



MUHAMMAD AL-XORAZMIY NOMIDAGI TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI FARG'ONA FILIALI



ANIQ VA TABIIY FANLARNI RIVOJLANTIRISHDA RAQAMLI TEXNOLOGIYALARNING O'RNI: MUAMMO VA INNOVATSION YECHIMLAR

mavzusidagi xalqaro ilmiy-texnik anjuman materiallari

РОЛЬ ЦИФРОВЫХ ТЕХНОЛОГИЙ В РАЗВИТИИ ТОЧНЫХ И ЕСТЕСТВЕННЫХ НАУК:
ПРОБЛЕМЫ И ИННОВАЦИОННЫЕ РЕШЕНИЯ

THE ROLE OF DIGITAL TECHNOLOGIES IN THE DEVELOPMENT OF EXACT AND
NATURAL SCIENCES: PROBLEMS AND INNOVATIVE SOLUTIONS



Farg'ona. 4-5-oktabr, 2024-yil.



**O‘ZBEKISTON RESPUBLIKASI OLIY TA’LIM, FAN VA
INNOVATSIYALAR VAZIRLIGI**

**O‘ZBEKISTON RESPUBLIKASI
RAQAMLI TEXNOLOGIYALAR VAZIRLIGI**

**MUHAMMAD AL-XORAZMIY NOMIDAGI
TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI
FARG‘ONA FILIALI**

**ANIQ VA TABIIY FANLARNI RIVOJLANTIRISHDA RAQAMLI
TEXNOLOGIYALARNING O‘RNI:
MUAMMOLAR VA INNOVATSION YECHIMLAR**

mavzusidagi xalqaro ilmiy-texnik konferensiya materiallar to‘plami
2-qism

4-5 oktyabr 2024 yil

Farg‘ona-2024

KONFERENSIYA TASHKILY QO'MITA TARKIBI:

Muxtarov F.M. — rais, TATU Farg'ona filiali direktori, t.f.b.f.d., dotsent.

Daliyev B.S. — rais o'rinbosari, tabiiy fanlar kafedrası mudiri, f-m.f.n.

Tashkiliy qo'mita a'zolari:

Polvonov B.Z. — TATU Farg'ona filiali ilmiy ishlar va innovatsiyalar bo'yicha direktor o'rinbosari, f-m.f.b.f.d., professor.

Abdullaev T.M. — TATU Farg'ona filiali o'quv ishlari bo'yicha direktor o'rinbosari, t.f.b.f.d., dotsent.

Zokirov S.I. — TATU Farg'ona filiali Ilmiy tadqiqotlar, innovatsiyalar va ilmiy-pedagogik kadrlar tayyorlash bo'limi boshlig'i, t.f.b.f.d., dotsent.

Otaqulov O.X. — TATU Farg'ona filiali Telekommunikatsiya muhandisligi va kasbiy ta'lim fakulteti dekani, t.f.n., dotsent.

Norinov M.U. — TATU Farg'ona filiali Kompyuter injiniringi va sun'iy intellekt fakulteti dekani, dotsent.

Sotvoldiev X.I. — TATU Farg'ona filiali dasturiy injiniring va kiberxavfsizlik fakulteti dekani.

To'xtasinov D.F. — TATU Farg'ona filiali sirtqi bo'lim boshlig'i.

Nurmatov M.M. — TATU Farg'ona filiali bosh hisobchisi.

Bozarov B.I. — TATU Farg'ona filiali Tabiiy fanlar kafedrası dotsenti.

Sabirov S.S. — TATU Farg'ona filiali Tabiiy fanlar kafedrası professor, t.f.n.

Movlonov P.I. — TATU Farg'ona filiali Tabiiy fanlar kafedrası katta o'qituvchisi.

Nasriddinov O.U. — TATU Farg'ona filiali Tabiiy fanlar kafedrası katta o'qituvchisi.

TAHRIRIYAT KENGASHI:

dotsent Sabirov S.S., dotsent Daliyev B.S., dotsent Bozarov B.I.,
dotsent To'xtasinov D.F., dotsent Umarov Sh.A., dotsent Zokirov S.I.

DASTUR QO‘MITA TARKIBI:

Maxkamov B.Sh.	— Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti rektori
Zaynobiddinov S.	— O‘zbekiston Respublikasi FA akademigi
Mamadalimov A.T.	— O‘zbekiston Respublikasi FA akademigi
Igamberdiev X.Z.	— O‘zbekiston Respublikasi FA akademigi
Turabekov B.T.	— Qozog‘iston Respublikasi FA akademigi
Tashev K.A.	— TATU II va IB prorektori
Muxtarov F.M.	— TATU Farg‘ona filiali direktori
Rasulov A.M.	— TATU Farg‘ona filiali professori
Rahmatov R.	— TATU professori
Salomov O.R.	— FarPI rektori, professor
Shermammedov B.Sh.	— FarDU rektori, professor
Ergashev S.F.	— FarPI professori
Rasulov R.Ya.	— FarDU professori
Otajonov S.	— FarDU professori
Sabirov S.S.	— TATU Farg‘ona filiali professori
Daliyev B.S.	— TATU Farg‘ona filiali, PhD
Abdullaev J.S.	— TATU Farg‘ona filiali dotsenti
To‘xtasinov D.F.	— TATU Farg‘ona filiali PhD
Bozarov B.I.	— TATU Farg‘ona filiali dotsenti;
Otajonov M.	— TATU Farg‘ona filiali PhD
Oxunjonov U. Yu.	— TATU Farg‘ona filiali PhD
Raimjanova O.S.	— TATU Farg‘ona filiali dotsenti
Ibroximov N.I.	— TATU Farg‘ona filiali dotsenti
Obidova G.K.	— TATU Farg‘ona filiali dotsenti
Asraev M.	— TATU Farg‘ona filiali PhD
Ismoilov M	— TATU Farg‘ona filiali PhD
Fayzullaev I.	— FarPI dotsenti

Kompyuterda chop etishga tayyorgarlik:

Nasriddinov O.U., Maniyozov O.A., Satvoldiev I.A., Bakirov E.V., Tolipov N.I.,
Qodirov X.A.

UZLUKSIZ NUTQNI ANIQLASH TIZIMLARIDA O'ZBEK TILINI MODELLASHTIRISH.

Mamatov Narzillo Solidjonovich

“TIQXMMI” Milliy tadqiqot universiteti, t.f.d., professor

Nuritdinov Nurbek Davlataliyevich

Namangan muhandislik-qurilish instituti, tayanch-doktorant

Annotasiya: Ushbu maqolada O'zbek tilining o'ziga xos xususiyatlarini hisobga oladigan til modellari tavsifi keltirilgan. Nutqni avtomatik aniqlash (ASR) tizimlarining aniqligi va ishlashini oshirishda til modellashtirishning roli katta ahamiyatga ega. Doimiy nutqdagi sintaktik va semantik tuzilmalarni to'g'ri tanlash orqali nutqni matnga aylantirishni yaxshilashdir. Ushbu maqsadga erishish uchun statistik va neyron tarmoqqa asoslangan til modellari, shu jumladan n-gramlar va takroriy neyron tarmoqlari (RNN) va transformator modellari kabi chuqur o'rganish arxitekturalaridan foydalanilgan.

Kalit so'zlar: Nutqni avtomatik aniqlash (ASR), Takroriy neyron tarmoqlari (RNN), n-gramm, Trigram tili, Faktoriy til.

Kirish.

Zamonaviy texnologiyalar rivoji inson-kompyuter o'zaro muloqotini sezilarli darajada osonlashtirdi, bunda nutqni avtomatik aniqlash tizimlari (ASR) muhim o'rin tutadi. Ushbu tizimlar foydalanuvchining nutqini avtomatik ravishda matnga aylantirish orqali kompyuterlar bilan tabiiy til orqali muloqot qilish imkonini beradi. Uzluksiz nutqni aniqlash tizimlarining asosiy vazifasi – inson nutqini aniq va samarali qayta ishlash hamda uni tushunarli matn shaklida ifodalashdir. Bu jarayonda tilni modellashtirishning ahamiyati katta bo'lib, tizimga nutqning sintaktik va semantik tuzilishini anglash, kontekstni to'g'ri talqin qilish imkonini beradi.

Adabiyot sharhi.

Nutqni avtomatik aniqlash (ASR), til modellarini qurish nutqni tanib olish, nutqni matnga aylantirish, nutqni matnga aylantirish va mashina tarjimasini usullari bo'yicha adabiyotlar va maqolalarni tahlil qildik.

Nutqni avtomatik aniqlash va o'zbek tilini modellashtirish.

Nutqni avtomatik aniqlash tizimlarining asosiy elementlaridan biri til modelidir. Ko'pgina mavjud tizimlar n-gramm so'zlarga asoslangan statistik modeldan foydalanadi. Ushbu model matnda n ta so'z ketma-ketligi zanjirining hosil

bo'lish ehtimolini baholashga imkon beradi va so'z tartibi qat'iy bo'lgan tillarda samarali. Biroq, bu so'z tartibi biroz bo'sh bo'lgan tillar, jumladan, o'zbek tili uchun kutilgan natijani bermaydi. Chunki o'zbek tilida statistik modellar samaradorligini pasaytiradigan bir qator o'ziga xos xususiyatlar mavjud. O'zbek tilida so'z tartibi nisbatan erkindir. O'zbek tili boy morfologiyaga ega bo'lgan agglyutinatativ til bo'lgani uchun nutqni aniqlashning bu tizimi lug'at hajmini sezilarli darajada oshiradi va n-gramm til modellarining noaniqlik koeffitsientini oshiradi. Til modeli-bu suhbatlarni amalga oshirish, tezkor qidiruv javoblarini taqdim etish, tarjima sifatini yaxshilash va ijtimoiy media xabarlarining ma'nosini baholash kabi turli xil ovozli yordamchi xizmatlarini taqdim etishda keng qo'llaniladigan so'zlar ketma-ketligining ehtimollik taqsimoti. Til modeli qanchalik yaxshi bo'lsa, natija shunchalik yaxshi bo'ladi. Buni ko'p yillar oldin mavjud bo'lgan nutqni aniqlash va tarjima tizimlarini zamonaviylari bilan taqqoslash orqali bilish mumkin.

Hozirgi vaqtda keng ko'lamli kontekstni yoki so'zlar orasidagi uzoq muddatli munosabatlarni modellashtirish uchun statistik til modellarining bir nechta turlari mavjud. Bunday turlardan biri trigger naqshlari bo'lib, ularda asosiy so'zning paydo bo'lishi maqsad deb ataladigan boshqa so'z bilan bog'liq bo'lgan so'zning paydo bo'lish ehtimolini oshiradi.

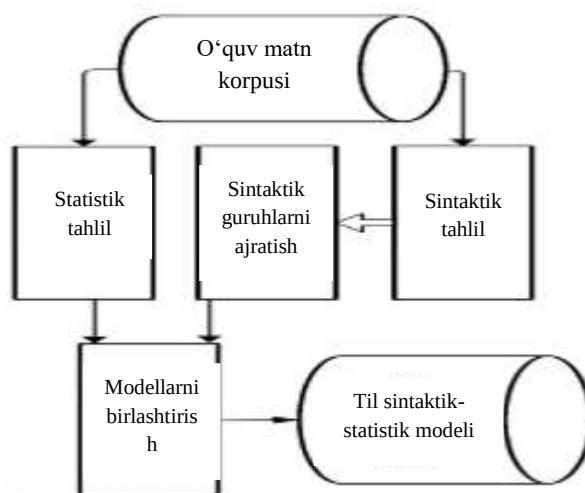
Trigger juftlarining soddalashtirilgan shakli kesh modeli bo'lib, u so'z tarixida qanchalik tez-tez sodir bo'lganiga qarab so'zning paydo bo'lish ehtimolini oshiradi, chunki ma'lum bir so'zni ishlatgandan so'ng, ma'ruzachi uni qayta ishlatishi taxmin qilinadi, bu so'z ma'lum bir mavzu bilan bog'langan yoki ma'ruzachi o'z lug'atida ushbu so'zni qo'llash tendentsiyasini bildiradi.

Uzoq masofali trigramma modeli taklif etiladi, bu trigramma modeli bo'lib, so'z ehtimolini nafaqat oldingi so'zlar uchun, balki bashorat qilingan so'zdan uzoqroqda joylashgan so'zlar uchun ham bashorat qiladi. U "grammatikaga" asoslangan bir nechta ajratuvchi so'zlar orqali bir-biriga bog'lanishi mumkin bo'lgan so'z juftlari majmuasini hosil qiladi.

Gapdagi uzoq masofali munosabatlarni modellashtirish imkonini beruvchi til modelining yana bir turi sintaktik-statistik modeldir. Bunday modelni yaratish uchun avvalo matn korpusining statistik tahlili o'tkaziladi va n-grammli so'zlar ro'yxati tuziladi. So'ngra sintaktik tahlil o'tkaziladi, bunda matnda ajratilgan, ya'ni grammatik jihatdan bog'langan so'z juftlari (sintaktik guruhlar) aniqlanadi. Bunday sintaktik guruhlar matn korpusining statistik tahlili natijasida olingan n grammli so'zlar qatoriga qo'shiladi. Tilning sintaktik-statistik modelini yaratish sxemasi 1. 1-rasmda keltirilgan.

Zarrachalarga asoslangan modellar morfologiyasi boy tillar uchun, masalan, flektiv tillar uchun ishlatiladi [86]. Bu yerda w so'zi, funksiya yordamida ma'lum

($L(w)$) bo'laklarga (morfemalarga) $U: w \rightarrow u^1, u^2, \dots, u^{L(w)}, u^i \in \Psi$ bo'linadi, bu erda Ψ so'z qismlari to'plami. So'zlarni morfemalarga bo'lishning ikki xil usuli mavjud: lug'at va algoritmik.



1.1-rasm. Tilning sintaktik-statistik modelini yaratish jarayoni

Algoritmik usullarning afzalligi shundaki, ular faqat matn tahliliga asoslanadi va hech qanday qo'shimcha bilimlardan foydalanmaydi, bu esa istalgan tildagi matnni tahlil qilish imkonini beradi, lekin so'zlar psevdomorfemik birliklarga bo'linadi. Lug'at usullarining afzalligi shundaki, ular so'zlarning morfemalarga to'g'ri bo'linishlarini olish imkonini beradi, bu esa keyingi iboralarni aniqlash gipotezalarida ishlov berish darajasida qo'llanilishi mumkin.

Boy morfologiyaga ega tillarga qo'llanilishi mumkin bo'lgan yana bir model bu arab tilini modellashtirish uchun til omil modeli (FML).

Ushbu model so'zning turli xususiyatlarini (omillarni) birlashtiradi va so'z $Y_i = (F_i^1, F_i^2, \dots, F_i^k)$ omil vektori sifatida ifodalanadi. Omil so'zning shakli, gap bo'lagi, ildiz, o'zak va boshqa morfologik va grammatik belgilar bo'lishi mumkin.

Til modeli sun'iy neyron tarmoqlari asosida qurilishi mumkin, xususan, bu ishda birinchi marta takroriy neyron tarmog'i (RNN) ishlatilgan; Ushbu modelning afzalligi shundaki, yashirin qatlam ko'rib chiqilayotgan so'zdan oldingi barcha kontekstni saqlaydi. Tarmoq kirish qatlamiga, yashirin qatlamga (kontekst qatlami yoki holat deb ham ataladi) va chiqish qatlamiga ega. Neyron tarmog'ini o'rgatgandan so'ng chiqish qatlami oldingi vaqt bosqichida yashirin qatlam holatini hisobga olgan holda oldingi so'z va keyingi so'zning ehtimollik taqsimotini ta'minlaydi. Yashirin qatlamning o'lchami odatda eksperimental ravishda tanlanadi.

Natijalar va muhokama.

Ushbu maqolada o'zbek tilining omil va neyrotarmoq modelini yaratish jarayoni va bunday modellardan eng yaxshi tanib olish gipotezalarini (N-eng yaxshi ro'yxat) avtomatik uzluksiz nutqni aniqlash tizimlarida qayta baholash bosqichida foydalanish yoritilgan.

O'zbek tilining trigramma modeli. O'rganish davomida o'zbek tilining modelini yaratish maqsadida ko'plab internet saytlari, elektron jurnallar, gazeta va kitoblardan matnlar korpusi to'planib, ularga dastlabki ishlov berish ishlari olib borildi.

Natijada, umumiy hajmi 250 million so'z bo'lgan korpus yaratildi, u tilning original (trigram) modelini, shuningdek, faktorial va neyrotarmoq modellarini o'rgatish uchun foydalanildi.

Trigram tili modelini yaratish uchun veb-saytlardan matnlarni avtomatik ravishda ajratib olish, qayta ishlash va tasniflash kabi funktsiyalarga ega bo'lgan maxsus dasturiy vosita yaratildi. Turli til modellari yordamida o'zbek tilida uzluksiz nutqni avtomatik tanib olish bo'yicha o'tkazilgan dastlabki tajribalarda 120 000 so'z shakllari lug'atiga ega til modelidan foydalanganda noto'g'ri tan olingan so'zlar soni eng kam bo'lishi mumkinligi aniqlandi.

O'zbek tilining yaratilgan modelini baholash o'quv korpusida mavjud bo'lmagan matnli elektron axborot resurslaridan iborat matnli korpus bo'yicha o'tkazildi va trigramma tili modelining noaniqlik koeffitsienti 524 ga teng bo'ldi.

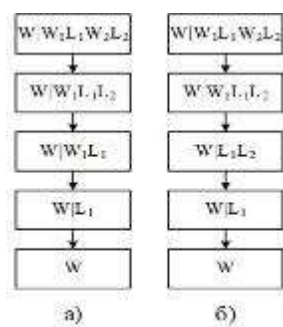
O'zbek tilining omilli modeli. O'quv matnlari korpusining morfologik tahlili maxsus ishlab chiqilgan dastur yordamida amalga oshirildi.

Bunda so'z shakli, lemma, o'zak, gap bo'lagi, morfologik belgilar tegi qo'llanilgan va undagi barcha so'zlar o'z omillari bilan almashtirilgan. Masalan, "uy" so'zi "W-uylar, L-uylar, S-uylar, P-ot, M" ga o'zgaradi, bunda W so'z shakli, L - lemma, S - o'zak, P - o'zak. gap qismi, M esa so'z haqidagi barcha grammatik ma'lumotlar, ma'lumotlarni umumlashtiruvchi morfologik belgilarning tamg'asi. Ikki faktorli modellar so'z shakli va matn korpusi uchun yuqorida sanab o'tilgan boshqa omillardan biri asosida ishlab chiqilgan.

O'quv matni korpusida n-gramma bo'lmasa yoki uning paydo bo'lish chastotasi juda past bo'lsa, u almashtiriladi. tomonidan (n-1)-gramm ehtimollik, qaytish tezligi bo'yicha y ga ko'paytiriladi . N-grammli modellarda qaytarish eng uzoq so'zni tashlash orqali amalga oshiriladi. Buning uchun birinchi navbatda eng uzoq so'z, keyin oldingi so'z va hokazo tashlanadi.

Faktoriy til modellarida teskari yo'lning ikkita varianti ikkita omil yordamida, ya'ni, birinchi navbatda, uzoq so'z shakli va omili, so'ngra yaqin so'z shakli va omilini yo'q qilish (2.3a-rasm) yoki birinchi navbatda so'zni saralash mumkin. shakllar masofa tartibida, keyin esa omillarda bir xil tartibda aniqlanadi (2.3b-rasm).

Yakuniy ishda ana shu yondashuv asosida o'zbek tilining omilli modeli qurildi.



2.3-rasm. Soʻz va lemma omillari bilan belgilanadigan FNMning qaytish yoʻllari:

a) birinchi qaytish yo‘li; b) ikkinchi qaytish yo‘li

Yaratilgan FNM uchun noaniqlik koeffitsientlari 2.1-jadvalda keltirilgan. Orqa yo‘l 1 bo‘lgan modellar, ishlatiladigan omillardan qat’i nazar, noaniqlik koeffitsientlari pastroq edi. Shu bilan birga, "so‘z shakli" va "lemma" omillari bo‘lgan model noaniqlik koeffitsientining eng past qiymatiga ega edi va uning teskari yo‘li 1 ga teng edi.

O'zbek tilining neyron tarmoq modeli. RNN bo'yicha trening RNNLM Toolkit (Recurrent Neural Network Language Modeling Toolkit) bepul dasturiy moduli asosida amalga oshirildi, unda neyron tarmoqni o'rganish tezligini pasaytirish maqsadida chiqish qatlami faktorlarga ajratildi.

Sinflar soʻz chastotasi bilan belgilanadi. Birinchidan, sinflar boʻyicha ehtimollik taqsimoti hisoblab chiqilgan, soʻngra tegishli sinfga tegishli soʻzlar uchun ehtimollik taqsimoti hisoblangan. 200, 500, 800 yashirin qatlam elementlari va 200 va 800 sinflarga ega modellar yaratilgan.

2.1-jadval

FNM noaniqlik koeffitsientlari

Faktorlar	Noaniqlik omili	
	1 yo‘l	2 yo‘l
V.M.	457	561
W.L.	580	591
W.P.	682	736
W.S.	562	687

Ishlab chiqilgan modellarning noaniqlik koeffitsientlarining qiymatlari 2.2-jadvalda keltirilgan.

2.2-jadval

O'zbek tilining neyron tarmoq modellarining noaniqlik koeffitsientlari

Sinflar soni	Yashirin qatlamdagi elementlar soni
--------------	-------------------------------------

	200	500	800
200	960	991	8660
800	2080	860	880

O'zbek tilida uzluksiz nutqni aniqlashning statistik til modeli. Akustik model sifatida chapdan o'ngga uchta holatga ega yashirin Markov modellari olindi va bu modellar HMMT (Hidden Markov Model Toolkit) yordamida o'zbek tilida 60 ta so'zlovchining uzluksiz nutqidan hosil bo'lgan nutq korpusi asosida yaratilgan. Ishning umumiy hajmi 60 Gb, audioyozuvlar davomiyligi esa 560 soatdan ortiq. Nutqni aniqlash tizimi audio kitoblardan olingan uzluksiz og'zaki iboralar yordamida sinovdan o'tkazildi. Undagi jumlar 3-30 so'zdan iborat bo'lib, har bir ma'ruzachi uchun yozib olish vaqti 30 daqiqagacha davom etdi. Bunday holda, sof nutq uchun vaqt 10-25 minut.

Test korpusining umumiy hajmi 1200 MB audio ma'lumotni tashkil etadi va o'zbek tilidagi uzluksiz nutqni avtomatik tanib olish tizimi transformator neyron tarmog'i modeli asosida ishlab chiqilgan. Tanib olish tizimi noto'g'ri tan olingan so'zlar ulushi (WER - Word Error Rate) bo'yicha nutqni aniqlashning sifat mezoni asosida baholandi. Nutqni dekodlash bosqichida trigram tili modelidan foydalanildi va natija WER = 27,8% ni tashkil etdi va gipotezalar soni 10, 20, 40 ni tashkil etdi. Keyin gipotezalar yaratilgan omil va til neyron tarmog'i modellari yordamida qayta baholandi.

Faktoriy til modeli va turli interpolyatsiya koeffitsientlari bilan tayanch til modeli bilan interpolyatsiya qilingan faktorial model uchun eng yaxshi tan olish gipotezalari ro'yxatini qayta baholashdan so'ng tan olish natijalari 2.3-jadvalda keltirilgan. Til modeli interpolyatsiyasi - har bir modelning og'irligini hisobga olgan holda turli modellardan olingan so'z ehtimollarining chiziqli birikmasi. Agar tilning faqat faktoriy modeli ishlatilsa, interpolyatsiya koeffitsienti 1 ga teng, aks holda 1 dan kichik.

2.3-jadval

Noto'g'ri tan olingan so'zlar soni (WER, %) eng yaxshi FNM tan olinishi gipotezalari ro'yxatini qayta baholashdan keyin

Til modeli	N=10		N=20		N=40	
	1 yo'l	2 yo'l	1 yo'l	2 yo'l	1 yo'l	2 yo'l
WM+3-gr.	25.81	24.4	23.4	24.7	24.6	24.76
WL+3-gr.	25.8	25.1	25.5	25.3	25.52	25.38
WP+3-gr.	25.4	25.5	25.7	25.4	25.05	25.36
WS+3-gr.	25.2	26.2	25.87	25.63	25.87	26.12

1-orqa yo'l bilan yaratilgan trigram modeli yordamida interpolyatsiya qilingan so'z shakllari va morfologik teglar omillar sifatida ishlatilgan 20 ta faraziy omil modellari ro'yxatini qayta baholashda olingan.

Tilning neyron tarmog'i modellaridan foydalangan holda o'tkazilgan tajribalarda, faktor modellarida o'tkazilgan tajribalarda bo'lgani kabi, yaxshi tanib olish gipotezalarining bir xil ro'yxati ishlatilgan va neyron tarmoq va trigramma modellarining interpolyatsiyasi ham amalga oshirilgan va tajribalar natijalari 2.4-jadvalda keltirilgan. Tilning neyron tarmog'i modelidan foydalanish so'zni aniqlash xatosini kamaytirishga imkon berdi.

Tajribalarda 200-sinfdagi RNN-dan foydalanish 500-sinfdagi RNN-dan foydalanishga qaraganda yaxshiroq tanib olish natijalarini ko'rsatdi, sinflar soni 100 tani tashkil etdi va 800 ta elementdan iborat yashirin qatlamning RNNsi foydalanishda eng yaxshi natijani berdi (WER = 23,7%) interpolyatsiya koeffitsienti 0,6 bo'lgan trigram modeli asosida interpolyatsiya qilingan RNNga asoslangan til modeli.

2.4-jadval

Turli N eng yaxshi farazlar ro'yxatini qayta baholagandan so'ng noto'g'ri tan olingan so'zlar soni (WER, %)

Til modeli	Interpolatsiya koeffitsienti, l	N=10	N=20	N=40
Yashirin qatlamda 200 elementli RNN va trigramma NM mavjud	1.0	27.2	27.6	26.8
	0,6	26.3	25.6	25.9
	0,5	26.1	24.9	24.9
	0.4	25.0	24.7	24.6
Yashirin qatlamda 400 elementli RNN va trigramma NM mavjud	1.0	26.1	26.3	26.4
	0,6	25.8	25.3	24.1
	0,5	25.5	25.4	24.8
	0.4	25.3	24.9	24.2
Yashirin qatlamda 800 elementli RNN va trigramma NM mavjud	1.0	26.5	24.6	24.7
	0,6	23.7	24.7	22.9
	0,5	23.5	24.1	22.7
	0.4	23.8	24.2	23.6

Xulosa. O'zbek tilidagi uzluksiz nutqni aniqlash tizimlari uchun o'zbek tilining uch xil statistik modellari (trigramma, omil va neyron tarmoq) chuqur o'rganilgan. FNM ning n-gramm modellaridan afzalligi shundaki, til modeli qo'shimcha lingvistik ma'lumotlarni qamrab oladi va boy morfologiyaga ega tillar, jumladan, o'zbek tili uchun nutqni aniqlash tizimining sifatini yaxshilaydi. RNN-ga asoslangan modellar boshqa modellardan ustundir, chunki ular ixtiyoriy til kontekstini saqlaydi. O'zbek tilida uzluksiz nutqni tan olish bo'yicha olib borilgan izlanishlar shuni ko'rsatdiki, eng yaxshi tan olish gipotezalarini qayta baholash

bosqichida ishlab chiqilgan omil va neyron tarmoq modellaridan foydalanish noto'g'ri tan olingan so'zlarni nisbatan kamaytirish imkonini beradi. Tilning trigramma modelidan foydalangan holda olingan natijalar bilan taqqoslaganda, noto'g'ri tan olingan so'zlar sonining nisbiy qisqarishi omil modelidan foydalanganda 10% va neyron tarmoq modelidan foydalanganda 18,5% ni tashkil etdi.

Foydalanilgan adabiyotlar.

1. Niyozmatova, N. & Mamatov, Narzillo & Tulaganova, Sh & Samijonov, Abdurashid & Samijonov, B.. (2023). *Tanish tizimlarida o'zbek nutqining nutq faolligini aniqlash usullari*. 050019. 10.1063/5.0145438.
2. Mamatov, N.S., Niyozmatova, N.A., Yo'ldoshev, Y.S., Abdullaev, S.S., Samijonov, A.N. (2023). *Diqqat mexanizmiga asoslangan neytral tarmoqda nutqni avtomatik aniqlash*. In: Zayniddinov, H., Singh, M., Tiwary, US, Singh, D. (tahrirlar) *Intelligent Human Computer Interaction. IHCI 2022. Kompyuter fanlari bo'yicha ma'ruza matnlari*, 13741-jild. Springer, Cham. https://doi.org/10.1007/978-3-031-27199-1_11
3. NS Mamatov, NA Niyozmatova, AN Samijonov va BN Samijonov, *"O'zbek tili uchun til modellari qurilishi"*, 2022 *Xalqaro axborot fanlari va kommunikatsiya texnologiyalari konferentsiyasi (ICISCT)*, Toshkent, O'zbekiston, 2022, s. 1-4, doi: 10.1109/ICISCT55600.2022.10146788.
4. Mamatov, Narzillo & Niyozmatova, N. & Abdullaev, Sh & Samijonov, Abdurashid & Erejepov, K.. (2021). *Transformator neyron tarmoqlariga asoslangan nutqni aniqlash*. 1-5. 10.1109/ICISCT52966.2021.9670093.
5. Mamatov, N., Niyozmatova, N., Samijonov, A. 2021. *Ovozli signallarni oldindan qayta ishlash uchun dasturiy ta'minot*. *Xalqaro amaliy fanlar va muhandislik jurnali*, 18, 2020163. [https://doi.org/10.6703/IJASE.202103_18\(1\).006](https://doi.org/10.6703/IJASE.202103_18(1).006)
6. Wiedecke, Bernd & Mamatov, Narzillo & Payazov, Mirabbos & Samijonov, Abdurashid. (2019). *Akustik signalni tahlil qilish va aniqlash*. *Innovatsion texnologiyalar va tadqiqot muhandisligi xalqaro jurnali*. 8. 2440-2442. 10.35940/ijitee. J9522.0881019.
7. Narzillo, M., Abdurashid, S., Parakhat, N., & Nilufar, N. (2019). *Vektor kvantlash usuliga asoslangan ovoz bilan karnayni avtomatik aniqlash*. *Innovatsion texnologiyalar va tadqiqot muhandisligi*

xalqaro jurnali, 8(10), 2443–2445.

<https://doi.org/10.35940/ijitee.J9523.0881019>

8. Mamatov N. S., Nuritdinov N. D. NEYRON TO 'RLARI NEGIZI VA ULARNING TURLARINI TAHLILI //Educational Research in Universal Sciences. – 2023. – T. 2. – №. 8. – C. 76-83.
9. Mamatov N. S., Nuritdinov N. D. ODAMNING RANGLARNI FIZIOLOGIK KO 'RISH XUSUSIYATLARI //Educational Research in Universal Sciences. – 2023. – T. 2. – №. 7. – C. 104-109.