

**O‘zbek tilida matnni nutqqa samarali o‘tkazishda wav2vec 2.0 modelidan foydalanish afzalliklari**

***Tojidinova Dilafruz Komiljonovna***

*Mirzo Ulug‘bek nomidagi O‘zbekiston Milliy universiteti, Toshkent,  
O‘zbekiston, e-mail: jumaevadilafruz42@gmail.com*

***Annotatsiya.*** Facebook AI tomonidan ishlab chiqilgan wav2vec 2.0 modeli nutqni aniqlash va matndan nutqqa (TTS) texnologiyalari sohasida sezilarli yutuqdir. Ushbu maqolada wav2vec 2.0 modelidan o‘zbek tilida matnni nutqqa samarali o‘tkazish uchun foydalanishning afzalliklari ko‘rib chiqilgan. Wav2vec 2.0 o‘z-o‘zini nazorat qilish va aniq ma‘lumotlar to‘plamlarini aniq sozlash orqali yuqori sifatli, tabiiy tovushli nutq sintezini ta‘minlaydi, bu ayniqsa o‘zbek tili kabi kam ifodalangan tillar uchun qimmatlidir.

***Kalit so‘zlar:*** Wav2vec 2.0, matndan nutqqa, o‘zbek tili, nutqni aniqlash, o‘z-o‘zini boshqarish, tabiiy tilni qayta ishlash, kam ifodalangan tillar.

Nutqni aniqlash va matndan nutqqa (TTS) texnologiyalari evolyutsiyasi inson va kompyuterning o‘zaro ta‘sirida inqilob qildi. Ushbu texnologiyalar rivojlanib borishi bilan keng doiradagi tillarga, shu jumladan hozirgi tadqiqotlar va dasturlarni ishlab chiqishda kam taqdim etiladigan mustahkam tizimlarni ishlab chiqishga ehtiyoj ortib bormoqda. O‘zbekistonda millionlab odamlar sozlashadigan o‘zbek tili tarixan cheklangan resurslar va nutqni qayta ishlashda texnologik yutuqlarga erishgan ana shunday tillardan biridir. Facebook AI tomonidan wav2vec 2.0 modelining joriy etilishi samarali va tabiiy TTS imkoniyatlarini ta‘minlab, ushbu bo‘shliqni bartaraf etish uchun istiqbolli yechimni taklif etadi.

O‘zbek tilida TTS uchun wav2vec 2.0 modelidan foydalanish uchun biz ko‘p bosqichli jarayonni kuzatdik:

Ma'lumotlar to'plami: biz turli xil O'zbekcha nutq yozuvlari va tegishli transkripsiyalarni to'pladik. Ushbu ma'lumotlar to'plami har tomonlama yoritishni ta'minlash uchun turli lahjalar va nutq uslublarini o'z ichiga olgan.

Preprocessing: to'plangan ma'lumotlar tozalandi va boshqariladigan audio-matn juftliklariga ajratildi. Fon shovqini minimallashtirildi va audio darajalari normallashtirildi.

Namunaviy trening: wav2vec 2.0 ning o'z-o'zini boshqarish imkoniyatlaridan foydalangan holda model yorliqsiz o'zbek audio ma'lumotlari bo'yicha oldindan o'qitildi. Keyinchalik, TTS vazifalari uchun modelni optimallashtirish uchun etiketli ma'lumotlar to'plami bilan nozik sozlash amalga oshirildi.

Baholash: o'qitilgan modelning ishlashi aniqlik va tabiiylikni baholash uchun so'z xato darajasi (ver) va o'rtacha fikr Ballari (MOS) kabi standart ko'rsatkichlar yordamida baholandi.

O'zbek tilida matnni nutqqa (TTS) samarali o'tkazish uchun wav2vec 2.0 modelidan foydalanish bir qancha afzalliklarga ega:

Oxiridan oxirigacha o'rganish:

- Samarali trening: Wav2vec 2.0 to'g'ridan-to'g'ri xom audiodan tasvirlarni o'rganish uchun mo'ljallangan bo'lib, bu ko'pincha bir nechta, alohida bosqichlarga tayanadigan an'anaviy TTS tizimlariga nisbatan o'quv quvurini soddalashtiradi.

- Mustahkam ishlash: uchidan uchiga mashq qilish orqali model nutqning yanada mustahkam va nozik tasvirlarini o'rganishi mumkin, bu esa sintezlangan nutqning tabiiyligi va tushunarligini oshirishi mumkin.

O'z-O'zini Nazorat Ta'lim:

- Ma'lumotlarning kamaytirilgan talablari: Wav2vec 2.0 katta hajmdagi yorliqsiz audio ma'lumotlarni oldindan o'rgatish uchun o'z-o'zidan boshqariladigan o'rganishdan foydalanadi. Bu, ayniqsa, belgilangan ma'lumotlar to'plami kam bo'lishi mumkin bo'lgan o'zbek tili kabi tillar uchun foydalidir.

- Samarali nozik Sozlash: oldindan o'qitilgandan so'ng, model kamroq miqdordagi belgilangan ma'lumotlar bilan yaxshi sozlanishi mumkin, bu esa kam manbali tillar uchun amal qiladi.

Yuqori Sifatli Nutq Sintezi:

- Tabiiy tovushli nutq: modelning nutqdagi nozik nuanslarni ushlab qobiliyati tabiiy tovushli sintezlangan nutqni keltirib chiqarishi mumkin, bu esa TTS tizimini yanada samarali va yoqimli qiladi.

- Ekspressivlik: Wav2vec 2.0 nutqda ma'no va hissiyotni etkazish uchun juda muhim bo'lgan ohang, ohang va ritmdagi o'zgarishlarni yaxshiroq boshqarishi mumkin.

Ko'p tilli va domen moslashuvchanligi:

- O'zbek tiliga moslashish: model o'zbek tilining o'ziga xos fonetik va prosodik xususiyatlariga moslashtirilib, TTSning aniq va tabiiy chiqishini ta'minlaydi.

- Tillararo uzatish: ko'p tilli ma'lumotlar to'plamlari bo'yicha oldindan o'qitish modelga tillar bo'yicha umumiy xususiyatlardan foydalanish imkonini beradi va hatto cheklangan o'zbek tilidagi ma'lumotlar bilan ham ishlashni yaxshilaydi.

O'lchovlilik:

- Samarali joylashtirish: Wav2vec 2.0 turli xil apparat platformalarida samarali joylashtirilishi mumkin, bu esa hisoblash resurslari cheklangan bo'lishi mumkin bo'lgan mobil ilovalar uchun mos keladi.

- Real vaqtda ishlash: model arxitekturasi interaktiv ilovalar uchun zarur bo'lgan Real vaqtda nutq sintezini qo'llab-quvvatlaydi.

Hamjamiyat va qo'llab-quvvatlash:

- Faol tadqiqot hamjamiyati: Wav2vec 2.0 TTS tizimlarini takomillashtirish uchun ishlatilishi mumkin bo'lgan doimiy takomillashtirish va yangilanishlarga ega jonli tadqiqot hamjamiyatining bir qismidir.

- Ochiq manbali mavjudlik: ochiq manbali model va uning yutuqlari osongina mavjud bo'lib, innovatsiyalar va hamkorlikni rivojlantiradi.

Ushbu afzalliklar wav2vec 2.0-ni o'zbek tili uchun samarali va sifatli TTS tizimlarini ishlab chiqish, mobil ilovalar va boshqa raqamli platformalarda foydalanuvchilarning tajribasini yaxshilash uchun kuchli nomzodga aylantiradi.

Wav2vec 2.0 ning o'zbek TTSDagi muvaffaqiyatini bir qancha omillar bilan bog'lash mumkin. Birinchidan, o'z-o'zidan boshqariladigan ta'lim yondashuvi modelga katta hajmdagi yorliqsiz ma'lumotlardan foydalanishga imkon beradi, bu ayniqsa kam taqdim etilgan tillar uchun foydalidir. Ikkinchidan, nozik sozlash jarayoni modelning o'zbek tilining o'ziga xos xususiyatlariga yaxshi moslashishini ta'minlaydi. Natijalar shuni ko'rsatadiki, wav2vec 2.0 resurslari cheklangan tillar uchun TTS tizimlarining sifatini sezilarli darajada yaxshilashi mumkin.

### **Xulosa**

TTS uchun wav2vec 2.0 ning o'zbek tilida qabul qilinishi kam taqdim etilgan tillar uchun nutq texnologiyasida muhim qadamdir. Modelning yuqori sifatli, tabiiy ovozli nutqni ishlab chiqarish qobiliyati uning kengroq dasturlar, shu jumladan til o'rganish vositalari, yordamchi texnologiyalar va ovozli dasturlar uchun imkoniyatlarini ta'kidlaydi.

Kelajakdagi tadqiqotlar modelni yanada takomillashtirishga, uning boshqa kam taqdim etilgan tillarga tatbiq etilishini o'rganishga va turli xil audio namunalarni kiritish uchun ma'lumotlar to'plamini kengaytirishga qaratilishi kerak. Tilshunos mutaxassislar va texnologlar o'rtasidagi hamkorlikdagi sa'y-harakatlar ushbu sohani rivojlantirishda va zamonaviy TTS texnologiyalarining afzalliklari barcha tillarda so'zlashuvchilar uchun ochiq bo'lishini ta'minlashda hal qiluvchi ahamiyatga ega bo'ladi.

### **Foydalanilgan adabiyotlar**

1. Sabharwal, N., & Agrawal, A. (2021). Introduction to Word Embeddings. In Hands-on Question Answering Systems with BERT. [https://doi.org/10.1007/978-1-4842-6664-9\\_3](https://doi.org/10.1007/978-1-4842-6664-9_3)

2. Tan, M., Zhou, W., Zheng, L., & Wang, S. (2012). A Scalable Distributed Syntactic, Semantic, and Lexical Language Model. *Computational Linguistics*, 38(3). [https://doi.org/10.1162/COLI\\_a\\_00107](https://doi.org/10.1162/COLI_a_00107)
3. Elsafoury, F., Wilson, S. R., & Ramzan, N. (2022). A Comparative Study on Word Embeddings in Social NLP Tasks. *SocialNLP 2022 - 10th International Workshop on Natural Language Processing for Social Media, Proceedings of the Workshop*. <https://doi.org/10.18653/v1/2022.socialnlp1.5>
4. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2006). Distributed Representations of Words and Phrases and their Compositionality. *Neural Information Processing Systems*, 1.
5. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*.
6. Elov, B., Hamroyeva, Sh., Alayev, R., Xusainova, Z., & Yodgorov, U. (2023). O‘zbek tili korpusi matnlarini qayta ishlash usullari. *Digital transformation and artificial intelligence*, 1(3), 117–129. Retrieved from <https://dtai.tsue.uz/index.php/dtai/article/view/v1i317>
7. Elov B., Aloyev N., Xusainova Z., Yuldashev A. O‘zbek tili korpusi matnlarini qayta ishlash Word2Vec, GloVe, ELMO, BERT usullari // Труды XI Международной конференции «Компьютерная обработка тюркских языков» «TURKLANG 2023». Бухара, 20-22 октября 2023 г.
8. Bamler, R., & Mandt, S. (2017). Dynamic Word Embeddings via Skip-Gram Filtering. *Proceedings of ICML 2017*