

NUTQNI AVTOMATIK TANIB OLIISHNING NEYRON TARMOQLARI

Niyozmatova Nilufar A'loxonovna,

Mamatov Narzullo Solidjonovich

“Toshkent irrigatsiya va qishloq xo'jaligini mexanizatsiyalash
muhandislari” Milliy tadqiqot universiteti

Samijonov Abdurashid Narzullo o'g'li

Toshkent axborot texnologiyalari universiteti

Samijonov Boymirzo Narzullo o'g'li

Sejong universiteti, Janubiy Koreya, Seul

m_narzullo@mail.ru

Annotatsiya: Tadqiqot nutqni avtomatik tanib olish (NATO)ning nutqni avtomatik tanib olishning neyron tarmoqlariga bag'ishlangan. O'tkazilgan tajriba ma'lum test to'plami tanlanmalaridan foydalanib, nutqni tanib olishda bir nechta ilovalar taqqoslangan. Ma'lumotlar to'plami har bir tizim tomonidan Python dasturlash tili ilovalari yordamida tahlil qilindi, chiqish ma'lumotlari normallashtiriladi va WER standartiga muvofiq oldindan transkripsiya qilingan etalon ma'lumotlari bilan taqqoslanadi. Sinov natijalari tahlili o'tkazilgan, nutqni avtomatik tanib olish tizimining samaradorligi uning elementlarini optimallashtirish va kerakli ma'lumotlar to'plamidan foydalangan holda o'qitishga bog'liqligi to'g'risida xulosalar chiqarilgan.

Kalit so'zlar: nutqni tanib olish, akustik modellashtirish, tilni modellashtirish, yashirin Markov modellari, chuqur o'qitish, neyron tarmoqlari.

Nutqni avtomatik tanib olish (NATO) – bu ko'plab zamonaviy ilovalarda qo'llaniladigan texnologiyadir. Nutqni avtomatik tanib olishning maqsadlaridan biri mashinalar va insonlar o'rtasida tabiiy, jonli aloqa o'xshashligini yaratishdir. Zamonaviy nutqni tanib olish va sintez qilish tizimlari ancha rivojlangan bo'lib, ular so'nggi texnologik ishlanmalardan foydalangan holda, ishlash aniqligi bo'yicha yuqori natijalarni ko'rsatmoqda.

NATO tizimlarining rivojlanishidagi sezilarli yutuqlarga qaramay, avtomatik nutqni tanib olish mukammal ishlamaydi. Har xil turdagi neyron tarmoqlari NATO tizimlarining bir qismiga aylanib, ularni ishini aniqligini oshirmoqda. Biroq, bu tendensiya mukammal emas, kompaniyalar nutqni tanib olish xizmati ichidagi neyron tarmoqlar ishini cheklamoqdalar.

Bugungi kunda bozorda turli xil algoritmlarga asoslangan statistik modellardan tortib neyron tarmoqlargacha bo'lgan NATO tizimlari mavjud. Ushbu tizimlarning ba'zilari aniqlik bo'yicha yuqori natijalarni ko'rsatmoqda hamda ularni har biri akustik va tilni modellashtirish yordamida tilga ishlov berishni turlicha yondashuvidan foydalanadi.

Nutqni avtomatik tanib olishda neyron tarmoqlar

Nutqni avtomatik tanib olish sohasida bir necha turdagi neyron tarmoqlar qo'llaniladi. Oldinga tarqalish neyron tarmoqlari – bu ma'lum bir tuzilishga, kirish va chiqish tugunlari soniga ega bo'lgan model bo'lib, signallarni to'playdi va ularni chiqish ma'lumotlariga aylantiradigan sensorlar to'plamidir. Ushbu turdagi o'qitilgan

neyron tarmog'i kirishida doim bir xil chiqish qiymatini chiqaradi. Boshqacha qilib aytganda, bu vaqtga bog'li bo'lmagan mustaqil modeldir. Ushbu turga o'ramli neyron tarmoqlari (O'NT) kiradi, ular umumiy signal timsollarini tanib olish masalalarida foydali vosita hisoblanadi. Nutqni avtomatik tanib olishda ular dastlabki ishlov berish bosqichida, shuningdek, yashirin Markov modellari bilan tandemda akustik va tilni modellashtirishda qo'llanilishi mumkin. Umuman olganda, o'ramli neyron tarmoqlari NATOda zaif samaraga ega va nutqni tanib olishda faqat ikkinchi darajali vazifalar uchun qo'llaniladi.

O'z tabiatiga ko'ra nutqni avtomatik tanib olish masalasi o'ramli neyron tarmog'i hal qilishi kerak bo'lgan vazifalardan sezilarli darajada farq qiladi. Birinchidan, NATOda kirish ma'lumotlari joriy uzun ketma-ketligiga ko'ra, kirish tugunlari to'plamini joylashtirish mumkin emas. Chunki audio fayl millionlab ma'lumotlar birliklarini qamrab olishi mumkin. Shu bilan birga, kirishni ma'lum bir qismi konteksti undan oldin va keyin ishlov berilgan kirish bilan bog'liq bo'lishi ham mumkin. Tabiiyki, bu masalalarni yechish uchun tayyor vositalardan foydalaniladi, mavjud neyron tarmoq texnologiyalari to'ldiriladi. Bunday neyron tarmoqlari rekurent (RNT) tarmoqlari deb ataladi. Ular muayyan kiruvchi to'plami uchun chiqish ma'lumotlari ham oldingi kirishga asoslanadi. Albatta, bunday tuzilma kamchiliklari ham bor va ko'plab yangi muammolarni keltirib chiqaradi. Masalan, chiqishni generatsiyalashda tizim qanchalik oldingi qiymatlarga qarashi kerakligi aniq emas. Uzoq qisqa muddatli xotirali neyron tarmoqlari ushbu masalani so'nggi kirishni oldingi kirishlar bilan aralashtiradigan maxsus tugunlarni kiritish orqali yechishga harakat qiladi.

Vaqtga bog'liq bo'lgan barcha neyron tarmoqlar, ya'ni chiqish ma'lumotlarini generatsiyalashda oldingi kirishlarga ishlov beruvchilar nutq va matn bilan ishlash uchun eng yaxshi neyron tarmoqlari hisoblanadi. Biroq, rekurent neyron tarmoqlar bunday modellarni eng katta sinfi sifatida ajralib turadi. Neyron tarmoqlari chiziqli va deyarli chiziqli bo'lmagan funksiyalarni modellashtirishda samarali hisoblanadi. Biroq, ular faqat qisqa vaqt ichida signallarga ishlov berishda to'g'ri ishlaydi. Agar uzluksiz tabiiy nutq haqida gap ketsa, neyron tarmoqlari ushbu masalani bilan deyarli yecha olmaydi. Shunga qaramay, neyron tarmoqlari statistik modeli usullar (SMU)ga asoslangan nutqni tanib olish tizimlari uchun dastlabki ishlov berish vositasi sifatida keng qo'llaniladi.

Neyron tarmoqlar yaxlitligi bo'lganligi uchun boshqa usullarga nisbatan ayrim kamchiliklarga ega emas. SMU ga asoslangan modellar chiqish ma'lumotlariga ishlov berishni talab qiladi, chunki NATO natijalari tizim ichida maxsus vakolatga ega. Chuqur o'qitish (ChO') esa o'z navbatida, kirish ma'lumotlari ketma-ketligini to'g'ridan-to'g'ri chiqish ma'lumotlari, matn bilan birlashtirishga imkon beradi. Biroq, neyron tarmoqlari leksik ma'lumotlar bilan ishlashda statistik modellardan ortda qoladi. SMU larda qo'shimcha til modellari mavjud bo'lib, ular tizim uchun leksik ma'lumotlar to'plamini yaratishga mo'ljallangan. Chuqur o'qitish modellarida bu leksika faqat qo'lda transkripsiya qilingan matnlarning o'quv ma'lumotlari bilan beriladi, shuning uchun modelning qamrovi cheklangan bo'ladi, bu esa lingvistik jihatdan xilma-xil tanlanma bilan ishlashda muammolarga olib keladi.

NATOdagi neyron tarmoqlarning SMUdan farqli o'laroq kamchiliklari sifatida

aksariyat xizmatlar uchun deyarli tarmoqni boshidan ishlab chiqishni talab qilishi bilan bog'liq, kompaniyalar faqat neyron tarmoqlarini joriy etish uchun yangi nutqni tanib olish algoritmlarini yaratishga vaqt sarflamaydi, balki mavjud tarmoq mexanizmlarini yaxshilash ob'ektiv ravishda yaxshi natijani ko'rsatishi mumkin, bundan tashqari bu butun tizimni yangi algoritmlarga optimallashtirishni talab qilmaydi, bitta komponentni yangilash strukturaga ta'sir qilmaydi.

NATO sohasida nutqni tanib olish tizimlarining ishlashini solishtirishda aniqlikni o'lchash uchun bir nechta turli standartlar mavjud, ulardan biri WER (word error rate, xato so'zlarning chastotasi) – nutqni tanib olish va avtomatik tarjima tizimi ishlash aniqligining umumiy standart o'lchovi. Ushbu standart og'zaki asosda ishlaydigan Levenshtein tahrir masofasiga asoslangan. WER konsepsiyasi oddiy va intuitivdir: xato tanib olingan so'zlar to'g'ri so'zlarga bo'linadi.

Xato so'zlarni o'lchashda bir xil standart doirasida xato natija foizi nutqga ishlov berish jarayonida o'chirilgan, almashtirilgan yoki qo'shilgan so'zlar soni bilan hisoblanadi. Algoritm ushbu so'zlarning yig'indisini matnda mavjud bo'lishi kerak bo'lgan so'zlar soniga bo'ladi. Ishlov berilgan matn to'g'riligini aniqlashda inson tomonidan aytilgan audio yozuv etalon rasshifrovkasi talab qilinadi.

$$WER = \frac{o'chirish + almashtirish + qo'shish}{kul\ rasshifrovkasidagi\ so'zlar\ soni}$$

WERni bir nechta kamchiliklari mavjud, masalan, u tinish belgilari bilan, ma'ruzachi nutqi xususiyatlari bilan ishlamaydi, xizmatchi so'zlari bilan bog'liq xatolarga u mustaqil ma'noli so'zlar bilan bir qatorda ishlov beradi, matndagi har qanday xatolar qat'iy aniqlaydi. Shunga qaramay, WER turli xil nutqga ishlov berish tizimlarini taqqoslashda qimmatli va foydali vosita hisoblanadi.

Tajribalar

Tajribalarda 1 dan 10 daqiqagacha davom etgan 20 ta audio fayl tanlab olindi. Tanlanma turli xil sharoitlarda yozib olingan o'zbek tilini turli dialektlari va aksentlarini jonli tabiiy nutqi namunalari mobil qurilma orqali yozib olingan audiolardan iborat.

Ma'lumotlar to'plamini har bir elementi qo'lda transkripsiyalangan audio bilan matn fayli yaratilgan. Sinov uchun o'xshash ma'lumotlarga ega bo'lish juda muhim, shuning uchun matnni transkripsiyalashda tizimni o'ziga xos xususiyatlarini inobatga olinib normallashtirildi, ularda tinish belgilari, katta harflar va raqamlar yo'q.

Nutqni avtomatik tanib olish uchun API ilovalari bilan ishlashda namunadagi audio namunalarga Python dasturi yordamida ishlov berildi.

Tanlanmada keltirilgan barcha ilovalar tanib olish uchun til dialektini tanlash imkoniyatiga ega emas, shuning uchun tajriba sofligi uchun o'zbek tili uchun standart variantdan foydalangan.

Tanib olish natijalari mos yozuvlar matnlari kabi transkripsiyalanadi. Yuqorida aytib o'tilganidek, NATOni barcha xizmatlari turlicha ishlaydi, ya'ni ba'zilar raqamlarni so'zlar kabi, boshqalari esa raqamlar deb tanib oladi. Barcha matnlar qo'lda tekshiriladi va tahrirlanadi, ma'lumotlar o'xshash shaklga keltiriladi, transkripsiyalangan so'zlar esa o'zgartirilmaydi, NATO xizmatlari tomonidan yo'l qo'yilgan xatolarni saqlab qolish juda muhimdir.

Pythonda import qilingan WER modulidan foydalanib, transkripsiyalangan matnlar berilgan ma'lumotlar bilan taqqoslanadi. Olingan natijalar birlashtirilgan jadvalga kiritiladi, unda qiymatlarning o'rtacha arifmetik qiymati ham hisoblab chiqiladi, tizimlarning umumiy aniqligi baholanadi. Qulaylik uchun test natijalari foiz sifatida taqdim etiladi.

Natijalar va muhokamalar

NATO tizimlarini sinovdan o'tkazish natijalariga ko'ra, nutqni avtomatik tanib olish usullaridan qaysi biri samaraliroq ekanligini aniqlash mumkin. Umumiy aniqlik SMU asosidagi til modellariga ega tizimlar NATOning eng samarali vositasi hisoblanadi. Neyron tarmoq akustik modeli tilni turli dialektlari va aksentlarini tanib olishda yaxshi natija ko'rsatadi. Yashirin Markov modellari akustik modellashtirishda chuqur o'qitish usuliga nisbatan past natija berdi. Tizimlarni past natijasiga o'zbek tilini turli shevalaridagi variantlarida erishildi, biroq, adabiy tilini tanib olishda yaxshi ishlamoqda. Bu modellarni dialekt xilma-xilligi bilan cheklangan ma'lumotlar tanlanmasida o'qitish zarur.

Xulosalar

Taqqoslash natijalarini tahlili akustik modellashtirish neyron tarmoqlar ishlashiga, lingvistik modellashtirish esa SMUga asoslanganda eng samarali ekanligini ko'rsatdi. Umuman olganda, neyron tarmoq til modellari statistik modellarga nisbatan kam samarali bo'lib chiqdi.

Avtomatik nutqni tanib olish tizimi aniqligi elementlarini alohida va birgalikda ishini optimallashtirishga juda bog'liq degan xulosaga kelish mumkin. Avtomatik nutqni taib olish tuzilmasida optimallashtirishni mavjud emasligi, tizim akustik va til modellashtirishni muvaffaqiyatli amalga oshirishiga qaramay, dastur natijalari VER standarti bilan qat'iy taqqoslaganda yuz foiz tanib olish noaniqligini ko'rsatishi mumkinligiga olib keldi. Chuqur o'qitish, yashirin Markov modellari va gibril usullar agar dastur to'g'ri o'qitilgan bo'lsa, samarali bo'ladi va bunda o'quv tanlanmasi muhim rol o'ynaydi. Nutqni tanib olish tizimlari qiyosiy tahlili rivojlanishni ushbu bosqichida neyron tarmoqlar, xususan chuqur o'qitish texnologiyasi NATODA lingvistik va akustik modellashtirishga yashirin Markov modellariga nisbatan samarasiz ekanligini ko'rsatdi.

Foydalanilgan adabiyotlar ro'yxati:

1. Abdel-Hamid O. et al. Convolutional neural networks for speech recognition // IEEE/ACM Transactions on audio, speech, and language processing. - 2014. -vol. 22. - M. 10. - Pp. 1533-1545.
2. Audhkhasi K. et al. Building competitive direct acoustics-to-word models for English conversational speech recognition // 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). - IEEE, 2018. - C. 4759-4763.
3. Chavan R. S. , Sable G. S. An overview of speech recognition using HMM // International Journal of Computer Science and Mobile Computing. - 2013. - vol. 2. - NY. 6. - Pp. 233-238.
4. Deng, L, Liu Y. (Ed.). Deep Learning in Natural Language Processing. — Springer, 2018.
5. Hinton G. et al. Deep neural networks for acoustic modeling in speech

recognition // IEEE Signal processing magazine. 2012. - T. 29.

6. Mohri, M. , Pereira F , Riley M. Weighted finite-state transducers in speech recognition // Computer Speech & Language. _2002. - vol. 16. -M.1.-Pp. 69-88.

7. l l. Wang, Y. Y. , Acero A., Chelba C. Is word error rate a good indicator for spoken language understanding accuracy // 2003 IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat. No. 03EX721). - IEEE, 2003. - Pp. 577-582.

LONG SHORT-TERM MEMORY (LSTM) ALGORITMI

t.f.n., dots. Kabildjanov Aleksandr Sabitovich¹,

Pulatov G'iyos Guforjonovich²,

“TIQXMMI” Milliy tadqiqot universiteti^{1,2}

Pulatova Gulxayo Azamjon qizi³

Toshkent axborot texnologiyalari universiteti Farg'ona filiali³

giyospulatov1987@gmail.com

Annotatsiya: Long Short-Term Memory (LSTM) modeli bir tur neyron tarmoq modellar bo'lib, matematikaviy model hisoblanadi. U o'quvni o'zlashtiradi va murakkab, o'quvni modellar yaratish va aniqlash uchun ishlatiladi. LSTM modeli o'quvning uzluksiz, qo'shimcha o'rganish modelini o'zlashtirishda samarali bo'lib, xotirada turli holatlarni saqlash imkonini beradi.

Kalit so'zlar: Long Short-Term Memory (LSTM), TensorFlow, Open Source, PyTorch.

Long Short-Term Memory (LSTM) algoritmi, matematik asoslangan, murakkab qoidalarni o'rganadigan va turli turlarini san'atlarini o'z ichiga olgan o'quvni modellar va sizni tasdiqlovchi amallarni bajaradigan bu xalqaro yangi tarmoqda o'rnatilgan qavatdir. LSTM qo'shimcha qo'shimcha asoslangan o'quvni modelni yaratish uchun ishlatiladi va texnikaviy tarmoqlar, ma'lumotlar tahlili va boshqa ilovalar yaratishda o'rtacha, to'plam rivojlanish jarayonini sodda va samarali qilish uchun foydalaniladi.

LSTM algoritmi o'quv tizimini o'zlashtiradi va avval o'rganilgan ma'lumotlar o'z ichiga oladi, ulardan foydalanish va natijalarni saqlash imkoniyatini beradi. Algoritmida murakkab modellar qurilgan va bu modellar ko'rib chiqish va to'g'ri javoblarni bajarishda foydalaniladi.

LSTM modelida har bir qadamning o'z bir xotirasi mavjud. U xotiraga o'zgaruvchilarni, turli holatlarni va o'qib chiqarishni saqlaydi. Buning natijasida, bu algoritm murakkab va uzluksiz tizimlarni aniqlash, ma'lumotlar prediktsiyasini qilish va ish olib boryapti. LSTM modeli, matematikaviy mashinalar o'quvining eng muhim qismini tashkil etadi.

LSTM algoritmi amaliyotda, ma'lumotlar tahlili, ma'lumotlar o'quvchi sistemalar va avtomatlashtirilgan tijoratda foydalaniladi. Uni barcha turlarda, qiziqarli dasturlash va so'nggi qurilmalarni yaratishda ishlatish mumkin. LSTM algoritmi model