham ko'proq uch o'lchovli fazoda bir nuqtadan ikkinchisiga o'tish uchun turli yo'llar ishlatilishi mumkin, bu lingvistik modellar doirasidagi asosiy o'zgaruvchanlikni aks ettiradi

Monografiyada TmS modelining tavsifi yetti bobda tuzilgan. 1-bob umumiy "metodlar-lingvistik voqelikning modeli" bo'lib, ular ichida to'plangan lingvistik kompetensiya va lingvistika o'rtasidagi qarama-qarshilik sifatida o'z o'rniga ega bo'lar edi. Turli lingvistik

vaziyatlarda birining va ikkinchisining modellari orasida harakat qilish. Shu nuqtayi nazardan qaraganda, til kompetensiyasi modeliga asoslanishi kerak, ularning har birida o'z tilshunoslik to'plami darajalari. Bu qayta tiklash uchun asosiy imkoniyat yaratadi.

Bunday darajalarning bir o'lchovli tizimidan ko'p o'lchovli va mos keladigan tizimga o'tish maxsus til kompetensiyasi modelini ko'p o'lchovli sifatida qurish – ichida uch o'lchovli til makonining minimal versiyasi. Chiziqli uch o'qga nisbatan talqin qilish va tushuntirish imkoniyatlari bir qator umume'tirof etilgan lingvistik tushunchalarning gistik fazosi doirasida fikr yuritiladi.

Morka Pronchai tomonidan 2022-yil 13-dekabrda chop etilgan, "Til modellari va uning turlarini tushuntirish" ga bag'ishlangan maqolasida, tilning evolyutsiyasi uni hayratda qoldirganligini ta'kidlar ekan, til modeli haqida shunday fikr yuritadi: "Til modeli so'zlarga ehtimolliklarni belgilash uchun mashinani o'rganishdan foydalanadi, so'zlarni bashorat qilishda foydalaniladi. Oldingi yozuvga asoslangan jumladagi keyingi so'z. Til modellari matndan o'rganiladi va ular manba matnini yaratish, matndagi keyingi so'zni bashorat qilish, nutqni aniqlash, optik belgilarni aniqlash va qo'l yozuvini tanib olish uchun ishlatilishi mumkin". Muallif mulohazalarini davom etdirib, Mashina qanday qilib jurnalistik maqolaga taqlid qiladigan maqola yaratishi mumkin? Til modeli nima? Shu kabi savollarga javob berishi bilan bir qatorda, til modellarining bir qancha turlari haqida to'xtalib ularni tushuntirib o'tadi: "Til modeli amalda, bu so'zlarning ma'lum bir ketma-ketligi "haqiqiy" bo'lish ehtimolini beradi. Ushbu kontekstdagi haqiqiylik grammatik aniqlikka ishora qilmaydi. Buning o'rniga, bu odamlar qanday yozishga o'xshashligini anglatadi va til modeli shunga o'rgatilgan. Bu muhim nuqta. Til modelida boshqa mashina o'rganish modellaridagi kabi sehr yo'q, ayniqsa chuqur neyron tarmoqlari, bu shunchaki namunadan tashqari kontekstda qayta ishlatilishi mumkin bo'lgan katta hajmdagi ma'lumotlarni siqilgan shaklga kiritish vositasidir.

Ehtimolli til modeli N-gramm ehtimolliklarni hisoblash yo'li bilan oddiy ehtimolli til modeli quriladi. N-gramma n so'zdan iborat ketma-ketlik bo'lib, bu yerda n noldan katta butun sonidir. N-gramm ehtimoli - n-grammning oxirgi so'zi ma'lum bir n-1 grammdan keyin kelishining shartli ehtimolligi (oxirgi so'zga e'tibor bermaslik). Bu oxirgi so'zning n-1 grammdan keyin oxirgi so'z bo'lmagan holatlarining nisbati. Bu kontsepsiya Markovning taxminidir. Berilgan n-1 gramm (hozirgi), n-gramm (kelajak) ehtimolliklari n-2, n-3 va hokazo grammlarga (o'tmish) bog'liq emas. Bu yondashuvning

aniq kamchiliklari mavjud. Eng muhimi, faqat oldingi n soʻz keyingi soʻzning ehtimollik taqsimotiga taʼsir qiladi. Murakkab matnlar keyingi soʻzni tanlashga hal qiluvchi taʼsir koʻrsatishi mumkin boʻlgan chuqur kontekstga ega. Bu oʻrinda olim mavzu doirasidagi bir qancha muammolar va ularning yechimlari borasidagi qarashlarini keltirib oʻtadi.

Keyingi soʻz oldingi n soʻzdan aniq boʻlmasligi mumkin, garchi n 20 yoki 50 boʻlsa ham. Bu atama oldingi soʻzni tanlashga taʼsir qiladi: “United” soʻzidan keyin Davlatlar boʻlsa, ehtimol ancha yuqori. Bundan tashqari, Hajmi (n) oshgani sayin, mumkin boʻlgan almashtirishlar soni keskin ortadi, garchi ularning aksariyati matnda hech qachon koʻrinmaydi va barcha hosil boʻlgan ehtimolliklarni (yoki N-grammlarning barcha raqamlarini) hisoblash va saqlash kerak. Bundan tashqari, yuzaga kelmaydigan N-grammlar kamlik muammosini keltirib chiqaradi, chunki ehtimollik taqsimotining granularligi juda past boʻlishi mumkin. Soʻz ehtimollari bir necha xil maʼnoga ega, shuning uchun koʻpchilik soʻzlar bir xil ehtimolga ega.

Neyral tarmoqqa asoslangan til modellari: Neyron tarmoqqa asoslangan til modellari kirishlarni kodlash yoʻli bilan siyraklik muammosini osonlashtiradi. Soʻzni joylashtirish qatlamlari har bir soʻzning semantik munosabatlarini oʻz ichiga olgan ixtiyoriy oʻlchamdagi vektorini yaratadi. Ushbu uzluksiz vektorlar keyingi soʻzning ehtimollik taqsimotida juda zarur boʻlgan granularlikni yaratadi. Bundan tashqari, til modeli funksiyadir, chunki barcha neyron tarmoqlar juda koʻp matritsali hisob-kitoblarga ega, shuning uchun keyingi soʻzning ehtimollik taqsimotini ishlab chiqarish uchun barcha n-gramm sonlarini saqlash shart emas. Viktor Lavrenko Til modellarning evolyutsiyasi Neyron tarmoqlar siyraklik muammosini hal qilsa ham, kontekst muammosi saqlanib qolmoqda. Birinchidan, kontekst muammosini yanada samaraliroq hal qilish uchun til modellari ishlab chiqildi - ehtimollik taqsimotiga taʼsir qilish uchun tobora koʻproq kontekstli soʻzlarni keltirdi. Ikkinchidan, maqsad modelga qaysi kontekstdagi soʻzlar boshqalardan koʻra muhimroq ekanligini oʻrganish imkoniyatini beruvchi arxitekturani yaratish edi. Til modeli nima qila oladi? Kontekstdan soʻz ehtimoli haqida xulosa chiqarish uchun zarur boʻlgan tabiiy tilni mavhum tushunish bir qator muammolarni hal qilish uchun ishlatilishi mumkin. Lemmatizatsiya yoki stemming qaratilgan soʻzni eng asosiy shakliga qisqartirish va shu bilan tokenlar sonini keskin kamaytirish. Agar soʻzning nutq qismi maʼlum boʻlsa, bu algoritmlar yaxshi ishlaydi. Feʼl postfikslari ot postfikslaridan farq qilishi mumkin, shuning

uchun til modeli uchun umumiy vazifa bo'lgan nutq qismini teglash (yoki POS teglari) uchun mantiqiy asos. Yaxshi til modeli bilan biz matnlarni ekstraktiv yoki mavhum umumlashtirishni amalga oshirishimiz mumkin. Agar bizda turli tillar uchun modellar mavjud bo'lsa, mashina tarjimasini tizimini osongina qurish mumkin. Kamroq oddiy foydalanish holatlari savollarga javob berishni o'z ichiga oladi (kontekstli yoki kontekstsiz)

Til modellari nutqni aniqlash, optik belgilarni aniqlash, qo'l yozuvini aniqlash va boshqalar uchun ham ishlatilishi mumkin. Imkoniyatlarning butun doirasi mavjud. Men ilgari aytib o'tgan birinchi model zich (yoki yashirin) qatlam va tepada joylashgan chiqish qatlamidir. Uzlusiz so'zlar sumkasi (CBOW) Word2Vec modeli. CBOW Word2Vec modeli so'zni kontekstdan taxmin qilishga o'rgatilgan. Skip-Gram Word2Vec modeli buning aksini qiladi, so'zdan kontekstni taxmin qiladi. Amalda, CBOW Word2Vec modeli uni o'rgatish uchun quyidagi strukturaning ko'plab misollarini talab qiladi: kirishlar n so'zdan oldin va/yoki so'zdan keyin, ya'ni chiqishdir. Ko'rishimiz mumkinki, kontekst muammosi hali ham o'zgarmagan.

Taktirilgan neyr tarmoqlari (RNN) Takroriy neyron tarmoqlari (RNN) bu masala bo'yicha yaxshilanishdir. RNNlar uzoq qisqa muddatli xotira (LSTM) yoki gated recurrent unit (GRU) uyali tarmoq bo'lishi mumkinligi sababli, keyingi so'zni tanlashda oldingi barcha so'zlarni hisobga oladi. AllenNLP ning ELMo dasturi bu tushunchani bir qadam oldinga olib boradi va ikki tomonlama LSTM dan foydalanadi, u so'zlarni hisoblashdan oldin va keyin kontekstni hisobga oladi.

Transformers rnn-ga asoslangan arxitekturalarning asosiy kamchiligi ularning ketma-ketligidan kelib chiqadi. Natijada, mashg'ulot vaqtlari uzoq ketma-ketliklar uchun ko'tariladi, chunki parallellashtirish imkoniyati yo'q. Bu muammoning yechimi transformator arxitekturasidir. OpenAI va Google BERT ning GPT modellari transformator arxitekturasidan ham foydalanadi. Ushbu modellar, shuningdek, "Diqqat" deb nomlangan mexanizmni qo'llaydi, uning yordamida model ma'lum holatlarda qaysi kirishlar boshqalarga qaraganda ko'proq e'tiborga loyiqligini bilib oladi. Model arxitekturasini nuqtayi nazaridan, asosiy kvant sakrashlari birinchi navbatda RNNlar, xususan, LSTM va GRU edi. siyraklik muammosi va disk maydonini kamaytiradigan til modellari, keyinchalik transformator arxitekturasini, parallellashtirishni mumkin va diqqat mexanizmlarini yaratadi. Ammo arxitektura til modeli ustun bo'lishi

mumkin bo'lgan yagona jihat emas. GPT-1 arxitekturasini bilan solishtirganda, GPT-3 deyarli hech qanday yangilikka ega emas. Lekin bu juda katta. U 175 milliard parametrga ega va u umumiy sudralib yurishda o'rgatilgan eng katta korpusda o'qitilgan. Bu qisman til modelining yarim nazorat ostida o'qitish strategiyasi tufayli mumkin. Ba'zi so'zlar tushirib qoldirilgan matndan o'quv namunasi sifatida foydalanish mumkin. GPT-3ning aqlbovar qilmaydigan kuchi shundaki, u so'nggi yillarda internetda paydo bo'lgan barcha matnlarni ko'proq yoki kamroq o'qigan va u o'z ichiga olgan tabiiy tilning murakkabligini aks ettirish qobiliyatiga ega. Ilgari, til modellari standart NLP vazifalari uchun ishlatilgan, masalan, nutq qismini (POS) teglash yoki engil o'zgartirishlar bilan mashina tarjimai. Bir oz qayta tayyorlash bilan BERT tabiiy tilning asosiy tuzilishini tushunishning mavhum qobiliyati tufayli POS-tegger bo'lishi mumkin. T5 bilan NLP vazifalari uchun hech qanday o'zgartirishga hojat yo'q. Agar u <M> tokenlari bo'lgan matnni olsa, u bu tokenlar tegishli so'zlar bilan to'ldirish uchun bo'shliqlar ekanligini biladi. Shuningdek, u savollarga javob berishi mumkin. Agar u savollardan keyin qandaydir kontekstni olsa, u javob uchun kontekstni qidiradi. Aks holda, u o'z bilimidan javob beradi. Qiziqarli fakt: u trivia viktorinasida o'z ijodkorlarini mag'lub etdi.

Yangi boshlanuvchilar uchun NLP. Til modellarining kelajagi bo'yicha to'liq qo'llanma. Ushbu soha sun'iy intellektni yaratishga eng yaqin bo'lgan sohamiz. AI atrofida juda ko'p shov-shuvlar mavjud va ko'plab oddiy qaror tizimlari va deyarli har qanday neyron tarmoq AI deb ataladi, ammo bu asosan marketing. Ta'rifga ko'ra, sun'iy intellekt mashina tomonidan amalga oshiriladigan insonga o'xshash aql qobiliyatlarini o'z ichiga oladi. Transfer o'rganish kompyuterni ko'rish sohasida porlaydi va AI tizimi uchun uzatishni o'rganish tushunchasi juda muhim bo'lsa-da, bir xil model NLP vazifalarining keng doirasini bajarishi va kiritilgan ma'lumotlardan nima qilish kerakligi haqida xulosa chiqarishi haqiqatdir. Yuqoridagi fikirlarni tahlil qilish natijasida shuni aytish mumkinki, demak, til modeli gapdagi kontekst va atrofdagi so'zlarga qarab, til model kontekstli yoki kontekstsiz savollarga javoblaringizni o'zgartirish mumkin. Ular javoblarni turli yo'llar bilan taqdim etishlari mumkin, masalan, aniq iboralarini olish, javobni tarjima qilish yoki ro'yxatga olish variantlarini tanlang. *Matnni umumlashtirish.* Til modellari hujjatlar, maqolalar, podkastlar, videolar va boshqalarni avtomatik ravishda eng muhim qismlarga yuklash uchun ishlatish mumkin. Modellar ikki usulda yaratish mumkin: manba matnidan eng muhim ma'lumotlarni

ajratib olish yoki asl tilni takrorlamaydigan yukni taqdim etish. *Hissiyot tahlili*. Tilni modellash tirishni hissiyotlarni tahlil qilish vositalari uchun yaxshi imkoniyatlardir, chunki u matnlarni ovoz ohangini va semantik yo'nalishini aniqlay oladi.

Suhbatdosh sun'iy intellekt tizimlarining bir qismi sifatida til modellari rivojlanishiga mos matn javoblarini taqdim etishi mumkin. *Mashina tarjimasini*. Mashinani o'rganishga samarali til modellarining uzoq kontekstlarni samarali umumlashtirish uchun mashina tarjimasini yaxshilash berdi. Matnni so'zma-so'z tarjima qilish o'rniga, til modellari kirish va chiqish ketma-ketliklarining tasvirlarini o'rganishi va ishonchli tuzatishni berishi mumkin.

Kodni to'ldirish. Yaqinda keng ko'lamlil til modellari kodni belgilab, tahrirlash va joriyning ta'sirchanligini namoyish qildi. Ular faqat oddiy dasturlash vositalarini bajarishlari mumkin, ko'pincha kodga tarjima qilishlari yoki xatoliklarni bajarishlari mumkin. Til modellarining turlari til modellari turli xil bo'lib, ularning ikki darajaga bo'lishi mumkin: statistik modellar va chuqur neuron tarmoqlarga barqaror modellar. Statistik til modellari

Statistik til modellari so'zlarning ma'lum ketma-ketligini bashorat qilish uchun ma'lumotlardagi statistik naqshlardan qaror topgan model turidir. Ehtimoliy til modelini olishning asosiy belgilarini N-gramm bilib olishlarini aytishdir.

N-gramma so'zlar ketma-ketligi bo'lib, n noldan katta sonidir. Tilning oddiy probabilistik modelini kiritish uchun siz matnda turli xil N-grammlar (so'z birikmalari) paydo bo'lishi mumkinligini hisoblaysiz. Bu har bir necha marta kelganini va oldingi so'zning paydo bo'lishi soniga bo'lish orqali amalga oshirish natijasida bajariladi. Bu fikr Markov kontseptsiyasiga parallel ravishda unda so'z birikmasi (kelajak) bundan boshqa (o'tmish) so'zlarga emas, balki faqat oldingi so'zga (hozirgi) bog'liqligini bildiradi. N grammalar nisbatan sodda va samarali, ular ketma-ketlikda so'zlarning uzoq muddatli kontekstini yuklab olish mumkin shu bilan birga til modelining quydagi turi haqida fikr yuritmoqchimiz.

Neyron tili modellari

Neyron tili modellari, nomidan ko'rinib turibdiki, so'zlar ketma-ketligi ehtimolini bashorat qilish uchun neyron tarmoqlardan foydalanadi. Ushbu modellar katta hajmdagi matnli ma'lumotlarga o'rgatilgan va tilning asosiy tuzilishini o'rganishga qodir. Ikkita yashirin qatlamli neyron tarmoq arxitekturasini.

Ular katta lug'atlarni boshqarishi va taqsimlangan vakilliklardan foydalangan holda noyob yoki noma'lum so'zlar bilan

shug'ullanishi mumkin. NLP vazifalari uchun eng ko'p qo'llaniladigan neyron tarmoq arxitekturalari takrorlanuvchi neyron tarmoqlari (RNN) va transformator tarmoqlaridir (bularni keyingi bo'limda ko'rib chiqamiz).

Neyron tili modellari kontekstni an'anaviy statistik modellarga qaraganda yaxshiroq ishlay oladi. Bundan tashqari, ular murakkabroq til tuzilmalarini va so'zlar orasidagi uzoqroq bog'liqlikni boshqarishi mumkin. Keling, RNN va transformator kabi neyron tili modellari buni qanday qilishini aniq tushunamiz. Tabiiy tilni qayta ishlash kontekstida statistik model oddiyroq til tuzilmalarini boshqarish uchun etarli bo'lishi mumkin. Biroq, murakkablik oshgani sayin, bu yondashuv samarasiz bo'ladi. Misol uchun, juda uzun statistik model matnlari bilan ishlaganda, aniq bashorat qilish uchun zarur bo'lgan barcha ehtimollik taqsimotlarini eslab qolish qiyin bo'lishi mumkin.

Buning sababi, 100 000 so'zdan iborat matnda model 100 000 ehtimollik taqsimotini eslab qolishi kerak. Va agar model ikki so'zni orqaga qaytarishi kerak bo'lsa, eslashi kerak bo'lgan taqsimotlar soni 100 000 kvadratgacha oshadi. Bu yerda RNN kabi murakkabroq modellar rol o'ynaydi. Takroriy neyron tarmoqlari (RNN) keyingi kirishlarni qabul qilishda oldingi natijalarni eslay oladigan neyron tarmoq turidir. Bu kirish va chiqishlar bir-biridan mustaqil bo'lgan an'anaviy neyron tarmoqlardan farqli o'laroq. RNNlar, ayniqsa, jumladagi keyingi so'zni taxmin qilish kerak bo'lganda foydalidir, chunki ular jumladagi oldingi so'zlarni hisobga olishlari mumkin. RNN ning asosiy xususiyati yashirin holat vektori bo'lib, ketma-ketlik ma'lumotlarini eslab qoladi. Ushbu "xotira" RNN ga barcha hisoblangan ma'lumotlarni kuzatib borish va bu ma'lumotlardan bashorat qilish uchun foydalanish imkonini beradi. Yashirin holat tarmoqdagi yashirin qatlam tomonidan saqlanadi. Biroq, RNN hisoblash qimmat bo'lishi mumkin va juda uzoq kirish ketma-ketligi uchun yaxshi o'lchovga ega emas. Gap uzunroq bo'lganda, boshlang'ich so'zlardan olingan ma'lumotlar ko'chiriladi va gapning qolgan qismi bilan birga uzatiladi. RNN jumlaning oxirgi so'ziga yetib borgunga qadar, birinchi so'zdan olingan ma'lumotlar nusxaning nusxasiga aylandi va ko'p marta o'zgartiriladi.

Lingvistika hamda kompyutor-texnologiyalari sohasida 2020-yilda yana bir ajoyib sun'iy intellekt yaratildi GPT-3 deb nomlangan va San-Fransiskoda OpenAI tomonidan ishlab chiqilgan bo'lib, u kitoblar, maqolalar va veb-saytlardan milliardlab so'zlarni o'zlashtirgandan so'ng ravon matn yaratishga qodir "katta til modeli". GPT-3 shu qadar rivojlanganki, ko'pchilik model tomonidan

yaratilgan yangiliklarni inson mualliflari yozgan yangiliklardan farqlashda qiynalgan. GPT-3-da suhbat vazifari uchun maxsus sozlangan ChatGPT bolalar versiyasi mavjud. Ushbu yutuqlar bilan tilni modellashtirish kontsepsiyasi butunlay yangi davrga kirdi. OpenAI ChatGPT Til modellari tabiiy tilni qayta ishlashning (NLP) asosiy komponentidir, chunki ular mashinalarga inson tilini tushunish, yaratish va tahlil qilish imkonini beradi. Ular asosan kitoblar yoki maqolalar to'plamlari kabi matnli ma'lumotlarning katta to'plamidan foydalangan holda o'qitiladi. Keyin modellar ushbu o'quv ma'lumotlaridan olingan belgilardan jumladagi keyingi so'zni bashorat qilish yoki grammatik jihatdan to'g'ri va semantik jihatdan mos keladigan yangi matn yaratish uchun foydalanadilar.

Tilshunoslik hamda kompyuter texnologiyalarining integrallashuvi jahon tilshunosligida ko'plab amaliy va nazariy tadqiqotlarning olib borilishiga asos bo'lib xizmat qilmoqda. Mana shunday tadqiqotlar sohaning yana-da rivojlanishini ta'minlaydi. Linivistik modellar tilshunoslikka doir masalalarni kompyuter dasturlari yordamida hal etilishida muhim vositadir ana shunday lingvistik dasturlar esa tilni modellashtirish asosida yaratiladi.

Til modellarini qo'llash natijasida Google Gboard va Microsoft SwiftKey klaviaturalaridagi matnli xabarlarini yozishda jumalarni yakunlash bo'yicha avtomatik takliflarni taqdim etuvchi aqlli funksiyalar Ishlab chiqildi. Bundan tashqari kontent yaratish ham shular jumlasidandir. Til belgilarining aniq namoyon bo'ladigan sohalardan biri kontent yaratishdir. Bu odamlar tomonidan taqdim etilgan ma'lumotlar va atamalardan to'liq matnlar yoki matn qismlarini yaratishni o'z ichiga oladi. Kontent yangiliklar maqolalari, press-revizlar va blog postlaridan tortib onlayn-do'kon mahsulot tavsiflarini qamrab oladi. Shu bilan birga til modellari POS yorlig'i vazifalarida eng so'nggi natijalarga erishish uchun keng qo'llaniladi. POS teglash - bu matndagi har bir so'zni ot, fe'l, sifat va boshqalar kabi nutqning tegishli qismi bilan belgilash jarayonidir.

Tabiiy tilni qayta ishlash til modellarini kishlab chiqish jahon tilshunosligida barcha rivojlangan tillar qatorida fransuz tilshunosligini ham tadqiqotlar bilan boyitish zaruriyatini tug'dirdi.

Mana shunday amaliy hamda nazariy tadqiqotlar natijasida avtoregressiv til modellarining hajmini oshirish va tayyorlash nol va past bosqichli o'rganish yordamida tabiiy tilni qayta ishlash muammolarini hal qilishning yangi usullarini topishga imkon berdi. GPT-3 kabi ekstremal miqyosdagi til modellari ko'p tilli imkoniyatlarni taklif qilsa-da, ingliz tilidan boshqa tillar deyarli o'rganilmagan.

Fransuzlar tomonidan fransuz tili uchun maxsus tayyorlangan katta ochiq manbali avtoregressiv til modeli taqdim etiladi. Natijalar shuni ko'rsatadiki, Cedille mavjud fransuz modellaridan ustundir va GPT-3 bilan bir qator fransuz ko'rsatkichlari bo'yicha raqobatlashadi. Bundan tashqari, ushbu modellar tomonidan ko'rsatilgan toksiklikni chuqur taqqoslashni taqdim etadi, bu Sedille ma'lumotlar to'plamini filtrlash tufayli til modellarining xavfsizligi yaxshilanishini ko'rmoqda. Camembert Fransuz til modeli CamemBERT - yangi ko'p tilli OSCAR korpusining fransuz sub-korpusida oldindan tayyorlangan RoBERTa arxitekturasiga asoslangan fransuz tili uchun eng zamonaviy til modeli. Bu CamemBERTni fransuz tili uchun to'rt xil quyi oqim vazifasi bo'yicha baholaydi. Nutqning bir qismini (POS) teglash, bog'liqlik tahlili, nomli obyektlarni aniqlash (NER) va tabiiy til xulosasi (NLI); oldingi monolingual va ko'p tilli yondashuvlar bilan solishtirganda ko'pgina vazifalar bo'yicha eng zamonaviy texnologiyalarni yaxshilash, fransuz tili uchun oldindan tayyorlangan katta til modellarining samaradorligini tasdiqlaydi.

CamemBERT Lui Martin, Benjamin Myuller, Pedro Ortiz Suares, Yoan Dupon, Loran Romari, Erik Villemonte de la Klerjeri, Jamet Sedda va Benoit Sago tomonidan o'rganilgan va baholangan.

Tilshunoslikda modellashtirishning asosiy maqsadi shaxsning yaxlit lingvistik qobiliyatini modellashtirishdir. Hozirgi tilshunoslikda "model" atamasining mazmuni asosan "nazariya" atamasi bilan ilgari yoritilgan, ayniqsa, Yelmslev tomonidan. Faqatgina bunday nazariya model nomiga loyiq deb hisoblanadi. Bu juda aniq ifodalangan va yetarlicha rasmiylashtirilgan (ideal holda, har bir model kompyuterda amalga oshirish imkonini berishi kerak) Modelni loyihalash nafaqat lingvistik hodisalarni aks ettirish vositalaridan biri, balki til hodisalari haqidagi bilimlarning haqiqatini tekshirishning obyektiv amaliy mezonidir. Til o'rganishning boshqa usullari bilan birlikda modellashtirish nutq faoliyatining yashirin mexanizmlari, uning nisbatan ibtidoiy modellardan tilning mohiyatini to'liqroq ochib beruvchi yanada mazmunli modellarga o'tishi haqidagi bilimlarni chuqurlashtirish vositasi bo'lib xizmat qiladi. Tizim sifatida til tarkibida mavjud.

Modellashtirish prinsipi: uning ayrim quyi tizimlari boshqalarni modellashtiradi, Modellashtirish usuli odatda ishora tizimlariga tayanadi, lekin tilning o'zi ishora tizimidir, yani. So'zlarni so'zlar yordamida modellashtiriladi. Har qanday model, jumladan, lingvistik model ham aniq bo'lishi kerak. Model, agar u ularning bayonotlarini bog'laydigan boshlang'ich obyektlarni

va ular bilan ishlash qoidalarini (yangi obyektlar va bayonotlarni shakllantirish yoki identifikatsiya qilish qoidalari) aniq ko'rsatilishi lozim. Formallik, to'g'rilik, noaniqlik nazariya taqdim etilayotgan tilning xossasidir. Bu xususiyat o'z-o'zidan rasmiy nazariyaning bashoratlari obyektiv eksperimental ma'lumotlar bilan mos kelishini ta'minlamaydi. Formal model eksperimental ma'lumotlar bilan bitta yoki boshqa talqin. Modelni talqin qilish deganda model obyektlari (ramzlari) o'rniga ma'lum bir predmet sohasi obyektlarini, masalan, tilni almashtirish uchun ehtimollik yoki qat'iy qoidalarini ko'rsatish tushuniladi.

Tilshunoslikdagi modellar tizimi tilshunoslikda tilni va uning alohida tomonlarini (fonologik, grammatik, leksik va boshqa tizimlar) tavsiflash uchun lingvistik tushunchalar va ular o'rtasidagi aloqalarni aniqlashtirish uchun qo'llaniladi, bu til hodisalarining cheksiz xilma-xilligi asosida yotgan tuzilmalarni aniqlashga yordam beradi. Qo'llash sohasiga qarab, modellar fonologik, morfologik, sintaktik va semantiklarga bo'linadi. Matematikani qurishda matematik tilshunoslikning vositalari va usullaridan foydalaniladi. Har qanday Modelda quyidagilar qayd etiladi: bevosita kuzatish ma'lumotlariga mos keladigan obyektlar — tovushlar, so'zlar, gaplar to'plami; tadqiqotchi tomonidan tavsiflash uchun qurilgan obyektlar (konstruksiyalar) qat'iy cheklangan toifalar, xususiyatlar, elementar semantik tuzilmalar va boshqalar to'plamidir.

Har bir model butun tilni emas, balki uning ma'lum bir sohasini yoki hatto alohida toifasini tavsiflaganligi sababli, tilning aniq tavsifi tilning bir sohasiga tegishli turli xil modellardan bir vaqtning o'zida foydalanishni o'z ichiga oladi [Советов, Яковлев, 2001]. Tilshunoslarni insonning til qobiliyati qanday tuzilganligi, tillar qanday rivojlanishi va ular o'zaro qanday munosabatlarda mavjudligi, shuningdek, tabiiy tillardagi matnlar qanday xususiyatlarga ega ekanligi ko'proq qiziqtiradi. Ushbu savollarni o'rganish uchun tilshunoslar ko'pincha matematik usullardan foydalanishlari kerak.

1950-1960-yillarda matematik tilshunoslikning boshlanishi-da, matematik mantiq usullari eng ommabop bo'lgan va hozirda, kompyuterlarning sohaga kirib kelishi hamda kompyuter va lingvistikaning integrallashuvi natijasida, turli xil statistik algoritmlar tobora muhim rol o'ynamoqda. Matematik mantiqning g'oyalari va usullari til *hodisalarining* tavsifini yanada formal, shuning uchun yanada qat'iyroq qilish imkonini beradi va miqdoriy usullarni katta hajmdagi lingvistik ma'lumotlarga qo'llash bizga birinchi qarashda

ko'rinmaydigan aloqalar va belgilarni aniqlash imkonini beradi. Shuningdek, matnlarni tabiiy tilda qayta ishlash: masalan, matn terish xatolarini tuzatish yoki matnlarni bir tildan boshqa tilga tarjima qilish. Lingvistik model hamda modellashtirish yuzasidan soha doirasida zamonaviy tilshunoslikning turli yo'nalishlaridagi muammolar va ularni hal qilinishida matematik usullar asos bo'lib xizmat qiladi.

Bir so'z bilan aytganda tabiiy tilni qayta ishlashda lingvistik modellar asos sifatida qaralar ekan, jahon tilshunosligida tilni modellashtirish masalasiga bo'lgan ilmiy nazariy hamda amaliy yondoshuvlari kompyuter lingvistikasi doirasida o'rganilayotgan har bir sohasi dolzarb ekanligini belgilash bilan bir qatorda uning yanada rivojlanishiga zamin yaratadi.

Foydalanilgan adabiyotlar

- Баранов А.Н. Введение в прикладную лингвистику. - М.: Изд-во ЛКИ, 2007. 360 – с.
- Лосев А.Ф. Введение в общую теорию языковых моделей: учебное пособие. - М., 2004.
- Советов Б.Я., Яковлев С.А. Моделирование систем. - М.: Высшая школа, 2001.
- Лихачев Д. С. Концептосфера русского языка // Русская словесность: антология. – М.: Academia, 1997. – С. 280–287.
- Попова З.Д., Стернин И.А. Понятие «концепт» в лингвистических исследованиях. – Воронеж, 1999. – С. 144.
- Бабушкин А. П. Типы концептов в лексико-фразеологической системе языка. – Воронеж: Изд-во ВГУ, 1996. – 104 с.
- Арутюнова Н. Д. Язык и мир человека. – М.: Языки русской культуры, 1999. – 896 с.
- Алефиренко Н. Ф. Проблемы вербализации концепта: теоретическое исследование. – Волгоград: Перемена, 2003. – 96 с.

THEORY AND PRACTICE OF NATURAL LANGUAGE MODELING IN WORLD LINGUISTICS

Sabura Xudayarova¹

Abstract. Several decades have passed since the advent of technology, which brought to life the formation and rapid development of computer technology. The development of such a non-traditional field of linguistics as computer has created the need to model linguistic knowledge and processes. This article focuses on the differences and similarities of the research conducted in the world linguistics on natural language processing and linguistic modeling.

Key words: *Model, modeling, deductive, analytical, synthetic, meta-language, N-gram, ELM program, Open AT, Neuron.*

References

- Baranov A.N. Vvedenie v prikladnuyu lingvistiku. - M.: Izd-vo LKI, 2007. 360 – s.
- Losev A.F. Vvedenie v obshchuyu teoriyu yazykovyx modeley: uchebnoe posobie. - M., 2004.
- Sovetov B.Ya., Yakovlev S.A. Modelirovanie sistem. - M.: Vysshaya shkola, 2001.
- Lixachev D. S. Kontseptosfera russkogo yazyka // Russkaya slovesnost': antologiya. – M.: Academia, 1997. – S. 280–287.
- Popova Z. D., Sternin I. A. Ponyatie «kontsept» v lingvisticheskix issledovaniyax. – Voronej, 1999. – S. 144.
- Babushkin A. P. Tipy kontseptov v leksiko-frazeologicheskoy sisteme yazyka. – Voronej: Izd-vo VGU, 1996. – 104 s.
- Arutyunova N. D. Yazyk i mir cheloveka. – M.: Yazyki russkoy kul'tury, 1999. – 896 s.
- Alefirenko N. F. Problemy verbalizatsii kontsepta: teoreticheskoe issledovanie. – Volgograd: Peremena, 2003. – 96 s.

¹Xudayarova Sabura Shuxrat qizi - Master of degree. Alisher Navo'i Tashkent State University of Uzbek Language and Literature.

E-mail: xudayorovasabura@navoiy-uni.uz

ORCID: 0000-0001-7102-0824

Jurnal 2017-yil 26-oktyabrda O'zbekiston Respublikasi Matbuot
va axborot agentligi tomonidan 0936-raqam bilan ro'yxatdan o'tgan.

Jurnal O'zbekiston Respublikasi Oliy Attestatsiya Komissiyasi tomonidan
filologiya fanlari bo'yicha falsafa doktori (PhD) va fan doktori (DSc)
dissertatsiyalari asosiy ilmiy natijalari chop etilishi lozim bo'lgan ro'yxatga
kiritilgan (30.10.2021. № 308/6).

Tahririyatga kelgan maqolalar mualliflarga qaytarilmaydi.

Manzil: Toshkent shahri, Yakkasaroy tumani, Yusuf Xos
Hojib ko'chasi 103-uy.
Telefonlar: +99871 281-45-11, +99871 281-41-93.
Website: compling.tsuull.uz
E-mail: kompling@navoiy-uni.uz

Bosishga 25.12.2023-yilda ruxsat etildi.
Bichimi 70x100 1/16, Ofset bosma. "Cambria" garniturasida.
Shartli b.t. 7,51. Nashr b.t. 7,62.

"O'zbekiston: til va madaniyat" jurnali tahririyatida
tayyorlandi va sahifalandi.
"YASHNOBOD NASHR" bosmaxonasida chop etildi.
Adadi 300 nusxa. Buyurtma №2.
Bosmaxona manzili: Toshkent shahar Yashnobod tumani,
58-a harbiy shaharcha.