

UDK 004.093

OVOZLI BUYRUQLARNI TANIB OLIISH USULLARI*Nurimov P.B.¹*

¹ Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti
Nukus filiali, Nukus, O'zbekiston
paranur87@gmail.com

Annotatsiya. *Ushbu maqola ovozli buyruqlarni tanib olish masalasi uchun ovozli buyruqlar modelini qurish uchun SVM va GMM usullaridan foydalanish ko'rib chiqilgan. Ovozli buyruqlarni tanib olish texnologiyasi keng ko'lamlilovalarda mavjud, jumladan ovozli yordamchilar, boshqaruv tizimlari va boshqalar. Maqolada ovozli buyruqlarni modellashtirish va tasniflash uchun SVM va GMM usullaridan foydalanishga asoslangan yondashuv taqdim etilgan. Maqolada qoraqalpoq tilidagi ovozli buyruqlar ma'lumotlar to'plami bo'yicha tajribalar taqdim etilgan. Tanib olish aniqligi va uning belgilar to'plamini ajratib olish algoritmlariga asoslangan umumiy samaradorlik nuqtai nazaridan SVM va GMM ko'rsatkichlarining qiyosiy tahlillari o'tkazildi. Eksperimental natijalar shuni ko'rsatadiki, SVM va GMM usullari ham ovozli buyruqlarni tanib olishda yuqori aniqlikga erishishi mumkin. Xulosa qilib aytganda, ushbu maqolada ovozli buyruqlarni tanib olish masalasida SVM va GMM usullarining taqqoslanishi keltirilgan.*

Kalit so'zlar: *ovoz signali, ovozli buyruqlarni tanib olish, belgilar to'plamini ajratish, tayanch vektorlar mashinalari, Gaus aralashma modeli.*

I. KIRISH

Bugungi kunda ovozli buyruqni tanib olish nutqni qayta ishlash va sun'iy intellekt sohasida jadal rivojlanayotgan soha hisoblanadi. U turli sohalarda, jumladan, ovozli yordamchilar, boshqaruv tizimlari va boshqa sohalarda keng ko'lamlilovalarga ega. Siri, Google Assistant va Amazon Alexa kabi ovozli yordamchilar kundalik hayotimizning ajralmas qismiga aylanib bormoqda [1, 2]. Ular foydalanuvchilarga ovozli buyruqlar orqali qurilmalar va ilovalar bilan o'zaro aloqada bo'lish imkonini beradi, bu esa jarayonlarni yanada qulay va samarali bo'lishini ta'minlaydi. Ovozli buyruqni tanib olish ushbu ovozli yordamchilarning ishlashidagi asosiy komponenti bo'lib, ularga foydalanuvchi tomonidan aytilgan buyruqlarni tushunish va bajarish imkonini beradi.

Bundan tashqari, ovozli buyruqni tanib olish aqlli uy tizimlari, avtomobil tizimlari, xavfsizlik tizimlari va boshqalar kabi boshqaruv tizimlarida ham keng qo'llaniladi. Foydalanuvchilar turli funksiyalar va qurilmalarni oddiygina ovozi yordamida boshqarishi mumkin, bu esa texnologiya bilan o'zaro aloqani yanada tabiiy va qulayroq qilish imkonini beradi.

Ovozli buyruqni tanib olish mashinali o'qitish (Machine Learning) algoritmlari va audio ma'lumotlarning statistik tahliliga asoslangan. SVM (Support Vector Machine) va GMM (Gaussian Mixture Model) modellari bu sohada qo'llaniladigan ikkita mashhur usullaridan hisoblanadi [2].

Ushbu maqolada biz SVM va GMM modellarining qoraqalpoq tilidagi ovozli buyruqlarni tanib olish vazifasidagi samaradorligini tahlil qilamiz va taqqoslaymiz. Biz har bir usul qanday ishlashini, ularning afzalliklari va cheklovlarini ko'rib chiqamiz va turli ko'rsatkichlar asosida qiyosiy tahlil qilamiz.

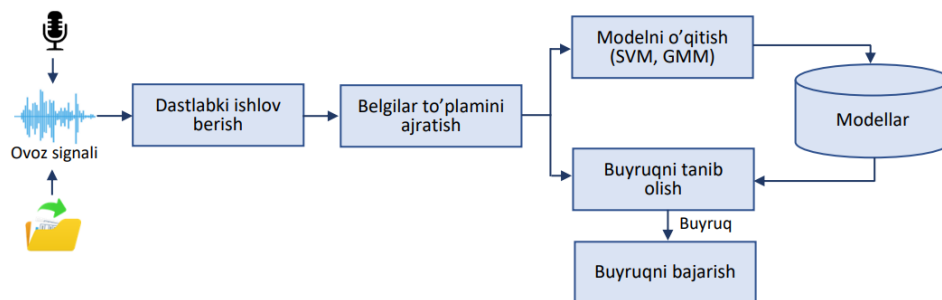
II. ASOSIY QISM

Ovozli buyruqlarni tanib olishning dolzarbligi ovozli boshqaruv va avtomatlashtirishga talab ortib borayotganligi sababli o'sishda davom etmoqda. Ovozli boshqaruv tizimlari smartfonlar, aqlli uylar, avtomobillar va boshqa qurilmalarda tobora keng tarqalgan. Ular foydalanuvchilarga klaviatura yoki ekrandan foydalanmasdan qurilmalarni boshqarish imkonini berib, qulaylik va samaradorlikni taqdim etadi.

Bundan tashqari, ovozli buyruqni tanib olish tibbiyot sohasida ham katta imkoniyatlarga ega. U diktant va transkripsiya tizimlarini ishlab chiqish, tibbiy yozuvlarni avtomatlashtirish va nogironlarga yordam berish uchun ishlatilishi mumkin. Avtomobil sanoatida ovozli buyruqni tanib olish tizimlari navigatsiya, multimedia va iqlim nazorati kabi turli xil avtomobil funktsiyalarini boshqarish uchun ishlatilishi mumkin. Bu avtomobilni boshqarish xavfsizligi va qulayligini oshirishi mumkin. Ovozli buyruqlarni tanib olish tizimlari ta'lim sohasida ham qo'llanilishi mumkin, bu nutq yoki eshitish qobiliyati zaif odamlarga muloqot qilish va o'rganishga yordam beradi.

Umuman olganda, ovozli buyruqlarni tanib olish tizimlari keng ko'lamli ilovalarga ega va qurilmalar va vazifalarni yanada qulay va samarali boshqarish talab qilinadigan bugungi dunyoda dolzarb bo'lib qolmoqda.

Ovozli buyruqni tanib olish - ovozli signalni kompyuter yoki boshqa qurilma tushuna oladigan matn shakliga aylantirish jarayoni hisoblanadi. Bu foydalanuvchilarga tugmalarni terish yoki bosish o'rniga o'zlarining ovozli buyruqlari yordamida qurilmalar bilan o'zaro muloqot qilish imkonini beruvchi texnologiyadir [2]. Ovozli buyruqni tanib olish ovozning balandligi, intonatsiyasi, nutq tezligi va so'z talaffuzi kabi akustik xususiyatlarini tahlil qilish va talqin qilishga asoslangan. Ovozni



1-rasm. Ovozli buyruqni tanib olish tizimi tuzilishi.

1. Ovozni yozib olish: Bu komponent foydalanuvchining ovozli signalini yozib olish uchun javobgardir. Bu mikrofon yoki boshqa audio qurilmalar yordamida amalga oshirilishi mumkin [2, 3].

2. Ovoz signaliga dastlabki ishlov berish: Ushbu komponentda ovozli signal shovqinni filtrlash, signalni kuchaytirish va signal sifatini yaxshilash uchun boshqa usullarni o'z ichiga olgan oldindan ishlov berish o'tkaziladi [2, 4, 5].

3. Belgilar to'plamini ajratib olish: Ushbu komponentda ovozli signal tanib olish uchun ishlatilishi mumkin bo'lgan belgilar to'plamiga aylantiriladi. Bu chastotalar va amplitudalar kabi spektral xususiyatlarni, shuningdek, tovushlar va pauzalar davomiyligi kabi vaqtinchalik xususiyatlarni o'z ichiga oladi [2, 6].

4. Modellashtirish va o'qitish: Ushbu komponent ovozli buyruqlarni taniy oladigan modellarni yaratish uchun mashinaviy o'qitish algoritm-laridan foydalanadi. Modellar turli xil ovozli buyruqlarni tanib olishni o'rganish uchun oldindan yozib olingan ovozli ma'lumotlarga o'qitiladi.

5. Buyruqni tanib olish: Ushbu komponentda modellar ovozli buyruqlarni tanib olish va ularning ma'nosini izohlash uchun ishlatiladi. Natijada ovozli signal kompyuter yoki boshqa qurilma tushunadigan va tegishli qaror qabul qilishi

tanib olish tizimlari foydalanuvchining ovozli buyruqlarini tanib olish uchun odatda katta hajmdagi ovozli ma'lumotlarga o'qitilgan algoritm va modellardan foydalanadi.

Ovozli buyruq tanib olingandan so'ng, tizim ilovalarni ishga tushirish, xabarlarini yuborish, internetda ma'lumot qidirish va boshqa ko'p ishlarni bajarishi mumkin. Ovozli buyruqni tanib olish tobora ommalashib bormoqda va smart-fonlar, aqlli uylar, avtomobillar va boshqa aqlli qurilmalar kabi qurilmalarda keng qo'llanilmoqda.

Ovozli buyruqni tanib olish tizimi odatda quyidagi komponentlardan iborat va uning tuzilishi quyidagi 1-rasmga keltirilgan:

mumkin bo'lgan ovozli buyruqning matn shakliga aylantiriladi.

6. Buyruqning bajarilishi: Ovozli buyruqni tanib olgandan so'ng, tizim dasturni ishga tushirish, xabar yuborish yoki internetda qidirish kabi boshqada tegishli amallarni bajarishi mumkin.

Bu komponentlarning barchasi foydalanuvchilarning ovozli buyruqlarini samarali va aniq tanib olish uchun birgalikda ishlaydi.

Ovoz signalidan belgilar to'plamini shakllantirish

Tanib olish tizimlari samaradorligi belgilar to'plami qanday tanlanganligiga bog'liq. Boshlang'ich belgilar fazosi qanchalik yaxshi tanlansa, tanib olish sifati shunchalik yuqori bo'ladi. Ovozli buyruqlarni tanib olish masalasi ham nutq signalidan belgilar to'plamini shakllantirishdan boshlanadi.

Vaqt bo'yicha aniqlikni olish uchun ovoz signalini bir-birining ustiga tushadigan kadrlarga ajratish talab etiladi va u oyna funktsiyasi yordamida ajratiladi, odatda Xemming oynasi funktsiyasi qo'llaniladi [2].

$$w[n] = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{N}\right), & 0 \leq n \leq N \\ 0, & \text{boshqa hollarda} \end{cases} \quad (1)$$

Kadrlar odatda 20-30 ms uzunlikda bo'ladi, shuning uchun audio butun kadr davomida o'zgarmaydi. Odatda 25 ms kadr davomiyligi va 16 kGts diskretlash chastotasi audio faylning 400 ta namunasini beradi.

Diskret Fure almashtirishi (DFA) signalda qanday chastotalar mavjudligini tavsiflaydi, lekin ularni o'z vaqtida lokallashtirmaydi. Nutq signali uchun qisqa muddatli Fure almashtirishi [2] ifoda orqali amalga oshiriladi:

$$H(n, k) = \sum_{n=1}^N x(n) w(n) e^{-i2\pi kn} \quad (2)$$

bu yerda $w(n)$ –Xemming oynasi.

Keyingi qadam har bir kadr uchun energiyani hisoblashdir. Bu chiqish signalini quvvat spektrining vaqt-chastota sohasiga solishtirish orqali hisoblanadi. Quvvat spektri chastota komponentlarini quyidagi ifoda bo'yicha kvadratga oshirish yo'li bilan quyidagicha hisoblanadi.

$$S(n, k) = |H(n, k)|^2 \quad (3)$$

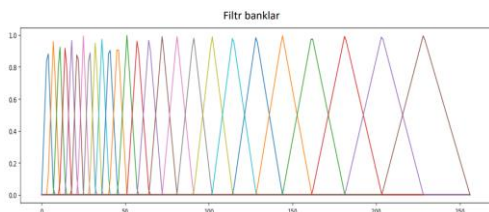
Avtomatik nutqni identifikatsiyalash tizimlarida (4) formulada ko'rsatilganidek, logarifmik spektrogramma hosil qilish uchun har bir vektorning logarifmidan foydalaniladi:

$$S(n, k) = 20 \log_{10} |H(n, k)|^2 \quad (4)$$

Ovozli buyruqlarni tanib olish tizimlarida mel masshtablangan filtr banklaridan foydalanilganda odatda 26 dan 60 gacha filtrlar banklari yaratiladi [2, 6].

$$M(f) = 1125 \ln \left(1 + \frac{f}{700} \right) \quad (5)$$

Filtrlar uchburchak shaklida bo'lib, har bir filtr chap qo'shni filtrning markazidan boshlanadi. Keyin filtr chiziqli ravishda 1 ga ko'tariladi va keyin yana chiziqli ravishda 0 ga tushadi, bu esa 2-rasmda ko'rsatilgan boshqa qo'shni filtrning markazi bo'ladi. Uchburchak filtrlarni quyidagi formuladagi kabi modellashtirish mumkin [6].



2-rasm. Mel-filtr banklari. Filtrlar soni 26 ta

Filtrlar hosil qilingandan so'ng, mel masshtablangan filtr banklari va quyida formulada ko'rsatilgan spektrogramma-ning skalyar ko'paytmasi logarifmini olish yo'li bilan hisoblanadi.

$$MSFB = 20 \log_{10} (S(n, k) \cdot F_m(k))$$

Mel chastotali kepstral koeffitsientlari (MFCC) ko'pincha nutqni avtomatik tanib olishda qo'llaniladi [2, 6] va diskret kosinus almashtirishini asosida filtr banklarini korrelyatsiyalash orqali hosil qilinadi. Mel-masshtabli filtr banklar yuqori darajada korrelyatsiyaga ega, chunki uchburchak filtrlar bir-birini qisman qoplagan bo'ladi. Filtrlar quyida formulada berilgan diskret kosinus almashtirishini qo'llash orqali dekorrelyatsiyalanadi:

$$S_{MFCC_i} = \sum_{k=1}^M X_k \cos \left[i \left(k - \frac{1}{2} \right) \frac{\pi}{M} \right]. \quad (6)$$

Diskret kosinus almashtirishini har bir filtr banki uchun qo'llash Mel-chastotali kepstral koeffitsientlarini (MFCC) beradi.

Nutq signaliga ishlov berishning boshqa sohalarida bo'lgani kabi, delta-belgilar deb ataladigan spektral farqlar odatda asosiy belgilar to'plamiga qo'shiladi. Bundan tashqari, odatda delta-belgilarining farqlari qo'shiladi va ular delta-delta belgilar deb ataladi. Delta Δ va delta-delta $\Delta\Delta$ belgilarini hisoblash formulasi quyida keltirilgan.

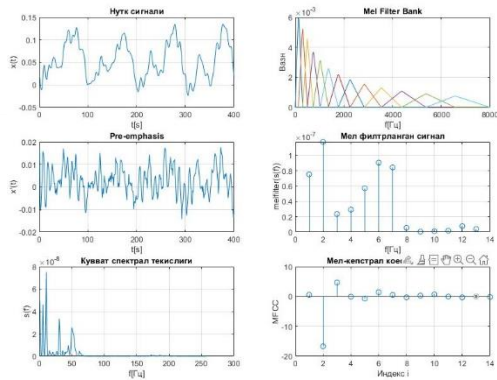
$$\Delta_i = MFCC_{i+1} - MFCC_i; \Delta\Delta_i = \Delta_{i+1} - \Delta_i; \quad (7)$$

Ovoz signalidan Mel-chastotali kepstral koeffitsientlari (MFCC)ni ajratib olish jarayoni quyidagi 3-rasmda keltirilgan.

Modellashtirish usullari

Ovoz signalidan belgilar to'plamini shakllantirib olgandan so'ng, ovozli buyruqlarni tanib olish uchun ularni ma'lumotlar bazasida saqlanadigan mos etalon belgilar to'plami bilan solishtirish zarur. Olingan belgilarni solishtirish qanday amalga oshiriladi va birinchi navbatda belgilar to'plamidan murakkabroq model yaratiladimi yoki yo'qmi. Buni ushbu bosqichda tanlash talab etiladi.

Har bir shaxs ovozi uchun foydalaniladigan belgilar to'plami soniga va ushbu belgilar to'plamini qanday modellashtirishni hohlayotganligiga ko'ra, belgilarlarni moslashtirish va modellashtirishni bir nechta turlaridan foydalanish mumkin. Quyida belgilar to'plamini modellashtirishning eng keng tarqalgan usullari keltirib o'tilgan.



3-рasm. Mel-chastrotali kepstral koeffitsientlari (MFCC)ni ajratib olish jarayoni

SVM (Support Vector Machines - Tayanch vektorlar mashinalari) - bu tasniflash va regressiya muammolari uchun ishlatiladigan mashinaviy o'qitish usuli. SVM ning asosiy printsipti ko'p o'lchovli fazoda turli sinflardan ma'lumotlarni ajratib turadigan optimal giperplanni topishdir [1, 7, 8].

SVM ning ishlash printsiptini quyidagicha tushuntirish mumkin:

Tayanch vektorlarini qurish: SVM ajratuvchi giperplanga eng yaqin ma'lumotlar nuqtalarini ifodalovchi tayanch vektorlarini qidiradi. Tayanch vektorlar giperplanning holati va shaklini aniqlashda asosiy rol o'ynaydi.

Optimal giperplanni topish: SVM ma'lumotlar sinflari orasidagi bo'shliqni maksimal darajada oshiradigan giperplanni topishga intiladi. Bo'shliq giperplan va har bir sinfnig eng yaqin tayanch vektorlari orasidagi masofani ifodalaydi. SVM ning maqsadi modelni eng yaxshi umumlashtirish qobiliyatiga erishish uchun eng katta bo'shliqqa ega giperplanni topishdir.

Ma'lumotlarni turlantirish: Agar ma'lumotlarni chiziqli ravishda ajratish mumkin bo'lmasa, SVM ma'lumotlarni chiziqli ravishda ajratish mumkin bo'lgan yuqori o'lchamga aylantirish imkonini beruvchi yadro hiylasidan foydalanishi mumkin. Yadrolarga misol sifatida polinom yadro, Gauss yadrosi (RBF) va sigmoid yadro kiradi.

SVM keng doiradagi tasniflash muammolarida, jumladan timsollarni tanib olish, nutqni tanib olish, matnni tahlil qilish va boshqalarda qo'llaniladi. Yuqori o'lchamli ma'lumotlar bilan ishlash va sinflarni samarali ajratish qobiliyati tufayli SVM yaxshi umumlashtirish qobiliyatiga ega va yuqori tasniflash aniqligini ta'minlaydi.

Biroq, SVM ham ba'zi cheklavlarga ega, jumladan, giperparametr tanlashga sezgirlik va katta ma'lumotlar to'plamini o'qitishdagi qiyinchiliklar. Biroq, ushbu cheklavlarni tegishli modelni sozlash va samarali optimallashtirish

algoritmalaridan foydalanish orqali bartaraf etish mumkin.

GMM (Gaussian Mixture Model - Gauss aralashmasi modeli) - bu ma'lumotlarning ehtimollik taqsimotini modellashtirish uchun foydalaniladigan statistik model hisoblanadi. U bir nechta Gauss taqsimotlarining kombinatsiyasiga asoslangan bo'lib, ularning har biri aralashmaning tarkibiy qismini ifodalaydi [9, 10].

GMM ning ishlash printsipti Gauss taqsimotlarining parametrlarini va ularning aralashmadagi nisbatlarini baholashdan iborat. Bu maksimal ehtimollik usuli yoki Expectation-Maximization (EM) usuli yordamida amalga oshiriladi. EM algoritmi ikki bosqichdan iborat: E-bosqich, bunda har bir aralashma komponentiga tegishli har bir kuzatishning aposterior ehtimolliklari hisoblab chiqiladi va M-bosqichida bunda aralashma komponentlarining parametrlari shu keyingi ehtimollar asosida yangilanadi.

GMM ma'lumotlarni klasterlash, ma'lumotlarning ehtimollik zichligini modellashtirish va ma'lumotlarni tasniflash kabi turli xil muammolarni hal qilishda qo'llanilishi mumkin.

Nutq ma'lumotlarini modellashtirish kontekstida GMM nutqni tanib olish uchun ishlatiladi. U nutq tovushlarini modellashtirish va nutq signalidagi fonemalar yoki so'zlarning eng ehtimoliy ketma-ketligini aniqlash uchun ishlatilishi mumkin. GMM nutqni tanib olishning yuqori aniqligiga erishish uchun Hidden Markov Models (HMM) kabi boshqa algoritmalar bilan birgalikda ishlatiladi.

Umuman olganda, GMM ma'lumotlarning ehtimollik taqsimotini modellashtirish uchun kuchli vosita bo'lib, timsollarni tanib olish, nutqni qayta ishlash, kompyuter ko'rishi va ma'lumotlarni tahlil qilish va modellashtirish talab qilinadigan boshqa sohalarda keng qo'llaniladi.

III. TAJRIBAVIY TADQIQOTLAR

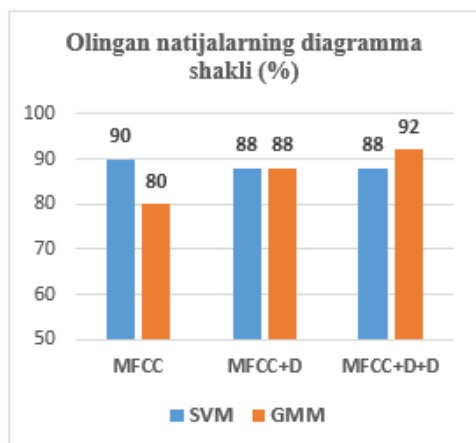
Mazkur tadqiqot ishida ovozli buyruqni tanib olish masalasini yechish uchun 5 kishi va har bir kishi 15 tadan 0 dan 9 raqamlarni aytgan ovoz ma'lumotlari asosida *python* dasturlash tilidan foydalanib tajribaviy tadqiqot o'tkazildi. *Python* dasturlash tilida *sklearn* [11] va *python_speech_features* [12] modullaridan foydalanildi.

Ovoz signallaridan belgilar to'plamini ajratib olish uchun MFCC algoritmidan foydalanilgan bo'lib, har bir ovozli buyruqning modelini qurish uchun yuqorida aytib o'tilgan SVM va GMM usullari qo'llanilgan. Modellarini solishtirish orqali qaror qabul qilish k-yaqin qo'shnilar usuli orqali

amalga oshirilgan. Olingan natijalar quyidagi 1-jadvalda keltirilgan.

1-jadval. Olingan natijalar

Belgilar tuplamini ajratish	SVM	GMM
MFCC (20 ta)	90,0%	80,0%
MFCC+Δ (40 ta)	88,0%	88,0%
MFCC+Δ+Δ (60 ta)	88,0%	92,0%



4-rasm. Olingan natijalar

Bu olingan natijalardan ko‘rish mumkin, SVM usuli ovoz signallaridan MFCC (20 ta) belgilar to‘plami ajratilganda GMM usuliga nisbatan yaqshiroq natija ko‘rsatdi, biroq, MFCC (60 ta) belgilar to‘plami ajratilganda GMM usuli SVM usuliga nisbatan yaqshiroq natija ko‘rsatdi va eng yaqshi 92% natijaga erishildi.

IV. XULOSA

O‘tkazilgan tadqiqot natijalaridan shuni xulosa qilish mumkin, agarda ovoz signalidan kamroq belgilar to‘plami ajratilganda SVM usuli samaraliroq va ko‘proq belgilar to‘plami ajratilganda GMM usuli yaqshi natija ko‘rsatadi. Demak, ovozli buyruqlarni tanib olishda MFCC (60 ta) belgilar to‘plamini ajratish algoritmi va ovozli buyruqlar modelini qurishda GMM usuli samarali hisoblanadi.

Xulosa qilib aytadigan bo‘lsak, ushbu ish natijalari SVM va GMM ning ovozli buyruqni tanib olish vazifasida samaradorligini tasdiqlaydi. Yondashuvni tanlashda aniq muammo va dastur talablarini hisobga olinishi kerak.

SVM va GMM asosida ovozli buyruqni tanib olish turli sohalarda, jumladan ovozli yordamchilar, boshqaruv tizimlari, nutqni avtomatik tanib olish va boshqa sohalarda qo‘llanilishi mumkin.

Ushbu usullar qoraqalpoq tilidagi ovozli interfeyslarni rivojlantirishga yordam beradi va odamlarning kompyuter tizimlari bilan o‘zaro aloqasi qulayligi va samaradorligini oshiradi.

ADABIYOTLAR

- [1] Huang, C. L., & Hsiao, Y. T., Voice command recognition using support vector machines with kernel-based feature extraction. *Applied Soft Computing*, 36, 177-184., 2015
- [2] H. Beigi. *Fundamentals of speaker recognition*. Springer US, 2011.
- [3] W. M. Campbell, D. E. Sturim, D. A. Reynolds, Support vector machines using GMM supervectors for speaker verification, *IEEE signal processing letters* 13 (5) (2006) 308–311.
- [4] Рабинер Л., Шафер Р. Цифровая обработка речевых сигналов. – М.: Радио и связь, 1981. – 496 с.
- [5] Маматов Н.С., Нуримов П.Б., Самижонов А.Н. Нутқ сигналларида овоз фаоллигини аниқлаш алгоритмлари. «Ахборот коммуникация технологиялари ва дастурий таъминот яратишда инновацион ғоялар» Республика илмий-техник конференцияси 17-18 май 2021 йил.
- [6] П.Б. Нуримов, А.Н. Самижонов, Ш.И. Фозилов, З.М. Сулаймонов Шахсни нутқи асосида таниб олишда белгилар фазосини шакллантириш. *Информатика ва энергетика муаммолари Ўзбекистон Журнали*, №4, 2020
- [7] Burges C., “A tutorial on support vector machines for pattern recognition”, In “Data Mining and Knowledge Discovery”. Kluwer Academic Publishers, Boston, 1998, (Volume 2).
- [8] Osuna E., Freund R., and Girosi F., “Support Vector Machines: Training and Applications”, A.I. Memo No. 1602, Artificial Intelligence Laboratory, MIT, 1997.
- [9] Desai D, Joshi M., Speaker Recognition Using MFCC and Hybrid Model of VQ and GMM. *Recent Advances in Intelligent Informatics* 235: 53-63.
- [10] W. M. Campbell, D. E. Sturim, D. A. Reynolds, Support vector machines using GMM supervectors for speaker verification, *IEEE signal processing letters* 13 (5) (2006) 308–311.
- [11] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, ..., Scikit-learn: Machine Learning in Python, <https://arxiv.org/abs/1201.0490>

[12] James Lyons et al. (2020, January 14). jameslyons/python_speech_features:

release v0.6.1 (Version 0.6.1). Zenodo. <http://doi.org/10.5281/zenodo.3607820>

Поступила в редакцию 24.09.2023

Citation: Nurimov P.B. (2023). Ovozli buyruqlarni tanib olish usullari. Raqamli texnologiyalarning nazariy va amaliy masalalari xalqaro jurnali. 4(6). – B. 47-52.

METHODS OF VOICE COMMANDS RECOGNITION

Nurimov P.B.¹

¹ Nukus branch of Tashkent University of information technologies named after Muhammad al-Khwarizmi, Nukus, Uzbekistan
paranur87@gmail.com

Abstract. This article discusses the use of SVM and GMM methods to build voice command models to solve the problem of voice command recognition. Voice recognition technology is available in a wide range of applications, including voice assistants, control systems and more. The article presents an approach based on the use of SVM and GMM methods for modeling and classifying voice commands. The article presents experiments on a dataset of Karakalpak voice commands. A comparative analysis of the performance of SVM and GMM was conducted in terms of recognition accuracy and overall performance based on feature set extraction algorithms. Experimental results show that SVM and GMM methods can achieve high accuracy in recognizing voice commands. Therefore, this paper presents a comparison between SVM and GMM methods for voice command recognition.

Keywords: voice signal, recognition voice commands, feature extraction, Support vector machines, Gaussian mixture model.

МЕТОДЫ РАСПОЗНАВАНИЯ ГОЛОСОВЫХ КОМАНД

Нуримов П.Б.¹

¹ Нукусский филиал Ташкентского университета информационных технологий имени Мухаммада ал-Хорезми, Нукус, Узбекистан
paranur87@gmail.com

Аннотация. В данной статье рассматривается использование методов SVM и GMM для построения моделей голосовых команд для решения задачи распознавания голосовых команд. Технология распознавания голоса доступна в широком спектре приложений, включая голосовых помощников, системы управления и многое другое. В статье представлен подход, основанный на использовании методов SVM и GMM для моделирования и классификации голосовых команд. В статье представлены эксперименты на наборе данных каракалпакских голосовых команд. Был проведен сравнительный анализ производительности SVM и GMM с точки зрения точности распознавания и общей производительности на основе алгоритмов извлечения набора функций. Результаты экспериментов показывают, что методы SVM и GMM могут обеспечить высокую точность распознавания голосовых команд. Таким образом, в этой статье представлено сравнение методов SVM и GMM для распознавания голосовых команд.

Ключевые слова: голосовой сигнал, распознавание голосовых команд, извлечение признаков, машины опорных векторов, модель Гауссовой смеси.