O‘ZBEK TILIDAGI TOVUSH VA MATINLARGA RAQAMLI ISHLOV BERISH ORQALI ULAR ORASIDAGI MOSLIK DARAJASINI ANIQLASH VA BAXOLASH ALGARITIMLARI.

Baxromov Alisher Baxtiyor o‘g‘li
Sharof Rashidov nomidagi Samarqand davlat universiteti tayanch doktaranti
E-mail: alisherbaxromov91921@gmail.com

Anatatsiya. So‘nggi yillarda sun’iy intellekt asosidagi nutqni matnga aylantirish (ASR) texnologiyalari tez sur’atlarda rivojlanib bormoqda. Bu jarayon, ayniqsa, ta’lim sohasida test savollariga og‘zaki javoblarni avtomatik tarzda yozib olish va baholash imkonini yaratadi. Tadqiqotning asosiy maqsadi — talaba yoki foydalanuvchining og‘zaki javobi ovoz yozuvi sifatida qabul qilinadi, keyin STT (Speech-to-Text) modellaridan foydalanib matnga aylantiriladi, hosil bo‘lgan matn esa etalon javob bilan solishtiriladi va avtomatik tarzda baholanadi. Baholash jarayonida klassik ko‘rsatkichlar — WER (Word Error Rate), CER (Character Error Rate) hamda zamonaviy semantik o‘xshashlik metrikalari qo‘llaniladi. Tadqiqot natijalari shuni ko‘rsatadiki, an’anaviy xatolik ko‘rsatkichlari bilan bir qatorda semantik o‘xshashlikni baholash foydalanuvchining javobini yanada adolatli tahlil qilish imkonini beradi. Ushbu yondashuv ta’limda test baholash tizimlarini avtomatlashtirish, shuningdek, apellyatsiya jarayonlarida talabalarning huquqlarini ta’minlashda samarali vosita bo‘lib xizmat qilishi mumkin.

Kalit so‘zlar: Nutqni matnga aylantirish, avtomatik baholash, WER, CER, semantik o‘xshashlik, til modeli, test tizimlari, apellyatsiya.

Abstract. In recent years, artificial intelligence-based speech-to-text (ASR) technologies have been developing rapidly. This process, especially in the field of education, allows for the automatic recording and evaluation of oral responses to test questions. The main goal of the study is that the oral response of a student or user is received as a voice recording, then converted into text using STT (Speech-to-Text) models, and the resulting text is compared with a reference response and automatically evaluated. The evaluation process uses classic indicators - WER (Word Error Rate), CER (Character Error Rate), as well as modern semantic similarity metrics. The results of the study show that, along with traditional error indicators, the assessment of semantic similarity allows for a more fair analysis of the user's response. This approach can serve as an effective tool for automating test assessment systems in education, as well as ensuring the rights of students in appeal processes.

Keywords: Speech-to-text conversion, automatic evaluation, WER, CER, semantic similarity, language model, testing systems, appeal.

Аннотация. В последние годы технологии преобразования речи в текст (ASR), основанные на искусственном интеллекте, стремительно

развиваются. Этот процесс, особенно в сфере образования, позволяет автоматически записывать и оценивать устные ответы на тестовые вопросы. Основная цель исследования заключается в том, чтобы устный ответ студента или пользователя был получен в виде записи голоса, затем преобразован в текст с использованием моделей STT (Speech-to-Text), а полученный текст сравнивался с эталонным ответом и автоматически оценивался. В процессе оценки используются классические показатели – WER (Word Error Rate), CER (Character Error Rate), а также современные метрики семантического сходства. Результаты исследования показывают, что, наряду с традиционными показателями ошибок, оценка семантического сходства позволяет более объективно анализировать ответ пользователя. Данный подход может служить эффективным инструментом для автоматизации систем оценки результатов тестов в образовании, а также для обеспечения прав студентов при подаче апелляций.

Ключевые слова: Преобразование речи в текст, автоматическая оценка, WER, CER, семантическое сходство, языковая модель, системы тестирования, апелляция.

Krish

Ushbu maqolada Audioni matinga aylantrish va aylantrib olingan matinni tayyor javob matni bilan solishtirib natijani baxolab berish bo’yicha olimlar va ilmiy tadqiqotchilar maqolalarini o’rganib chiqish va taxlil qilishdan iborat

1. Fred Jelinek. IBM Thomas J.atson Research markazida ishlagan. U nutqni avtomatik tanishda ehtimollik modellarini joriy qilgan birinchi olimlardan biri. Fred Jelinekning *Statistical Methods for Speech Recognition* kitobi tadqiqotlari hammasi jamlangan va asosiy g‘oyasi shundan iborat bo‘lganki, nutqni tanish jarayonida statistik til modellari eng muhim ro‘l o‘ynagan Nutqni tanish jarayonini matematik va statistik modellashtirish orqali amaliy tizimlarni samaraliroq qilishga erishgan. Uning ishlari bugungi zamonaviy Nutiqni matinga aylantirish(Speech- to -Text STT) tizimlarining nazariy asosini tashkil etmoqda.

2. Lawrence Rabiner va Biing-Hwang Juang. 1980–1990-yillarda HMM (Hidden Markov Models) asosida nutq tanish texnologiyalarini ishlab chiqqan. Rabinerning 1989-yilda chop etilgan *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition* maqolasi ushbu sohada asosiy manba hisoblangan. Ularning metodlari Bell Labs laboratoriyalarida qo‘llanilib, telefon tizimlarida nutqni tanishga asos yaratgan. **Ularning ishlari** o‘nlab yillar davomida nutqni matnga aylantirishda HMM eng keng qo‘llanilgan model bo‘lib hizmat qilgan.

3. James L. Flanagan. Bell Labsda nutq akustikasi, signalni kodlash va nutqni sintezlash ustida ishlagan. U ovozli aloqa tizimlari uchun kodlash

ISSN 3030-3907

algoritmlarini ishlab chiqqan. *Speech Analysis, Synthesis and Perception* kitobi bu sohaning ilk asosiy manbalaridan biri hisoblangan. Nutq signalining fizik asoslarini tadqiq etib, STT tizimlari uchun mustahkam akustik poydevor yaratgan.

4. Geoffrey Hinton. 2000-yillar boshida chuqur neyron tarmoqlarni nutqni tanishda qo‘llash g‘oyasini ilgari surdi. 2012-yilda Google bilan hamkorlikda nutqni tanish tizimlariga DNN (Deep Neural Networks) ni tatbiq etdi. DNN’lar STT aniqligini 20–30% yaxshilagan. Bugungi kunda barcha yirik STT tizimlari (Google Speech-to-Text, Amazon Transcribe, Siri va boshqalar) aynan Hinton yo‘nalishini davom ettirmoqda.

5. Alex Graves. Recurrent Neural Networks (RNN), LSTM (Long Short-Term Memory), va **CTC (Connectionist Temporal Classification)** metodini ishlab chiqqan. CTC algoritmi tufayli nutq segmentlarga ajratilmagan bo‘lsa ham, model uni matnga aylantira oladi. Bu real vaqtda STT tizimlari uchun muhim burilish bo‘ldi. Hozirgi kunda ko‘plab STT tizimlari (DeepSpeech, wav2vec va boshqalar) CTC yoki unga o‘xshash yondashuvlarga tayanadi.

6. Junyoung Chung, Caglar Gulcehre, Kyung Hyun Cho, Yoshua Bengio larning *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling* nomli maqolasida esa takrorlanuvchi neyron tarmoqlar (RNN) uchun turli arxitekturalar oddiy **tanh-unit, LSTM (Long Short-Term Memory)** va **GRU (Gated Recurrent Unit)** modellarining ketma-ketliklarni modellashtirish vazifalaridagi samaradorligi real ma’lumotlar va tajribalar bilan sinovdan o‘tkazib baholangan.

Tadqiqotda polifonik musiqa va nutq signali kabi real ketma-ketlik ma’lumotlari ustida tajribalar o‘tkazilgan. Natijalari shuni ko‘rsataganki, oddiy RNN modellarida gradientning yo‘qolishi muammosi sezilarli bo‘lib, uzoq muddatli bog‘lanishlarni ushlashda muvaffaqiyatsizlikka olib keladigan. Unga qarshi LSTM va GRU arxitekturalari uzoq muddatli ma’lumotlarni samarali eslab qolib, yuqori aniqlikka erishganlar. Ayniqsa, GRU modeli soddaroq tuzilishi va tezroq o‘qitilishi bilan ajralib turadigan, ko‘pgina hollarda LSTM bilan teng yoki undan ham yuqori natijalar bergan. Va natijada Gated Recurrent Neural Network (GRNN) modellarining ketma-ketliklarni modellashtirishda ustunligi isbotlangan.

7. Zhiyuan Tang, Ying Shi, Dong Wang, Yang Feng, Shiyue Zhang larning *Memory Visualization for Gated Recurrent Neural Networks in Speech Recognition* nomli maqolasida GRNN(Gated Recurrent Neural Networks), ya’ni **LSTM** va **GRU** modellarining oxirgi vaqtlarda avtomatik nutqni tanib olishda ASR(bu inson nutqini audio sifatida qabul qilib matinga aylantiradi) samaradorligi yuqori ekanligi isbotlangan. Shunga qaramay, bu modellar ichida ma’lumot qanday saqlanishi, qachon o‘chirib tashlanishi va qanday chiqarilishi ko‘pincha ichki qaor qabul qilishni tushuntrib bo‘lmaydi. Bu maqolada **GRNN(**Gated Recurrent Neural

ISSN 3030-3907

Network) **modellarining xotira faoliyatini vizualizatsiya qilish** usulini taklif etgan.

Tajribalari davomida **LSTMdagi input, forget va output eshiklari** hamda **GRU'dagi update va reset eshiklari** vaqt bo'yicha kuzatishdi va nutq signalidagi turli vaziyatlarda (pauza, shovqin, cho'zilish) qanday reaksiya qilishini tahlil qilgan. Natijalari shuni ko'rsatdiki, GRU va LSTM turli xil xotira boshqarish mexanizmlariga ega bo'lsa-da, ikkalasi ham nutqni tanishda samarali ishlaydi. Shu bilan birga, **vizualizatsiya tahlili** ushbu modellarni yaxshiroq tushunishga yordam beradi hamda kelajakda yanada samarali arxitekturalarni ishlab chiqish uchun asos yaratadi.

8. Andrew Ng va Baidu Research. Ishlari: "Deep Speech" 20214-2015 yillarda yaratilgan loyihalari orqali chuqur neyron tarmoqlar asosida STT tizimlari yaratildi. Ushbu loyihada chuqur neyron tarmoqlar DNN (Deep Neural Networks) asosida end-to-end avtomatik nutqni tanish tizimi taqdim etgan. An'anaviy STT(aoudioni matinga alantirish) tizimlari akustik model, til modeli va nutq tovushlarini yozuvdan aniq ifodalash kabi alohida bosqichlarga tayangan. Deep Speech esa bu murakkab bosqichlarni bitta neyron tarmoq modeli o'rqali o'rnini bosishga urinagan.

Tizimning yangiliklari quydagicha bo'lgan: GPU-klasterlardan foydalangn holda, juda katta hajmdagi nutq va matn ma'lumotlari ustida modelni o'qitish, end-to-end o'qitish kiritilgan audio malumotlrni birdaniga matnga aylantiriladi, RNN (Recurrent Neural Network ketma ketlikdagi malumotlar bilan ishlashga mo'ljallangan dastur) va CTC (Connectionist Temporal Classification muommoni gal qilish uchun ishlab chiqilgan yo'qotish funksiyasi) arxitekturasidan foydalangan. Deep Speech 2 ingliz va xitoy tillarida ishlash uchun kengaytirilib, ko'p tilli STT tizimlari uchun asos yaratgan. Va hozirgi kunda jida mashkur **OpenAI jamoasi Whisper** 680 ming soatdan ortiq ko'p tilli nutq ma'lumotlarida o'qitilgan ochiq manbali STT modeli. 100 dan ortiq tillarda, jumladan **o'zbek tilida ham** ishlaydi. Shovqinli nutqni ham yaxshi tanish xususiyatiga ega. **Facebook wav2vec 2.0 va XLS-R.** Annotatsiya qilinmagan audio ma'lumotlarda **self-supervised learning** metodini qo'llab STT modellarini yaratgan. O'zbek tili kabi resursi kam tillar uchun juda mos, chunki kichik korpuslar bilan ham yuqori aniqlik beradi. **Mozilla Deep Speech** Andrew Ng'ning Deep Speech loyihasi asosida ochiq manbali STT modelini yaratdi. Dasturchilar va tadqiqotchilar uchun erkin foydalanish imkonini berdi. Common Voice (2017-yilda boshlangan ochiq manbali nutq ma'lumotlar bazasini yaratish) loyihasi orqali turli tillardgi katta hajimdagi audio fayillarni matinli fayilga aylantirib yig'ilgan.

O'zbek olimlari va tadqiqotchilar

ISSN 3030-3907

1. Muhammadjon Musayev, Saida Musaxo‘jayeva, Ilyos Hujayorov, Yerbo‘lat Xasanov Mannon Ochilov va Huseyin Atakan Vorillar tomonidan chiqarilgan *An Open-Source Uzbek Speech Corpus and Initial Speech Recognition* **nomli** maqolasi O‘zbek tili uchun erkin foydalanish mumkun bo‘lgan kata hajimdagi ovozli yozuvlarni matinga aylantrish taqdim etilgan va chuqur neyro‘n tarmog‘ining yashirin Markov madeli (DNN, HMM) va end-to-end(E2E) arxitekturasidan foydalangan holda nutiqni aftamatik aniqlash ASR(Automatic Speech Recognition)ning dastlabki natijalari haqida xabar bergan.tadqiqotda O‘zbek tilidagi kata hajimdagi ovozli yozuvlarni matinga aylantrish USC(Uzbek Speech Corpus)da jami 105 soatlik transkripsiyalangan audio yozuvlarga ega bo‘lgan 958 xil maruzachilarni o‘z ichiga olgan. Ekistrimental natijalari USC ning ASR uchunqo‘llanilishi uchun istiqbolli natijalarni ko‘rsatgan, tasdiqlash va test test to‘plamlarida mos ravishda 18,1% va 17,4% so‘z xatosi darajasiga erishilgan. Bu O‘zbek tilida ilmiy asoslangan birinchi yirik STT resursi, Ko‘plab tadqiqotlar uchun ochiq manba sifatida xizmat qilmoqda.

2. Abduqahhor Saidov, Sardorbek Yusupov, Anvar Narzullaev, Alisher Matyakubov, Saidmurod Radjabov, Hafiz Ibrahim, Jasur Temirovlarning *A 60 Hour Uzbek Read Speech Corpus with Punctuation, Casing, and Context* maqolasida **akademik tadqiqot aqsadlarida erkin foydalanish mumkun bo‘lgan kiril va lotin alifbolarida transkriplarni oz ichiga olgan o‘zbek tilining o‘qiladigan nutqiy ko‘rpusi FeruzaSpeech bilan tanishtirilgan. Bu maqolaga Toshken shaxridan kelgan bitta ona tilida so‘zlashuvchi ayolning 60 soatlik yuqori sifatli yozuvini oz ichiga olgan** Lotin va Kirill yozuvlari bilan. Ushbu yozuvlar kitobdan va BBC Newsdan qisqa parchalardan iborat. Maqolada CommonVoice 16.1 ning o‘zbekcha ma’lumotlari, O‘zbek nutqi korpusi ma’lumotlari va FeruzaSpeech integratsiyasi natijasida FeruzaSpeech ma’lumotlaridagi Word xatosi darajasi (WERs) yaxshilanishi muhokama qilingan.

3. Mukhamadiyev, A.; Mukhiddinov, M.; Khujayarov, I.; Ochilov, M.; Cho, J larning *Development of Language Models for Continuous Uzbek Speech Recognition System* nomli maqolasida katta lug‘atga ega bo‘lgan avtomatik nutqni aniqlash tizimlari va boshqa tabiiy tillarni qayta ishlash dasturlari til modelisiz ishlay olmayganligi aniqlangan. Oldindan o‘rgatilgan til modellari bo‘yicha tadqiqotlarning aksariyati ingliz, xitoy va turli Yevropa tillari kabi mashhurroq tillarga qaratilgan, biroq ommaga ochiq o‘zbek nutqi ma’lumotlar to‘plami yo‘q. Shuning uchun kam resursli tillarning til modellarini o‘rganish va yaratish zarur. tadqiqotlarining maqsadi o‘zbek tili uchun kam resursli til modelini ishlab chiqish va lingvistik hodisalarni tushunish orqali ushbu cheklovni bartaraf etishdan iborat.

Tadqiqotda O‘zbek tilining o‘ziga xos xususiyatlarini hisobga oladigan statistik va neyron-tarmoqqa asoslangan til modellarining ishlashini o‘rganib,

ISSN 3030-3907

UzLM nomli o‘zbek tili modelini taklif qildgan. O‘zbek tiliga xos lingvistik taqdimotda avvalgi tadqiqotlarida qo‘llanganidek, bir xil yoki kamroq o‘rgatuvchi so‘zlardan foydalangan holda turli manbalardan 80 million so‘zdan foydalangan holda yanada mustahkam UzLM yaratish imkonini bergan. Ushbu korpusni yaratish uchun oltmish sakkiz mingga yaqin turli so‘zlar va 15 million jumlalar to‘plangan. O‘zbek tili nutqini uzluksiz tanib olish bo‘yicha o‘tkazgan testlar eksperimental natijalari shuni ko‘rsatganki, qo‘lda kodlash bilan solishtirganda, neyron-tarmoqqa asoslangan til modellaridan foydalanish belgilar xatosi darajasini 5,26% gacha qisqartirgan.

4. Salayeva Maxliyo, Kuriyozov Elmurod, Salayev Ulug‘beklarning *Uzbek automatic speech recognition models using deep learning techniques* nomli maqolasida o‘zbek tili uchun avtomatik nutqni aniqlash (ASR) modellaridan foydalanish bo‘yicha tadqiqot taqdim etilgan. O‘zbek tili uchun avvaldan mavjud bo‘lgan ASR modellarining ishlashi baholanadi va taqqoslanadi, shuningdek, kam resursli o‘zbek tili uchun ASR modellarini yaratish usullari. Tadqiqot kam resursli tillarda ASR modellarining ishlashini yaxshilash uchun uzatishni o‘rganish va ma’lumotlarni ko‘paytirish usullaridan foydalanishni o‘z ichiga oladi. Bundan tashqari, kam resursli tillarda ASR modellarini o‘qitish uchun nazoratsiz o‘rganish algoritmlarining samaradorligi ham muhokama qilingan. Ushbu tadqiqot natijalari kam ta’minlangan tillar uchun ASR texnologiyasini rivojlantirish uchun muhim ahamiyatga ega.

Matnlarni solishtirish va semantik baholash

1. Thennal D K, Jesin James, Deepa P Gopinath, Muhammed Ashraf Klarning *Advocating Character Error Rate for Multilingual ASR Evaluation* nomli maqolasida nutqni avtomatik aniqlash (ASR) tizimlari an’anaviy ravishda inglizcha ma’lumotlar to‘plamidan foydalangan holda baholangan, so‘z xatosi darajasi (WER) ustun ko‘rsatkich bo‘lib xizmat qilgan. WERning soddaligi va talqin qilish qulayligi uning keng tarqalishiga, ayniqsa ingliz tiliga yordam bergan. Biroq, ASR tizimlari ko‘p tilli kontekstlarga kengayganligi sababli, WER turli yo‘llar bilan muvaffaqiyatsizlikka uchraygan, ayniqsa morfologik jihatdan murakkab tillar yoki aniq so‘z chegaralari bo‘lmagan tillarda. maqolada baholash ko‘rsatkichi sifatida WER cheklovlarini hujjatlashtiradi va ko‘p tilli ASR baholashda asosiy ko‘rsatkich sifatida belgilar xatosi darajasini (CER) himoya qilagan. Maqolada shuni ko‘rsatganki, CER WER duch keladigan ko‘plab qiyinchiliklardan qochadi va yozish tizimlarida ko‘proq izchillikni namoyish etadi.

ASR transkripsiyalarini inson tomonidan baholangan uchta tilda: malayalam, ingliz va arab tillarida o‘tkazish orqali taklifimizni qo‘llab-quvvatlagan, ular alohida morfologik xususiyatlarga ega. Tadqiqot shuni ko‘rsatadiki, CER inson mulohazalari bilan WERga qaraganda, hatto ingliz tiliga

ISSN 3030-3907

ham yaqinroq bog‘liq. Keyingi tadqiqotlarini osonlashtirish uchun bu tadqiqotda ASR ko‘rsatkichlarini kelajakda taqqoslash uchun insoniy baholash ma’lumotlar to‘plamini chiqaradi. Topilm shuni ko‘rsatadiki, turli tillarning o‘zgaruvchan lingvistik xususiyatlarini hisobga olish uchun ko‘p tilli ASR baholashlarida CERga ustunlik berish yoki hech bo‘lmaganda to‘ldirish kerak bo‘ladi. Ishda baholash ko‘rsatkichi sifatida WER cheklovlarini hujjatlashtiradi va ko‘p tilli ASR baholashda asosiy ko‘rsatkich sifatida belgilar xatosi darajasini (CER) himoya qiladi va CER WER duch keladigan ko‘plab qiyinchiliklardan qochadi va yozish tizimlari bo‘ylab ko‘proq muvofiqlikni namoyish etadi.

Shu yerda WER va CER haqida to‘xtalib o‘tamiz Rev kompaniyasinining saytida WER shunday tushuntirilgan manba: https://www.rev.com/resources/what-is-wer-what-does-word-error-rate-mean. Nutqni avtomatik aniqlash (ASR) texnologiyasi nutq tilini aniqlash va qayta ishlash uchun mashinalar va dasturlardan foydalanadi. Bundan tashqari, ovoz orqali odamning shaxsini tasdiqlash uchun ham foydalanish mumkin. Ushbu texnologiya so‘nggi yillarda sezilarli darajada rivojlangan, ammo har doim ham mukammal natijalarni bermaydi.Nutqni tanib olish va uni matn shakliga o‘tkazish jarayonida ba’zi so‘zlar qolib ketishi yoki noto‘g‘ri tarjima qilinishi mumkin. Agar siz ASR dan ma’lum darajada foydalangan bo‘lsangiz, ehtimol siz "so‘z xatosi darajasi" (WER) iborasiga duch kelgansiz. Keling, WER g‘oyasini, uni qanday hisoblashni va nima uchun muhimligini ko‘rib chiqaylik. So‘z xatosi darajasi qanday ko‘rib chiqiladi, bu murakkab fikr bo‘lib tuyulishi mumkin, ammo asosiy WERni hisoblash usuli aslida juda oddiy. Asosan, WER - bu umumiy so‘zlarga bo‘lingan xatolar soni hisoblanadi. Bundan shuni xulosa qilib aytish mumkunki ASR texnalogiyasidan qilingan kamchiliklarni WER texnalogiyasi yordamida ancha aniqlikga erishish mumkun.

CER esa WER dan farqli o‘laroq Har darajasidagi hatolar nisbati tushiniladi, CER ham WER kabi hisoblanadi so‘zlar o‘rniga harflar hisoblanadi bu ayniqsa qiyin so‘zlarni, texnik atamalarni isimlarni yoki kam ishlatiladigan so‘zlarni baholashda juda samarali bo‘ladi. Shunga qaramay bugungi kunda WER va CER yetarli bo‘mayabdi chunki so‘zlar sino‘nim bilan aytilsa WER katta bo‘lib ketadi shu sababli semantik o‘xshshlik(semantic similarity) ishlatiladi. Steffen Gerboldning S*emantic similarity prediction is better than other semantic similarity measures* nomli maqolasida shu borada ishlar olib borilgan. Tabiiy til matnlari o‘rtasidagi semantik o‘xshashlik odatda quyi ketma-ketliklar o‘rtasidagi o‘zaro bog‘liqlikni ko‘rib chiqish (masalan, BLEU) yoki o‘rnatish (masalan, BERTScore, S-BERT) yordamida o‘lchanadi. Ushbu maqolada faqat semantik o‘xshashlikni o‘lchashga qiziqqanlarida, bunday vazifa uchun nozik sozlangan model yordamida o‘xshashlikni to‘g‘ridan-to‘g‘ri taxmin qilish yaxshiroq ekanligini ta’kidlagan.

ISSN 3030-3907

GLUE benchmarkidagi semantik matn o‘xshashligi benchmark vazifalari (STS-B) uchun nozik sozlangan modeldan foydalanib, biz STSScore yondashuvini aniqlaymiz va natijada paydo bo‘lgan o‘xshashlik boshqa yondashuvlarga qaraganda mustahkam semantik o‘xshashlik o‘lchovi bo‘yicha bizning kutganlarimiz bilan yaxshiroq mos kelishini ko‘rsatamiz.

Hila Gonen1 Shauli Ravfogen, Yanai Elazar, Yoav Goldberg, larning “*It’s not Greek to mBERT: Inducing Word-Level Translations from Multilingual BERT*” nomli maqolasida esa quydagi ishlar amalga oshirilgan. Ko‘p tilli BERT (mBERT) tillar bo‘ylab uzatish imkonini beruvchi boy tillararo vakilliklarni o‘rganadi. Maqolada mBERT-ga kiritilgan so‘z darajasidagi tarjima ma’lumotlarini o‘rganilgan va hech qanday nozik sozlashsiz ajoyib tarjima imkoniyatlarini ochib beradigan ikkita oddiy usulni taqdim etgan. Natijalari shuni ko‘rsatganki, ushbu ma’lumotlarning aksariyati chiziqli bo‘lmagan tarzda kodlangan, ba’zilari esa faqat chiziqli vositalar yordamida tiklanishi mumkin. Tahlillarining bir qismi sifatida mBERT til kodlash komponenti mavhum va o‘zaro tillar komponentini o‘z ichiga olgan tasvirlarni o‘rganishi haqidagi giptozani sinovdan o‘tkazadi va mBERT ko‘rinishlari ichida empirik til-identifikatsiyasi pastki fazosini aniq aniqlaydi.

Mikhail Evtikhiev, Egor Bogomolov, Yaroslav Sokolov, Timofey Bryksinlarning “*Out of the BLEU: how should we assess quality of the Code Generation models?*” nomli maqolasida esa so‘nggi yillarda tadqiqotchilar juda ko‘p sonli kod ishlab chiqarish modellarini yaratdilar va joriy qildilar deyilgan. Shu bois har bir yangi model versiyasini inson tomonidan baholash mumkin emasligi sababli, hamjamiyat insoniy mulohazalar natijalarini yaqinlashtirish uchun BLEU kabi avtomatik baholash ko‘rsatkichlarini qabul qilganlar. Ushbu ko‘rsatkichlar mashina tarjimasi domenidan kelib chiqadi va ular kod ishlab chiqarish vazifalari uchun qo'llanilishi yoki bu vazifani inson baholashiga qanchalik mos kelishi noma’lum. Kodning o'xshashligini baholash uchun ishlab chiqilgan CodeBLEU va RUBY kabi boshqa ko'rsatkichlar ham mavjud bo'lib, ular manba kodining xususiyatlarini hisobga oladi. Biroq, bu ko'rsatkichlar bo'yicha ularning inson bahosi bilan kelishilganligi bo'yicha deyarli hech qanday tadqiqotlar mavjud emas. Shunga qaramay, metrik ballardagi minimal farqlar so'nggi maqolalarda ba'zi kod ishlab chiqarish modellarining boshqalardan ustunligini da'vo qilish uchun ishlatilgan.

Maqolada oltita ko'rsatkich - BLEU, ROUGE-L, METEOR, ChrF, CodeBLEU va RUBY - kod ishlab chiqarish modellarini baholash uchun qo'llanilishi bo'yicha tadqiqotni taqdim etgan. Ikki xil kod ishlab chiqarish ma'lumotlar to'plami bo'yicha tadqiqot o'tkaziladi va ushbu ma'lumotlar to'plamida ishlaydigan barcha modellarning sifatini baholash uchun inson annotatorlari(tajriba ishtirokchilari)dan foydalanilgan. Natijalar shuni ko'rsatganki, Python bir-

ISSN 3030-3907

laynerlarining CoNaLa ma'lumotlar to'plami uchun ko'rsatkichlarning hech biri, agar model ballaridagi farq 5 balldan kam bo'lsa, >95% ishonch bilan qaysi model yaxshiroq ekanligi haqidagi inson fikrini to'g'ri taqlid qila olmaydi. Muayyan strukturaning sinflaridan iborat HearthStone ma'lumotlar to'plami uchun model ballaridagi farq kamida 2 ball bo'lishi bir modelning boshqasidan ustunligini da'vo qilish uchun etarli bo'lgan. Mikhail Evtikhiev, Egor Bogomolov, Yaroslav Sokolov, Timofey Bryksinlarning topilmalari shuni ko'rsatadiki, ChrF ko'rsatkichi kod ishlab chiqarish modellarini baholash uchun keng tarqalgan BLEU va CodeBLEU-ga qaraganda yaxshiroq mos keladi.

Zhanibek Kozhirbayevning *Kazakh Speech Recognition: Wav2vec2.0 vs.* Whisper nomli maqolasida Qozoq nutqini facebukning Wav2Vec2.0, Wav2Vec2-XLS-R madeli va Open AI ning Whisper madellari yordamida matinga aylantirib ASR ni sinovdan o'tkazgan hamda ularning samaradorligini aniqlagan. Ushbu tadqiqotning maqsadi qozoq tilidagi nutqni transkripsiyalashda ushbu modellarning samaradorligini o'rganish va ularning ishlashini mavjud nazorat qilinadigan avtomatik nutqni aniqlash (ASR) tizimlari bilan solishtirishdir. Tadqiqot shuningdek, boshqa tillardagi ma'lumotlardan oldingi tayyorgarlik uchun foydalanish imkoniyatini o'rganish va maqsadli til ma'lumotlarini nozik sozlash model ishlashini yaxshilash mumkinligini tekshirishga qaratilgan. Shunday qilib, ushbu ish resurslari kam bo'lgan til sozlamalarida oldindan tayyorlangan ko'p tilli modellardan foydalanish samaradorligi haqida tushuncha berishi mumkin. Wav2vec2.0 modeli sinov to'plamida 2,8 belgilar xatosi darajasiga (CER) va 8,7 so'z xatosi darajasiga (WER) erishilgan, bu uchdan-end Transformer modeli erishgan eng yaxshi natijaga yaqindan mos keladi. Katta shivirlash modeli sinov majmuasida taxminan 4 CERga erishadi.

Ibragimova S.N. Turayev B.Sh. Abdullayeva M.I.larning Solving the problems ofnormalization of non-standardwords in the text of the uzbeklanguage nomli maqolasida nostandart so'zlar (raqamsan, sonlar, qisqartmalar va h.k)larni aniqlash va ularni normatizatsiya qilish bo'yicha ishlar amalga oshirilgan shular bo'yicha qisqacha tanishib chiqamiz. Matnni normallashtirish matndan nutqqa (TTS) tizimining muhim tarkibiy qismi bo'lib, matnni normallashtirishning qiyinligi nostandart so'zlarni farqlashda yotadi. Bu ishda o'zbek nutqi asosidagi nostandart so'zlar taksonomiyasi ishlab chiqilib, nostandart so'zlarni aniqlashning ikki bosqichli strategiyasi taklif qilingan. Maqolada o'zbekcha matinning normalizatsiyasi har tomonlama o'rganilgan, o'zbekcha so'zlarning nostandart taksinomasi ishlab chiqilgan tizimning ishlashi Masalan bugun 23-sentyabr 2025-yil degan matin kiritilsa uni quydagicha to'g'irlab beradi bugun yigirma uchinchi sentyabr ikki ming yigirma beshinchi yil ko'rinishida chiqaradi buni takomillashtirib nutiqda tinish belgilar ishlatilmaydi tinish belgilarni o'z orniga qoyadigan tizimniham qo'shsa bo'ladi.

Baholash algoritmi

ISSN 3030-3907

Yuqorida ko’rsatib o’tilgan texnikalardan foydalanib audioni matinga aylantrib olinadi va tayyor javob matni bilan solishtiriladi bizni ishimizni asosiy qismi mana shu baxolash qismi hisoblanadi .

1. Agar javob matni asl matin bilan solishtirilganda 85% dan 100% gacha aniqlikda chiqarilsa baxoho 5 yoki bal ko’rinishida chiqariladi
2. Agar javob matni asl matin bilan solishtirilganda 70% dan 84% gacha mos chiqsa 4 baxo yoki ko’rsatilgan ball chiqariladi.
3. Agar javob asl matin bilan solishtirilganda 56% dan 69% gaccha chiqsa 3 baxo yoki ko’rsatilgan ball chiqariladi
4. Agar javob matni asl matin bilan solishtirilganda 55% va undan past chiqsa umuman to’g’ri javob berilmadi degan xulosa berilishi kerak.

Apellyatsiya mexanizmi

Apelatiya mexanizmihham xuddi shu tarzda sihlaydi javob berayotgan shaxsning audiyosi yozib olingan holda saqlab qo’yiladi isim familiya ochistva iloji bo’lsa paspo’rt malumotlari bilan, javob bergan shaxs o’z balidan norozi bo’lib apelatsiya berilgan taqdirda yozib olingan audiyo fayil matinga aylantirilib asil javob matni bilan solishtirilib baxolash mexanizimida ko’rsatilgan bandlarga amal qilgan holda qayta ko’rich chiqiladi va oxirigi hulosa chiqarilib bal beriladi. Nazoratchi xatosi bilan past baxo qo’yilgan bo’lsa tizim tomonidan ko’rib chiqilib yuqori baxo berilsa oxirgi najija shu tizim baxosi asosida qo’yiladi.

Tadqiqot uchun mavjud ishlar

Bunday yondashuvlar chet ellda “Automated Spoken Short Answer Scoring”, “ASR + NLP for Assessment”, yoki “Computer-Assisted Language Testing (CALT)” nomlari bilan ishlangan:

ETS (Educational Testing Service, TOEFL) – talabalar nutqini yozib olib, matnga aylantirib, NLP yordamida baholash ishlari olib borilgan.

SemEval (Semantic Evaluation tasks) – sinonimlar, parafraz va semantik yaqinlikni o‘lchash uchun ishlatiladigan benchmarklar mavjud.

Automatic Short Answer Grading – qisqa javoblarni NLP yordamida baholash bo‘yicha ko‘plab maqolalar bor.

Tavsiya qilinadigan texnologiyalar

ASR uchun: Whisper, wav2vec2.0 (o‘zbek uchun fine-tune qilinishi mumkin).

Semantik solishtirish uchun: Sentence-BERT (SBERT) - javob va etalon orasidagi cosine similarity. UzbekBERT / mBERT - o‘zbek tilida semantik solishtirish uchun.

Baholash mezonlari:

WER/CER (aniqlik uchun). Semantic Similarity Score (ma’no uchun). Final score = $\alpha \cdot$WER + $\beta \cdot$Semantic similaritylar ishlatiladi.

ISSN 3030-3907

Xulosa: O‘rganish natijalari shuni ko‘rsatadiki, nutqni matnga aylantirish (ASR) texnologiyalarini test savollari javoblarini avtomatik baholash jarayoniga qo‘llash katta imkoniyatlarga ega. Klassik baholash ko‘rsatkichlari (WER va CER) orqali modelning texnik aniqligini o‘lchasa, semantik o‘xshashlik metrikalari yordamida javobning mazmuniy mosligi baholanishi mumkin. Ushbu ikki yondashuvni birgalikda qo‘llash natijasida foydalanuvchining og‘zaki javobini adolatli va aniq baholash imkoniyati yaratiladi. O‘rganishdan kelib chiqqan holda, bunday tizimlar ta’lim jarayonida avtomatlashtirilgan baholashni samarali tashkil etishda, shuningdek, talabalarning bilimini tezkor va xolisona tekshirishda muhim ahamiyat kasb etadi.

Adabiyotlar

1. Frederick Jelinek. “Statistical Methods for Speech Recognition”, Cambridge 1997.
2. Lawrence Rabiner va Biing-Hwang Juang. nutq tanish texnologiyalarini. “*A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*”, 1989.
3. James L. Flanagan. Bell Labs. *Speech Analysis, Synthesis and Perception* 1972.
4. Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, Yoshua Bengio. *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling* 2014-yil.
5. Zhiyuan Tang, Ying Shi, Dong Wang, Yang Feng, Shiyue Zhang. *Memory Visualization for Gated Recurrent Neural Networks in Speech Recognition* 2017.
6. Andrew Ng va Baidu Research. Deep Speech: Scaling Up End-to-End Speech Recognition.” *arXiv preprint arXiv 2014.*
7. Muhammadjon Musayev, Saida Musaxo‘jayeva, Ilyos Hujayorov, Yerbo‘lat Xasanov Mannon Ochilov va Huseyin Atakan Vorillar. *An Open-Source Uzbek Speech Corpus and Initial Speech Recognition* 2021.
8. Abduqahhor Saidov, Sardorbek Yusupov, Anvar Narzullaev, Alisher Matyakubov, Saidmurod Radjabov, Hafiz Ibrahim, Jasur Temirov. *A 60 Hour Uzbek Read Speech Corpus with Punctuation, Casing, and Context 2024.*
9. Mukhamadiyev, A.; Mukhiddinov, M.; Khujayarov, I.; Ochilov, M.; Cho,J. *Development of Language Models for Continuous Uzbek Speech Recognition System* 2023.
10. Salayeva Maxliyo, Kuriyozov Elmurod, Salayev Ulug‘beklarning *Uzbek automatic speech recognition models using deep learning techniques* 2023.
11. Thennal D K, Jesin James, Deepa P Gopinath, Muhammed Ashraf Klarning *Advocating Character Error Rate for Multilingual ASR Evaluation*
12. Hila Gonen1 Shauli Ravfogen, Yanai Elazar, Yoav Goldberg, larning “*It’s not Greek to mBERT: Inducing Word-Level Translations from Multilingual BERT* November 2020.
13. Mikhail Evtikhiev, Egor Bogomolov, Yaroslav Sokolov, Timofey Bryksin. *Out of the BLEU: how should we assess quality of the Code Generation models?* May 2023.

ISSN 3030-3907

14. Zhanibek Kozhirbayev. *Kazakh Speech Recognition: Wav2vec2.0 vs. Whisper January 2023.*
15. Ibragimova S.N. Turayev B.Sh. Abdullayeva M.I.larning Solving the problems ofnormalization of non-standardwords in the text of the uzbeklanguage November r 2023.

ISSN 3030-3907